

RESEARCH ARTICLE

RNA sequencing of intestinal mucosa reveals novel pathways functionally linked to celiac disease pathogenesis

Maureen M. Leonard^{1,2,3}, Yu Bai⁴, Gloria Serena^{2,3}, Kourtney P. Nickerson², Stephanie Camhi², Craig Sturgeon⁵, Shu Yan², Maria R. Fiorentino², Aubrey Katz¹, Barbara Nath⁶, James Richter⁶, Matthew Sleeman⁴, Cagan Gurer⁴, Alessio Fasano^{1,2,3*}

1 Mass General Hospital for Children and Division of Pediatric Gastroenterology and Nutrition, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Center for Celiac Research and Treatment, Mucosal Immunology and Biology Research Center, Boston, Massachusetts, United States of America, **3** Celiac Research Program, Harvard Medical School, Boston, Massachusetts, United States of America, **4** Regeneron Pharmaceuticals, Tarrytown, New York, United States of America, **5** Graduate Program in Life Sciences, University of Maryland, Baltimore, Maryland, United States of America, **6** Department of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts, United States of America

* afasano@mgh.harvard.edu



OPEN ACCESS

Citation: Leonard MM, Bai Y, Serena G, Nickerson KP, Camhi S, Sturgeon C, et al. (2019) RNA sequencing of intestinal mucosa reveals novel pathways functionally linked to celiac disease pathogenesis. *PLoS ONE* 14(4): e0215132. <https://doi.org/10.1371/journal.pone.0215132>

Editor: Karol Sestak, Tulane University, UNITED STATES

Received: July 17, 2018

Accepted: March 27, 2019

Published: April 18, 2019

Copyright: © 2019 Leonard et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data has been deposited to SRA under the accession number PRJNA528755.

Funding: This work was supported partially by The National Institutes of Health (NIH) grants DK048373 and DK104344 to A. F. and by Regeneron Pharmaceuticals Inc.

Competing interests: Y.B., C.G., M.S. are employees of Regeneron Pharmaceuticals Inc. A.F. is a consultant for AbbVie, stock holder at Alba Therapeutics, serves as a consultant for Innovate

Abstract

Background & aims

The early steps in the pathophysiology of celiac disease (CD) leading to loss of tolerance to gluten are poorly described. Our aim was to use RNA sequencing of duodenal biopsies in patients with active CD, CD in remission, and non-CD controls to gain insight into CD pathophysiology, identify additional genetic signatures linked to CD, and possibly uncover targets for future therapeutic agents.

Methods

We performed whole transcriptome shotgun sequencing of intestinal biopsies in subjects with active and remission CD and non-CD controls. We also performed functional pathway analysis of differentially expressed genes to identify statistically significant pathways that are up or down regulated in subjects with active CD compared to remission CD.

Results

We identified the upregulation of novel genes including IL12R, ITGAM and IGSF4 involved in the immune response machinery and cell adhesion process in the mucosa of subjects with active CD compared to those in remission. We identified a unique signature of genes, related to innate immunity, perturbed exclusively in CD irrespective of disease status. Finally, we highlight novel pathways of interest that may contribute to the early steps of CD pathogenesis and its comorbidities such as the spliceosome, pathways related to the innate immune response, and pathways related to autoimmunity.

Biopharmaceuticals, is a consultant for Ubiome, has a research agreement with Takeda Pharmaceuticals, and has a speaker agreement with Mead Johnson Nutrition. MML performs sponsored research with Takada Pharmaceuticals and Glutenostics. Other authors have declared that no conflict of interest exists. This does not alter our adherence to all PLOS ONE policies on sharing data and materials.

Abbreviations: CD, Celiac Disease; CLDN, Claudin; EGF, Epidermal Growth Factor; GPCR, G protein coupled receptor; GSEA, Gene Set Enrichment Analysis; HLA, Human Leukocyte Antigen; IEL, Intraepithelial Lymphocyte; ICAMs, Intercellular Adhesion Molecules; TCR, T cell receptor; anti-tTG2, Tissue Transglutaminase; T1D, Type 1 diabetes.

Conclusions

Our study confirmed previous findings based on GWAS and immunological studies pertinent to CD pathogenesis and describes novel genes and pathways that with further validation may be found to contribute to the early steps in the pathogenesis of CD, ongoing inflammation, and comorbidities associated with CD.

Introduction

Celiac disease (CD) is a chronic systemic autoimmune disease that occurs globally in genetically predisposed individuals in response to ingestion of gluten-containing grains [1]. While the pathognomonic damage occurs in the small intestine, clinical manifestations are varied and include both intestinal and extra-intestinal symptoms[2]. CD is a unique autoimmune disease in that there is a strong genetic component, notably human leukocyte antigen (HLA) DQ2 and DQ8[3]. During active disease, the production of auto-antibodies against tissue transglutaminase (anti-tTG2) can be measured in the blood to indicate the presence of disease. While HLA DQ2 and/or DQ8 and gluten ingestion are necessary to develop CD, they are not sufficient as most individuals with compatible HLA do not develop CD. Furthermore, CD can develop at any age[4] but what leads ultimately to this loss of tolerance to gluten and development of intestinal inflammation is unknown.

Despite the well described pathogenic role of the adaptive immune response in CD, a complete understanding of the pathophysiology of CD, particularly concerning the early steps leading to loss of tolerance to gluten, are poorly defined[2]. Additionally, biomarkers confirming mucosal recovery for patients with CD on a gluten free diet are lacking. For this reason, we sought to use whole transcriptome shotgun sequencing of duodenal biopsies for an analysis to gain more insight on CD pathophysiology and possibly identify biomarkers suggestive of active disease and mucosal recovery, including pathways which may lead to the discovery of future targets for therapeutic interventions. Our work provides a base from which further investigation may validate and mechanistically link these findings to previously unknown aspects of the pathogenesis of CD.

More than 57 non-HLA associated genetic loci contributing to CD have been identified, which together with the HLA contribution, explains less than 50% of CD heritability[5–7]. Molecular analysis of the pathogenic mechanisms in CD have assessed gene expression profiles using small intestinal biopsies and specific cell types[8–15]. Previous work has utilized RNA sequencing of whole duodenal biopsies to investigate expression of genes associated with tight junctions in patients with active CD, remission CD, and control patients [15]. In this study, we performed whole transcriptome shotgun sequencing in duodenal biopsies in subjects with active CD compared to those with CD in remission and compared these subjects to individuals with CD-compatible HLA genotypes who are not affected by CD. This approach allowed us to gain insight regarding specific pathways possibly involved in the cascade of events upstream to the adaptive immune response that contribute to the loss of tolerance to gluten. Furthermore, through a focus on molecular signatures that are conserved in patients with CD regardless of the disease activity and those signatures that differentiate active CD from CD in remission, we shed insight on constitutively-involved pathways that may be at play not only in CD but also in other chronic inflammatory diseases.

Results

Comparison of gene expression in patients with active CD, CD in remission, and non-CD control patients

A heatmap comparing the three study groups shows clear gene expression differences (Fig 1A). The profound differences between groups in this analysis is testimonial of the strict adherence to subject enrollment criteria, quality of sample collection, and analysis. Our analysis identified 945 genes significantly different when comparing subjects with active CD to non-CD control subjects (S1 Fig). We identified 290 genes significantly different between subjects with CD in remission and non-CD control subjects. We identified 538 genes significantly different between subjects with active CD and CD in remission. Finally, 117 genes were common to subjects with active and remission CD (S2 Fig). A venn diagram (Fig 1B) shows the unique signature of genes perturbed exclusively in active CD (471) and those perturbed exclusively in remission CD (134). There are an additional 117 genes, mainly involved in innate immune response, that are changed constitutively in subjects with CD irrespective of the disease status (active or remission). Due to the large number of genes differentially expressed between subjects with active CD and remission CD, we sought to further analyze these differentially expressed genes by functional pathway analysis.

To validate the RNAseq analysis, we performed reverse transcriptase (RT) real time polymerase chain reaction (PCR) for a subset of genes that showed significantly altered expression in active CD patients. Our data confirmed the RNA seq analysis with IFN γ ($p \leq 0.05$), IL10 ($p \leq 0.05$), MCM2 ($p \leq 0.005$), CDC45 ($p \leq 0.001$) and CXCL11 ($p \leq 0.05$) significantly over-expressed in the active celiac populations as compared to control subjects. However, RT PCR for CXCR4, CADM1 and CLDN10 did not confirm RNA seq analysis (S3 Fig). We also performed immunohistochemistry on duodenal biopsies from healthy control patients, patients with active CD, and patients with CD in remission (S4 Fig). Our data confirmed the RNA seq analysis with MCM2 and CDC45 over-expressed in patients with active CD compared to patients with CD in remission and healthy control subjects. Please note that there is unspecific staining in the figure and CDC45 and MCM2 specific staining is seen in the crypt and crypt nuclei respectively.

Pathway analysis in subjects with active CD, CD in remission, and non-CD control subjects

We used two approaches to analyze pathways associated with patients with active CD and CD in remission. First, to analyze the enriched pathways in the unique genetic signatures related to active CD, and in those signatures common to active CD and CD in remission, we applied the Pathway Enrichment module of NextBio as described in the methods [16]. These signatures were filtered based on the fold change and p-value thresholds followed by segmentations via the Venn Diagram, and thus were no longer genome-wide. Only specific types of enrichment tests including the running Fisher algorithm offered by NextBio were applicable. Other methods such as Gene Set Enrichment Analysis (GSEA) were not suitable for the non-genome-wide signature lists. Using this approach, we identified 169 significant pathways unique to subjects with active CD compared to CD subjects in remission or non-CD control subjects (S1 Table). The top perturbed pathways exclusively upregulated in patients with active CD were overwhelmingly related to the cell cycle which was noted in 6 of the top 10 pathways. Other highly significant pathways included PLK-1 signaling event, FOXM1 transcription factor network, retinol metabolism, and other drug metabolism.

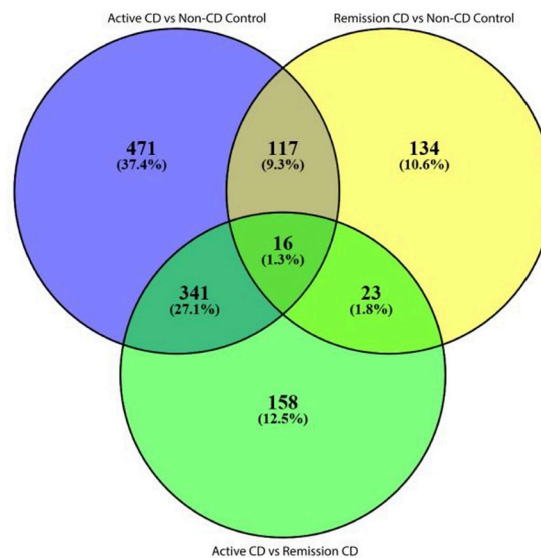
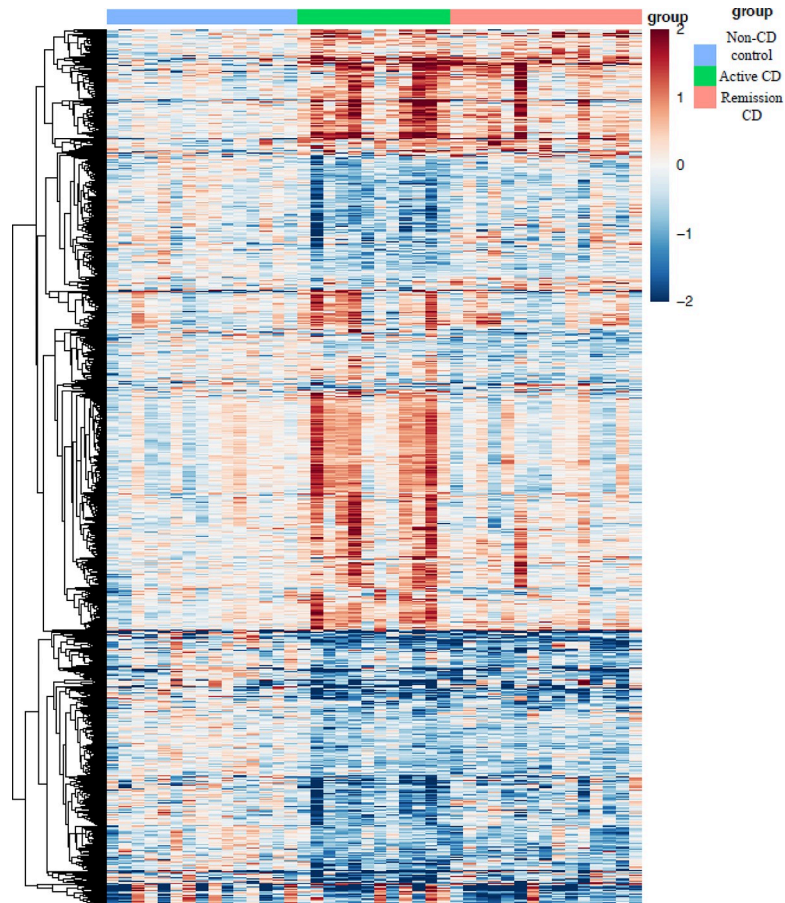


Fig 1. Comparison of gene expression in non-cd control, active CD, and remission CD subjects. Heatmap shows non-CD control subjects (blue), active CD subjects (green) and remission CD subjects (pink) with each subject's data represented vertically. The color scale show gene expression with dark blue indicating downregulation, light blue indicating decreased expression, light red indicating increased expression, and red indicating upregulation. (A) Gene expression in non-CD control, active CD, and remission CD subjects is notable for the distinct differences between

each group. (B) A venn diagram of unique and overlapping gene signatures shows genes exclusively altered in active CD (purple) (471) and remission CD (yellow) (134). An additional 117 genes are constitutively changed in subjects with CD irrespective of the disease status but not non-CD controls.

<https://doi.org/10.1371/journal.pone.0215132.g001>

We also identified 8 pathways that were constitutively changed in subjects with CD, irrespective of their disease state, that were distinct from non-CD control subjects. Intriguingly all pathways were upregulated and most of them, such as NK cell mediated cytotoxicity and chemokine signaling pathway, listed in [Table 1](#), are related to the innate immune system. Other notably upregulated pathways, such as interferon signaling link the innate immune system to the adaptive immune system. Antigen processing and presentation, a well described key to CD pathogenesis is also upregulated [2]. These results suggest that individuals with CD have a constitutive upregulation of pathways related to innate immune surveillance.

We ultimately chose to focus on the pathways that were significantly different between subjects with active CD and those in remission. To do this we took a second approach to analyze pathways in an effort to minimize bias that may be introduced by using the empirically defined fold change and p-value cutoffs. Therefore, we chose to perform gene set enrichment analysis (GSEA) for the active vs. remission analysis as it does not require any pre-selection and utilizes the genome-wide signature list. In this way, all genes were considered in the comparison of active CD compared to remission CD for analysis. KEGG pathway database was used for the GSEA analysis. We identified 48 KEGG pathways of significant interest ([Fig 2](#)). Of these, 35 pathways were significantly upregulated and 13 pathways were significantly downregulated in subjects with active CD compared to CD subjects in remission ([Fig 2](#)). Highly enriched pathways that were upregulated included DNA replication, valine, leucine, isoleucine biosynthesis, primary immunodeficiency, homologous recombination, and proteasome. Alternatively, enriched pathways which were downregulated included ascorbate and aldarate metabolism, retinol metabolism, pentose and glucuronate interconversions, sphingolipid metabolism, and fatty acid metabolism. When ranked by significance, the top three pathways, all of which were upregulated, included: cytokine-cytokine receptor interaction, chemokine signaling pathway, and cell cycle. Our results that follow highlight the top three perturbed pathways significantly different between subjects with active CD compared to those in remission as well as other novel pathways of interest to us which may shed light on the early events leading to the development of CD.

Immune response: Cytokine-cytokine receptor interaction and chemokine signaling pathway

Our results show the signaling pathways of cytokines ([S5 Fig](#)) and chemokines ([S6 Fig](#)) as two of the top three perturbed pathways upregulated in subjects with active CD compared to those

Table 1. Pathways common to active and remission CD but not non-CD controls.

Biogroup Name	Genes(#)	Direction	P-Value
Natural killer cell mediated cytotoxicity	6	up	1.00E-08
Ras-Independent pathway in NK cell-mediated cytotoxicity	3	up	3.90E-08
Antigen processing and presentation	5	up	5.10E-08
Genes involved in Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	4	up	9.30E-07
Graft-versus-host disease	3	up	1.30E-05
Genes involved in Interferon alpha/beta signaling	3	up	5.20E-05
Genes involved in Interferon Signaling	3	up	0.0007
Chemokine signaling pathway	3	up	0.001

<https://doi.org/10.1371/journal.pone.0215132.t001>

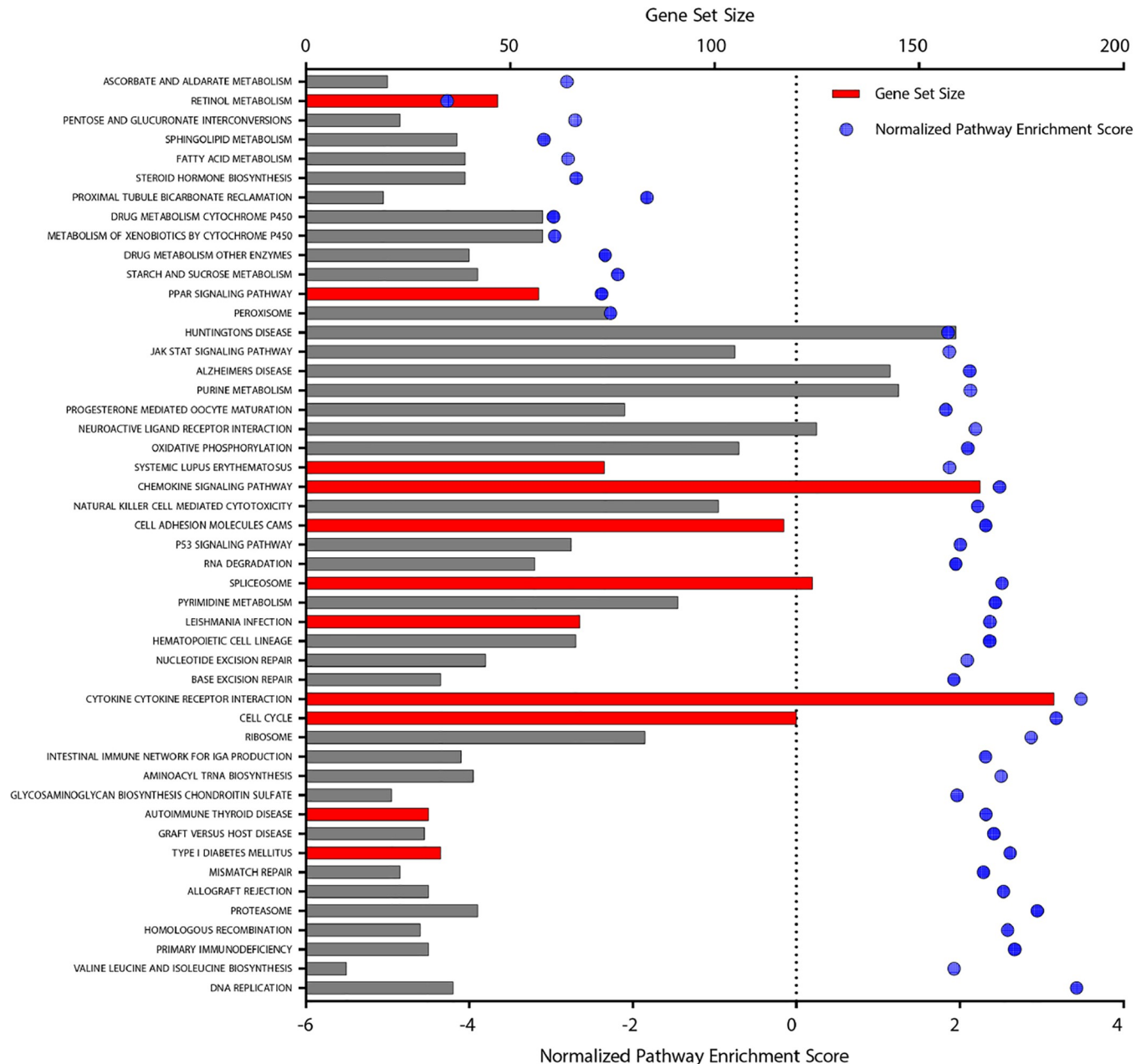


Fig 2. Significant pathways identified in active CD compared to remission CD subjects according to GSEA analysis using KEGG pathways. Pathways significantly altered in patients with active CD compared to remission CD per GSEA are shown horizontally. The gene set size is depicted by the size of the bar associated to each pathway per the vertical axis on the left. The normalized pathway enrichment score is shown on the right vertical axis and is depicted by the circles associated with each path. Pathways in red are discussed in detail.

<https://doi.org/10.1371/journal.pone.0215132.g002>

in remission. The identification of these previously described pathways in the innate and adaptive immune systems highlights the strength and validity of our analytical approach [17].

Neutrophil recruitment in the intestinal lamina propria represents one of the earliest steps in the cascade of events in response to gluten [18, 19]. We found an upregulation of neutrophils chemo attractants and their regulators (IL-8, CXCL1, CXCL2, CXCL3, CXCL6 and CXCR4) in subjects with active CD compared to those in remission. Furthermore, our data

also showed an increased expression of CXCR3 and its ligands (CXCL9, CXCL10, CXCL11). The interaction between gliadin and CXCR3 is known to trigger IL-8 secretion by epithelial and immune cells resulting in neutrophil migration [20, 21]. Our analysis confirmed that both EGF and IL-15 pathways were upregulated in subjects with active CD compared to those in remission and that IL-15 dependent macrophage-attracting chemokines (CCL20 and its receptor CCR6) were also upregulated. Similarly, upregulation of type -1 IFNs, IL1B, CCR2, CCR7 and CCR4 were found in active CD subjects supporting the previously described contribution of the innate immune component to the pathogenesis of the disease and, more specifically, the recruitment of macrophages and dendritic cells [22–25].

Our data further demonstrates the upregulation of Th1- and Th17-derived cytokine pathways such as IFN γ , TNF α , IL-2, IL-10, IL-17, and IL-21 are consistent with previous work which highlights the role of the adaptive immune response and the autoimmune nature characteristic of the chronic inflammation in subjects with active CD.

We also identified novel genes involved in the immune response machinery, IL-12RB1 and IL-12RB2, upregulated in the mucosa of patients with active CD which was consistent with our findings that IL-12 signaling was upregulated in subjects with active CD compared to those in remission.

Cell cycle

Our analysis shows upregulation of CDX genes related to the MAPK signaling pathway in active CD compared to CD in remission. Additional findings were the upregulation of molecules related to DNA replication (MCM helicase and Cdc45) and of c/myc oncogene (S7 Fig) [26].

Other pathways of interest

Cell adhesion molecules. The cell adhesion pathway analysis revealed changes in subjects with active CD compared to remission CD (S8 Fig). Specifically, the claudin (CLDN) family showed a down-regulation of the barrier forming CLDN 4, 12, 15 and up-regulation of the pore forming CLDN 2 and 10. Barrier forming CLDN 1 and 5 were upregulated along with CLDN18. C1orf106, a gene shown to regulate epithelial adherens junction stability, was down-regulated in subjects with active CD compared to subjects with CD in remission and non-CD control subjects again suggesting a role for increased intestinal permeability in subjects with active CD [27]. Additionally, we detected the up-regulation of IGSF4, a nectin-like molecule involved in a variety of cell-cell junctions, such as adherens junctions and tight junctions in epithelial cells [28].

Our analysis also revealed the upregulation of intercellular adhesion molecules (ICAMs), members of the immunoglobulin superfamily. Our analysis revealed upregulation of ICAM1-3, a molecule present on antigen presenting cells, both cytotoxic and helper T cells, and B cells. In addition, ITGAM, an adhesion molecule expressed primarily on monocytes and neutrophils, which is involved in activation, adherence, and migration of leukocytes has been found upregulated in samples from our active CD subjects.

Spliceosome. Spliceosome related pathways was one of the top 5 perturbed pathways in subjects with active CD compared to those in remission. Specifically, the expression of several genes coding for proteins necessary for the assembly (prp5), activation (prp2) and finally disassembly (prp22, prp43) of the spliceosome complex were significantly increased in the active CD patient group (S9 Fig)

IL-17, retinoic acid and associated autoimmune diseases. We found an upregulation in IL-17 signaling in subjects with active CD compared to subjects with CD in remission in

agreement with previous work[29, 30]. Our data showed a down-regulation of the vitamin A/retinoic acid pathways suggesting a possible correlation between the Th17 mediated immune response that characterizes active CD and alterations in retinol metabolism (S10 Fig). Additionally, the pathways related to systemic lupus erythematosus (S11 Fig), type 1 diabetes (T1D) (S12 Fig), and autoimmune thyroiditis (S13 Fig), were significantly upregulated in subjects with active CD compared to those in remission.

Discussion

Our analysis of subjects with active CD compared to those in remission and non-CD controls using whole transcriptome shotgun sequencing of duodenal biopsies revealed results consistent with many of the previously described pathways identified in the pathogenesis of CD, suggesting that our analysis is robust. Our finding that patients with CD, irrespective of disease status, have a constitutive upregulation of pathways related to innate immunity confirms the important role of the innate immune system in CD pathogenesis. By focusing on the statistically significant pathways that are up or down regulated in subjects with active CD compared to those in remission, we identified novel pathways of interest such as the spliceosome pathway that may contribute to the early steps involved in losing tolerance to gluten. We also identified novel genes involved in the immune response machinery (IL-12RB1 and IL-12RB2) and cell adhesion process (IGSF4 and ITGAM) in the mucosa of subjects with active CD compared to those in remission. Together, our findings support a possible link between the microbiome, innate immune response, and the development of CD and highlight possible associations that with future validation may lead to crucial knowledge of the steps leading to loss of tolerance to gluten. Finally, we also suggest possible molecular mechanisms for the association with other autoimmune diseases and persistent enteropathy despite a strict gluten free diet. Our data serve to act as a launching point to suggest areas of further investigation and validation for CD pathogenesis, treatment, and associations with other autoimmune disease.

Our findings that patients with active CD had an upregulation in cytokine-cytokine receptor interaction and the chemokine signaling pathway is expected given our understanding of CD pathogenesis and the inflammation associated with active CD. Our data have also provided possible insights into genes involved in the immune response. While upregulation of IFN γ in subjects with active CD is well described, our finding of IL-12 upregulation in the mucosa of subjects with active CD is novel. Previous work has identified elevated IL-12, specifically IL-12RB2, in the mucosa of patients with active Crohn's disease [31]. We found upregulation of both IL-12RB1 and IL-12RB2 in subjects with active CD compared to those in remission. While the molecular basis of Th1 differentiation is not well described, STAT-4 and IRF-1 have been associated with IL-12 and IFN γ signaling in T cells and involved in regulating cytokine production of Th1 cells at the transcriptional level. Our analysis found a significant upregulation of IRF-1 in subjects with active disease compared to controls.

Our analysis identified the upregulation of cell adhesion molecules such as IGSF4 and others involved in leukocyte transendothelial migration such as ITGAM (S8 Fig) in active CD. ITGAM which combines with the beta 2 chain (ITGB2, also found upregulated in our analysis) to form a leukocyte-specific integrin referred to as macrophage receptor 1 [32] is important in the adherence of neutrophils and monocytes to stimulated endothelium and favors the transmigration from circulation to the intestinal lamina propria. These findings could link the early steps of increased antigen trafficking from the intestinal lumen to the lamina propria with the innate immune response including neutrophil transmigration[21] from the bloodstream to the lamina propria, activation of NK cells and, ultimately, adaptive immune response involving both T and B cell activation.

An upregulation of spliceosome pathways has not been described before in CD. However, our group has previously shown that alternative splicing influences how T regulatory cells suppress the Th17 mediated immune response in active CD subjects [33]. Previous work suggests the microbiome-influenced microenvironment may directly alter Th17 expression [33–35]. Our data shows an upregulation of IL-17. IL-17 can be activated by gliadin, bacteria, and many other factors, and thus it is possible to interpret this finding as implicating the microbiome in the pathogenesis of CD and supports the possibility that early steps leading to loss of tolerance to gluten involve gliadin mistaken as a component of a bacterial pathogen [29, 36].

Villous atrophy, the histological hallmark of CD, has been hypothesized to be related to increased cell destruction/enhanced apoptosis and/or a defect in epithelial cell regeneration. We found an increased expression of CDX genes in active CD subjects compared to remission which could be consistent with induction of stem cell turnover and intestinal regeneration. Additionally, we found an upregulation in molecules related to DNA replication (MCM helicase and Cdc45), confirmed with RT PCR and immunohistochemistry, and of *c/myc* oncogene, which has been previously shown to be upregulated in CD patients exposed to gluten and associated with CD mucosal damage [26]. Thus, it is possible that the increased apoptosis in active CD [37] is only partially compensated by the accelerated turnover of stem cells and activation of DNA replication, a process that is influenced by the activation of specific oncogenes such as *c-myc*. This inefficient compensation by the stem cell compartment could explain why some patients with CD, despite strict adherence to the treatment, have persistent enteropathy.

The co-morbidity between CD and other autoimmune diseases is well established [38], however it remains unknown if this association is due to gene clusters increasing the risk of multiple autoimmune diseases, or complication of untreated CD [3, 39]. Vitamin A and its metabolites have a modulatory effect on Th17/Treg balance by improving the differentiation of CD4⁺ cells toward a regulatory phenotype and suppressing the production of Th17 derived proinflammatory cytokines [40]. Several autoimmune diseases characterized by Th17 mediated immune response have been associated with impairments in vitamin A modulatory function [41–43]. We found an upregulation in IL-17 signaling and down-regulation of the vitamin A/retinoic acid pathways in subjects with active CD compared to subjects with CD in remission. We also found evidence to suggest the risk of developing comorbid autoimmune conditions may be particularly high in the active phase of CD as pathways for lupus, T1D, and autoimmune thyroid disease, conditions linked to dysregulation of retinoic acid metabolism, were upregulated [44–46]. Our findings are supported by previous studies suggesting the early diagnosis of CD prevents the development of further autoimmune disease and that comorbid autoimmune conditions improve with CD treatment. [47–52].

In summary, our study confirmed previous findings based on GWAS and immunological studies pertinent to CD pathogenesis, but also suggests novel areas of investigation related to CD pathogenesis, mucosal repair, and the association with other autoimmune diseases that with further validation may fill knowledge gaps in the chain of events leading to break of tolerance to gluten and onset of CD autoimmunity in genetically-predisposed individuals. Given the large amount of data, we chose to highlight the top three pathways significantly altered in patients with active CD compared to those with remission CD. However, it is possible that differences observed in differentially expressed genes between these groups could be explained by the difference in the crypt depth-villous height ratio since patients with active CD will have shorter villi and therefore may express lower levels of villous genes and elevated expression of crypt genes. The data generated does not localize the altered genes and therefore future validation of these findings is warranted. Additionally, since duodenal biopsies represent a small sample of the intestine, it is possible that some patients found to have CD in remission could

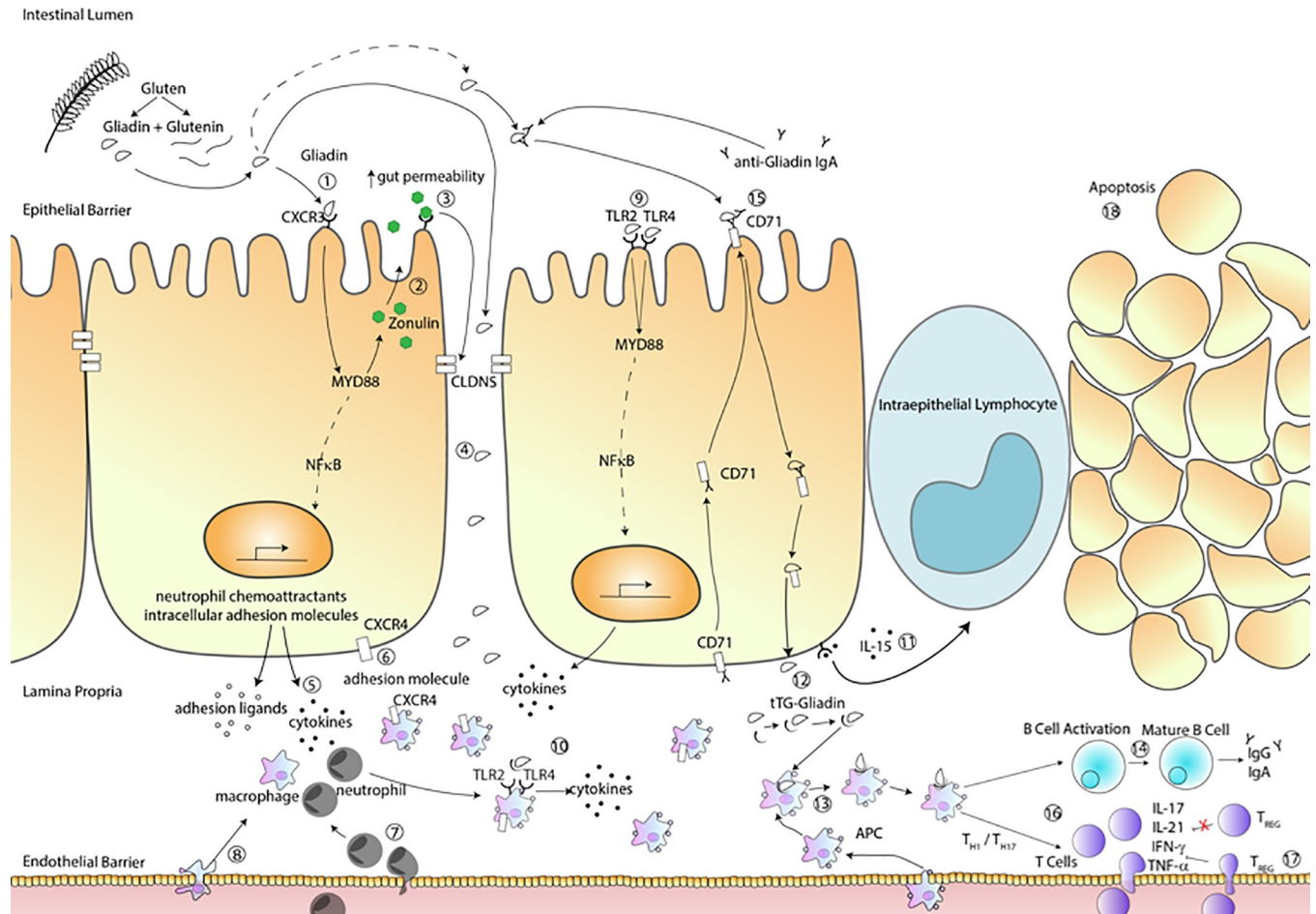


Fig 3. Possible pathogenetic steps leading to CD.

<https://doi.org/10.1371/journal.pone.0215132.g003>

have other areas of their intestine with active inflammation which could affect our findings. The power in using RNAseq is that it provides a broad scope of genes and associated pathways involved in the pathogenesis of CD. However, it requires further validation and our validation techniques supported several but not all our findings. Our goal was not to validate all findings but instead provide a roadmap for ourselves and other investigators to investigate new insights into the pathogenesis of CD. Based on these findings, our interpretation of this data combined with previously published work is that gluten exposure followed by increased gut epithelial permeability, coupled with increased endothelial permeability causing the migration of immune cells from the bloodstream to the intestinal lamina propria could represent early events for the initiation of the CD innate immune response (Fig 3). It is possible that this innate immune response to gluten may mimic similar responses triggered by the exposure to pathogens, such as *Leishmania* (pathway noted to be significantly upregulated in patients with active CD), suggesting that gluten may mistakenly be interpreted as a danger signal by the intestinal mucosal immune surveillance system. Following these early events, cellular insult causing release of intracellular tTG leads to deamidation of gliadin fragments present in the lamina propria following their paracellular passage. Binding of deamidated gliadin to HLA DQ2/8 molecules on the surface of APC and subsequent activation of a Th1/Th17 response, secondary to dysfunctional Tregs due to possible contribution from the spliceosome pathway

activation and activation of other adaptive immune cells, could lead to the CD autoimmune enteropathy due to increased apoptosis and defects in the stem cell regenerative compartment (Fig 3). The combination of this information offers novel potential targets for personalized therapies integrative to the gluten free diet and, hopefully, in the future, for possible disease interception (primary prevention) in genetically at-risk subjects. Furthermore, future work using single cell immune profiling to examine patients in the active and remission state of CD, as well as those who fail to respond to the gluten free diet, will be essential to understanding not only CD but other chronic immune based disorders.

Innate immune response. Gliadin binding to CXCR3(1) triggers zonulin release(2), changes in intercellular protein complexes gene expression and ultimately, increased gut permeability(3) leading to gliadin fragment passage into the lamina propria(4) (LP). Here, gliadin causes CXCR3-mediated IL-8 secretion(5) that, together with CXCR4/CXCL12 upregulation (6), leads to migration of neutrophils in the LP. This process is facilitated by increased neutrophils' expression of ITGAM aiding their migration from the bloodstream to the LP(7). B2integrin expression on macrophages and the endothelial ligand ICAM leads to macrophage migration to the LP(8). Gluten also initiates a cascade of anti-microbial responses, including the activation of TLR2/TLR4 with subsequent MyD88-dependent NFkB activation and transcription of pro-inflammatory cytokines(9). Increased expression of TLR2/TLR4 on macrophages leads to the release of anti-microbial molecules(10) secondary to aberration in alternative splicing machinery. The innate immune response is completed by gliadin-dependent IL-15 upregulation, leading to the recruitment intraepithelial lymphocytes which have cytotoxic activity on epithelial cells(11). **Adaptive Immunity** In the LP, gliadin is deamidated and binds to tTG forming a tTG-gliadin complex(12) recognized by antigen presenting cells (APCs) expressing HLA-DQ2/8(13). APC's presenting gliadin active T helper cells that trigger B cell activation and maturation resulting in production of anti-tTG autoantibodies(14). CD71 translocated to the apical side of intestinal epithelial cells to transport sIgA-gliadin complexes (15). APC also activate TH1 cells releasing proinflammatory cytokines and TH17 cells releasing IL-17 and IL21(16). T regulatory cells migrate from the blood to stop inflammation, however, due to dysfunctional spliceosome pathway, they are not capable to stop inflammation (17). **Ongoing Inflammation** Changes in apoptotic pathways and stem cell turnover-related genes suggest that the celiac enteropathy may be the results of increased apoptosis coupled with incomplete compensation by the stem cell compartment(18).

Methods

Study subjects

Seventy-four subjects were approached for inclusion in the study in our clinic or at the time of their clinically indicated endoscopy between June 2014 and August 2015. Of those, 42 subjects met criteria for inclusion into the study. All subjects were 18 years of age or older, carried HLA DQ2 or DQ8, were in good general health and were willing to provide informed consent. Twelve subjects with active CD, 15 subjects with remission CD, and 15 non-CD control subjects with compatible genetics and absence of small intestinal inflammation were included in the study. Characteristics of these patients are shown in Table 2. CD subjects were required to have a diagnosis according to previously described algorithms[53]. In addition to meeting inclusion criteria, subjects with active CD were required to have villous blunting consistent with Marsh 3 on duodenal biopsy and subjects with remission CD must have maintained a gluten free diet for at least one year and have a duodenal biopsy consistent with CD in remission (Marsh 0–2)[54]. Non-CD control subjects were in good health, carried compatible HLA genetics, and were undergoing an endoscopy for evaluation unrelated to CD such as

Table 2. Patient characteristics.

Diagnosis Stratification	Age (mean)	Gender (%F)	tTG IgA level (mean)	DQ2 Homozygous (%)	DQ2 Heterozygous (%)	DQ8 (%)	DQ2/DQ8 (%)
Active CD (N = 12)	36.1	64.3	143.3	33.0	50.0	17.0	0.0
Remission CD (N = 15) ^A	35.5	86.7	12.7	35.7	42.9	21.4	0.0
Non-CD Control (N = 15)	41.8	73.3	1.9	13.3	46.7	20.0	20.0

^AHLA type of one subject had too low a coverage at DQB1 locus via RNAseq so no allele typing was performed

<https://doi.org/10.1371/journal.pone.0215132.t002>

gastroesophageal reflux. Non-CD control subjects were required to have negative celiac serology, normal duodenal biopsies, and no diagnostic abnormalities identified during the clinically indicated endoscopy. Exclusion criteria included receiving systemic immunosuppressive therapy and/or a vaccination within the past 30 days, participation in any clinical research study evaluating an investigational drug or device within the past 30 days, history of other gastrointestinal disease such as inflammatory bowel disease, receipt of anti-coagulant therapy within the previous 30 days, and women who were pregnant or breastfeeding.

Study samples

Serology. All subjects underwent venipuncture at the time of the endoscopy procedure. A minimum of 55cc of blood were collected from each participant. Serum was evaluated for antibodies to IgA tissue transglutaminase (anti-tTG) QUANTA Lite Rh-tTG IgA ELISA (INOVA Diagnostics, San Diego, CA, USA) on the BioFlash platform.

HLA determination. Whole blood was collected and stored at -80 degree Celsius. HLA was determined using the DQ-CD Typing Plus (BioDiagne, Palermo, Italy) per the manufacturer's instructions for all but three subjects included in the analysis. The HLA for the additional three subjects who did not have a whole blood aliquot available for testing were determined by RNA sequencing.

Duodenal biopsy. Four duodenal biopsies were obtained during clinically indicated upper endoscopy for each subject. Two biopsies were immediately placed in RNA later and stored at -20° C for RNA sequencing. Two biopsies were placed in 4% paraformaldehyde for four hours at room temperature, washed with 70% ethanol, and stored at room temperature for future use. A histopathologic assessment was completed for all patients by a single pathologist. Results were confirmed by a second pathologist and reported using the modified Marsh scale[54].

RNA sequencing. The protocols for RNA extraction and sequencing library preparation were performed as described earlier[55]. Specifically, strand-specific RNA-seq libraries were prepared from 500 ng RNA using KAPA stranded mRNA-Seq Kit (KAPA Biosystems). The libraries were amplified by twelve-cycle PCR. Sequencing was performed on Illumina HiSeq2000 (Illumina) by multiplexed paired-read run with 100 cycles. The resulting FASTQ files were analyzed via FastQC to ensure sufficient data quality.

DNACore RNA extraction and RT-qPCR protocol. For RNA extraction tissue samples were transferred from RNAlater (ThermoFisher, Cat # AM7021) to 1–3 milliliters Qiazol reagent (Qiagen, Cat # 79306). Samples were homogenized on a custom Omni homogenizer (Omni, Cat #SKU: 51-000-1) at 20,000 RPM for 180 seconds. Lysates were phase separated with chloroform and the aqueous phase was purified on the Kingfisher flex (ThermoFisher, Cat # 5400630) with the MagMAX-96 for Microarrays Total RNA Isolation Kit (ThermoFisher, Cat # AM1839) with an additional DNase (Qiagen, Cat # 79254) step added between

the first and second washes. RNA was quantized on the Nanodrop (ThermoFisher, Cat # ND-8000-GL) according to manufactures protocol.

For RT-qPCR, between 37.5 nanograms and 375 nanograms of RNA, per RT-qPCR assay to be run, was mixed with SuperScript VILO Master Mix (ThermoFisher, Cat# 11755500) and cycled according to manufacturer instructions. The cDNA was diluted with nuclease free water to between 0.5 nanograms per microliter and 5 nanograms per microliter. RT-qPCR assays were made by combining water, mastermix (ThermoFisher, Cat #4370074 or Biorline, Cat #CSA-01113), and 20X assay mix. The assay mix is a commercially available mixture of forward and reverse primers combined with a fluorescently labeled and quenched probe sequence (ThermoFisher, Cat# 351372 or LGC BioSearch, Cat# DLO-RFBL-MIX). Samples were run in triplicate on a 384-Well plate (ThermoFisher, Cat# 4343370) by pipetting 5 microliters diluted cDNA and 10 microliters of appropriate RT-PCR assay into each well. The 384-well plate was covered with an optically clear seal (Agilent, Cat# 16985-001), spun down, and read on an ABI 7900HT Fast Real-Time PCR System with 384-Well Block Module and Automation Accessory (ThermoFisher, Cat# 4329002) for 40 cycles according to the specifics of the mastermix used.

Immunohistochemistry. Immunohistochemistry of paraffin embedded duodenal biopsies were performed on well oriented 5 μ m thick paraffin embedded samples from human duodenum of patients. Two samples from patients diagnosed with acute CD (Marsh 3a-c), patients with CD in remission (Marsh 0-1), and normal mucosa subjects (Marsh 0), were obtained from the Pathology Clinics at MGH for this purpose. Duodenal sections from control donors, donors with active celiac disease (CDA) or remission celiac disease (CDR) were stained for IHC using Vector Laboratories Vectastain Elite Kit and Vector DAB Kit per manufacturer's instructions. Briefly, sections were deparaffinized and hydrated using xylenes and a graded ethanol series. Antigen retrieval was performed using sodium citrate buffer at pH 6.0. Sections were blocked against endogenous peroxidase activity using 0.3% hydrogen peroxide in water for thirty minutes. After washing and blocking using diluted normal serum (Vector Laboratories), primary antibodies against MCM2 (Thermo Fisher clone 2BE, Cat #MA5-15893) and CDC45 (Sigma Prestige Antibodies, Cat# HPA00614) were used at a concentration of 1:400 and 1:150, respectively, overnight at 4 degrees. The next day, slides were washed and treated with diluted biotinylated secondary antibodies (Vector Laboratories) for 30 minutes at room temperature. Sections were then washed and incubated with Vectastain Elite ABC reagent followed by peroxidase substrate solution (Vector DAB kit) until developed. Sections were mounted and imaged using an inverted light microscope (Nikon Eclipse TE2000-5) at 10x magnification. Post-acquisition analysis was conducted using ImageJ [56] and the IHC toolbox plugin[57].

Statistical analysis

Derivation of differentially expressed gene signatures. The transcriptome sequencing reads were mapped to the hg19 genome reference using the OmicSoft ArrayStudio RNASeq analysis pipeline (Qiagen Inc) with two mismatches allowed. The resulting read counts summarized at gene level represented the raw gene expression measures. We normalized the raw gene expression in each sample by global scaling to match the median library size (i.e. the total number of mapped reads) as well as the upper quantile of the gene-level read counts across all samples, as described in previous studies[58]. An empirical minimum read count of 10 was applied to flag the "absence" and "presence" of genes in each sample. For each comparison between two groups of samples, we first eliminated genes that were not flagged as "presence" in at least 60% of the samples of the higher expressing group. Next, fold changes associated

with the comparison were calculated as the ratio between the arithmetic mean expressions in the two groups. The statistical significance (p-value) of the differential expressions was assessed under negative binomial distribution models using DESeq package version 1.6[59]. At the end, we selected genes with fold changes no less than 1.5 in either up or down directions with a p-values of at least 0.05 as the significantly perturbed gene signatures.

Gene set enrichment analysis(GSEA) and pathway analysis. In the active CD versus remission CD comparison we performed the Gene Set Enrichment Analysis (GSEA). The input contained all gene symbols detectable in the RNA sequencing, together with their fold changes between the active and remission CD subjects to provide the needed rank information in the GSEA. The fold changes smaller than 1 were converted to the negative reciprocal values. The analysis was performed using the build-in “GSEA” function in the R package “clusterProfiler”[60]. The gene sets used in the analysis were the curated KEGG canonical pathways in the MSigDB (version: c2.cp.kegg.v6.1.symbols.gmt). The pathways with adjusted p-value ≤ 0.05 were reported. To analyze the enriched pathways in the unique signatures to the active CD, and in the common signatures to the active and remission CD, we executed the Pathway Enrichment module of NextBio, which applies a running Fisher algorithm over the MSigDB curated canonical pathways[16]. Note that GSEA is not applicable here as the signatures were subjected to the differential gene expression selection criteria as well as Venn Diagram analysis (Fig 2), and thus no longer contain the genome-wide genes. Pathways with a running Fisher test p-value ≤ 0.001 and at least 3 genes occurring in the signatures were considered significant and reported.

Study approval. Written informed consent was obtained from subjects prior to inclusion in the study according to the standards outlined and approved by the Partners Human Resource Committee Institutional Review Board.

Supporting information

S1 Table. Pathways unique to patients with active CD.

(DOCX)

S1 Fig. Gene Expression Signatures Unique to Active CD. Heatmap shows non-CD control subjects (blue), active CD subjects (green) and remission CD subjects (pink) with each subject's data represented vertically. The color scale show gene expression with dark blue indicating downregulation, light blue indicating decreased expression, light red indicating increased expression, and red indicating upregulation.

(TIF)

S2 Fig. Gene Expression Common to Active and Remission CD. Heatmap shows non-CD control subjects (blue), active CD subjects (green) and remission CD subjects (pink) with each subject's data represented vertically. The color scale show gene expression with dark blue indicating downregulation, light blue indicating decreased expression, light red indicating increased expression, and red indicating upregulation.

(TIF)

S3 Fig. RTPCR Results in a subset of genes in healthy control patients, patients with active CD, and patients with CD in remission.

(TIF)

S4 Fig. Immunohistochemistry results in a subset of healthy control patients, patients with active CD, and patients with CD in remission.

(TIF)

S5 Fig. KEGG pathway cytokine-cytokine receptor interaction. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S6 Fig. KEGG pathway chemokine signaling pathway. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S7 Fig. KEGG pathway cell cycle. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S8 Fig. KEGG pathway cell adhesion molecules. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S9 Fig. KEGG pathway spliceosome. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S10 Fig. KEGG pathway retinol metabolism. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S11 Fig. KEGG pathway systemic lupus erythematosus. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S12 Fig. KEGG pathway type 1 diabetes mellitus. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

S13 Fig. KEGG pathway autoimmune thyroid disease. We performed a subtraction analysis using GSEA for patients with active CD compared to those with CD in remission. Genes identified as significant are highlighted in red against a background of other known genes in the pathway.

(TIF)

Author Contributions

Conceptualization: Cagan Gurer, Alessio Fasano.

Data curation: Maureen M. Leonard, Yu Bai, Kourtney P. Nickerson, Craig Sturgeon, Shu Yan, Aubrey Katz, Barbara Nath, James Richter.

Formal analysis: Yu Bai, Gloria Serena, Kourtney P. Nickerson, Maria R. Fiorentino, Matthew Sleeman, Alessio Fasano.

Funding acquisition: Alessio Fasano.

Investigation: Maureen M. Leonard, Yu Bai, Gloria Serena, Kourtney P. Nickerson, Maria R. Fiorentino, Matthew Sleeman, Cagan Gurer, Alessio Fasano.

Methodology: Kourtney P. Nickerson, Maria R. Fiorentino, Cagan Gurer, Alessio Fasano.

Project administration: Maureen M. Leonard, Stephanie Camhi, Craig Sturgeon, Shu Yan, Aubrey Katz, Barbara Nath, James Richter, Matthew Sleeman, Alessio Fasano.

Supervision: Alessio Fasano.

Visualization: Maria R. Fiorentino.

Writing – original draft: Maureen M. Leonard, Yu Bai, Gloria Serena, Kourtney P. Nickerson, Cagan Gurer, Alessio Fasano.

Writing – review & editing: Maureen M. Leonard, Yu Bai, Gloria Serena, Kourtney P. Nickerson, Stephanie Camhi, Craig Sturgeon, Shu Yan, Maria R. Fiorentino, Aubrey Katz, Barbara Nath, James Richter, Matthew Sleeman, Cagan Gurer, Alessio Fasano.

References

1. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clinical Gastroenterology and Hepatology*. 2018; 16(6):823–36. e2. <https://doi.org/10.1016/j.cgh.2017.06.037> PMID: 29551598
2. Leonard MM, Sapone A, Catassi C, Fasano A. Celiac Disease and Nonceliac Gluten Sensitivity: A Review. *JAMA*. 2017; 318(7):647–56. <https://doi.org/10.1001/jama.2017.9730> PMID: 28810029.
3. Lundin KE, Wijmenga C. Coeliac disease and autoimmune disease-genetic overlap and screening. *Nat Rev Gastroenterol Hepatol*. 2015; 12(9):507–15. <https://doi.org/10.1038/nrgastro.2015.136> PMID: 26303674.
4. Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, et al. Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Ann Med*. 2010; 42(7):530–8. <https://doi.org/10.3109/07853890.2010.514285> PMID: 20868314.
5. Gutierrez-Achury J, Zhernakova A, Pulit SL, Trynka G, Hunt KA, Romanos J, et al. Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat Genet*. 2015; 47(6):577–8. <https://doi.org/10.1038/ng.3268> PMID: 25894500; PubMed Central PMCID: PMC4449296.
6. Kuja-Halkola R, Lebowitz B, Halfvarson J, Wijmenga C, Magnusson PK, Ludvigsson JF. Heritability of non-HLA genetics in coeliac disease: a population-based study in 107 000 twins. *Gut*. 2016; 65(11):1793–8. <https://doi.org/10.1136/gutjnl-2016-311713> PMID: 27207974.
7. Ricano-Ponce I, Wijmenga C, Gutierrez-Achury J. Genetics of celiac disease. *Best Pract Res Clin Gastroenterol*. 2015; 29(3):399–412. <https://doi.org/10.1016/j.bpg.2015.04.004> PMID: 26060105.
8. Juuti-Uusitalo K, Maki M, Kaukinen K, Collin P, Visakorpi T, Vihinen M, et al. cDNA microarray analysis of gene expression in coeliac disease jejunal biopsy samples. *J Autoimmun*. 2004; 22(3):249–65. <https://doi.org/10.1016/j.jaut.2003.12.001> PMID: 15041046.
9. Diodado B, Wapenaar MC, Franke L, Duran KJ, Goerres MJ, Hadithi M, et al. A microarray screen for novel candidate genes in coeliac disease pathogenesis. *Gut*. 2004; 53(7):944–51. <https://doi.org/10.1136/gut.2003.018374> PMID: 15194641; PubMed Central PMCID: PMC1774096.
10. Bracken S, Byrne G, Kelly J, Jackson J, Feighery C. Altered gene expression in highly purified enterocytes from patients with active coeliac disease. *BMC Genomics*. 2008; 9:377. <https://doi.org/10.1186/1471-2164-9-377> PMID: 18691394; PubMed Central PMCID: PMC2533024.

11. Castellanos-Rubio A, Santin I, Martin-Pagola A, Irastorza I, Castano L, Vitoria JC, et al. Long-term and acute effects of gliadin on small intestine of patients on potentially pathogenic networks in celiac disease. *Autoimmunity*. 2010; 43(2):131–9. <https://doi.org/10.3109/08916930903225229> PMID: 19814655.
12. Bragde H, Jansson U, Jarlsfelt I, Soderman J. Gene expression profiling of duodenal biopsies discriminates celiac disease mucosa from normal mucosa. *Pediatr Res*. 2011; 69(6):530–7. <https://doi.org/10.1203/PDR.0b013e318217ecec> PMID: 21378598.
13. Plaza-Izurieta L, Fernandez-Jimenez N, Irastorza I, Jauregi-Miguel A, Romero-Garmendia I, Vitoria JC, et al. Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in celiac disease. *Eur J Hum Genet*. 2015; 23(8):1100–5. <https://doi.org/10.1038/ejhg.2014.244> PMID: 25388004; PubMed Central PMCID: PMC4795102.
14. Quinn EM, Coleman C, Molloy B, Dominguez Castro P, Cormican P, Trimble V, et al. Transcriptome Analysis of CD4+ T Cells in Coeliac Disease Reveals Imprint of BACH2 and IFN γ Regulation. *PLoS One*. 2015; 10(10):e0140049. <https://doi.org/10.1371/journal.pone.0140049> PMID: 26444573; PubMed Central PMCID: PMC4596691.
15. Jauregi-Miguel A, Fernandez-Jimenez N, Irastorza I, Plaza-Izurieta L, Vitoria JC, Bilbao JR. Alteration of tight junction gene expression in celiac disease. *J Pediatr Gastroenterol Nutr*. 2014; 58(6):762–7. <https://doi.org/10.1097/MPG.0000000000000338> PMID: 24552675.
16. Kupersmidt I, Su QJ, Grewal A, Sundaresh S, Halperin I, Flynn J, et al. Ontology-based meta-analysis of global collections of high-throughput public data. *PLoS One*. 2010; 5(9). <https://doi.org/10.1371/journal.pone.0013066> PMID: 20927376; PubMed Central PMCID: PMC2947508.
17. Garrote JA, Gomez-Gonzalez E, Bernardo D, Arranz E, Chirido F. Celiac disease pathogenesis: the proinflammatory cytokine network. *J Pediatr Gastroenterol Nutr*. 2008; 47 Suppl 1:S27–32. <https://doi.org/10.1097/MPG.0b013e3181818fb9> PMID: 18667914.
18. Lammers KM, Chieppa M, Liu L, Liu S, Omatsu T, Janka-Junttila M, et al. Gliadin Induces Neutrophil Migration via Engagement of the Formyl Peptide Receptor, FPR1. *PLoS One*. 2015; 10(9):e0138338. <https://doi.org/10.1371/journal.pone.0138338> PMID: 26378785; PubMed Central PMCID: PMC4574934.
19. Diosdado B, van Bakel H, Strengman E, Franke L, van Oort E, Mulder CJ, et al. Neutrophil recruitment and barrier impairment in celiac disease: a genomic study. *Clin Gastroenterol Hepatol*. 2007; 5(5):574–81. <https://doi.org/10.1016/j.cgh.2006.11.014> PMID: 17336591.
20. Lammers KM, Lu R, Brownley J, Lu B, Gerard C, Thomas K, et al. Gliadin induces an increase in intestinal permeability and zonulin release by binding to the chemokine receptor CXCR3. *Gastroenterology*. 2008; 135(1):194–204 e3. <https://doi.org/10.1053/j.gastro.2008.03.023> PMID: 18485912; PubMed Central PMCID: PMC2653457.
21. Lammers KM, Khandelwal S, Chaudhry F, Kryszak D, Puppa EL, Casolaro V, et al. Identification of a novel immunomodulatory gliadin peptide that causes interleukin-8 release in a chemokine receptor CXCR3-dependent manner only in patients with coeliac disease. *Immunology*. 2011; 132(3):432–40. <https://doi.org/10.1111/j.1365-2567.2010.03378.x> PMID: 21091908; PubMed Central PMCID: PMC3044909.
22. Jabri B, Kasarda DD, Green PH. Innate and adaptive immunity: the yin and yang of celiac disease. *Immunol Rev*. 2005; 206:219–31. <https://doi.org/10.1111/j.0105-2896.2005.00294.x> PMID: 16048552.
23. Westerholm-Ormio M, Vaarala O, Pihkala P, Ilonen J, Savilahti E. Immunologic activity in the small intestinal mucosa of pediatric patients with type 1 diabetes. *Diabetes*. 2003; 52(9):2287–95. PMID: 12941768.
24. Chladrkova B, Kamanova J, Palova-Jelinkova L, Cinova J, Sebo P, Tuckova L. Gliadin fragments promote migration of dendritic cells. *J Cell Mol Med*. 2011; 15(4):938–48. <https://doi.org/10.1111/j.1582-4934.2010.01066.x> PMID: 20406323; PubMed Central PMCID: PMC3922678.
25. Beitnes AC, Raki M, Brottveit M, Lundin KE, Jahnsen FL, Sollid LM. Rapid accumulation of CD14+CD11c+ dendritic cells in gut mucosa of celiac disease after in vivo gluten challenge. *PLoS One*. 2012; 7(3):e33556. <https://doi.org/10.1371/journal.pone.0033556> PMID: 22438948; PubMed Central PMCID: PMC3306402.
26. Ciclitira PJ, Stewart J, Evan G, Wight DG, Sikora K. Expression of c-myc oncogene in coeliac disease. *J Clin Pathol*. 1987; 40(3):307–11. PMID: 3549790; PubMed Central PMCID: PMC1140905.
27. Mohanan V, Nakata T, Desch AN, Levesque C, Boroughs A, Guzman G, et al. C1orf106 is a colitis risk gene that regulates stability of epithelial adherens junctions. *Science*. 2018; 359(6380):1161–6. <https://doi.org/10.1126/science.aan0814> PMID: 29420262.
28. Shingai T, Ikeda W, Kakunaga S, Morimoto K, Takekuni K, Itoh S, et al. Implications of nectin-like molecule-2/IGSF4/RA175/SgIGSF/TSLC1/SynCAM1 in cell-cell adhesion and transmembrane protein

- localization in epithelial cells. *J Biol Chem*. 2003; 278(37):35421–7. <https://doi.org/10.1074/jbc.M305387200> PMID: 12826663.
29. Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, Ramon Bilbao J. TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. *Autoimmunity*. 2009; 42(1):69–73. <https://doi.org/10.1080/08916930802350789> PMID: 19127457.
 30. Sapone A, Lammers KM, Mazzarella G, Mikhailenko I, Carteni M, Casolaro V, et al. Differential mucosal IL-17 expression in two gliadin-induced disorders: gluten sensitivity and the autoimmune enteropathy celiac disease. *Int Arch Allergy Immunol*. 2010; 152(1):75–80. <https://doi.org/10.1159/000260087> PMID: 19940509; PubMed Central PMCID: PMCPMC2956008.
 31. Parrello T, Monteleone G, Cucchiara S, Monteleone I, Sebkova L, Doldo P, et al. Up-regulation of the IL-12 receptor beta 2 chain in Crohn's disease. *J Immunol*. 2000; 165(12):7234–9. PMID: 11120856.
 32. Rosetti F, Mayadas TN. The many faces of Mac-1 in autoimmune disease. *Immunol Rev*. 2016; 269(1):175–93. <https://doi.org/10.1111/immr.12373> PMID: 26683153.
 33. Serena G, Yan S, Camhi S, Patel S, Lima RS, Sapone A, et al. Proinflammatory cytokine interferon-gamma and microbiome-derived metabolites dictate epigenetic switch between forkhead box protein 3 isoforms in coeliac disease. *Clin Exp Immunol*. 2017; 187(3):490–506. <https://doi.org/10.1111/cei.12911> PMID: 27936497; PubMed Central PMCID: PMCPMC5290237.
 34. Vodicka M, Ergang P, Hrcir T, Mikulecka A, Kvapilova P, Vagnerova K, et al. Microbiota affects the expression of genes involved in HPA axis regulation and local metabolism of glucocorticoids in chronic psychosocial stress. *Brain Behav Immun*. 2018; 73:615–24. <https://doi.org/10.1016/j.bbi.2018.07.007> PMID: 29990567.
 35. Sandquist I, Kolls J. Update on regulation and effector functions of Th17 cells. *F1000Res*. 2018; 7:205. <https://doi.org/10.12688/f1000research.13020.1> PMID: 29527301; PubMed Central PMCID: PMCPMC5820607.
 36. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*. 2009; 139(3):485–98. <https://doi.org/10.1016/j.cell.2009.09.033> PMID: 19836068; PubMed Central PMCID: PMCPMC2796826.
 37. Senger S, Sapone A, Fiorentino MR, Mazzarella G, Lauwers GY, Fasano A. Celiac Disease Histopathology Recapitulates Hedgehog Downregulation, Consistent with Wound Healing Processes Activation. *PLoS One*. 2015; 10(12):e0144634. <https://doi.org/10.1371/journal.pone.0144634> PMID: 26649570; PubMed Central PMCID: PMCPMC4674131.
 38. Lauret E, Rodrigo L. Celiac disease and autoimmune-associated conditions. *Biomed Res Int*. 2013; 2013:127589. <https://doi.org/10.1155/2013/127589> PMID: 23984314; PubMed Central PMCID: PMCPMC3741914.
 39. Fasano A, Catassi C. Clinical practice. Celiac disease. *N Engl J Med*. 2012; 367(25):2419–26. <https://doi.org/10.1056/NEJMcp1113994> PMID: 23252527.
 40. Bono MR, Tejon G, Flores-Santibanez F, Fernandez D, Roseblatt M, Sauma D. Retinoic Acid as a Modulator of T Cell Immunity. *Nutrients*. 2016; 8(6). <https://doi.org/10.3390/nu8060349> PMID: 27304965; PubMed Central PMCID: PMCPMC4924190.
 41. Tejon G, Manriquez V, De Calisto J, Flores-Santibanez F, Hidalgo Y, Crisostomo N, et al. Vitamin A Impairs the Reprogramming of Tregs into IL-17-Producing Cells during Intestinal Inflammation. *Biomed Res Int*. 2015; 2015:137893. <https://doi.org/10.1155/2015/137893> PMID: 26583087; PubMed Central PMCID: PMCPMC4637025.
 42. Raverdeau M, Breen CJ, Misiak A, Mills KH. Retinoic acid suppresses IL-17 production and pathogenic activity of gammadelta T cells in CNS autoimmunity. *Immunol Cell Biol*. 2016; 94(8):763–73. <https://doi.org/10.1038/icb.2016.39> PMID: 27089940.
 43. Handono K, Firdausi SN, Pratama MZ, Endharti AT, Kalim H. Vitamin A improve Th17 and Treg regulation in systemic lupus erythematosus. *Clin Rheumatol*. 2016; 35(3):631–8. <https://doi.org/10.1007/s10067-016-3197-x> PMID: 26852315.
 44. Shi Y, Wang H, Su Z, Chen J, Xue Y, Wang S, et al. Differentiation imbalance of Th1/Th17 in peripheral blood mononuclear cells might contribute to pathogenesis of Hashimoto's thyroiditis. *Scand J Immunol*. 2010; 72(3):250–5. <https://doi.org/10.1111/j.1365-3083.2010.02425.x> PMID: 20696023.
 45. Wang Y, Zhong YJ, Wang YY, Xing J, Wang ZM. All-trans retinoic acid prevents the development of type 1 diabetes by affecting the levels of interferon gamma and interleukin 4 in streptozotocin-induced murine diabetes model. *Genet Mol Res*. 2016; 15(1). <https://doi.org/10.4238/gmr.15017522> PMID: 27050967.
 46. Kluger MA, Nosko A, Ramcke T, Goerke B, Meyer MC, Wegscheid C, et al. RORgamma expression in Tregs promotes systemic lupus erythematosus via IL-17 secretion, alteration of Treg phenotype and suppression of Th2 responses. *Clin Exp Immunol*. 2017; 188(1):63–78. <https://doi.org/10.1111/cei.12905> PMID: 27880975; PubMed Central PMCID: PMCPMC5343349.

47. Ventura A, Magazzu G, Greco L. Duration of exposure to gluten and risk for autoimmune disorders in patients with celiac disease. SIGEP Study Group for Autoimmune Disorders in Celiac Disease. *Gastroenterology*. 1999; 117(2):297–303. PMID: [10419909](#).
48. Ventura A, Neri E, Ughi C, Leopaldi A, Citta A, Not T. Gluten-dependent diabetes-related and thyroid-related autoantibodies in patients with celiac disease. *J Pediatr*. 2000; 137(2):263–5. <https://doi.org/10.1067/mpd.2000.107160> PMID: [10931424](#).
49. Cosnes J, Cellier C, Viola S, Colombel JF, Michaud L, Sarles J, et al. Incidence of autoimmune diseases in celiac disease: protective effect of the gluten-free diet. *Clin Gastroenterol Hepatol*. 2008; 6(7):753–8. <https://doi.org/10.1016/j.cgh.2007.12.022> PMID: [18255352](#).
50. Amin R, Murphy N, Edge J, Ahmed ML, Acerini CL, Dunger DB. A longitudinal study of the effects of a gluten-free diet on glycemic control and weight gain in subjects with type 1 diabetes and celiac disease. *Diabetes Care*. 2002; 25(7):1117–22. PMID: [12087007](#).
51. Hansen D, Brock-Jacobsen B, Lund E, Bjorn C, Hansen LP, Nielsen C, et al. Clinical benefit of a gluten-free diet in type 1 diabetic children with screening-detected celiac disease: a population-based screening study with 2 years' follow-up. *Diabetes Care*. 2006; 29(11):2452–6. <https://doi.org/10.2337/dc06-0990> PMID: [17065683](#).
52. Narula P, Porter L, Langton J, Rao V, Davies P, Cummins C, et al. Gastrointestinal symptoms in children with type 1 diabetes screened for celiac disease. *Pediatrics*. 2009; 124(3):e489–95. <https://doi.org/10.1542/peds.2008-2434> PMID: [19706580](#).
53. Scanlon SA, Murray JA. Update on celiac disease—etiology, differential diagnosis, drug targets, and management advances. *Clin Exp Gastroenterol*. 2011; 4:297–311. <https://doi.org/10.2147/CEG.S8315> PMID: [22235174](#); PubMed Central PMCID: [PMCPMC3254208](#).
54. Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease: time for a standardized report scheme for pathologists. *Eur J Gastroenterol Hepatol*. 1999; 11(10):1185–94. PMID: [10524652](#).
55. Atanasio A, Decman V, White D, Ramos M, Ikiz B, Lee HC, et al. C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci Rep*. 2016; 6:23204. <https://doi.org/10.1038/srep23204> PMID: [26979938](#); PubMed Central PMCID: [PMCPMC4793236](#).
56. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012; 9(7):671–5. PMID: [22930834](#); PubMed Central PMCID: [PMCPMC5554542](#).
57. Shu J, Qiu G, Mohammad I, editors. A semi-automatic image analysis tool for biomarker detection in immunohistochemistry analysis. *Image and Graphics (ICIG), 2013 Seventh International Conference on*; 2013: IEEE.
58. Bullard JH, Purdom E, Hansen KD, Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics*. 2010; 11:94. <https://doi.org/10.1186/1471-2105-11-94> PMID: [20167110](#); PubMed Central PMCID: [PMCPMC2838869](#).
59. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010; 11(10):R106. <https://doi.org/10.1186/gb-2010-11-10-r106> PMID: [20979621](#); PubMed Central PMCID: [PMCPMC3218662](#).
60. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012; 16(5):284–7. <https://doi.org/10.1089/omi.2011.0118> PMID: [22455463](#); PubMed Central PMCID: [PMCPMC3339379](#).