Journal Pre-proof

Immunomodulatory effects of a short-term gluten-free diet on pediatric celiac disease: findings from a single-cell transcriptomics study

Rafael Martín-Masot , Nerea Correa-López , Marta Herrador-López , Víctor Manuel Navas-López , Francisco David Carmona , Teresa Nestares , Lara Bossini-Castillo

PII: S0955-2863(25)00225-6
DOI: https://doi.org/10.1016/j.jnutbio.2025.110063
Reference: JNB 110063

To appear in: *The Journal of Nutritional Biochemistry*

Received date: 25 April 2025
Revised date: 21 July 2025
Accepted date: 8 August 2025

Please cite this article as: Rafael Martín-Masot , Nerea Correa-López , Marta Herrador-López , Víctor Manuel Navas-López , Francisco David Carmona , Teresa Nestares , Lara Bossini-Castillo , Immunomodulatory effects of a short-term gluten-free diet on pediatric celiac disease: findings from a single-cell transcriptomics study, *The Journal of Nutritional Biochemistry* (2025), doi: https://doi.org/10.1016/j.jnutbio.2025.110063

Highlights

- Single-cell RNA-seq revealed that a gluten-free diet downregulated pro-inflammatory genes and upregulated immune-regulatory genes in peripheral blood mononuclear cells in pediatric celiac disease.

- Regulatory T cells showed increased *HLA-G* and decreased *DDX5* expression post gluten-free diet, indicating enhanced suppressive function.

- Novel non-invasive biomarkers for monitoring immune modulation in pediatric celiac disease were identified, but validation in larger studies is required due to the limited sample size.

Immunomodulatory effects of a short-term gluten-free diet on pediatric celiac disease: findings from a single-cell transcriptomics study

Rafael Martín-Masot[1,2,3], Nerea Correa-López[4], Marta Herrador-López[1], Víctor Manuel Navas-López[1], Francisco David Carmona[5,6], Teresa Nestares[2,4,†], Lara Bossini-Castillo[5,6,†].

1-) Sección de Gastroenterología y Nutrición Infantil. Hospital Regional Universitario de Málaga, Málaga, Spain.

2-) Instituto de Nutrición y Tecnología de los Alimentos "José Mataix Verdú" (INYTA), Centro de Investigación Biomédica (CIBM). Universidad de Granada, Granada, Spain.

3-) Departamento de Farmacología y Pediatría. Facultad de Medicina. Universidad de Málaga, Málaga, Spain.

4-) Departamento de Fisiología. Facultad de Farmacia. Universidad de Granada, Granada, Spain

5-) Departamento de Genética e Instituto de Biotecnología, Centro de Investigación Biomédica (CIBM). Universidad de Granada, Granada, Spain.

6-) Reproducción Humana y Enfermedades Hereditarias y Complejas (IBS-TEC14). Terapias Avanzadas y Tecnologías Biomédicas. Instituto de Investigación Biosanitaria de Granada, ibs.GRANADA, Granada, Spain.

† These authors share senior authorship

Correspondence to:

Lara Bossini-Castillo PhD

Departamento de Genética e Instituto de Biotecnología, Universidad de Granada, Centro de Investigación Biomédica (CIBM), Parque Tecnológico Ciencias de la Salud, Avenida del Conocimiento, s/n, 18016, Armilla (Granada), Andalucía, Spain.

Email: lbossinicastillo@ugr.es

Rafael Martín-Masot MD PhD

Departamento de Farmacología y Pediatría. Facultad de Medicina. Universidad de Málaga, Campus Teatinos, Blvr. Louis Pasteur, 32, Puerto de la Torre, 29010 Málaga, Spain.

Email: rafammgr@uma.es

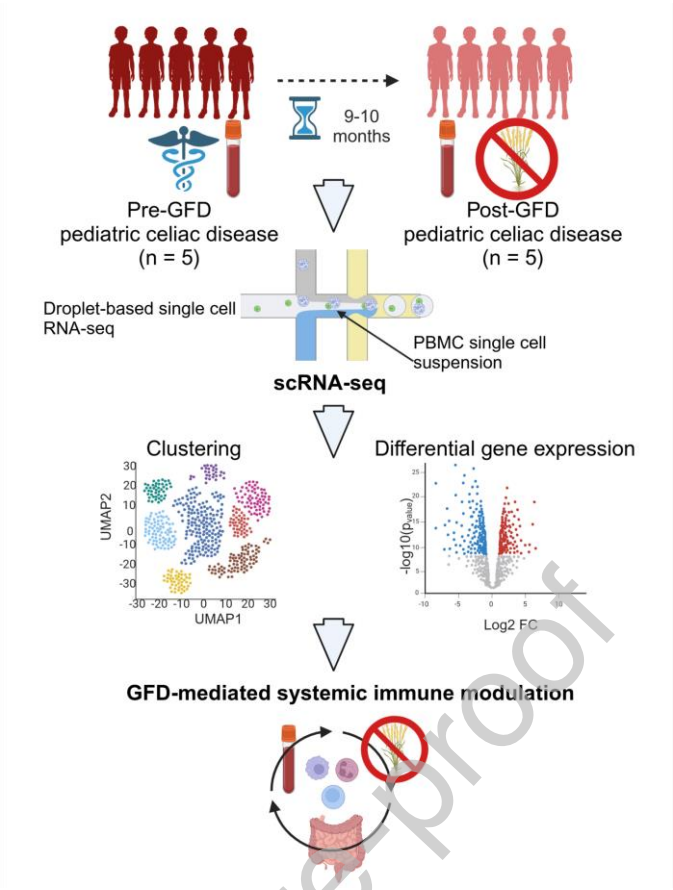
Keywords: celiac disease, Gluten-Free Diet, immune modulation, scRNA-seq, regulatory T cells**.**

Abbreviations: APCs (antigen-presenting cells); CD (celiac disease); DE (differential expression); DEGs (differentially expressed genes); GFD (Gluten-Free Diet); GIP (gluten immunogenic peptides); HLA (human leukocyte antigen); IEL (intraepithelial lymphocytes); PBMCs (peripheral blood mononuclear cells); QC (quality control); scRNA-seq (single-cell RNA sequencing); Th (T helper cell); Treg (regulatory T cell); UMAP (uniform manifold approximation and projection for dimension reduction).

Support: RMM, TN and LBC obtained funding for this study.

ABSTRACT

Celiac disease (CD) is an autoimmune disorder with a strong genetic component, triggered by gluten ingestion. Although a Gluten-Free Diet (GFD) is the standard treatment, its short-term effects on immune cell modulation in pediatric CD remain largely unexplored. This study aimed to investigate transcriptional changes in peripheral blood mononuclear cells (PBMCs) of pediatric CD patients following a strict GFD for 9–10 months, using single-cell RNA sequencing (scRNA-seq). An observational longitudinal study was conducted on five pediatric CD patients pre-GFD and post-GFD (confirmed by gluten immunogenic peptide determination in feces). PBMCs were analyzed using droplet-based scRNA-seq to identify cluster markers and differentially expressed genes (DEGs) between pre-GFD and post-GFD cells. Nineteen immune cell clusters encompassing a variety of classical immune cell subtypes were identified. Key findings included the downregulation of pro-inflammatory genes and the upregulation of immune-regulatory genes after a GFD in different immune cell subsets. Changes in macrophages and monocytes suggested improved immune balance, while T cells demonstrated a shift towards reduced effector activity. Notably, post-GFD regulatory T cells transitioned into a trajectory towards enhanced immunosuppressive profiles, as evidenced by increased *HLA-G* and decreased *DDX5* expression. A strict short-term GFD induced significant immune modulation in pediatric CD patients, highlighting potential biomarkers for disease monitoring. Nevertheless, due to the small sample size, results should be interpreted with caution, and larger cohort studies are needed for further confirmation and validation. These findings provide insights into the immunological mechanisms of GFD and suggest avenues for non-invasive diagnostic strategies to enhance patient management.



Graphical abtract. Experimental design schematic for peripheral blood mononuclear cells (PBMC) scRNA-seq analysis in pediatric patients with celiac disease before and after a short-term Gluten-Free Diet (GFD).

INTRODUCTION

Celiac disease (CD) is a systemic T-cell mediated autoimmune disorder that primarily affects the small intestine as a result of gluten exposure and unclear environmental factors in genetically predisposed individuals [1,2]. CD results in a reversible inflammatory process in the small intestinal mucosa than can lead to gastrointestinal symptoms such as pain, diarrhea, nausea, and vomiting, among others [3]. However, the manifestations of CD can vary widely, affecting multiple organs and ranging from asymptomatic cases to those with severe symptoms [4,5]. Although there is significant regional variability [6] combined analyses across different study groups showed that 1.4 % of the population have CD-associated autoantibodies (seroprevalence), while a 0.7 % of the population is diagnosed with CD via biopsy [6]. Additionally, this condition exhibits a sex-bias, affecting women more frequently than men (with a 2:1 ratio) [7]. A clear worldwide increase in incidence has also been documented [8]. Interestingly, CD has two characteristic onset peaks: the first occurs shortly after weaning, within the first two years of life, and the second occurs in adulthood, around the second or third decade of life. Moreover, the clinical manifestations differ between children and adults; for example, intestinal symptoms predominate in children under 3 years of age, whereas such symptoms are less common in adults [9].

In terms of genetic susceptibility, a strong association has been described between CD and the expression of certain variants in the human leukocyte antigen class II (HLA class II) system [10]. CD develops in individuals who frequently carry the HLA-DQ2 or HLA-DQ8 haplotypes, which are present in about 40% of the general population. This suggests that while these haplotypes are necessary for the development of the disease, they are not sufficient on their own [11]. Remarkably, these alleles seem to have an allele dose-dependent effect, with homozygous carriers of HLA-DQ2.5 being at higher risk of developing CD than heterozygotes [1].

CD is characterized by an inappropriate immune response that culminates in the destruction of enterocytes and, consequently, atrophy of the intestinal villi [1]. The activation of the innate immune

response, mediated by intraepithelial lymphocytes (IEL), and the adaptive immune response, mediated by CD4+ T cells and B cells, occurs in the small intestine. In CD the presence of IL-15 induces the activation of antigen-presenting cells (APCs) expressing HLA-DQ2 or HLA-DQ8, which also secrete this proinflammatory cytokine. IL-15 binds to CD4+ T cells and promotes their polarization into a T helper 1 (Th1) lymphocyte profile, characterized by the cytokine secretion, including interferon gamma (IFN-γ) and tumor necrosis factor alpha (TNF-α), contributing to inflammation and intestinal permeability. Likewise, B cells differentiate into plasma cells capable of secreting antibodies against tissue transglutaminase (tTG) [1]. On the other hand, regulatory T cells (Treg), whose primary function is maintaining immune regulation, exhibit impaired suppressor function, being unable to control disease pathogenesis effectively [12].

The only treatment currently available for CD, with the exception of refractory CD, is a lifelong GFD. Continuing gluten intake exacerbates clinical symptoms and intestinal damage, thus increasing the risk to certain gastrointestinal cancers. However, strict adherence to GFD leads to histological and clinical remission, reducing complications. Over time, serological parameters stabilize and intestinal villi can regenerate, alleviating symptoms such as malabsorption, diarrhea, and weight loss [13]. Interestingly, it has been reported that CD patients exhibit a disease-specific gene expression profiles in peripheral blood mononuclear cells (PBMCs) despite adherence to a GFD [14]. Moreover, short gluten challenges have been shown to alter the gene expression profiles of adult CD patients [15]. However, little is known about the short term effects of a GFD in pediatric patients or the mechanisms involved in immune balance restoration. The high cost of gluten-free substitute foods, social pressures, travel, meals away from home, etc., make following a GFD challenging for CD patients. Novel, non-invasive methods for detecting dietary transgressions, such as the determination of the presence of gluten immunogenic peptides (GIP) in stool or urine, have revealed that dietary transgressions (voluntary or involuntary) are common and must be managed [16–18]. Therefore, early diagnosis and continuous patient-follow-up are essential to prevent irreversible complications in pediatric patients.

In addition to GFD adherence monitorization, there is room for improvement in CD diagnosis and treatment. Diagnosis is often based on serological markers, such as anti-tTg antibodies, and anti-endomysial antibodies. In some cases, histopathological damage is evaluated using the Marsh classification of intestinal villi and lymphocyte infiltration through biopsy, which is invasive method [19,20]. However, seronegative cases and conditions with similar symptoms continue posing diagnostic challenges [21]. Delayed diagnosis can lead to prolonged immune activation, worsening the prognosis [22]. Consequently, there is a need for non-invasive biomarkers that can detect CD in early stages. In this line, blood is sampled using minimally invasive techniques, which would be very beneficial, especially for very young pediatric patients [23]. Bulk RNA analyses from PBMCs have successfully identified some disease biomarkers in the blood of pediatric patients, but failed to provide cell subtype-specific gene expression data [24]. In contrast, single-cell RNA sequencing (scRNA-seq) techniques, which allow whole transcriptome analysis at single-cell resolution, have revealed previously unknown tissue-specific features in human immune cells [25]. Such analyses have been proven useful in studying CD progression in the gut and identifying cell-specific quantitative trait loci and pre- or post- seroconversion biomarkers in the blood of CD pediatric patients [26–28]. However, this approach has not been applied to characterize the short term immune modulation induced by a GFD.

Therefore, we carried out the largest analysis of individual immune cells, using a droplet-based scRNA-seq technique, to date to characterize the peripheral blood immunological profile of pediatric CD patients undergoing a controlled strict GFD for 9-10 months. Identifying significant transcriptomic changes and immune regulation biomarkers may facilitate early diagnosis and patient monitoring through non-invasive methods, ultimately improving the quality of life of affected individuals.

PATIENTS AND METHODS

Study design and patient description

We designed an observational and longitudinal study in which the immunological gene expression profile of the same 5 pediatric CD patients was analyzed using scRNA-seq pre-GFD and after following a strict GFD for 9-10 months, as confirmed using GIP determination in feces (post-GFD), decreased autoantibody titers in routinary tests and expert clinical routine follow-ups.

GIP help to reliably determine whether the patients were consuming gluten 3-5 days before sampling, which allowed for the precise interpretation of the obtained results by attributing the observed effects to a well-implemented GFD. The presence of GIP in stool samples was assessed using the iVYCHECK GIP Stool kit (Biomedal S.L., Seville, Spain), which has previously shown a sensitivity range of 95–100% and a specificity of 100%, following the manufacturer's instructions [29,30]. Additionally, GIP were measured as an objective biomarker of recent gluten exposure. To improve sensitivity and capture potential loss in adherence we measured GIP twice, 4 months after initiation of the GFD and before the 9-10 month GFD blood extraction (Supplementary Table 1). Moreover, to consider variability between weekdays and weekends, in each GIP determination two separate fecal samples were analysed—one on a school day and one on a non-school day—thus aiming to cover the full behavioral spectrum. Moreover, dietary adherence was assessed using a structured 24-hour dietary recall over three days (including a non-working day), conducted and supervised by a trained dietitian. This allowed for detailed analysis of food intake, including portion sizes and brand-specific gluten-free products. Nevertheless, it should be considered that no current method can definitively confirm long-term dietary adherence with absolute certainty.

All patients were recruited at the Gastroenterology and Pediatric Nutrition Unit of Hospital Regional Universitario de Málaga (Málaga, Spain) by specialized clinicians and staff. The legal guardians of all patients signed an informed written consent form, and all samples were irreversibly anonymized as defined by the EU Directive 2001/20/EC and the applicable national requirements relating to data

protection. The study adhered to the standards set by the International Conference on Harmonization and Good Clinical Practice (ICH-GCP), and to the ethical principles originating in the Declaration of Helsinki. This study was approved by the Ethical Committee 'Comité de Ética de la Investigación Provincial de Málaga' (1977-N-21).

All individuals were of European origin, with an age at diagnosis ranging between 8-14 years. Female patients were overrepresented in the cohort (4 females and 1 male) (**Supplemental Table 1**). The patients fulfilled the ESPGHAN 2020 criteria indicating CD [20]. Exclusion criteria ensured that patients were not suffering from an acute infectious process at the time of sample extraction. In addition, patients with liver, lung, rheumatologic or chronic renal disease were excluded. Likewise, obese patients (according to the criteria of the International Task Force [31]), those with inflammatory bowel disease, diabetes, or those who did not sign the consent form to be included in the study were excluded. The samples were distributed into 2 groups: 1) pre-GFD: individuals at or close to the moment of their diagnosis with CD and, therefore, consuming a gluten-containing diet, and 2) CD group after 9-10 months of GFD (post-GFD): individuals following a strict GFD, with negative AATG-IgA, in addition to two negative determinations of fecal GIP to discard dietary transgressions [32].

Peripheral blood mononuclear cell suspension generation and cryopreservation

Thirty milliliters of peripheral blood from each participant were collected in EDTA tubes (Greiner #4550356) and processed for cryopreservation within 3h of extraction. PBMCs were isolated using Ficoll ® Paque Plus (Merk #GE 17-1440-02) density gradients following the manufacturer's protocol. Next, cell suspensions were cryopreserved in 10% DMSO (Merk D2438) and 90% fetal bovine serum (FBS, Gibco #10082-147) medium and frozen in a −80ºC ultrafreezer at a controlled rate for at least 24 h. Finally, cells were stored in liquid nitrogen for long term preservation.

Generation of scRNA-seq libraries

For the generation of scRNA-seq libraries, cryopreserved cells were thawed, washed and manually counted using a Neubauer counting chamber. Then, living cells underwent the established protocols

for the Next GEM technology by 10x Genomics. Samples were assayed following the manufacturer's instructions for the following kits: Chromium Next GEM Single Cell 5′ Library Chromium Next GEM Chip G Single Cell Kit (10x Genomics, PN-1000127) and Chromium Next GEM Single Cell 5′ Library and Gel Bead Kit v2 (PN-1000265). Pools of cells from randomly distributed samples were prepared to control for batch effects. The generated cDNA was sequenced on a NovaSeq 6000 sequencer. An average of more than 90 % of reads in cells was obtained, with an average of more than 60,000 reads per cell. The GRCh38 genome build was used as an alignment reference and unique molecular identifier (UMI) matrix counts were generated using the Cell Ranger Single Cell Software Suite (v3.0.0 by 10x Genomics) with default parameters. We identified an average of more than 2,200 genes per cell.

Generation of genome-wide genotype data

Genomic DNA samples obtained from the remnants of PBMCs from the scRNA-seq protocol were genotyped at the genome-wide level using the Infinium™ Global Screening Array-24 v3.0 (GSA) in an iScan System (Illumina, Inc), following the manufacturer's protocol. Genotype quality control (QC) was then performed as described elsewhere [33].

Analysis of scRNA-seq data

Each cell was assigned to a donor based on the genotype information described above and following the procedures implemented in the cellSNP and vireo packages for scRNA-seq donor deconvolution [34,35]. Unassigned cells or cells assigned to more than one donor were discarded from further analysis.

The analysis of the generated scRNA-seq dataset was performed using the gold-standard Seurat package v4 [36]. We performed cell-wise QC, and only those cells exceeding a minimum of 200 genes detected (threshold for low transcriptional complexity signal) and a maximum of 3,500 (putative doublet) and less than 5% of reads assigned to mitochondrial genes (cell apoptosis signal) continued in the analysis.

After QC, the matrices of the different pools were normalized, scaled (considering the number of detected genes, the percentage of reads in mitochondrial genes, pool and sex as covariates), and integrated using the Canonical Correlation Analysis method implemented in Seurat. Once all datasets were integrated, the different clusters were identified based on the 2,000 genes showing the most variable expression, the first 30 principal components and high resolution (resolution = 1.2), using the *FindNeighbors* and *FindClusters* (implementing the Louvain algorithm) functions included in Seurat and represented in a UMAP dimension (Uniform Manifold Approximation and Projection for Dimension Reduction) graph. Finally, we analyzed the cluster markers and the differentially expressed genes (DEG) between the post-GFD and pre-GFD time points in the different clusters using a Wilcoxon Rank Sum test with the *FindAllMarkers* and the *FindMarkers* functions, respectively. Each cluster was assigned to classical immunological cell types based on the expression of previously well-established cell markers. We confirmed that donor identity was not significantly influencing DEG by applying a logistic regression framework (test.use = "LR") including donor as a covariate. Only genes showing significant differential gene expression with an adjusted p-value lower than 0.1 were considered statistically significant.

Reclustering was performed by selecting the relevant cells using the *subset* function in Seurat and recalculating HVG, PCAs, distances and clusters considering only the subsetted cells.

Pseudotime analysis was performed using Monocle 3 to investigate the dynamic trajectories of Tregs across identified clusters [37,38]. After pre-processing and dimensionality reduction, cells were ordered along a pseudotemporal axis based on their transcriptional profiles. The root node was selected according to biologically relevant markers indicative of less suppressive Tregs (e.g. *CCR7*, *TCF7*). This approach enabled the reconstruction of potential differentiation pathways, revealing transcriptional changes associated with Treg functional states and highlighting key genes driving their progression along the pseudotime trajectory.

RESULTS

In total, more than 21,000 high quality single cell whole transcriptomes and more than 25,000 genes were analyzed from 5 pediatric patients with CD. We dissected the largest scRNA-seq dataset generated in pediatric CD patients to date (**Supplemental Table 1**) as described below.

The implementation of unsupervised clustering methods allowed us to identify a total of 19 clusters (**Figure 1A**, **Supplemental Table 2**). Moreover, using known cell markers, the main subtypes of immune cells that make up the PBMCs were manually annotated (**Figure 1B, Supplemental Figure 1, Supplemental Tables 3 - 4**). Although we observed that the youngest and HLA-DQ2.5 homozygous patient showed the most relevant effects in cell proportions post-GFD (**Supplemental Table 1**, **Supplemental Figure 2**), overall the cellular distribution and immune cell subtype proportions did not show significant differences before and after the introduction of the GFD or between individuals (**Figure 1C-D**, **Supplemental Figures 2-3**, **Supplemental Table 5**). Interestingly, when comparing gene expression in each of these cell subtypes between cells obtained pre-GFD and post-GFD, a total of 1,082 DEGs were identified, of which 94 corresponded to B lymphocytes, 493 to CD4+ T lymphocytes, 298 to CD8+ T lymphocytes, 79 to γδ T lymphocytes, 643 to macrophages, 80 to NK cells, 23 to NK-T cells, and 8 to regulatory T cells (Tregs) (**Supplemental Table 6**).

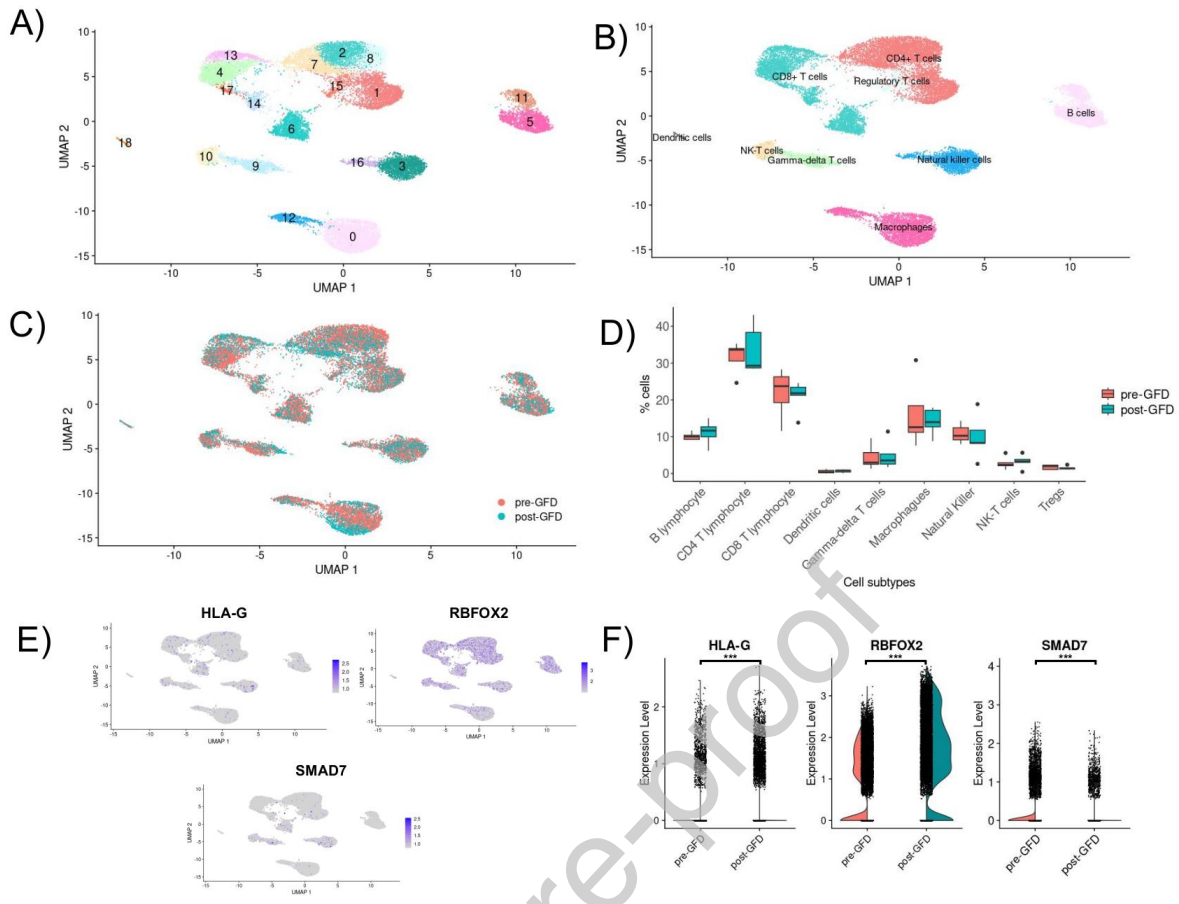


Figure 1. UMAP clustering scatter plot of peripheral blood mononuclear cells (PBMCs) colored by **A)** high definition non-supervised clustering, **B)** classical immunological cell subtypes, **C)** pre-gluten free diet (GFD) or post-GFD. **D)** Boxplots representing percentages of the different cell types pre-GFD and post-GFD. **E)** Uniform Manifold Approximation and Projection (UMAP) colored by single cell expression of 3 DEG between pre-GFD and post-GFD. **F)** Violin plots for 3 DEGs between pre-GFD and post-GFD. ***: Padj < 10^{-15}.

General gene expression changed in response to a Gluten-Free Diet

The comparison of DEGs between the set of all cells obtained from individuals pre-GFD and after following the GFD for 9-10 months yielded relevant results. Among the genes whose expression levels were reduced after the GFD (**Supplemental Table 7**), it is worth highlighting that we observed a decrease in *SMAD7* expression levels (**Figure 1E-F, Supplemental Table 7**). *SMAD7* encodes an intracellular inhibitor of transforming growth factor ß1 (TGF-ß1), which is implicated in the defective function of this factor and promotes the production of inflammatory cytokines in CD [39]. Other

notable repressed genes include *MAFB*, a transcription factor that regulates macrophage differentiation [40].

Furthermore, as shown in **Supplemental Table 7,** an increase in the expression of some genes was identified after the GFD. One of the most notable is *RBFOX2*, a master regulator of alternative splicing of a large number of transcripts involved in cell differentiation and development [41]. The upregulation of *RBFOX2* in different cell subtypes and clusters following the GFD (**Figure 1E-F**, **Supplemental Tables 6 and 8**) suggests a broad impact on the alternative splicing of key immune and cellular function genes, contributing to the modulation of the immune response and tissue recovery in celiac patients.

Another gene highly involved in immune system function and overexpressed after the GFD is *HLA-G*, a molecule capable of inhibiting the action of certain immune cells and inducing immune tolerance (**Figure 1E-F**, **Supplemental Table 7**). However, it is noteworthy that there was also an increase in the expression of genes such as *CCL3*, *CCL3L1*, *CCL4*, and *CCL4L2*, which are involved in the activation of IEL and in chemotaxis [42].

The expression of prediagnostic and active celiac disease markers was affected by a Gluten-Free Diet

One of the key findings of this study is the observation that the expression levels of certain genes, previously proposed as potential prediagnostic or disease activity markers in CD, were modulated following adherence to a GFD. However, it should be noted that these genes are not yet widely recognized or established as clinical biomarkers. For example, *TXNIP* (Thioredoxin-interacting protein), which was proposed as a pre-diagnostic marker for CD in CD4+ T cells and NK cells [27] (**Figure 2**, **Supplemental Table 6**). After GFD, we observed increased expression levels of *TXNIP* in CD4+, CD8+, gamma-delta T cells and B cells. *TXNIP* is a gene that acts as a key modulator of the immune system by influencing redox balance and immune cell function.

In the case of CD4+ T lymphocytes, after a GFD, there was a decrease in the expression of genes involved in cellular stress such as *GADD45B, DUSP1, PPP1R15A,* and *RGCC*. Notably, the expression of *IRF1,* which acts as an activator of genes involved in innate and adaptive immune responses, decreased, while the expression of the anti-inflammatory molecule *HLA-G* increased. Similarly, genes related to both pro-inflammatory and anti-inflammatory contexts, such as *NFKBIA, KLF6,* and *TNFAIP3*, were downregulated (**Supplemental Table 6**).

Regarding CD8+ T lymphocytes, we also observed a dual effect of the GFD (**Supplemental Table 6**). Post-GFD, there was a notable decrease in the expression of stress- and apoptosis-related genes such as *GADD45B, DUSP1, PPP1R15A, DDIT3, HSPA5*, and *ZFP36*, but also a reduction in the expression of anti-inflammatory molecules. *i.e. NFKBIA*, *YPEL5*. Additionally, the expression of *HLA-G* and *DDX17,* which is involved in altering RNA secondary structures and in the assembly of ribosomes and the spliceosome, increased after the cessation of gluten intake.

The NK cell compartment after GFD showed a clear gene expression alteration characterized by a decreased expression of pro-inflammatory or activation molecules such as the previously mentioned *SMAD7*, *TSC22D3* (which appears to mediate the anti-inflammatory and immunosuppressive effects of certain steroids and inhibit pro-inflammatory molecules), and *GNLY* (which encodes a protein localized within cytotoxic granules and is released upon antigen recognition). However, these cells also showed an increased expression of pro-inflammatory-, activation-, or migration-related genes, such as *JAK1, TLN1, CD69,* and *SKAP1* (**Supplemental Table 6**).

A similarly striking profile was observed in macrophages after the GFD, characterized by increased expression of genes related to the production of pro-inflammatory cytokines and inflammatory processes, such as *CCL4, CCL3, TNFAIP6, IL1B, TNF, CCL3L1, CCL4L2*, *NFKBIZ,* and *NLRP3*. However, among the upregulated DEGs post-GFD, we also observed anti-inflammatory *loci* such as *NFKBIA.* Moreover, several proinflammatory genes were also downregulated after GFD, for instance, *KLF2* (implicated in T-cell viability and inflammation), *PTGER4* (responsible for activating T-cell signaling factors), and *ANXA1* (associated with inflammation). Nevertheless, we observed a significant

decrease in *MAFB*, which might be related to reduced anti-inflammatory M2-polarization [40] (**Supplemental Table 6**).

In summary, the results of the analysis of classical immune cell subtypes revealed that adherence to a GFD significantly modulated the gene expression of various prediagnostic and active CD markers, even over a relatively short period. Specifically, the GFD impacted the regulation of genes involved in immune system function, affecting both pro-inflammatory and anti-inflammatory immune responses. However, while these genetic changes suggest modulation of the immune profile, the phenotypic consequences at the level of immune cells and their clinical impact within the context of CD still require further investigation. Consequently, we decided to analyze the contribution of high-definition clusters to identify specific sets of cells that might mediate the clinically observed immunomodulation associated with long-term GFD.

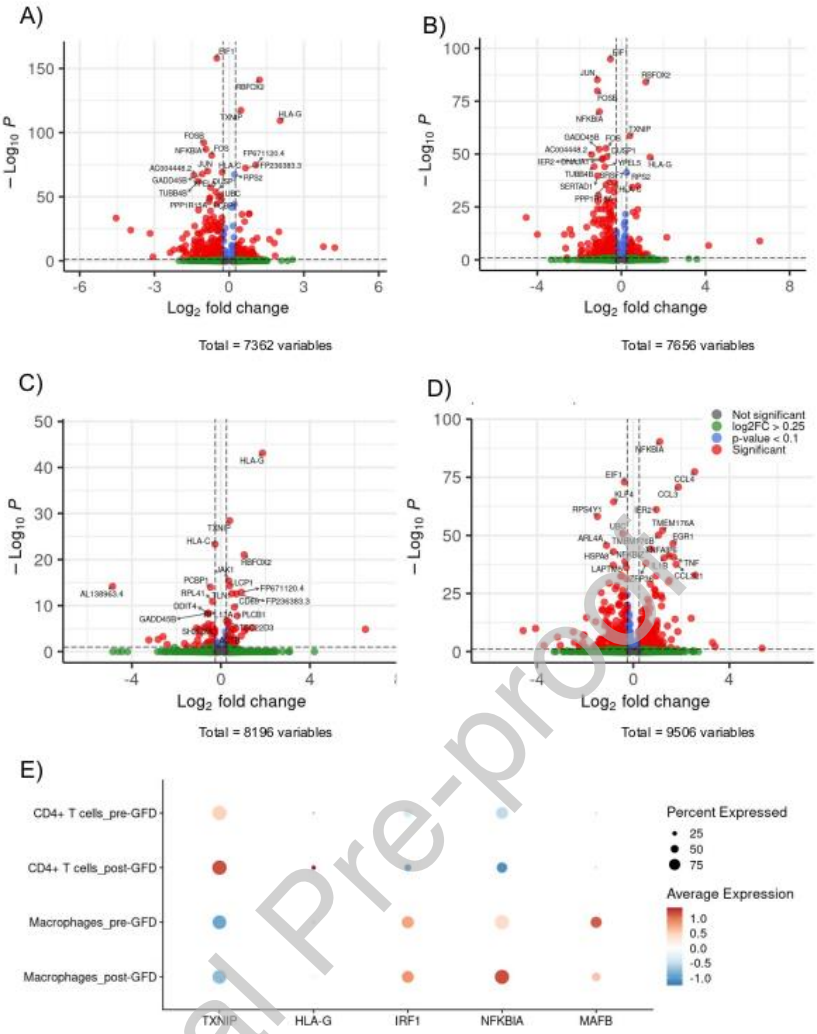


Figure 2. Volcano plots illustrating the DEGs between post-GFD and pre-GFD in **A)** CD4+ T cells, **B)** CD8+ T cells, **C)** Natural killer cells and **D)** Macrophages. **E)** Dotplot showing the expression of several DEGs post-GFD and pre-GFD in CD4+ T cells and macrophages.

Gluten-Free Diet mediates a shift towards a less effector phenotype in T cells

Although the overall frequency of cells in the different clusters was not significantly affected by the GFD (**Supplemental Figure 2**), we observed two CD4+ T cell clusters that showed the largest differences between pre-GFD and post-GFD: cluster 7 and cluster 8 (**Supplemental Figure 2**, **Supplemental Table 5**).

Both cell clusters shared the expression of naïve T cell markers such as *CCR7* and *TCF7* (**Supplemental Table 2**). However, we observed that, compared to cluster 8, cluster 7 was overrepresented post-GFD (**Figure 3A**). This overrepresentation of cluster 7 was particularly evident in the youngest and the only HLA-DQ2.5 homozygous donor (CEL_05), but the tendency toward a more populated cluster 7 and especially a reduced cluster 8 after a GFD was shared by the majority of the patients (**Figure 3A**). Cluster 7 exhibited a decrease in the expression of several relevant genes for CD and T cell activation and differentiation (**Figure 3B**, **Supplemental Table 9**). For example, the expression of early activation T cell markers, such as *CD69*, was reduced in this subset of CD4+ T cells. Moreover, this cluster showed a downregulation of *TNFAIP3*, which is a known genetic risk factor for CD with increased expression levels in adult untreated CD PBMCs and in early stages of differentiation predisposed to a pro-inflammatory Th17 phenotype [43-45]. We also observed that *CXCR4*, which is associated with TCR-mediated activation [46], and *CREM*, which is known to induce an inflammatory phenotype in oligoarticular juvenile idiopathic arthritis [47], had decreased expression after a GFD. Finally, the most significant gene expression reduction observed in cluster 7 corresponded to *NR4A3* (log2FC = -3.723), which has been reported as an early-responding gene after T cell receptor stimulation in gluten-specific CD4+ T cells [48].

In the CD8+ T cell compartment, we observed a shared tendency to increased proportions of a specific naïve CD8+ T cell cluster (cluster 13) after the short term GFD, which is consistent with a reduction in the numbers of effector cells. Again this tendency was distinctly evident for the youngest and HLA-DQ2.5 homozygous donor, but shared by the majority of the other patients (**Figure 3A**). Interestingly, the post-GFD cells showed a less activated phenotype with reduced expression levels of *FOS*, *FOSB*, and *KLF2* (**Figure 3C**, **Supplemental Table 8**).

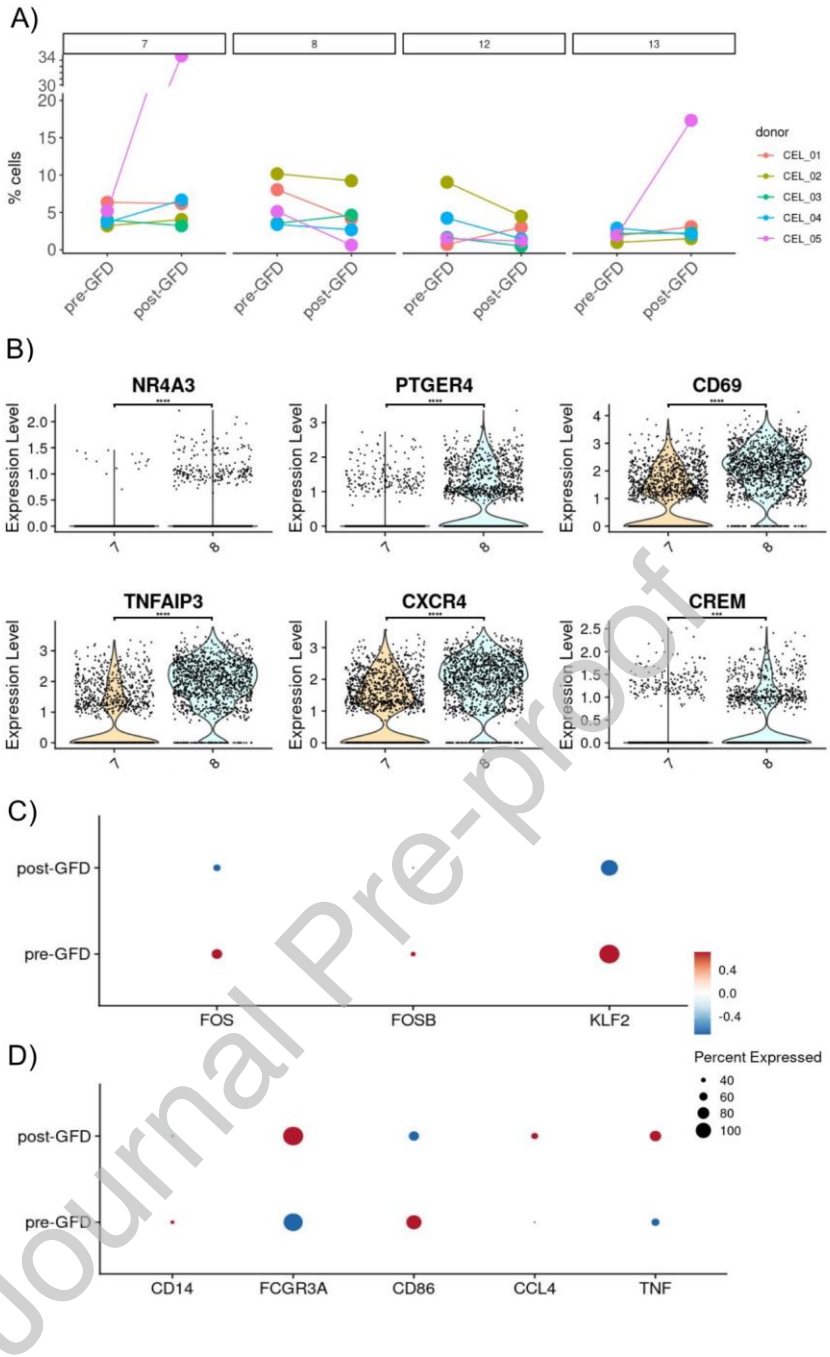


Figure 3. A) Scatter plot showing cell proportion changes in different clusters for each donor pre-GFD and post-GFD, **B)** Violin plots of DEGs in the comparison between clusters 7 and 8 (CD4+ T cells). Dot plots showing **C)** DEGs in cluster 13 comparison between pre-GFD and post-GFD (CD8+ T cells), and **D)** DEGs in cluster 12 comparison between pre-GFD and post-GFD (non-classical macrophages). ****: Padj < 10^{-90}; ***: Padj < 10^{-40}.

Non-classical monocyte frequency in peripheral blood is decreased post-GFD

We observed that the only monocyte cluster that showed a slightly decreased frequency after a GFD, shared by the majority of the rest was cluster 12 (**Figure 3A**). This monocytic cluster was predominantly composed of non-classical monocytes ($CD14^{low}$, $CD16^{high}$), which are known to be affected by lifestyle and dietary habits. We observed a significant reduction post-GFD (**Figure 3D**). Although the role of this monocyte subset is often associated with a patrolling and anti-inflammatory profile, increased proportions of non-classical monocytes in the blood of patients with several chronic immune-mediated disorders have been described [49,50]. Therefore, this decrease in non-classical monocyte numbers might be linked to a better immune balance in the monocyte compartment. Moreover, we observed that post-GFD the remaining non-classical monocytes showed increased expression of *CD16* and decreased expression of genes associated with activation, such as *CD86*. Nevertheless, we observed increased expression of migration markers, *i.e. CCL4* and *TNF* (**Figure 3D**, **Supplemental Table 8**) [51], which might reflect that these cells are actively migrating to affected locations to promote inflammation resolution.

Treg function might be enhanced by the Gluten-Free Diet

After high resolution clustering, we were able to identify a unique cluster (cluster 15) including almost 400 CD4+ CD25+ FOXP3+ Tregs in the dataset (**Figure 4A**, **Supplemental Figure 1**, **Supplemental Tables 2 and 4**). Remarkably, out of the observed pre-GFD versus post-GFD DEGs, *HLA-G* stands out as the most significantly upregulated gene and *DDX5* as the most downregulated after the GFD (**Figure 4B**, **Supplemental Table 8**). Although neither gene has been studied in the context of CD before, we consider that this post-GFD expression profile reflects a more immunoregulatory phenotype that would likely be linked with the amelioration of symptoms when gluten is absent in the diet. In fact, the expression of HLA-G in T cells and even a specific Treg subset not expressing FOXP3 but with high levels of HLA-G expression has been identified as a relevant player in pregnancy and immune-mediated diseases as they exert immunosuppression by different mechanisms [52,53]. Moreover, DDX5 has been shown to restrict the suppression function of $ROR\gamma t^{+}$ Tregs, which was directly linked with increased gut inflammation in mice [54].

Finally, we performed reclustering of Tregs and identified 3 different Treg clusters (**Figure 4C**). Although the cell proportion differences were not statistically significant after a GFD, we found a trend toward enrichment in clusters 0 and 2, which showed pro-Treg development and intermediate profiles, respectively (**Figure 4D**). On the contrary, we observed a decrease in cluster 1, which was related to the expression of anti-Treg development genes (**Figure 4E**). For example, cluster 0 was characterized by high levels of *IL32*, *CD74*, *CD52*, *DUSP4*, and a variety of HLA class II genes. Conversely, in cluster 1, we identified increased levels of *TCF7* (**Supplemental Table 10**). Interestingly, we observed that these cluster markers seemed to be expressed in a gradient that could be related to the suppressive functions of the Tregs. Finally, we discovered the existence of a trajectory underlying this dynamic process that, starting from cluster 1, transitioned through cluster 2, which would be associated with an intermediate suppressive state, and finalized in cluster 0 with an enhanced immuno-modulatory capacity (**Figure 4F**). Moreover, we observed that after a GFD, the cells that belonged to cluster 1 showed a tendency toward a more suppressive profile compared to their pre-GFD counterparts (**Figure 4G**).

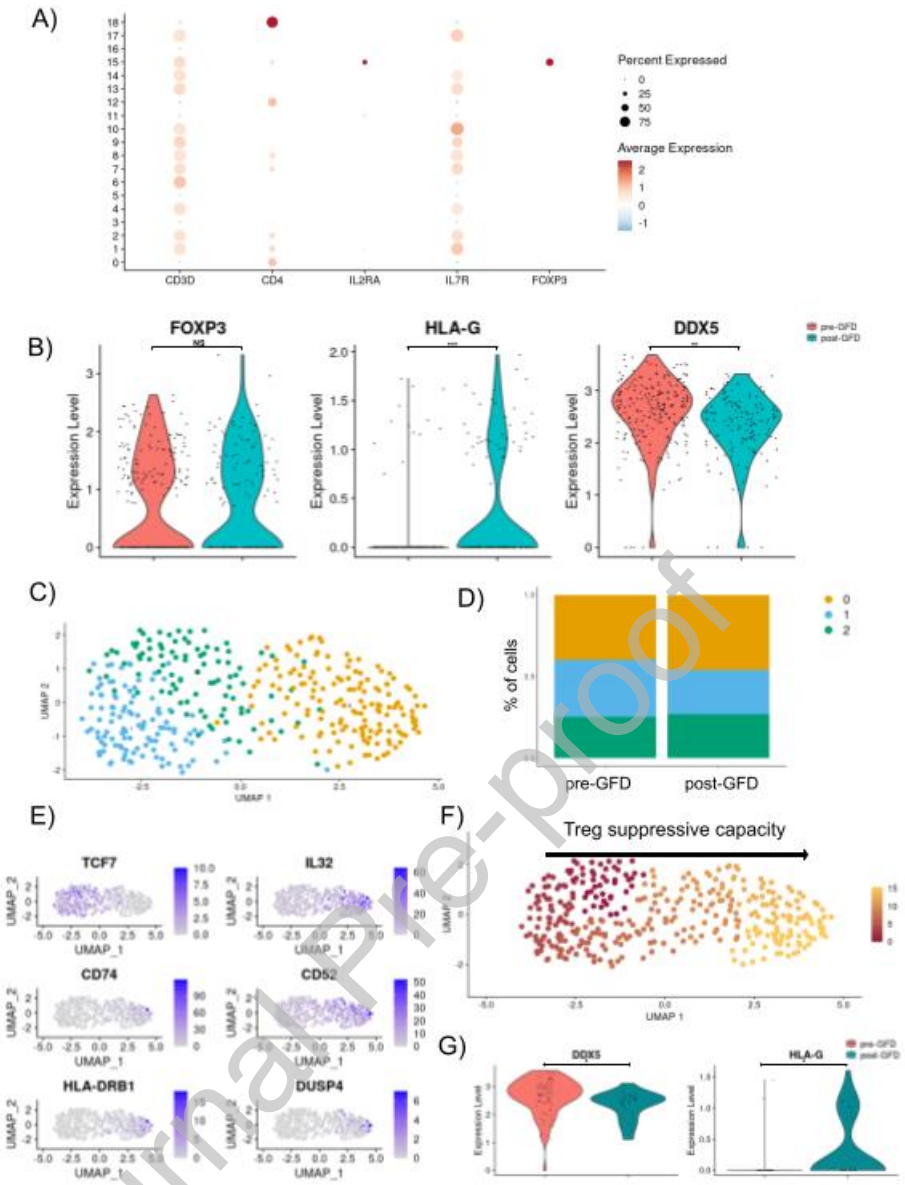


Figure 4. Regulatory T cell (Treg) characteristics in CD pre- and post-GFD after reclustering. **A)** Dot plot showing the expression of classical Treg markers in the 19 high-definition clusters, **B)** Violin plots for 3 DEGs between pre-GFD and post-GFD in Tregs, **C)** UMAP clustering scatter plot colored by Treg clusters after reclustering, **D)** Cell proportions of each Treg cluster pre-GFD and post-GFD, **E)** UMAP clustering scatter plots colored by single cell expression of 6 Treg cluster markers, **F)** UMAP clustering scatter plot colored by pseudotime, **G)** Violin plots for 2 DEG between pre-GFD and post-GFD in cluster 1 Tregs. ***: Padj < 10^{-3}; **: Padj < 0.1; *P < 10^{-3} but Padj > 0.1; NS: not significant.

DISCUSSION

The GFD is the only treatment for patients with CD, and numerous studies have demonstrated its long-term clinical benefits. However, the specific impact of GFD on immune responses in children shortly after diagnosis remained largely unexplored. Traditionally, the effects of gluten in patients have been studied through controlled gluten exposures (challenges), focusing primarily on the alterations in the gut or in gluten-specific T cells in peripheral blood. Our study is the first to analyze the effects of a GFD for 9-10 months after diagnosis under natural conditions. Remarkably, we showed that the application of scRNA-seq to CD PBMCs demonstrated the potential of this technology to identify biomarkers for monitoring immune responses and supporting the follow-up of CD patients through minimally invasive methods.

Among the various immune cell types analyzed, T cells exhibited the most pronounced transcriptional alterations. Although we observed a general downregulation of inflammation-related genes in T cells, indicating a potential reduction in inflammatory responses, it is noteworthy that certain pro-inflammatory genes (such as *CCL3*, *CCL4*, *TNF*, and *IL1B*) showed increased expression in specific immune cell subsets after GFD. This apparent paradox suggests that immune modulation following GFD is complex, involving both anti-inflammatory and residual or adaptive pro-inflammatory signals. We would like to highlight that the overexpression of *HLA-G*, observed across several groups, emerged as a potential key biomarker of GFD-mediated immunomodulation. Noteworthy, the expression of *HLA-G* in CD contexts has been observed to help restore tolerance to dietary gluten and different polymorphisms in the 3′ untranslated and 5′ upstream regulatory regions of this gene, which control its expression levels, have been associated with susceptibility to CD [55-57]. Moreover, HLA-G, present in both membrane-bound and soluble forms, is known to suppress inflammatory and immune responses by 1) inhibiting cytotoxic CD8+ T cells and NK cells, 2) inducing their apoptosis, 3) suppressing CD4+ T cell proliferation while promoting an immune-suppressive phenotype, 3) impairing the function of antigen-presenting cells and B cell differentiation, 4) shifting the immune response toward a Th2 profile, and 5) promoting the development of regulatory T cells and IL-10-secreting dendritic cells [58].

Recent studies further reinforce the central role of IL-10-producing regulatory cells in celiac disease. Passerini et al. demonstrated that IL-10-secreting dendritic cells (DC-10) and Tr1 cells are fundamental in controlling pathological T-cell responses to gluten, protecting the intestinal mucosa from damage and representing a marker of potential celiac disease. Their data suggest that the accumulation of IL-10-producing cells in the gut after a GFD contributes to the restoration of immune homeostasis, and that the IL-10 axis is a critical checkpoint in disease evolution. Moreover, innovative approaches using tolerogenic dendritic cells engineered to secrete IL-10 have shown efficacy in dampening antigen-specific T cell responses and inducing Tr1 cells with regulatory phenotypes, both in vitro and in preclinical models. These findings open new avenues for antigen-specific tolerance induction and support the translational potential of targeting IL-10 pathways in CD management [59-60].

Moreover, the upregulation of *RBFOX2* after GFD may have several functional consequences, for example: shifting the immune system towards a more balanced state, (reducing inflammation and allowing for tissue recovery in the gut), influencing the splicing of genes related to cell adhesion and migration (important in restoring the integrity of the intestinal epithelium), and modulating the expression of splicing variants related to stress and inflammatory signaling pathways (contributing to the reduced inflammatory profile observed after GFD). Our findings also show that the levels of CD-associated molecules such as *TXNIP* remain elevated after a short-term GFD [27]. We consider that the increase in *TXNIP* levels after-GFD might reflect a reduced effector capacity of T cells as naive human T cells exhibit higher *TXNIP* production than stimulated T cells [61]. Moreover, it has been reported that re-expression of *TXNIP* and consequently inhibition of Trx1 are important to restrain late T-cell expansion and that the absence of TXNIP leads to increased B cell response in mice [62]. Consequently, the transition to restored immune balance seems to be complex and disease-specific. Studying the balance between pro-inflammatory and anti-inflammatory signals in different dietary scenarios will be crucial to better understand the pathophysiology of CD, as our data suggest that even after a strict GFD, low-grade inflammation and heightened sensitivity to pro-inflammatory stimuli may persist. This may reflect intrinsic immune system features in CD that are not fully normalized by

dietary intervention alone, reinforcing the importance of personalized patient management and the development of targeted therapeutic strategies.

One of the most critical findings of this research is the downregulation of *DDX5*, a transcriptional corepressor involved in anti-inflammatory pathways. We observed that the expression of *DDX5* was reduced post-GFD in several lymphocyte populations, including B cells, CD8+ T cells, CD4+ T cells and Tregs. We hypothesize that this reduction might be particularly relevant in Tregs, as reduced expression of *DDX5* in RORγt Tregs has been associated with increase in their immunosuppressive capabilities [54]. Although we were unable to identify a distinct RORγt Treg cluster, likely due to the low numbers of these cells in PBMCs, the GFD-induced repression of *DDX5* expression in Tregs was evident. Given that Tregs are crucial in immune-mediated diseases by maintaining immune tolerance and preventing excessive immune responses [63,64], *DDX5* could represent a therapeutic target in CD. Interestingly, DDX5 has been implicated in the specification and function of the anti-inflammatory tuft cells to regulate microbial repertoire, suggesting its relevance also in gut-related contexts [37]. Exploring *DDX5* inhibition to enhance Treg function in CD patients and tracking it as a biomarker for CD activity or early detection warrant further investigation.

In addition to *DDX5*, we observed substantial transcriptional changes in Tregs following adherence to the GFD. There was a notable decrease in the expression of inflammation-related genes, while genes associated with immune tolerance and the inhibition of effector cell populations were upregulated. We identified the existence of a pseudotime associated with an increased suppressive capacity post-GFD, driven by the expression gradients of different genes that have been linked to Treg suppressive abilities. For example, *IL32* is highly expressed in tumor-infiltrating Tregs within esophageal squamous cell carcinoma, promoting *FOXP3* expression in CD4+ T cells [65]. Moreover, *CD74* is selectively overexpressed in tumor-infiltrating Tregs, where it plays a critical role in maintaining their function and stability, with its deletion impairing Treg cytoskeletal organization, Foxp3 expression, and tumor accumulation, ultimately accelerating tumor rejection in preclinical models [66]. Likewise, *CD52* and *DUSP4* have been identified as markers of highly suppressive Treg subsets [67,68]. Finally, the increasing gradient of HLA class II transcripts, known to be expressed in activated effector T cells

and Tregs [69], indicates enhanced immune regulation. Conversely, as the pseudotime advanced, we observed a decreasing gradient of *TCF7* expression, known to limit the development of Tregs [70], which is indicative of a shift toward a more tolerant immune environment. These findings indicate the generation of GFD-induced Treg subsets with increased suppressive power, which may help mitigate the harmful effects of gluten exposure in susceptible individuals.

Regarding macrophages, we observed a slight decrease in non-classical anti-inflammatory monocyte proportions and reduced M2 polarization markers. Although these proportion differences were not statistically significant, we found a significant increase in migration marker expression post-GFD in this cluster. This might indicate active migration of anti-inflammatory monocytes to sites of inflammation, contributing to immune balance restoration after a GFD. Such recruitment of disease-relevant monocytes to affected tissue and depletion of their populations in the blood to maintain immune homeostasis is well-established and has been observed in other immune-mediated diseases [71,72]. If confirmed in CD post-GFD in the long term, it could be used as a relevant biomarker patient monitorization in the future.

Despite the relevant findings reported here, our study has several limitations that should be considered. First, individual variability and a limited sample size challenge the generalization of results. Notably, the youngest and HLA-DQ2.5 homozygous donor displayed significant deviations in cellular profiles, especially in cluster 7. However, these differences might be due to age-related variability or other known (such as HLA-DQ alleles) or unknown parameters. But pairing pre-GFD and post-GFD samples from each donor likely ensured consistent tendencies by controlling for interindividual differences and the general tendencies and expression profiles were shared between all the patients. Secondly, the low proportion of Tregs limited the identification of specific clusters and insights into their heterogeneity. Larger and more homogeneous cohorts are needed to validate these findings. Additionally, a direct functional validation of the transcriptomic findings was beyond the scope of this study. While scRNA-seq provides valuable insights into potential regulatory mechanisms, functional assays are necessary to determine the actual impact of these gene expression changes on immune cell function. Future research should address this gap by combining transcriptomic profiling with targeted

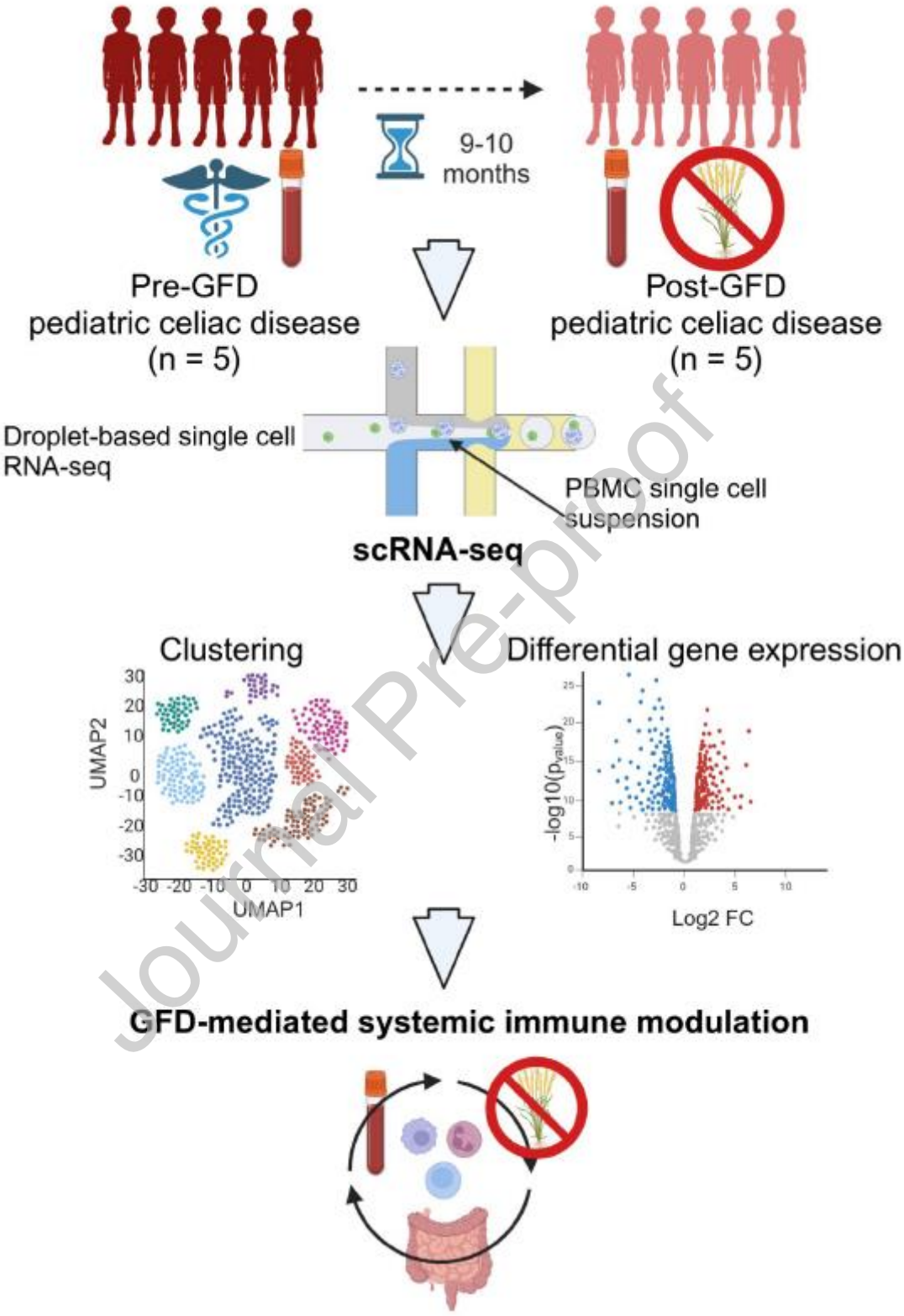
functional experiments to better elucidate the mechanisms underlying immune modulation after a GFD.

In summary, the application of scRNA-seq has uncovered significant transcriptional changes related with immune-modulation in immune cells following a GFD in patients with CD. Importantly, the observed upregulation of certain pro-inflammatory genes post-GFD nuances the overall immunoregulatory narrative, highlighting that immune restoration is a dynamic and potentially incomplete process in the short term. These findings underscore the need for cautious interpretation of transcriptomic changes and further studies to clarify the clinical significance of persistent or fluctuating inflammatory gene expression after dietary intervention. The identification of novel biomarkers for CD monitoring, such as *HLA-G* and *DDX5*, can improve tracking and reduce complications. Nevertheless, a critical assessment of the identified biomarkers highlights several challenges. Translational application will depend on demonstrating that these biomarkers outperform or complement existing serological and clinical tools. Furthermore, analytical validation (ensuring accurate and reproducible measurement) and clinical validation (proving association with meaningful clinical outcomes) are essential steps before clinical implementation. Finally, several necessary steps must be addressed to advance toward clinical utility: validation in larger, independent cohorts to confirm the reproducibility and robustness of the findings; inclusion of patients across a broader age range to ensure the biomarkers are applicable throughout the pediatric and potentially adult spectrum; assessment of different durations of adherence to a GFD, which would help determine the temporal dynamics and stability of the biomarkers; longitudinal studies to evaluate the performance of these biomarkers over time and in various clinical scenarios. Only through such comprehensive validation and extension of the study population can these candidate biomarkers be reliably translated into clinical practice for celiac disease monitoring and diagnosis.

Ultimately, our goal is to enhance the quality of life for CD patients by identifying novel biomarkers, such as *HLA-G* and *DDX5* among others, that reflect immune modulation following a GFD. The measurement of these biomarkers could support non-invasive monitoring of disease activity and remission, and help stratify patients according to their immunoregulatory response to dietary

intervention. Our findings suggest that a panel of transcriptomic markers, including *HLA-G*, may provide greater sensitivity and specificity for clinical applications. These results lay the groundwork for future translational research aimed at validating and integrating these biomarkers into personalized management and therapeutic strategies for CD.

FIGURE LEGENDS



Graphical abstract. Experimental design schematic for peripheral blood mononuclear cells (PBMC) scRNA-seq analysis in pediatric patients with celiac disease before and after a short-term Gluten-Free Diet (GFD).

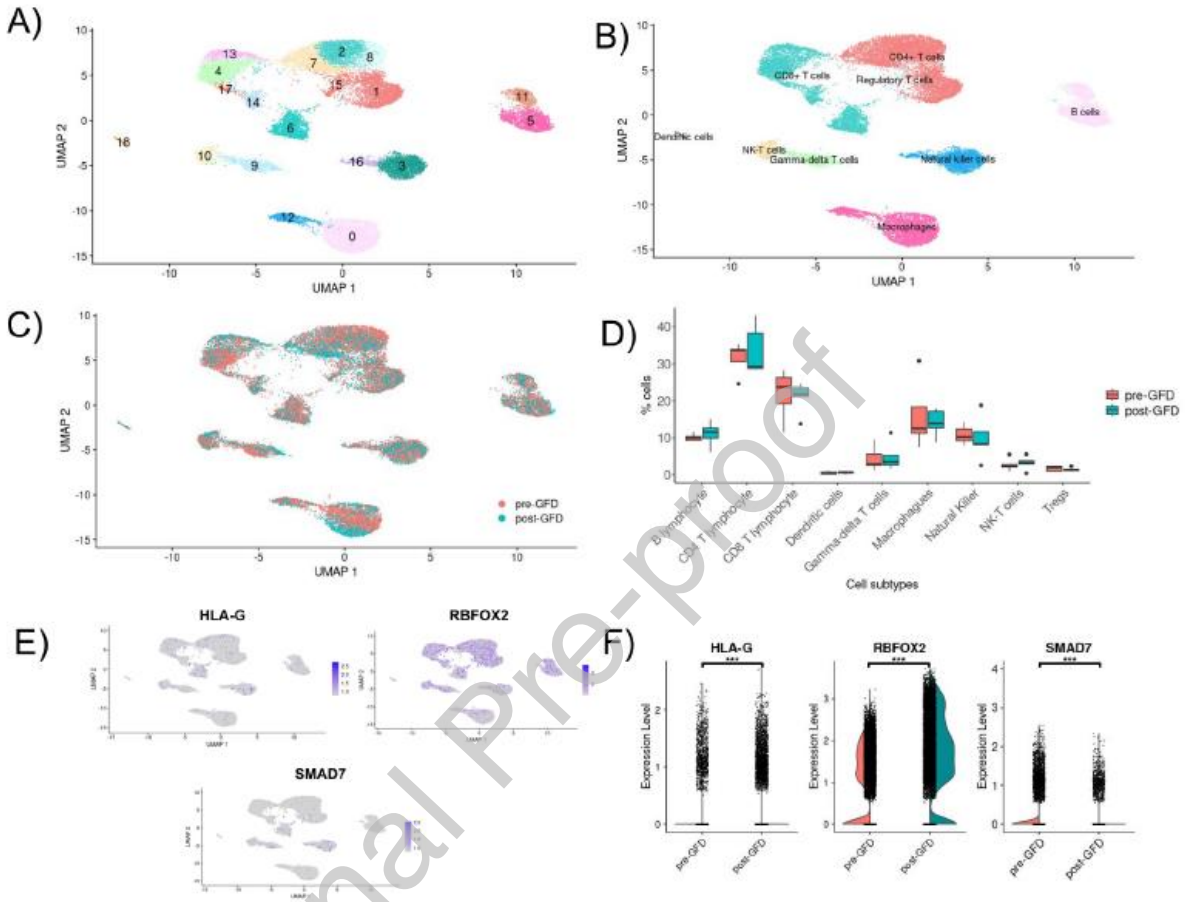


Figure 1. UMAP clustering scatter plot of peripheral blood mononuclear cells (PBMCs) colored by A) high definition non-supervised clustering, B) classical immunological cell subtypes, C) pre-gluten free diet (GFD) or post-GFD. D) Boxplots representing percentages of the different cell types pre-GFD and post-GFD. E) Uniform Manifold Approximation and Projection (UMAP) colored by single cell expression of 3 DEG between pre-GFD and post-GFD. F) Violin plots for 3 DEGs between pre-GFD and post-GFD. ***: Padj < 10-15.

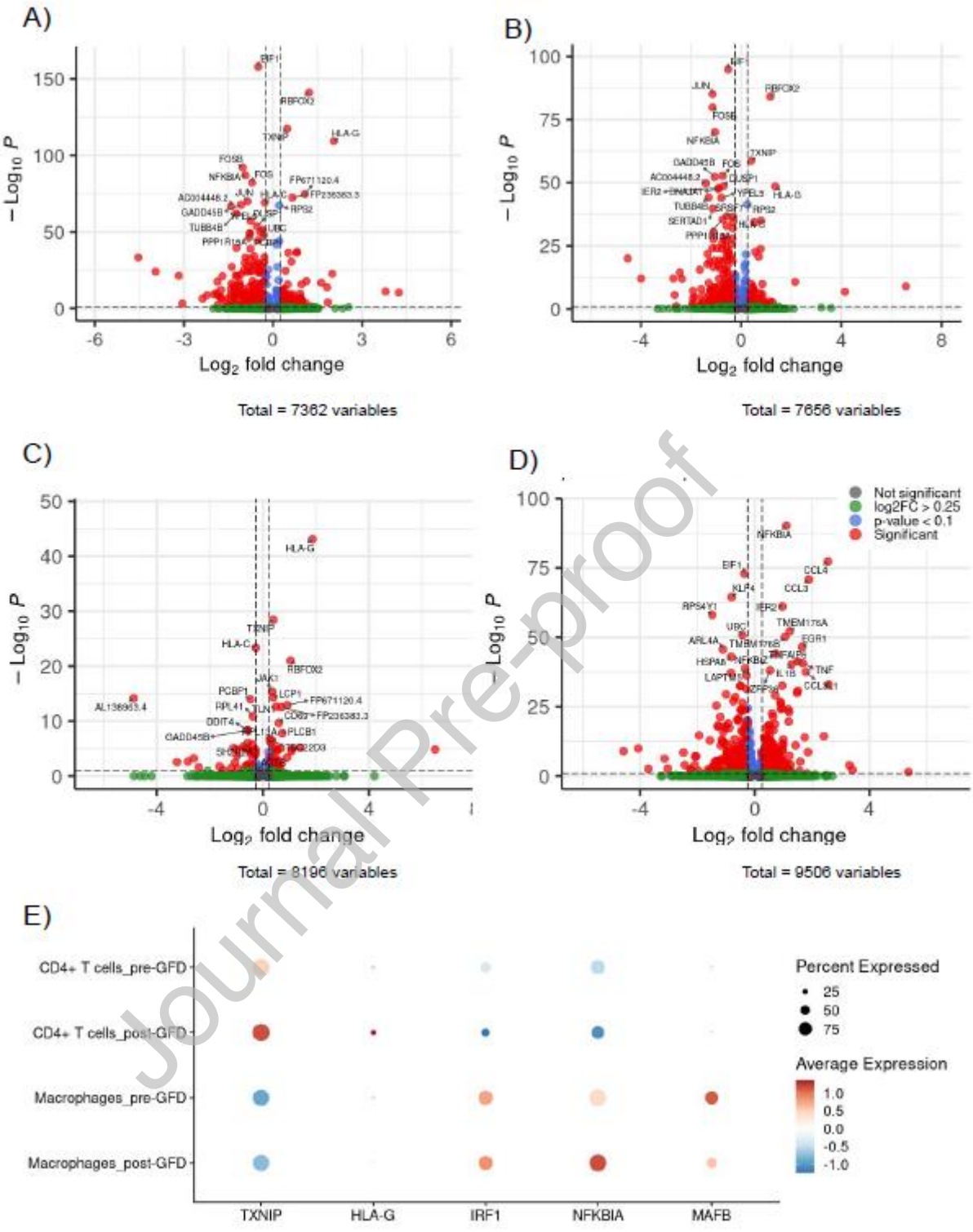


Figure 2. Volcano plots illustrating the DEGs between post-GFD and pre-GFD in A) CD4+ T cells, B) CD8+ T cells, C) Natural killer cells and D) Macrophages. E) Dotplot showing the expression of several DEGs post-GFD and pre-GFD in CD4+ T cells and macrophages.

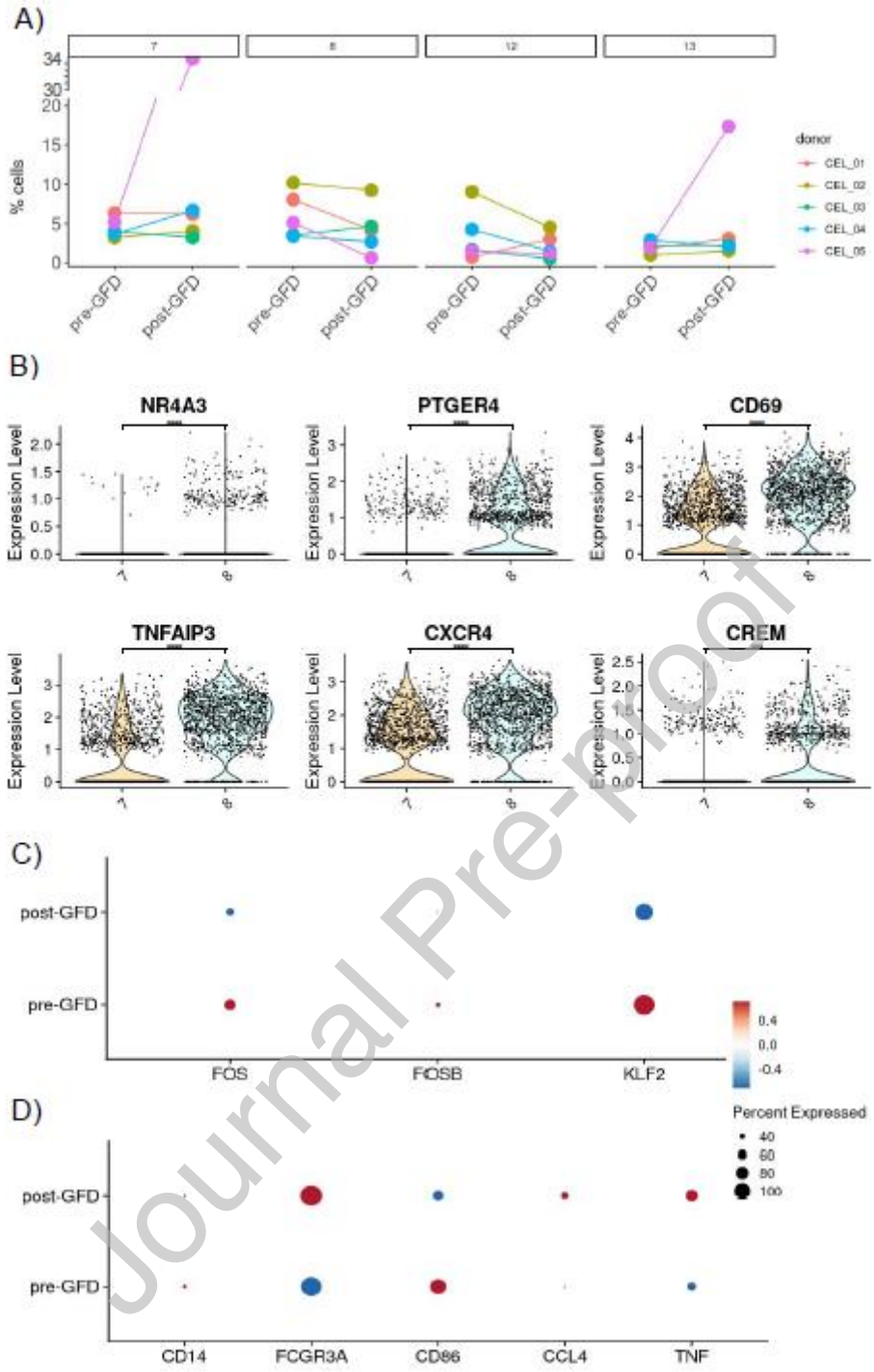


Figure 3. A) Scatter plot showing cell proportion changes in different clusters for each donor pre-GFD and post-GFD, B) Violin plots of DEGs in the comparison between clusters 7 and 8 (CD4+ T cells). Dot plots showing C) DEGs in cluster 13 comparison between pre-GFD and post-GFD (CD8+ T cells), and D) DEGs in cluster 12 comparison between pre-GFD and post-GFD (non-classical macrophages). ****: Padj < 10-90; ***: Padj < 10-40.

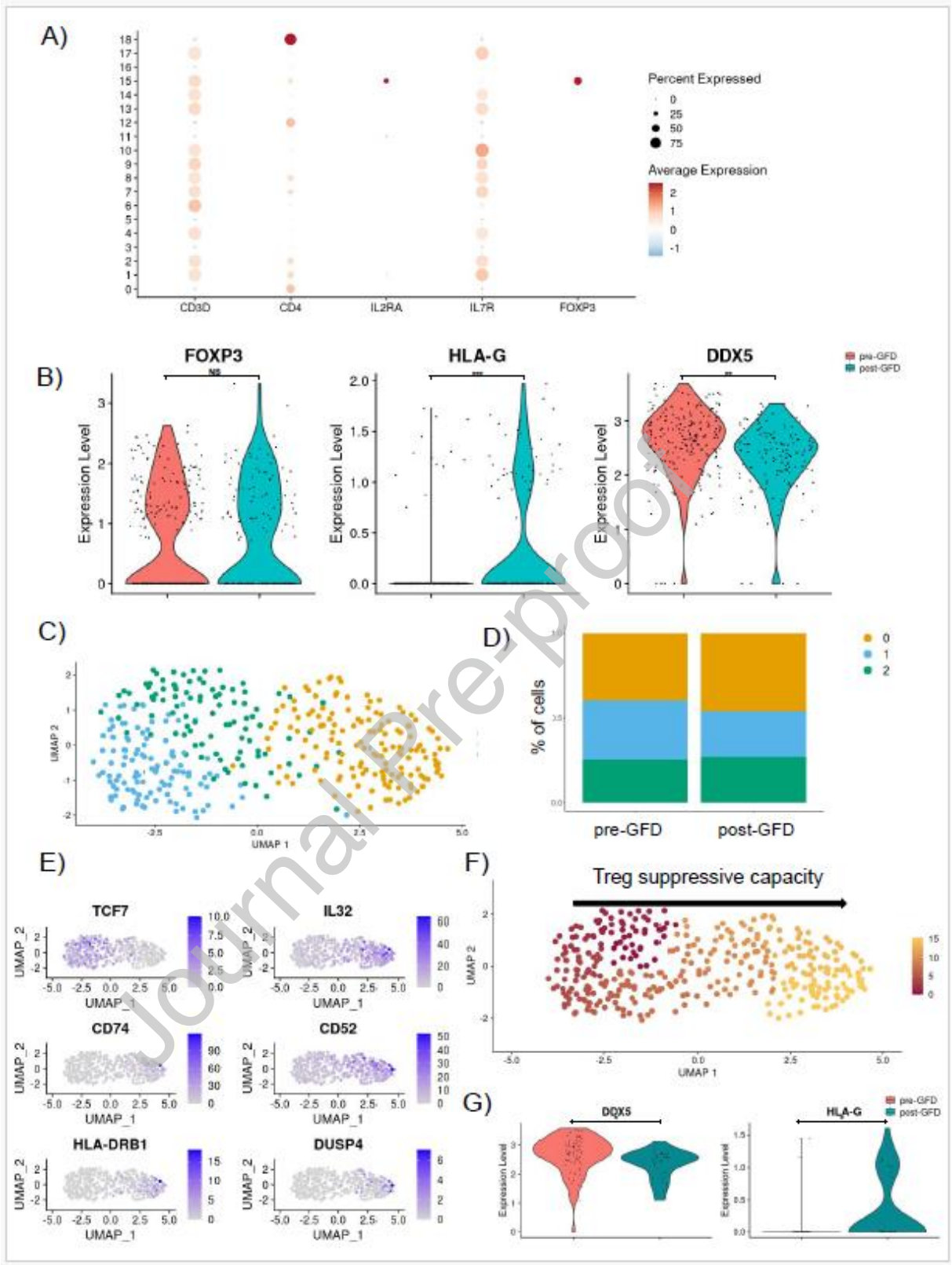


Figure 4. Regulatory T cell (Treg) characteristics in CD pre- and post-GFD after reclustering. A) Dot plot showing the expression of classical Treg markers in the 19 high-definition clusters, B) Violin plots for 3 DEGs between pre-GFD and post-GFD in Tregs. **C)** UMAP clustering scatter plot colored

by Treg clusters after reclustering, **D)** Cell proportions of each Treg cluster pre-GFD and post-GFD, **E)** UMAP clustering scatter plots colored by single cell expression of 6 Treg cluster markers, **F)** UMAP clustering scatter plot colored by pseudotime, **G)** Violin plots for 2 DEG between pre-GFD and post-GFD in cluster 1 Tregs. ***: Padj < 10-3; **: Padj < 0.1; *P < 10-3 but Padj > 0.1; NS: not significant.

Supplemental Figure 1. Heatmap representing the gene expression levels of the top 10 classical immune cell cluster markers stratified by cell type.

Supplemental Figure 2. Bar plots showing cell proportions of **A)** each high definition cluster stratified by pre-GFD or post-GFD samples, **B)** each high definition cluster stratified by donor and GFD status, **C)** each classical immune cell cluster stratified by pre-GFD or post-GFD samples, **D)** each classical immune cell stratified by donor and GFD status.

Supplemental Figure 3. UMAP clustering scatter plot colored by patient.

DATA AVAILABILITY

Data described in the manuscript and analytic code will be made available upon request pending approval from the corresponding author.

FUNDING

This work was supported by the Andalusian Government through the research projects of 'Financiación de la Investigación, Desarrollo e Innovación Biomédica y en Ciencias de la Salud en Andalucía' (ref. PI-0070-2024) and 'Proyectos de investigación Aplicada del Plan Propio de Investigación y Transferencia de la Universidad de Granada 2023' (ref. C-CTS-301-UGR23). This work has been partially funded by the research grant "Ayuda a la investigación" from the SPAO 2023 edition and the Beca de investigación Vicente Varea scholarship from the Sociedad Española de Gastroenterología, Hepatología y Nutrición Pediátrica (SEGHNP) 2023 edition.

ACKNOWLEDGEMENTS

The authors thank the patients for their participation in the current study. The pilot samples for this project were partially funded by 10x Genomics through a core lab award for the services carried out by the Genomics Unit at the Instituto de Parasitología y Biomedicina López-Neyra . We thank BioRender (through individual license at https://app.biorender.com) for the elements used in Graphical Abstract. Nerea Correa acknowledges the Ph.D. Program "Nutrición y Ciencias de los Alimentos" from the University of Granada; published results are part of her Doctoral Thesis.

CONFLICT OF INTEREST

Authors declare no conflict of interest.

REFERENCES

[1] Catassi C, Verdu EF, Bai JC, Lionetti E. Coeliac disease. Lancet 2022;399:2413–26.

[2] Jimenez J, Loveridge-Lenza B, Horvath K. Celiac Disease in Children. Pediatr Clin North Am 2021;68:1205–19.

[3] Goel G, Tye-Din JA, Qiao S-W, Russell AK, Mayassi T, Ciszewski C, et al. Cytokine release and gastrointestinal symptoms after gluten challenge in celiac disease. Sci Adv 2019;5:eaaw7756.

[4] Gandini A, Gededzha MP, De Maayer T, Barrow P, Mayne E. Diagnosing coeliac disease: A literature review. Hum Immunol 2021;82:930–6.

[5] Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PHR, et al. The Oslo definitions for coeliac disease and related terms. Gut 2013;62:43–52.

[6] Lebwohl B, Rubio-Tapia A. Epidemiology, Presentation, and Diagnosis of Celiac Disease. Gastroenterology 2021;160:63–75.

[7] Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2018;16:823–36.e2.

[8] King JA, Jeong J, Underwood FE, Quan J, Panaccione N, Windsor JW, et al. Incidence of Celiac Disease Is Increasing Over Time: A Systematic Review and Meta-analysis. Am J Gastroenterol 2020;115:507–25.

[9] Parzanese I, Qehajaj D, Patrinicola F, Aralica M, Chiriva-Internati M, Stifter S, et al. Celiac disease: From pathophysiology to treatment. World J Gastrointest Pathophysiol 2017;8:27–38.

[10] Iversen R, Sollid LM. The Immunobiology and Pathogenesis of Celiac Disease. Annu Rev Pathol 2023;18:47–70.

[11] Gnodi E, Meneveri R, Barisani D. Celiac disease: From genetics to epigenetics. World J Gastroenterol 2022;28:449–63.

[12] Cook L, Munier CML, Seddiki N, van Bockel D, Ontiveros N, Hardy MY, et al. Circulating gluten-specific FOXP3CD39 regulatory T cells have impaired suppressive function in patients with celiac disease. J Allergy Clin Immunol 2017;140:1592–603.e8.

[13] Aljada B, Zohni A, El-Matary W. The Gluten-Free Diet for Celiac Disease and Beyond. Nutrients 2021;13. https://doi.org/10.3390/nu13113993.

[14] Sangineto M, Graziano G, D'Amore S, Salvia R, Palasciano G, Sabbà C, et al. Identification of peculiar gene expression profile in peripheral blood mononuclear cells (PBMC) of celiac patients on gluten free diet. PLoS One 2018;13:e0197915.

[15] Yohannes DA, de Kauwe A, Kaukinen K, Kurppa K, Mäki M, Anderson RP, et al. Effects of Gluten Challenge on PBMC Gene Expression Profiles in Diet Treated Celiac Disease. Front Immunol 2020;11:594243.

[16] Moreno M de L, Rodríguez-Herrera A, Sousa C, Comino I. Biomarkers to Monitor Gluten-Free Diet Compliance in Celiac Patients. Nutrients 2017;9. https://doi.org/10.3390/nu9010046.

[17] Moreno M de L, Cebolla Á, Muñoz-Suano A, Carrillo-Carrion C, Comino I, Pizarro Á, et al. Detection of gluten immunogenic peptides in the urine of patients with coeliac disease reveals transgressions in the gluten-free diet and incomplete mucosal healing. Gut 2017;66:250–7.

[18] Martín-Masot R, Jiménez-Muñoz M, Herrador-López M, Flor-Alemany M, Navas-López VM, Nestares T. The Importance of an Early Evaluation after Establishing a Gluten-Free Diet in Children with Celiac Disease. Nutrients 2023;15. https://doi.org/10.3390/nu15071761.

[19] Villanacci V, Vanoli A, Leoncini G, Arpa G, Salviato T, Bonetti LR, et al. Celiac disease: histology-differential diagnosis-complications. A practical approach. Pathologica 2020;112:186–96.

[20] Husby S, Koletzko S, Korponay-Szabó I, Kurppa K, Mearin ML, Ribes-Koninckx C, et al. European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. J Pediatr Gastroenterol Nutr 2020;70:141–56.

[21] Schiepatti A, Sanders DS, Biagi F. Seronegative coeliac disease: clearing the diagnostic dilemma. Curr Opin Gastroenterol 2018;34:154–8.

[22] Fuchs V, Kurppa K, Huhtala H, Mäki M, Kekkonen L, Kaukinen K. Delayed celiac disease diagnosis predisposes to reduced quality of life and incremental use of health care services and medicines: A prospective nationwide study. United European Gastroenterol J 2018;6:567–75.

[23] Ramírez-Sánchez AD, Tan IL, Gonera-de Jong BC, Visschedijk MC, Jonkers I, Withoff S. Molecular Biomarkers for Celiac Disease: Past, Present and Future. Int J Mol Sci 2020;21. https://doi.org/10.3390/ijms21228528.

[24] Galatola M, Cielo D, Panico C, Stellato P, Malamisura B, Carbone L, et al. Presymptomatic Diagnosis of Celiac Disease in Predisposed Children: The Role of Gene Expression Profile. J Pediatr Gastroenterol Nutr 2017;65:314–20.

[25] Domínguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science 2022;376:eabl5197.

[26] van der Wijst MGP, Brugge H, de Vries DH, Deelen P, Swertz MA, LifeLines Cohort Study, et al. Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. Nat Genet 2018;50:493–7.

[27] Ramírez-Sánchez AD, Chu X, Modderman R, Kooy-Winkelaar Y, Koletzko S, Korponay-Szabó IR, et al. Single-Cell RNA Sequencing of Peripheral Blood Mononuclear Cells From Pediatric Coeliac Disease Patients Suggests Potential Pre-Seroconversion Markers. Front Immunol 2022;13:843086.

[28] Atlasy N, Bujko A, Bækkevold ES, Brazda P, Janssen-Megens E, Lundin KEA, et al. Single cell transcriptomic analysis of the immune cell compartment in the human small intestine and in Celiac disease. Nat Commun 2022;13:4920.

[29] Gerasimidis K, Zafeiropoulou K, Mackinder M, Ijaz UZ, Duncan H, Buchanan E, et al. Comparison of Clinical Methods With the Faecal Gluten Immunogenic Peptide to Assess Gluten Intake in Coeliac Disease. J Pediatr Gastroenterol Nutr 2018;67:356–60.

[30] Roca M, Donat E, Masip E, Crespo Escobar P, Fornes-Ferrer V, Polo B, et al. Detection and quantification of gluten immunogenic peptides in feces of infants and their relationship with diet. Rev Esp Enferm Dig 2019;111:106–10.

[31] Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes 2012;7:284–94.

[32] Comino I, Real A, Vivas S, Síglez MÁ, Caminero A, Nistal E, et al. Monitoring of gluten-free diet compliance in celiac patients by assessment of gliadin 33-mer equivalent epitopes in feces. Am J Clin Nutr 2012;95:670–7.

[33] Cerván-Martín M, Tüttelmann F, Lopes AM, Bossini-Castillo L, Rivera-Egea R, Garrido N, et al. Immune and spermatogenesis-related loci are involved in the development of extreme patterns of male infertility. Commun Biol 2022;5:1220.

[34] Huang X, Huang Y. Cellsnp-lite: an efficient tool for genotyping single cells. Bioinformatics 2021;37:4569–71.

[35] Huang Y, McCarthy DJ, Stegle O. Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. Genome Biol 2019;20:273.

[36] Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell 2019;177:1888–902.e21.

[37] Long T, Abbasi N, Hernandez JE, Li Y, Sayed IM, Ma S, et al. RNA binding protein DDX5 directs tuft cell specification and function to regulate microbial repertoire and disease susceptibility in the intestine. Gut 2022;71:1790–802.

[38] Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014;32:381–6.

[39] Sedda S, De Simone V, Marafini I, Bevivino G, Izzo R, Paoluzi OA, et al. High Smad7 sustains inflammatory cytokine response in refractory coeliac disease. Immunology 2017;150:356–63.

[40] Kim H. The transcription factor MafB promotes anti-inflammatory M2 polarization and cholesterol efflux in macrophages. Sci Rep 2017;7:7591.

[41] Zhou D, Couture S, Scott MS, Abou Elela S. RBFOX2 alters splicing outcome in distinct binding modes with multiple protein partners. Nucleic Acids Res 2021;49:8370–83.

[42] Mayassi T, Ladell K, Gudjonson H, McLaren JE, Shaw DG, Tran MT, et al. Chronic Inflammation Permanently Reshapes Tissue-Resident Immunity in Celiac Disease. Cell 2019;176:967–81.e19.

[43] Trynka G, Zhernakova A, Romanos J, Franke L, Hunt KA, Turner G, et al. Coeliac disease-associated risk variants in TNFAIP3 and REL implicate altered NF-kappaB signalling. Gut 2009;58:1078–83.

[44] Khalkhal E, Rezaei-Tavirani M, Asri N, Nobakht F, Jahani-Sherafat S, Haidari MH, et al. Introducing New Potential Biomarkers for Celiac Disease among the Genes Extracted from General Databases. Middle East J Dig Dis 2022;14:192–9.

[45] Kumar V, Gutierrez-Achury J, Kanduri K, Almeida R, Hrdlickova B, Zhernakova DV, et al. Systematic annotation of celiac disease loci refines pathological pathways and suggests a genetic explanation for increased interferon-gamma levels. Hum Mol Genet 2015;24:397–409.

[46] Kumar A, Humphreys TD, Kremer KN, Bramati PS, Bradfield L, Edgar CE, et al. CXCR4 physically associates with the T cell receptor to signal in T cells. Immunity 2006;25:213–24.

[47] Ohl K, Nickel H, Moncrieffe H, Klemm P, Scheufen A, Föll D, et al. The transcription factor CREM drives an inflammatory phenotype of T cells in oligoarticular juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2018;16:39.

[48] Bakker OB, Ramírez-Sánchez AD, Borek ZA, de Klein N, Li Y, Modderman R, et al. Potential impact of celiac disease genetic risk factors on T cell receptor signaling in gluten-specific CD4+ T cells. Sci Rep 2021;11:9252.

[49] Narasimhan PB, Marcovecchio P, Hamers AAJ, Hedrick CC. Nonclassical Monocytes in Health and Disease. Annu Rev Immunol 2019;37:439–56.

[50] Ma W-T, Gao F, Gu K, Chen D-K. The Role of Monocytes and Macrophages in Autoimmune Diseases: A Comprehensive Review. Front Immunol 2019;10:1140.

[51] Sindhu S, Kochumon S, Shenouda S, Wilson A, Al-Mulla F, Ahmad R. The Cooperative Induction of CCL4 in Human Monocytic Cells by TNF-α and Palmitate Requires MyD88 and Involves MAPK/NF-κB Signaling Pathways. Int J Mol Sci 2019;20. https://doi.org/10.3390/ijms20184658.

[52] Pankratz S, Ruck T, Meuth SG, Wiendl H. CD4(+)HLA-G(+) regulatory T cells: Molecular signature and pathophysiological relevance. Hum Immunol 2016;77:727–33.

[53] Al-Khunaizi NR, Tabbara KS, Farid EM. Is there a role for HLA-G in the induction of regulatory T cells during the maintenance of a healthy pregnancy? Am J Reprod Immunol 2020;84:e13259.

[54] Ma S, Yang Q, Chen N, Zheng A, Abbasi N, Wang G, et al. RNA binding protein DDX5 restricts RORγt T suppressor function to promote intestine inflammation. Sci Adv 2023;9:eadd6165.

[55] Torres MI, López-Casado MA, Luque J, Peña J, Ríos A. New advances in coeliac disease: serum and intestinal expression of HLA-G. Int Immunol 2006;18:713–8.

[56] Rizzo R, Bortolotti D, Bolzani S, Fainardi E. HLA-G Molecules in Autoimmune Diseases and Infections. Front Immunol 2014;5:592.

[57] Catamo E, Zupin L, Segat L, Celsi F, Crovella S. HLA-G and susceptibility to develop celiac disease. Hum Immunol 2015;76:36–41.

[58] Martín-Villa JM, Vaquero-Yuste C, Molina-Alejandre M, Juarez I, Suárez-Trujillo F, López-Nares A, et al. HLA-G: Too Much or Too Little? Role in Cancer and Autoimmune Disease. Front Immunol 2022;13:796054.

[59] Passerini L, Amodio G, Bassi V, Vitale S, Mottola I, Di Stefano M, et al. IL-10-producing regulatory cells impact on celiac disease evolution. Clin Immunol 2024; 260:109923.

[60] Passeri L, Andolfi G, Bassi V, Russo F, Giacomini G, Laudisa C, et al. Tolerogenic IL-10-engineered dendritic cell-based therapy to restore antigen-specific tolerance in T cell mediated diseases. J Autoimmun 2023;138:103051.

[61] Levring TB, Kongsbak-Wismann M, Rode AKO, Al-Jaberi FAH, Lopez DV, Met Ö, et al. Tumor necrosis factor induces rapid down-regulation of TXNIP in human T cells. Sci Rep 2019;9:16725.

[62] Muri J, Thut H, Kopf M. The thioredoxin-1 inhibitor Txnip restrains effector T-cell and germinal center B-cell expansion. Eur J Immunol 2021;51:115–24.

[63] Bossini-Castillo L, Glinos DA, Kunowska N, Golda G, Lamikanra AA, Spitzer M, et al. Immune disease variants modulate gene expression in regulatory CD4 T cells. Cell Genom 2022;2:None.

[64] Dominguez-Villar M, Hafler DA. Regulatory T cells in autoimmune disease. Nat Immunol 2018;19:665–73.

[65] Han L, Chen S, Chen Z, Zhou B, Zheng Y, Shen L. Interleukin 32 Promotes Foxp3 Treg Cell Development and CD8 T Cell Function in Human Esophageal Squamous Cell Carcinoma Microenvironment. Front Cell Dev Biol 2021;9:704853.

[66] Bonnin E, Rodrigo Riestra M, Marziali F, Mena Osuna R, Denizeau J, Maurin M, et al. CD74 supports accumulation and function of regulatory T cells in tumors. Nat Commun 2024;15:3749.

[67] Samten B. CD52 as both a marker and an effector molecule of T cells with regulatory action: Identification of novel regulatory T cells. Cell Mol Immunol 2013;10:456–8.

[68] Yan D, Farache J, Mingueneau M, Mathis D, Benoist C. Imbalanced signal transduction in regulatory T cells expressing the transcription factor FoxP3. Proc Natl Acad Sci U S A 2015;112:14942–7.

[69] Gootjes C, Zwaginga JJ, Roep BO, Nikolic T. Defining Human Regulatory T Cells beyond FOXP3: The Need to Combine Phenotype with Function. Cells 2024;13. https://doi.org/10.3390/cells13110941.

[70] Barra MM, Richards DM, Hansson J, Hofer A-C, Delacher M, Hettinger J, et al. Transcription Factor 7 Limits Regulatory T Cell Generation in the Thymus. J Immunol 2015;195:3058–70.

[71] Shi C, Pamer EG. Monocyte recruitment during infection and inflammation. Nat Rev Immunol 2011;11:762–74.

[72] Villanueva-Martin G, Acosta-Herrera M, Carmona EG, Kerick M, Ortego-Centeno N, Callejas-Rubio JL, et al. Non-classical circulating monocytes expressing high levels of microsomal prostaglandin E2 synthase-1 tag an aberrant IFN-response in systemic sclerosis. J Autoimmun 2023;140:103097.

Author Statement

Lara Bossini-Castillo: Conceptualization, Formal analysis, Writing - Original Draft, Writing - Review & Editing **Francisco David Carmona**: Resources, Writing - Review & Editing **Nerea Correa-López:** Investigation, Formal analysis, Writing - Review & Editing **Marta Herrador-López:** Investigation, Writing - Review & Editing **Rafael Martín-Masot**: Conceptualization, Investigation, Writing - Original Draft, Writing - Review & Editing **Víctor Manuel Navas-López**: Resources, Writing - Review & Editing **Teresa Nestares**: Conceptualization, Resources, Writing - Original Draft, Writing - Review & Editing.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: