이학석사 학위논문

전사체 분석 기반 크론병 특이 장간막 지방 조직 비후의 역할 분석

Transcriptomic analysis reveals the function of Creeping fat in Crohn′s disease

울산대학교 대학원

의과학과

박소정

전사체 분석 기반 크론병 특이 장간막 지방 조직 비후의 역할 분석

지도교수 이호수

이 논문을 이학석사학위 논문으로 제출함

2024 년 2 월

울산대학교 대학원
의과학과
박소정

박소정의 이학석사학위 논문을 인준함

심사위원 박 상 형 (인)

심사위원 예 병 덕 (인)

심사위원 이 호 수 (인)

울산대학교 대학원

2024 년 2 월

ABSTRACT

One of the pathologic features of Crohn's disease (CD) is a "creeping fat" (CF), mesenteric adipose tissue (MAT) wrapping the inflamed intestine. This study aims to define distinct transcriptional profile and identify cellular composition of CD-CF. We collected 59 MAT samples: 23 paired samples from CD patients with CF (CD-CF) and MAT around the uninflamed intestine (CD-MAT) and 13 MAT samples from non-CD patients (Con-MAT). Differentially expressed gene (DEG) analysis performed by DESeq2, cell deconvolution analysis performed by CIBERSORTx. We identified a total of 529 DEGs (|log2FoldChange|>1.5; false discovery rate <0.05). Of these, 323 genes showed an incremental pattern from Con-MAT to CD-MAT, and to CD-CF, while 105 genes showed a decremental pattern. Incremental pattern genes were related to immune cell responses, including B-cell and T-cell activation, while genes with a decremental pattern were involved in cell trafficking and migration. Cell deconvolution analysis revealed significant changes in cellular composition between the CD-CF and Con-MAT groups, with increased proportions of B-cells/plasma Cells (P-value=6.00×10^{-8}) and decreased proportions of preadipocytes (P-value=6.20×10^{-4}) in the CD-CF group. To conclude, dynamic transcriptional changes and altered cellular components may be crucial for the shift from Con-MAT to CD-MAT and to CD-CF, providing insight into the mechanisms behind CF and highlighting its possible role in CD pathogenesis.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AR	anterior resection
BP	biological process
CD	Crohn's disease
CD-CF	creeping fat of Crohn's disease
CD-MAT	mesenteric adipose tissue of Crohn's disease
Con-MAT	mesenteric adipose tissue of control patients
CRC	colorectal cancer
DEG	differentially expressed gene
FDR	false discovery rate
GSEA	gene set enrichment assay
GO	gene ontology
IBD	inflammatory bowel disease
ICR	ileocecal resection
ILCs	innate lymphoid cells
IMM	immunomodulator
KEGG	Kyoto encyclopedia of genes and genomes
LHC	left hemicolectomy
MAT	mesenteric adipose tissue
MNPs	mononuclear phagocytes
NES	normalized enrichment score
PCA	principal component analysis
RHC	right hemicolectomy
SB R&A	small bowel resection and anastomosis
scRNA	single-cell RNA
SNP	single nucleotide polymorphism
ssGSEA	single-sample GSEA
TPC	total proctocolectomy
TPM	transcripts per million
UC	ulcerative colitis
UMAP	uniform manifold approximation and projection

1. INTRODUCTION

Crohn's disease (CD) involves chronic, relapsing transmural inflammation and is not clearly understood.[1] Factors such as the immune system, changes in the microbial population, genetic susceptibility, and environmental factors are known to influence the onset of CD.[2]

The relationship between mesenteric adipose tissue (MAT) and the pathogenesis of CD has been widely supported by accumulating evidence.[3-6] The MAT is thought to be involved in energy-storage, endocrine function, and host defense mechanisms. MAT contributes to immune regulation activity through the secretion of adipokines, cytokines, and fatty acids.[7] Creeping fat (CF),[9] characterized by the hyperplastic expansion of MAT wrapping around the inflamed intestine, is a widely accepted pathognomonic feature of CD.[3, 4] CF has been linked to transmural inflammation, muscular hypertrophy, fibrosis, and stricture formation in patients with CD.[6, 10] MAT has been suggested to contribute to the progression of the disease in CD, some association proposed.[11] CF has also been linked to a higher risk of post-surgical recurrence of CD,[12] leading to the development of surgical techniques such as ileocecal resection with mesentery excision.[13]

Several studies have investigated the molecular and cellular features of CF in CD.[5, 6, 10, 14-16] In one study, CF led to a marked upregulation of plasma cell genes with abundant T-, B-, and plasma cell infiltration.[15] Together with a broad spectrum of both pro-inflammatory and anti-inflammatory mediators,[7] CF appeared to be involved in maintaining microenvironmental homeostasis, as a complex immunologic zone adjacent to the inflamed intestine. However, due to difficulty in obtaining MATs, studies investigating transcriptomic data of CF in patients with CD[5, 15, 17] are still insufficient to clarify the pathogenesis and relevant role of CF in CD. Therefore, we aimed to define a transcriptional signature and the cellular components of CF. To characterize the alteration of CF, we compared transcriptomic data of MAT samples of non-CD controls, and paired samples of CF and non-CF in the same patient with CD.

2. MATERIALS AND METHODS

2.1 Patient and sample acquisition

The overall study design is presented in Figure 1. In total, 23 patients with a small-bowel involving CD who underwent bowel resection, three patients with ulcerative colitis (UC)-associated colorectal cancer (CRC), and 10 patients with non-inflammatory bowel disease (IBD) CRC or high-grade dysplasia were enrolled in this study from Asan Medical

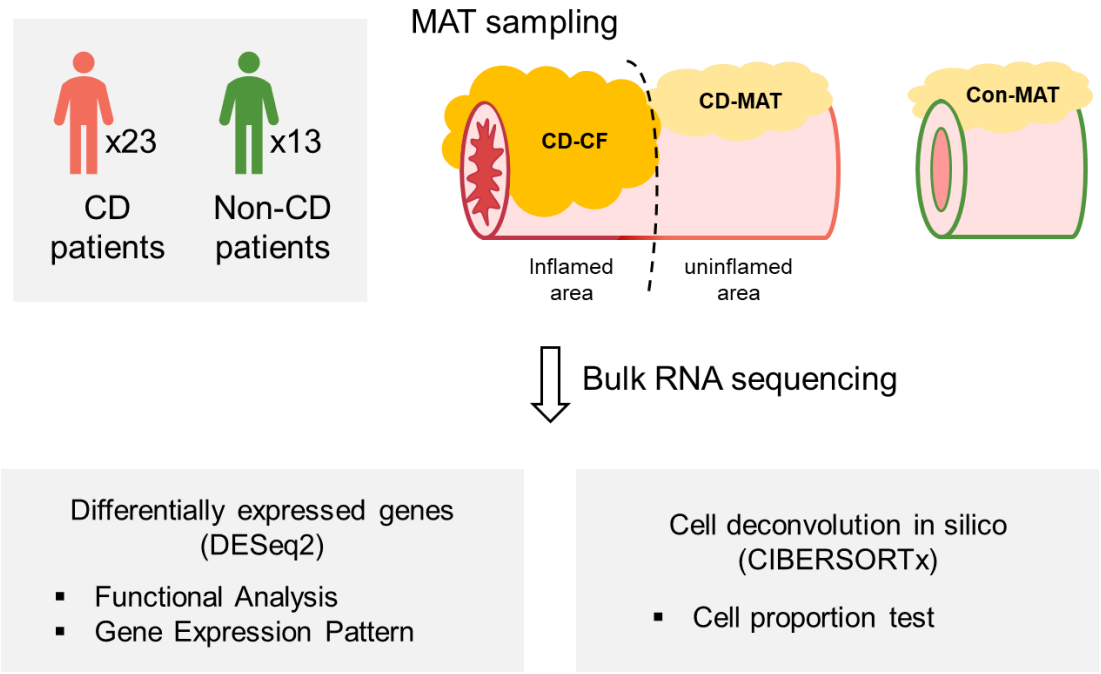

Figure 1. Overview of the sample collection and study design.

Center, a tertiary referral center in South Korea. During bowel resection, paired samples including CF wrapping the inflamed intestine, so-called "CD-CF," and non-creeping MAT adjacent to uninflamed intestine, so-called "CD-MAT," were collected from 23 patients with CD. As non-CD controls, MAT samples from non-CD controls were collected during surgery, and referred to as "Con-MAT." All surgically acquired samples were selected at the discretion of a surgeon, who has over 20 years of experience in both CRC and IBD-related surgery.

Baseline demographic data and clinical data including sex, smoking history, and age at surgery were collected from medical records. Among patients with CD, CD-related data included date of CD diagnosis, disease location/behavior based on the Montreal classification (L1–L3/B1–B3, perianal modifier),[1] previous medication history, and CD-related bowel resection history. Among patients with UC, UC-related data including date of UC diagnosis, disease extent based on the Montreal classification (E1–E3)[1] and previous medication history were also collected.

2.2 RNA extraction and sequencing

In total, 59 fat tissue samples were collected at the time of surgery and immediately placed in RNAlater solution (Ambion Inc, Austin, TX, USA) and kept at –80°C. The total RNA was extracted from all tissue samples using the RNeasy Plus Mini Kit (QIAGEN, Tokyo, Japan), according to the manufacturer's instruction. The quantity, purity, and integrity of the total RNA were measured using a 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Further deep sequencing only included RNA extractions with RNA integrity values ≥7. Preparation of whole-transcriptome sequencing libraries used TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Globin (Illumina, San Diego, CA, USA), and an Illumina HiSeq 2500 platform was used to generate 101-bp paired-end reads.

2.3 Quality control and alignment procedures

Quantification of gene expression levels was performed using RNA-SeQC[18] (version 2.1.2) and quality control of the data was performed using FastQC[19] (version 0.11.9) and MultiQC[20]. RNA sequencing reads were trimmed using Cutadapt[21] (version 3.7) in paired-end mode, and then filtered reads were mapped according to the reference genome GRCH38 using STAR software[22] (version 2.7.9a). The adapter sequences and low-quality reads were filtered out using Cutadapt with the parameters '-q 20 and -m 20'. We performed alignment for the trimmed reads using STAR (default settings) and GRCh38 reference genome in GENCODE[23]

release 40. We used RNA-SeQC[18] to confirm the unique mapping rate and ribosomal RNA rate, and also applied it to assess transcript abundance, expected read counts, and transcripts per million (TPM) reads for each gene by selecting the uniquely mapped reads with a mapping quality >255, and ≤ six mismatched bases to the reference genome.

2.4 Identification of differentially expressed genes (DEGs)

For low-expression gene filtering, raw genes containing reads less than 6 in more than half of all the samples and genes with a normalization value less than 0.1 in more than half of all samples were excluded. RNA sequencing data were further filtered to include protein-coding genes and long non-coding RNAs (lncRNAs). Filtered gene count and TPM lists were intersected, then 17,268 genes were found. Adjusted for sex, differential gene expression analysis was conducted using the "DESeq2" R package,[24] with a cutoff criteria of $|\log_2\text{FoldChange (FC)}| > 1.5$ and a false discovery rate (FDR) <0.05. DEGs were visualized with volcano plots using the "ggplot2" R package, and a heatmap was also generated using the "pheatmap" R package. Principal component analysis (PCA) plots and hierarchical clustering were used to identify outliers.

2.5 Functional pathway analysis

Gene set enrichment analysis (GSEA) was performed to identify functional enrichment of gene sets in CF using the "fgsea" R package.[25] The gene sets used in the analysis were MSigDB hallmark gene sets (h.all.v6.2.symbols.gmt).[26] gene ontology (GO) biological process (BP) pathways,[27] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway.[28] Gene sets with FDR <0.05 were considered significantly enriched in CF. In addition, single-sample GSEA (ssGSEA) via R package "SSGSEA" was performed to assess the immunological profile of each sample. GO enrichment analysis of a list of DEGs according to four expression patterns was performed with the publicly available tool, Enrichr.[29, 30]

2.6 Cell deconvolution analysis

To estimate cell subset proportions in the bulk RNA sequencing, we performed a cell-type deconvolution using CIBERSORTx11.[31] The reference matrix for deconvolution analysis was constructed based on single-cell RNA (scRNA) sequencing data from the Gene Expression Omnibus (GSE156110).[32] The performance of the predicted cell type proportions was evaluated using Pearson's correlation and root-mean-square error.

2.7 Statistical analyses

Values for continuous variables are expressed as median (range). For statistical comparisons between two unmatched groups, we used either the two-tailed *t*-test or the Mann-Whitney U test, depending on the data distribution. Paired *t*-test or Wilcoxon signed-rank test was used to compare paired samples from patients with CD, depending on the data distribution. To compare more than two groups, we used either one-way analysis of variance or the Kruskal-Wallis test. All tests were two-sided and a P-value <0.05 was considered statistically significant. All statistical analyses were performed using R (version 4.1.2; R Core Team, Vienna, Austria).

3. RESULTS

3.1 Baseline characteristics

Among 23 patients with CD, 16 (69.6%) patients were male, and the median age at diagnosis of CD was 25.0 years (range, 16.0 to 49.0 years). The median disease duration at surgery was 8.2 years (range, 0.2 to 21.3 years). In terms of disease location and behavior, 14 (60.9%) had ileocolonic location, 17 (73.9%) had penetrating behavior, and 13 (56.5%) had perianal disease modifier. With respect to medication history, 18 patients (78.3%) had been previously exposed to immunomodulators, and six patients were on maintenance therapy with biologic agents such as infliximab, adalimumab, or ustekinumab, and none had been exposed to vedolizumab. The main reason for surgery included intra-abdominal abscess formation (43.5%), or stricture of affected bowel (56.5%). Table 1 shows the baseline demographic data of the patients.

Among 13 non-CD controls, four (30.8%) patients were male and none had experienced previous bowel resection. Two patients underwent total proctocolectomy (patients with UC-associated CRC), and 11 patients underwent right hemicolectomy (one patient with UC, one patient with high-grade dysplasia irrelevant to IBD, and nine patients with non-IBD CRC). All three patients with UC had extensive colitis. The median age at UC diagnosis was 18.0 years (range, 13.0 to 26.0 years) and the median disease duration at surgery was 22.8 years (range, 15.0 to 31.9 years). Regarding medication history, none had prior or current exposure to biologic agents.

3.2 Distinct transcriptional profile of CD-CF and CD-MAT, when compared to Con-MAT

To define a gene expression profile related to CF, we identified DEGs among CD-CF, CD-MAT, and Con-MAT. Significant DEGs were defined as having an FDR <0.05 and a |log2FoldChange| >1.5. Using PCAs (Figure 2A–C) and heatmaps (Figure 2D–F), three groups were separated by their gene expression profiles. A total of 463 DEGs (386 protein-coding genes and 77 lncRNAs, Figure 2G) were found between CD-CF and Con-MAT, 130 DEGs (120 protein-coding genes and 10 lncRNAs, Figure 2H) were found between CD-MAT and Con-MAT, and 101 DEGs (92 protein-coding genes and nine lncRNAs, Figure 2I) were found between CD-CF and CD-MAT.

An upregulation of B-cell/plasma cell marker genes and downregulation of cell proliferation-related genes was observed across the different comparisons. For instance, *JSRP1* was the most upregulated gene in both CD-CF vs Con-MAT ($log_2FC = 9.39$, FDR = 7.34×10^{-19})

Table 1. Baseline demographic data of patients.

Variables		CD (n=23)	Non-CD control	
			UC (n=3)	CRN (n=10)
Age at IBD diagnosis, median (range), yr		25 (16.0–49.0)	18 (13.0–26.0)	-
Age at surgery, median (range), yr		32 (19.0–53.0)	45 (34.0–49.0)	74.5 (54.0–79.0)
Male, (%)		16 (69.6)	0	4 (40.0)
Disease duration, median (range), yr		8.2 (0.2–21.3)	22.8 (15.0–31.9)	-
Smoking (%)	Never smoker	10 (76.9)	3 (100)	7 (70.0)
	Current or Ever smoker	3 (23.1)	0	3 (30.0)
Previous bowel resection (%)		11 (47.8)	0	0
CD location[a] (%)				
Ileal		9 (39.1)	-	-
Ileocolonic		14 (60.9)	-	-
CD behavior[a] (%)				
Stricturing		6 (26.1)	-	-
Penetrating		17 (73.9)	-	-
Perianal disease modifier (%)		13 (56.5)	-	-
UC extent[a] (%)				
Extensive		-	3 (100)	-
Previous exposure to biologics (%)		6 (26.1)	0[b]	-
Previous exposure to IMM (%)		18 (78.3)	0	-
Type of surgery (%)	Colectomy (ICR, RHC, LHC, AR, TPC)	16 (69.6)	3 (100)	10 (100)
	SB R&A	7 (30.4)	0	0
Indication of surgery (%)	Abscess	10 (43.5)	-	-
	Stricture	13 (56.5)	-	-
	Colorectal malignancy or dysplasia	-	3 (100)	10 (100)

AR, anterior resection; CD, Crohn's disease; CRN, colorectal neoplasm; yr, year; ICR, ileocecal resection; IMM, immunomodulator; LHC, left hemicolectomy; RHC, right hemicolectomy; SB R&A, small bowel resection and anastomosis; TPC, total proctocolectomy; UC, ulcerative colitis.

[a]Based on the Montreal classification.

[b]Data of previous exposure to biologics was missing in one patient.

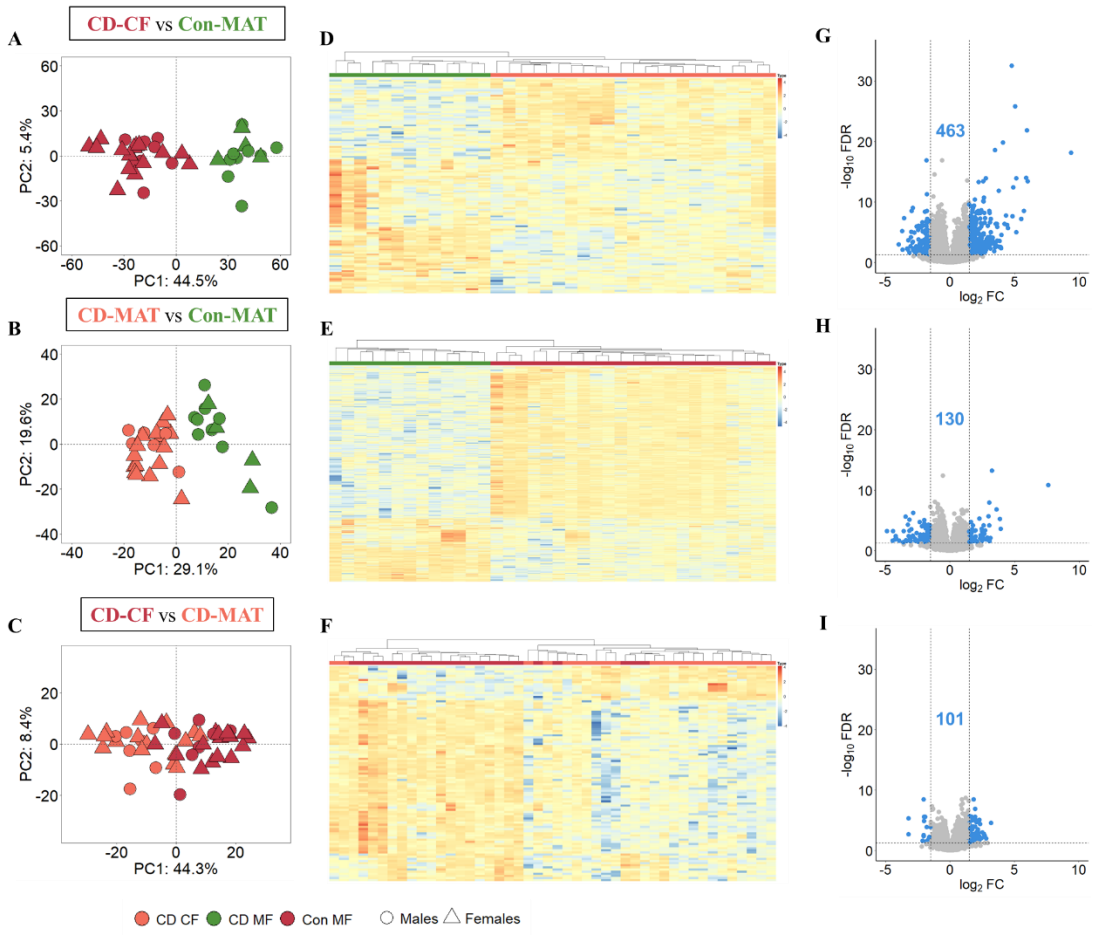

Figure 2. Analysis of DEGs among CD-CF, CD-MAT, and Con-MAT.

463 DEGs of CD-CF vs Con-MAT were presented as (A) PCA plot and (D) heatmap plot. 130 DEGs of CD-MAT vs Con-MAT were presented as (B) PCA plot and (E) heatmap plot. 101 DEGs of CD-CF vs CD-MAT were shown as (C) PCA plot and (F) heatmap plot. In the PCA and heatmap plot, dot shapes indicate sex of patients, and the colors red, orange, and green indicate CD-CF, CD-MAT, and Con-MAT, respectively. (G-I) In volcano plots, blue dots on the right side and left side represent significantly upregulated genes and downregulated genes, respectively, while grey dots represent genes without significance. Blue text in volcano plot indicates the count of DEGs.

and CD-MAT vs Con-MAT (log_2FC = 7.63, FDR = 1.56×10^{-11}). In DEGs of CD-CF vs CD-MAT, the most upregulated gene was *CD79A* (log_2FC = 3.19, FDR = 2.98×10^{-5}). Meanwhile, the most downregulated genes were *HOXA10* (log_2FC = -3.95, FDR = 9.85×10^{-4}), *ITLN1* (log_2FC = -5.52, FDR = 1.66×10^{-7}), and *MYOC* (log_2FC = -3.21, FDR = 5.28×10^{-6}) in CD-CF vs Con-MAT, CD-MAT vs Con-MAT, and CD-CF vs CD-MAT, respectively.

The intersection of the DEGs with IBD-susceptibility loci showed that 24 DEGs between CD-CF and Con-MAT, five DEGs between CD-MAT and Con-MAT, and five DEGs between CD-CF and CD-MAT were located within ± 250 kilobases of genome-wide associated variants (Table 2).[33] This suggests a possible relationship between these genetic variants and the gene expression profile of CF in CD.

3.3 Complex interplay between immune and lipid metabolism in the development and progression of CF

To identify significantly enriched gene sets, GSEA was performed using the Wald-statistic values from DEGs. Based on the hallmark database analyses (Figure 3), three DEG sets were commonly enriched with genes related to immune response, such as “allograft rejection,” “interferon gamma response,” and “interferon alpha response,” while metabolic pathways such as “adipogenesis” showed a negative relationship. Additional GSEA based on gene ontology biological process (GO BP) and Kyoto encyclopedia of genes and genomes (KEGG) pathways (Figure 4) also showed concordant results. Using ssGSEA, we further compared enrichment scores of co-expressed gene sets in each sample, focusing on the shared pathways across three comparisons, “adipogenesis,” “allograft rejection,” “interferon alpha response,” and “interferon gamma response” (Figure 5). The results revealed contrasting enrichment trends between CD-CF and Con-MAT, with immune response pathways enriched and metabolic pathway negatively correlated in CD-CF, while the opposite was seen in Con-MAT.

Each DEG was overlapped and presented as a Venn diagram (Figure 6A), which included separate color shades according to the expression pattern. There were 16 genes commonly shared among the three DEG sets, including 11 marker genes for B-cell/plasma cells: *JSRP1*, *FCRL5*, *DERL3*, *FCRL2*, *MZB1*, *POU2AF1*, *JCHAIN*, *BHLHA15*, *TNFRSF17*, *HLA-DOB*, and *SLAMF7*, and five genes not specific for B-cell/plasma cells: *LINC02362*, *SDC1*, *AMPD1*, *IFNG-AS1*, and *ITLN1*. The heatmap of log_2FC values of these 16 genes is displayed based on the DEG sets (Figure 6B). Additionally, the expression patterns of DEGs were divided into four groups – incremental, decremental, down-bent, and up-bent – based on the comparison results of CD-MAT vs CD-CF and of CD-MAT vs Con-MAT. The expression levels of 323

Table 2. The list of intersection of the DEGs with IBD-susceptibility loci.

Gene Symbol	CD-CF vs Con-MAT		CD-MAT vs Con-MAT		CD-CF vs CD-MAT		SNP[a]	Distance (bp)[b]
	log_2FC	FDR	log_2FC	FDR	log_2FC	FDR		
CD19	**3.50**	**3.48×10^{-03}**	1.43	0.37	2.06	0.06	rs28449958	403187
TAS1R3	**2.95**	**6.23×10^{-05}**	1.51	0.09	1.44	3.65×10^{-02}	rs12103	83786
HLA-DQA2	**2.79**	**8.55×10^{-03}**	1.27	0.36	1.52	0.12	rs6927022	666724
IL23R	**2.77**	**3.69×10^{-03}**	1.92	NA	0.85	0.38	rs11581607	568783
CCR6	**2.73**	**4.55×10^{-08}**	1.48	1.10×10^{-02}	1.25	7.15×10^{-03}	rs1819333	261740
MEI1	**2.70**	**2.63×10^{-07}**	1.20	0.06	**1.50**	**1.79×10^{-03}**	rs727563	167874
IKZF3	**2.64**	**1.93×10^{-11}**	0.99	4.09×10^{-02}	**1.65**	**2.62×10^{-06}**	rs12946510	1845341
SP140	**2.42**	**2.16×10^{-09}**	1.16	1.52×10^{-02}	1.26	6.10×10^{-04}	rs6716753	893505
LTB	**2.07**	**0.02**	0.77	0.53	1.30	0.10	rs6927022	666724
RORC	**1.86**	**9.43×10^{-03}**	0.85	0.37	1.01	0.13	rs4845604	4391
THEMIS	**1.85**	**6.47×10^{-04}**	0.99	0.13	0.86	0.10	rs9491891	569079
CD6	**1.83**	**9.84×10^{-04}**	0.57	0.46	1.25	1.18×10^{-02}	rs11230563	195471
KIF21B	**1.73**	**5.64×10^{-08}**	0.58	0.16	1.15	4.79×10^{-05}	rs7554511	91828
ZNF831	**1.69**	**8.99×10^{-07}**	0.69	0.11	1.00	1.35×10^{-03}	rs259964	1299072
LIME1	**1.68**	**5.15×10^{-06}**	0.72	0.12	0.96	4.24×10^{-03}	rs6062496	1407184
ITGAL	**1.66**	**4.19×10^{-07}**	0.48	0.28	1.18	4.79×10^{-05}	rs11150589	9836
STAT4	**1.63**	**6.31×10^{-07}**	1.25	5.87×10^{-04}	0.39	0.26	rs1517352	901888
IL7R	**1.59**	**2.58×10^{-04}**	0.57	0.33	1.03	9.35×10^{-03}	rs3194051	23579
RASGRP1	**1.59**	**3.84×10^{-06}**	0.36	0.48	1.23	4.53×10^{-05}	rs16967103	411087
LY75	**1.51**	**2.51×10^{-06}**	0.50	0.25	1.02	3.85×10^{-04}	rs4664304	990653
MYRF	-0.26	0.81	**-1.87**	**2.53×10^{-02}**	**1.61**	**1.67×10^{-02}**	rs4246215	188337
CFB	-0.65	0.39	**-2.02**	**2.20×10^{-03}**	1.37	1.29×10^{-02}	rs6927022	666724
SLC39A8	-1.12	4.42×10^{-02}	**-2.15**	**8.83×10^{-05}**	1.03	3.18×10^{-02}	rs13126505	614224
CCL21	**-1.51**	**1.17×10^{-03}**	0.00	1.00	**-1.51**	**1.43×10^{-04}**	rs9408254	27153
HOXA9	**-2.65**	**1.01×10^{-02}**	-1.50	NA	-1.15	0.25	rs10486483	278152
ITLN1	**-2.82**	**8.98×10-^{03}**	**-5.52**	**1.66×10^{-07}**	**2.70**	**3.55×10^{-03}**	rs4656958	19576
HOXA10	**-3.95**	**9.85×10^{-04}**	**-3.18**	**1.77×10^{-02}**	-0.77	0.56	rs10486483	278152

CD-CF, creeping fat of Crohn's disease; CD-MAT, mesenteric adipose tissue of Crohn's disease; Con-MAT, mesenteric adipose tissue of control; FDR, false discovery rate; log_2FC, log_2FoldChange; NA, not available; SNP, single nucleotide polymorphism.

Each DEG set was indicated in bold.

[a]SNPs with susceptibility to inflammatory bowel disease in European (de Lange KM et al. *Nat Genet* 2017, Ref 33).

[b]Distance: the distance from SNP to the transcription start site (TSS) of the gene.

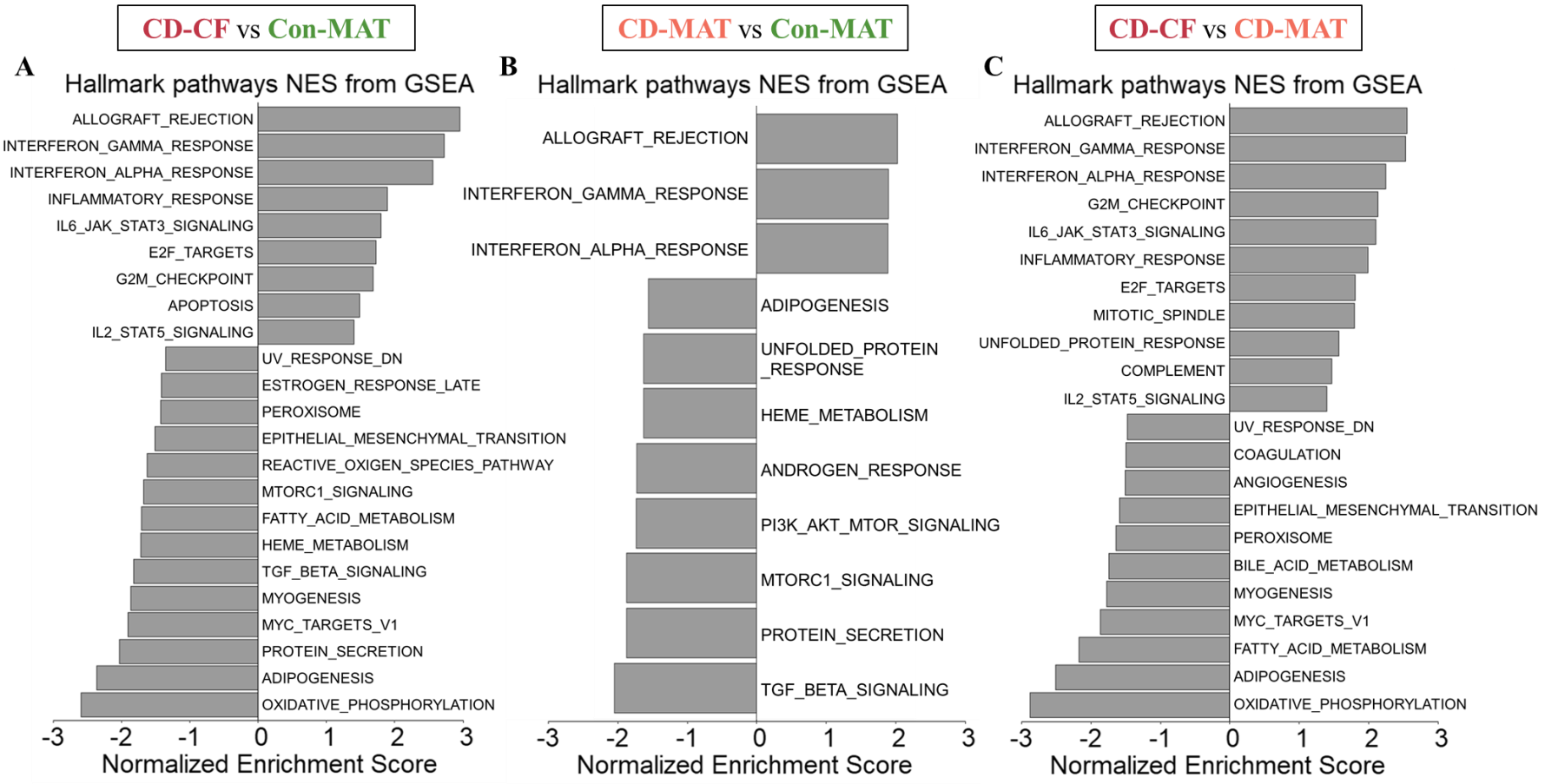

Figure 3. Functional analysis via Gene set enrichment analysis (GSEA).

GSEA querying hallmark database (h.all.v6.2.symbols.gmt) was performed using the Wald-statistic values from DEGs of (A) CD-CF vs Con-MAT, (B) CD-MAT vs Con-MAT, and (C) CD-CF vs CD-MAT. The indicated pathways satisfy an FDR <0.05.

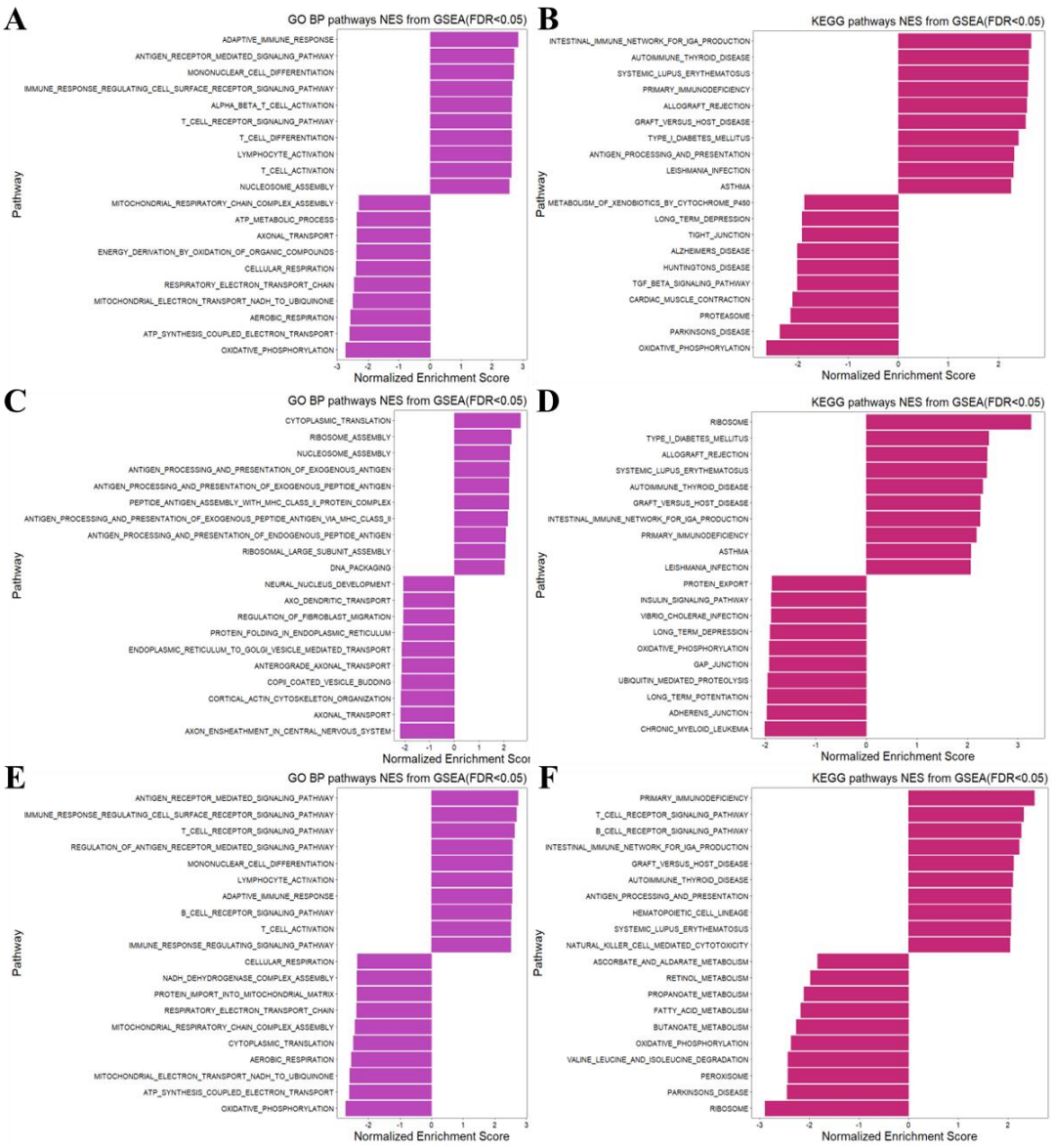

Figure 4. Gene set enrichment analysis based on GO BP and KEGG.

GO BP (left) and KEGG (right) pathways showing enriched gene sets (FDR<0.05). (A–B) A positive normalized enrichment score (NES) value indicates enrichment in CD-CF, and a negative NES indicates enrichment in Con-MAT. (C–D) A positive NES value indicates enrichment in CD-MAT, and a negative NES indicates enrichment in the Con-MAT. (E–F) A positive NES value indicates enrichment in CD-CF, and a negative NES indicates enrichment in the CD-MAT.

Figure 5. Single-sample GSEA of hallmark pathway.

ssGSEA showing the separate enrichment scores of pathways, "adipogenesis," "allograft rejection," "interferon alpha response," and "interferon gamma response," from each sample. (A) Heatmap showing NES per sample for each pathway, with a positive NES indicating a positive enrichment, while a negative NES value indicates a negative relationship. (B) Box plots showing NES of pathways, "adipogenesis," "allograft rejection," interferon alpha response," and "interferon gamma response," according to CD-CF, CD-MAT, and Con-MAT. Each box plot contains the P-value from the ANOVA.

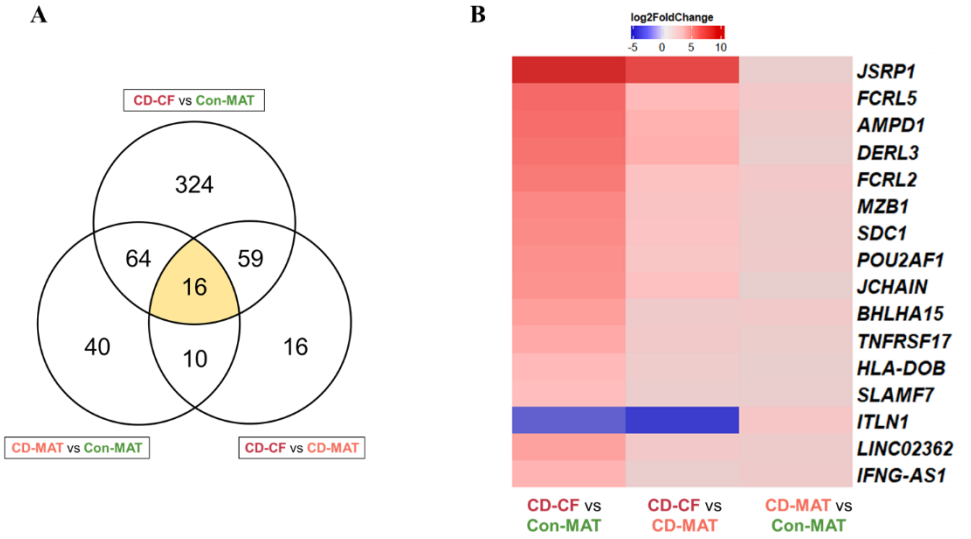

Figure 6. Comparative analysis of three analytical combinations.

(A) A Venn diagram of 529 unique DEGs among CD-CF, CD-MAT, and Con-MAT. Each region was marked with different color shades. (B) A heatmap of log2FC values of 16 common DEGs among CD-CF, CD-MAT, and Con-MAT.

genes were incremental from Con-MAT to CD-MAT and from CD-MAT to CD-CF, while 105 genes showed a decremental pattern (Figure 7 and Figure 8). In addition, 65 genes showed the lowest expression level and 36 genes showed the highest expression level in CD-MAT in comparison to CD-CF and Con-MAT. Genes with an incremental pattern were related to immune cell responses including B-cell and T-cell activation, and genes with a decremental pattern were involved in cell trafficking and migration. Genes with a down-bent pattern were related to cell-cell junction and cytoskeletal trafficking, and genes with an up-bent pattern involved genes related to muscle contraction.

3.4 Altered cellular composition of CD-CF

Cellular composition of bulk tissue samples was estimated through cell deconvolution analysis. The analysis was based on scRNA sequencing data from GSE156110.[32] We identified six cell types B-cells/plasma cells, T-cells, mononuclear phagocytes (MNPs), preadipocytes, endothelial cells, muscle cells, innate lymphoid cells (ILCs) (Figure 9). The relative cell type percentages are presented as a stacked bar plot the most common cell type in all three groups was muscle cells, followed by preadipocytes (Figure 10). CD-CF had a higher proportion of immune-related cells (B-cells/plasma cells) compared to CD-MAT and Con-MAT (Figure 11 and Table 3). B-cell/plasma cells showed a significant difference between CD-MAT and Con-MAT (P-value=6.00×10^{-8}). These results suggested that the distinct gene expression of CD-CF was possibly mediated by an alteration of cellular composition that differed from those of both Con-MAT and CD-MAT.

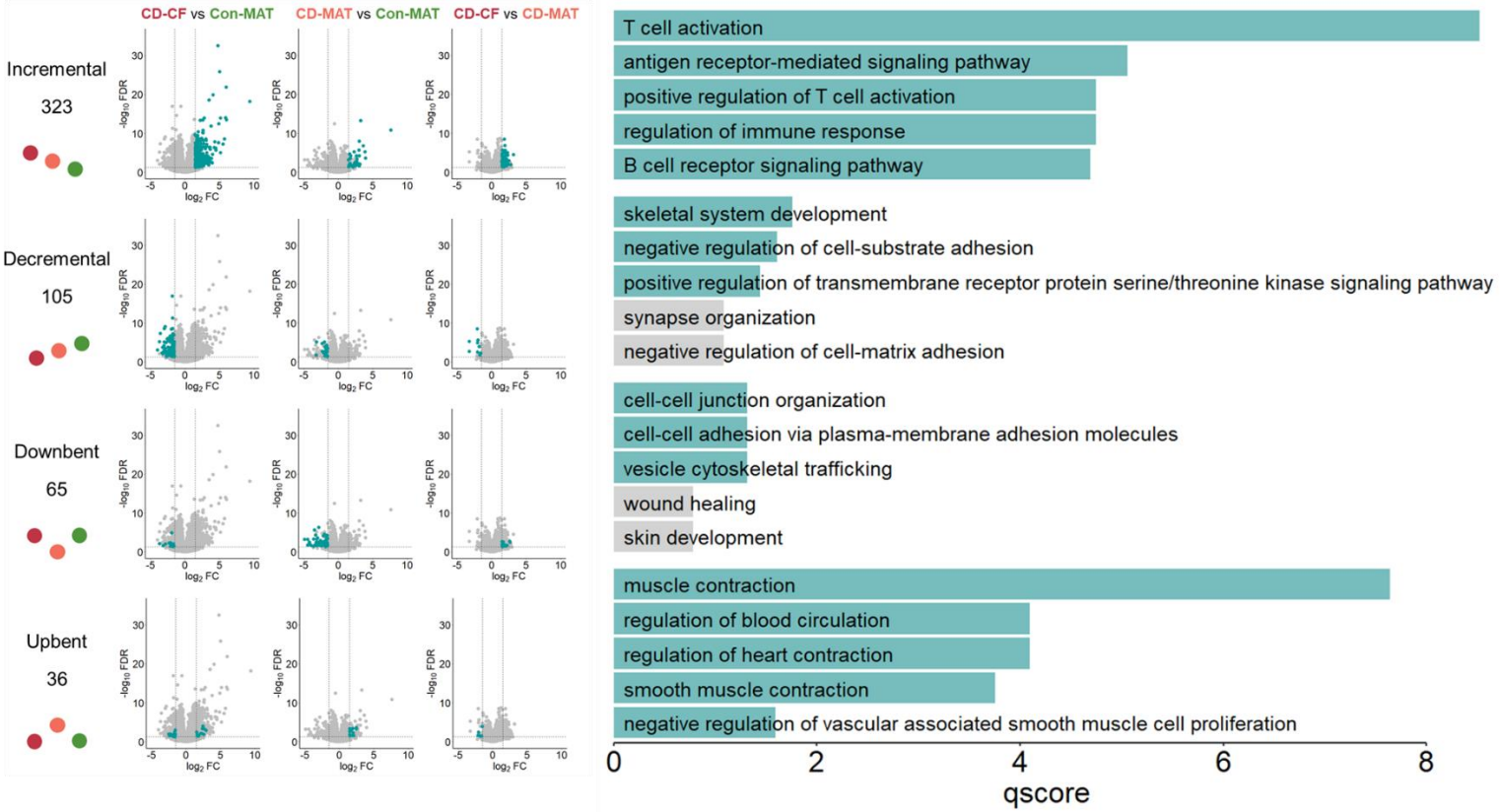

Figure 7. Classification of DEGs according to gene expression patterns.

The red, orange, green dots represent expression patterns. In the volcano plot, the points indicated where the each DEGs are located. Additional GO BP analysis using Enrichr was performed on the genes according to four expression patterns. Blue bar represents significantly enriched pathways (FDR <0.05). The qscore is calculated from $\log_{10}$FDR.

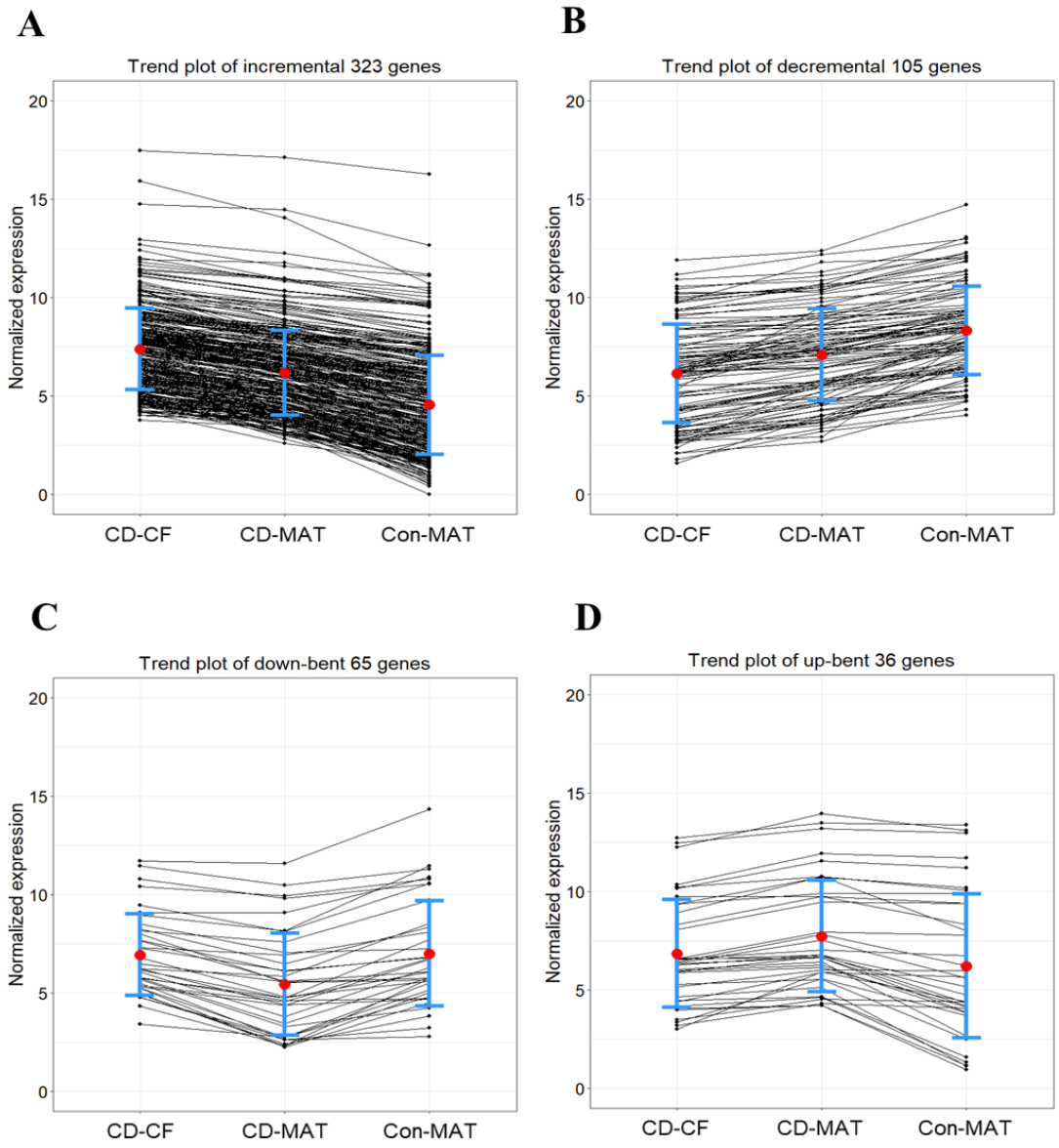

Figure 8. Expression value line plots of DEGs according to expression patterns groups.

Line plots of normalized gene expression values of DEGs according to four expression patterns in CD-CF, CD-MAT, and Con-MAT. (A) 323 genes with incremental pattern, (B) 105 genes with decremental pattern, (C) 65 genes with down-bent pattern, and (D) 36 genes with up-bent pattern. The means and standard deviations of the mean are indicated with red dots and blue error bars, respectively.

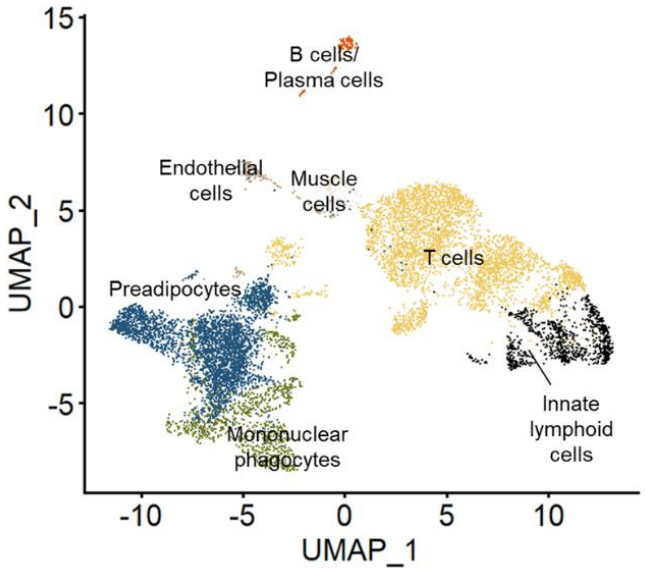

Figure 9. Single cell reference uniform manifold approximation and projection (UMAP) plot.

Publicly available single-cell RNA sequencing data (GSE156110, Ref. 32) was processed with the R package "Seurat" (version 4.3.0). Using cell-type specific markers, cluster annotation was performed.

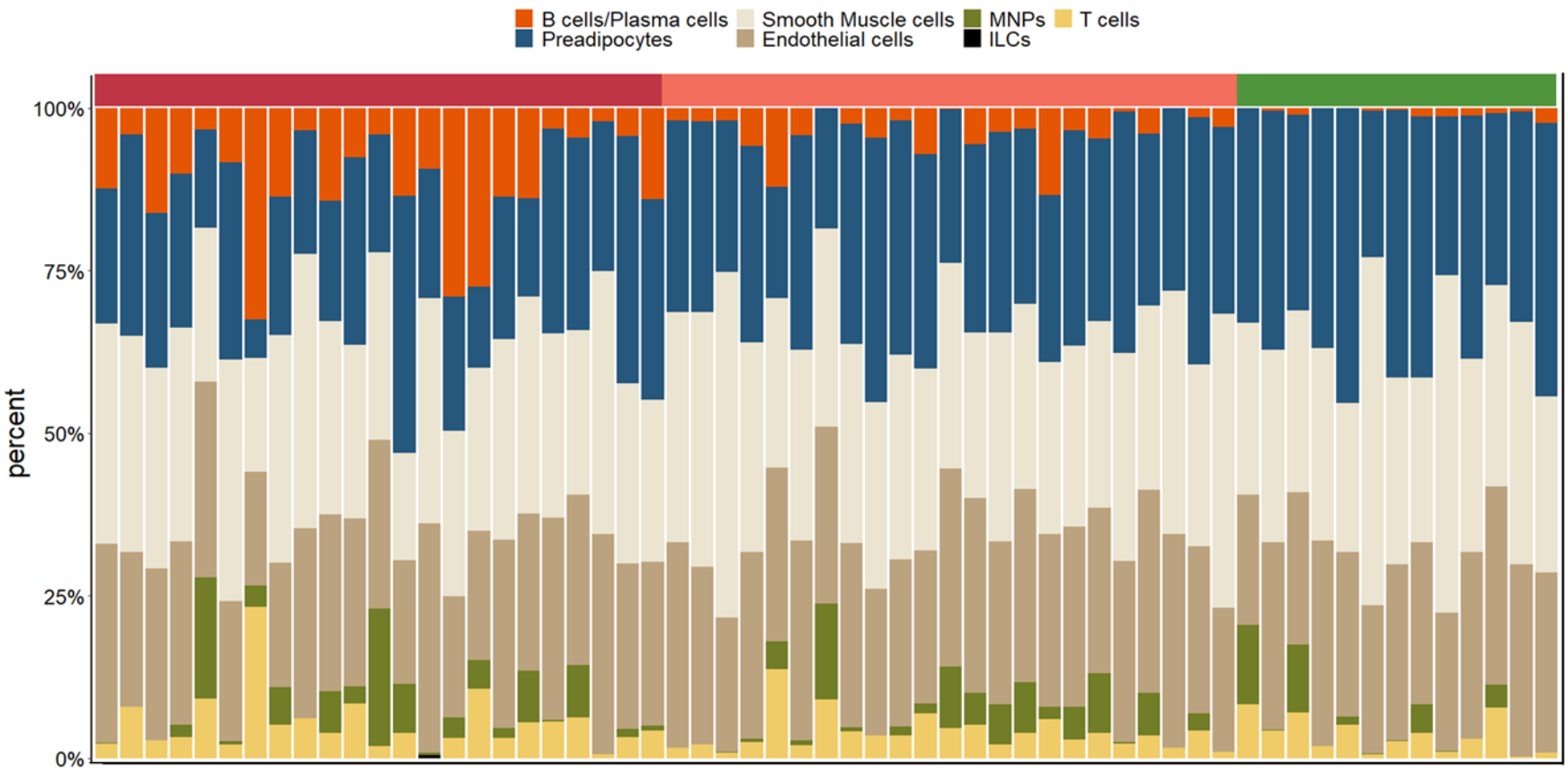

Figure 10. Deconvolution analysis for cell compositions.

The composition of six cell subsets in each patient sample. The profiling of cells was inferred by deconvolution analysis based on scRNA sequencing data from the Gene Expression Omnibus (GSE156110, Ref. 32) with adequate performance measured by Pearson's correlation of 0.83 and root-mean-square error of 0.57. The proportions of different cell types are shown in the stacked bar plot.

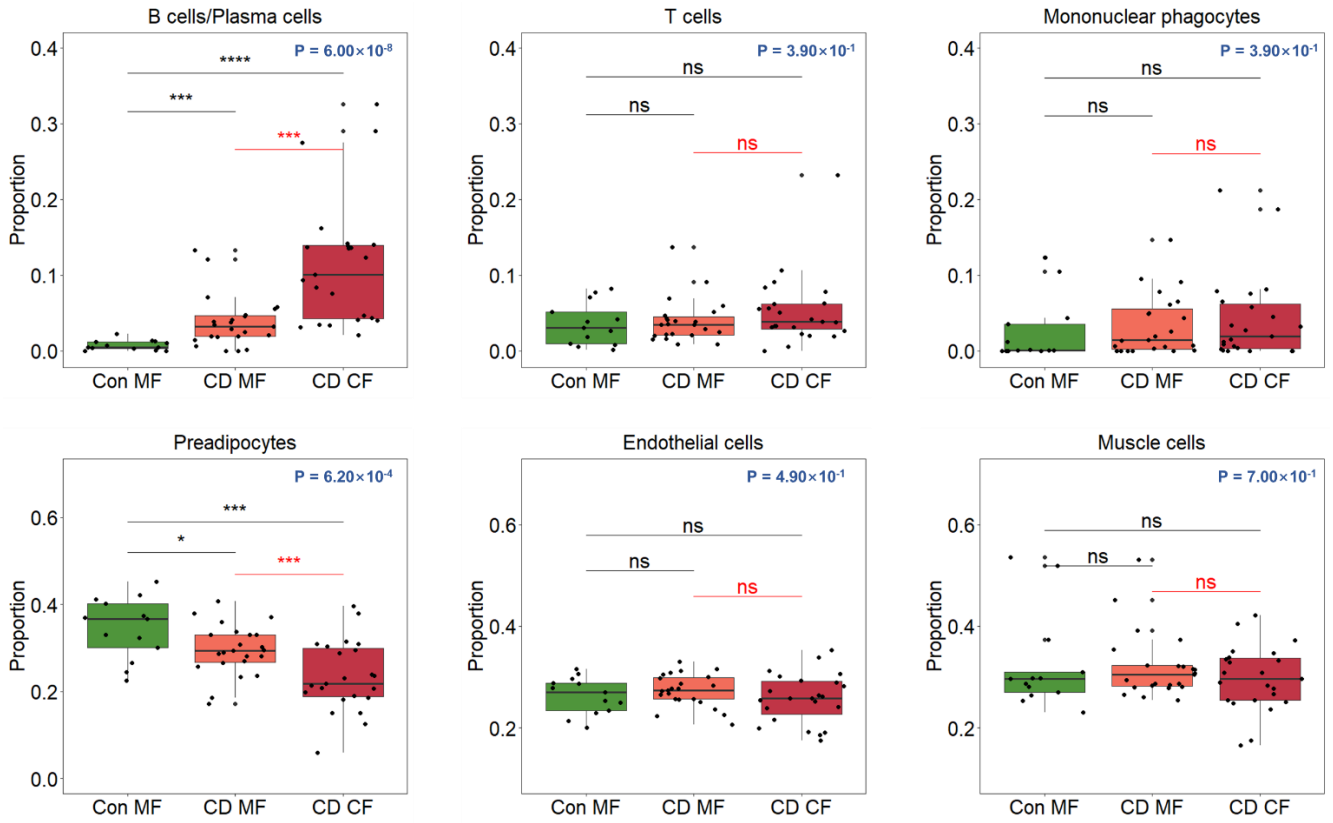

Figure 11. Deconvolution analysis proportion plot.

Boxplots comparing the proportion of the cell types among CD-CF, CD-MAT, and Con-MAT. P-values were depicted on each boxplot (blue: Kruskal-Wallis test, black: two-tailed T-test, red: paired T-test, *P-value <0.05; **P-value <0.01; ***P-value <0.001).

Table 3. Statistical analysis of proportions of deconvoluted cell types[a].

	T-test[b]		paired T-test[b]	Kruskal Wallis test[c]
Cell type	Con MF - CD CF	Con MF - CD MF	CD MF - CD CF	
B cells / Plasma cells	**9.71×10^{-07}**	**1.04×10^{-04}**	**1.61×10^{-04}**	**5.71×10^{-07}**
T cells	0.54	0.73	0.05	0.52
Mononuclear phagocytes	0.59	0.89	0.39	0.69
Endothelial cells	NA	NA	NA	NA
Preadipocytes	**3.05×10^{-03}**	0.10	**1.01×10^{-03}**	**1.48×10^{-03}**
Innate lymphoid cells	NA	NA	NA	NA

CD-CF, creeping fat of Crohn's disease; CD-MAT, mesenteric adipose tissue of Crohn's disease; Con-MAT, mesenteric adipose tissue of control; NA, not available.

[a]Deconvolution reference was adopted from publicly available single-cell RNA cluster matrix (Hildreth A.D. et al. *Nat Immunol* 2021, Ref. 32).

[b]For the comparison of two groups, Student's t-test was used for unmatched groups, and paired t-test was used for paired groups, depending on the statistical distribution of data.

[c]Kruskal-Wallis test was used to compare more than two groups, as data were not normally distributed.
Bold indicates statistically significant difference with a p-value <0.05.

4. Discussion

In our study, distinct transcriptomic profiles were identified for each group with a different MAT status, through DEG analysis of paired MAT samples from CD individuals and non-CD controls. Especially, a shift in gene expression was identified, from Con-MAT to CD-MAT and CD-CF. Among 529 DEGs, 323 genes (61.1%) showed an incremental expression pattern, and marker genes for B-cell/plasma cell and T-cell were upregulated with enrichment of immune response pathways. A significant alteration of cellular composition was observed in CD-CF compared to both CD-MAT and Con-MAT, and this may result in a unique transcriptomic profile of CD-CF.

We compared the gene expression profiles of CD-MAT and Con-MAT and found that CD-MAT may readily respond to local inflammation. Despite enrichment of immune response pathways in CD-MAT, only B-cell/plasma cell composition was significantly increased. This implicates the existence of a preceding transcriptomic change rather than a change in cell composition in CD-MAT, as a primed status for inflammatory responses. This finding raises the possibility of MAT serving as a source of immunoglobulin to protect from intestinal microorganisms.[34] This possibility is partially supported by the findings of Ha et al.[5] who showed that CF expansion was promoted by bacterial translocation in CD. Further studies investigating the serial change of MAT according to disease activities or severities may provide a deeper understanding of the pathogenesis of CF in CD.

Notably, about 20% of the DEGs showed up-bent or down-bent expression patterns. Among 16 common DEGs, *ITLN1* was the only gene with a down-bent pattern. *ITLN1* is a calcium-dependent lectin mainly expressed in goblet cells and is involved in innate immunity with a protective role in the intestinal epithelium.[35-37] Interestingly, rs4656958, which is 2kb 5' of *ITLN1*, was identified as an IBD susceptibility locus in a previous genome-wide association study.[33] Thus, the downregulation of *ITLN1* in CD-CF and CD-MAT could be interpreted as a transcriptome-level change in the pathological state of CD. One explanation could be that *ITLN1* expression is relatively upregulated in MAT to deal with active inflammation. This may support a protective role of CF in local inflammation control. However, the reason for slightly higher expression of *ITLN1* in CD-CF compared to CD-MAT still requires further investigation.

The relationship of CF with intestinal fibrosis, and excessive extracellular matrix (ECM) production is thought to be key for the pathogenesis of CD.[6, 10, 16] Mao et al.[6] revealed that activated muscle cells could produce ECM scaffold, and the fibronectin within the ECM promoted migration of preadipocytes out of the mesenteric fat via integrin α5β1 and they

subsequently differentiated into adipocytes to form CF. However, our study found no significant overexpression of ECM-related genes or integrin genes such as *ITGA5* and *ITGB1* (Table 4–5 and Figure 12–13). The switch from *ITGA5* to *ITGA6* during adipocyte differentiation[38] may explain the downregulation of *ITGA5* and *ITGB1* and slightly upregulated *ITGA6* in mature, adipocyte-rich CF in this study. Instead, the expression of *ITGA4* and *ITGAB7*, which encode integrins α4 and β7 was higher in CD-CF compared to Con-MAT. This finding implies that CF could serve as a reservoir for systemically recruited immune cells, and supports the idea that surgical removal of mesenteric tissue could be a useful approach in the treatment of CD.[13] Whether integrins α4 and β7 are downregulated in CF of CD patients who have been treated with vedolizumab is a subject that deserves further investigation, but it was impossible for us to determine this as our study contained no patients who had been exposed to vedolizumab.

The functional analyses of DEGs showed that genes associated with immune responses were enriched, while genes involved in lipid metabolism such as adipogenesis were diminished in CD-CF and CD-MAT. Caputa et al.[39] investigated the role of adipose tissue on bacterial function in mice and found that the nucleotide binding oligomerization domain containing 1-mediated detection of intracellular bacteria and interferon-γ signaling rewired the transcriptional and metabolic features of perinodal adipocytes from lipid metabolism to immune response, leading to enhanced nitric oxide production for antibacterial functions. As discussed by Han et al., [8] though the actual mechanism of the unique link between immunity and metabolism is unclear, this reverse relationship of immune and metabolic response observed in CF might be an ideal strategy for rapidly controlling invasive microbes.

Our study has several strengths, including a comparison of CF and non-CD controls, as well as paired samples of CF and non-CF from the same individuals. This allows a better understanding of the distinct gene expression profiles and cellular composition of CF, compared to both CD-MAT and Con-MAT. Additionally, the present study also highlights tissue specific transcriptome profiles in macroscopically normal-looking CD-MAT, providing insights into the pathogenesis of CD. Furthermore, as non-CD controls included MAT from UC patients, of which the transcriptomic profile and cellular composition showed more similarity with non-IBD controls, rather than that of CD in the same category of IBD, this supports the idea that CF is a pathognomonic feature of CD, and it may be a clue for the complementary diagnostic tool for CD.

However, our study also had several limitations. Unfortunately, we could not discern whether this transcriptomic alteration was derived from a quantitative or qualitative

Table 4. Expression level of ECM-related genes identified from DEG analysis.

Gene Symbol	Mean normalized value			CD-CF vs Con-MAT[a]		CD-MAT vs Con-MAT		CD-CF vs CD-MAT[b]	CD-CF vs CD-MAT vs Con-MAT[c]	
	CD-CF	CD-MAT	Con-MAT	M-W *P*	T-test *P*	M-W *P*	T-test *P*	paired T-test *P*	K-W *P*	ANOVA *P*
COL6A3	16.01	15.79	15.27		**7.43×10^{-05}**		**2.17×10^{-03}**	0.22		**1.61×10^{-03}**
TGFB1	8.81	7.82	7.81		**8.34×10^{-04}**		0.98	**4.90×10^{-05}**		**6.50×10^{-05}**
TGFBR1	12.05	12.01	12.58		**2.76×10^{-07}**		**1.42×10^{-07}**	0.56		**1.16×10^{-07}**
TGFBR2	13.86	13.82	14.17		**7.77×10^{-03}**		**2.14×10^{-03}**	0.40		**1.62×10^{-03}**
TGFBR3	12.79	13.2	13.44		**2.53×10^{-04}**		0.10	**1.89×10^{-05}**		**4.55×10^{-04}**
FN1	13.88	14.4	14.62		**1.31×10^{-02}**		0.41	**2.35×10^{-02}**		1.09×10^{-02}
COL1A1	12.69	11.59	11.12	**1.32×10^{-03}**		0.38		**1.51×10^{-02}**		2.80×10^{-03}
COL6A1	13.71	13.44	13.22	**5.72×10^{-03}**		0.34		0.11		0.06
TGFB1I1	7.3	7.26	7.03	0.52		0.43		0.87	0.69	
TGFB2	8.47	8.7	8.18		0.29		0.05	**4.82×10^{-02}**		0.12
TGFB3	9.84	10.02	9.9		0.70		0.48	0.19		0.46
TGFBI	12.79	12.88	13.18		0.06		0.11	0.47		0.11
TGFBRAP1	9.75	9.58	9.8		0.62		**4.82×10^{-02}**	0.10		0.07
ACTA2	12.31	13.12	12.09		0.65	**4.00×10^{-02}**		**3.83×10^{-03}**		1.52×10^{-02}
MMP9	7.63	5.79	5.35		**3.21×10^{-02}**		0.66	**4.72×10^{-04}**	1.87×10-02	
TIMP2	11.74	11.68	12.12	**1.96×10^{-02}**		**2.51×10^{-03}**		0.70		3.15×10^{-02}
TIMP3	13.84	14.12	14.36		**1.16×10^{-02}**		0.16	**3.73×10^{-03}**		2.21×10^{-02}
TIMP4	11	11.43	12.31		**1.77×10^{-03}**		**2.06×10^{-02}**	0.05		4.40×10^{-03}

ANOVA, analysis of variance; CD-CF, creeping fat of Crohn's disease; CD-MAT, mesenteric adipose tissue of Crohn's disease; Con-MAT, mesenteric adipose tissue of control; K-W; Kruskal-Wallis test; M-W, Mann-Whitney U test; *P*, *P* value.

[a]For the comparison of unmatched two groups, Student's t-test or Mann-Whitney U test was used, depending on the statistical distribution of data.

[b]Paired t-test was used to compare CD-CF and CD-MAT, as differences of all the mean normalized values between paired groups were normally distributed.

[c]ANOVA or Kruskal-Wallis test was used to compare more than two groups, depending on the statistical distribution of data.

Bold indicates a statistically significant difference with a p-value <0.05, and bold, italic text indicates a significant difference with a Bonferroni-adjusted p-value <0.0028.

Table 5. Expression level of integrins identified from differentially expressed gene (DEG) analysis.

Gene Symbol	Mean normalized value			CD-CF vs Con-MAT[a]		CD-MAT vs Con-MAT[a]		CD-CF vs CD-MAT[a]		CD-CF vs CD-MAT vs Con-MAT[b]	
	CD-CF	CD-MAT	Con-MAT	M-W *P*	T-test *P*	M-W *P*	T-test *P*	Wilcoxon *P*	paired T-test *P*	K-W *P*	ANOVA *P*
ITGA4	11.41	10.40	9.89		2.16×10^{-06}		4.06×10^{-02}		1.40×10^{-06}		6.93×10^{-07}
ITGA6	12.43	12.15	11.86		5.53×10^{-04}		0.06		2.13×10^{-02}		6.33×10^{-04}
ITGA6-AS1	4.31	3.37	1.83		5.83×10^{-03}		0.07		0.21	1.84×10^{-02}	
ITGA8	9.96	10.07	8.69	1.96×10^{-02}		1.62×10^{-02}			0.62	3.17×10^{-02}	
ITGAL	9.84	8.69	7.98		3.44×10^{-06}		2.81×10^{-02}		3.20×10^{-06}		4.13×10^{-08}
ITGB1BP1	12.37	12.83	13.33		8.24×10^{-06}		4.24×10^{-03}		5.36×10^{-04}		6.45×10^{-06}
ITGB2-AS1	5.99	5.21	2.41	6.49×10^{-05}			1.71×10^{-03}		1.64×10^{-03}	6.30×10^{-05}	
ITGB5	12.60	12.89	13.31		1.89×10^{-04}		9.82×10^{-03}		5.78×10^{-03}		2.56×10^{-04}
ITGB7	8.29	7.25	6.12		2.72×10^{-07}		1.18×10^{-03}		5.16×10^{-05}		4.49×10^{-08}
ITGA1	13.04	13.10	13.40	1.33×10^{-02}		2.59×10^{-02}			0.60	3.42×10^{-02}	
ITGA10	5.21	5.77	4.25	0.77		0.43			0.07	0.54	
ITGA11	9.65	10.20	9.91		0.32		0.24		7.95×10^{-03}		2.83×10^{-02}
ITGA2	8.48	8.60	8.63	0.36			0.91		0.42		0.75
ITGA2B	5.72	5.34	5.22		0.20		0.77		0.37	0.31	
ITGA3	8.31	8.20	8.38		0.84	0.95			0.55		0.81
ITGA5	11.23	11.38	11.30		0.73		0.71		0.37		0.68
ITGA7	12.21	12.45	12.97		8.89×10^{-03}		0.05		0.13		1.92×10^{-02}
ITGA9	10.99	11.21	11.34		4.53×10^{-02}		0.44		0.09	0.07	
ITGA9-AS1	8.40	8.30	8.29		0.56		0.96		0.56	0.88	
ITGAD	6.02	4.23	5.14		0.11		0.13		3.24×10^{-04}	3.98×10^{-03}	
ITGAE	9.07	9.24	9.44	4.34×10^{-02}		0.20			0.05		4.10×10^{-02}
ITGAM	10.80	10.68	10.96		0.44		0.19		0.39		0.36
ITGAV	13.52	13.61	13.75		2.69×10^{-02}		0.17		0.26		0.14
ITGAX	9.37	8.74	8.50		0.07		0.60		1.23×10^{-02}		0.06
ITGB1	15.58	15.91	15.87		0.06		0.76	1.77×10^{-03}			1.85×10^{-02}
ITGB1-DT	4.51	5.16	3.43		0.21		4.15×10^{-02}		0.29	3.27×10^{-02}	
ITGB2	12.23	11.87	11.48		5.20×10^{-03}		0.13		8.46×10^{-03}		1.37×10^{-02}
ITGB3	9.67	10.02	10.33		1.07×10^{-03}		0.10		3.97×10^{-02}		5.29×10^{-03}
ITGB3BP	9.21	9.12	8.93		0.12		0.29		0.39		0.23
ITGB4	8.66	9.01	8.85		0.44		0.49		1.43×10^{-02}		0.21
ITGB8	10.41	9.50	10.34		0.86		0.05		4.59×10^{-03}		5.35×10^{-03}
ITGB8-AS1	4.54	3.97	2.90		0.08		0.26		0.34	0.35	
ITGBL1	4.92	4.98	5.86		0.23		0.27		0.93	0.32	

ANOVA, analysis of variance; CD-CF, creeping fat of Crohn's disease; CD-MAT, mesenteric adipose tissue of Crohn's disease; Con-MAT, mesenteric adipose tissue of control; K-W; Kruskal-Wallis test; M-W, Mann-Whitney U test; Wilcoxon; Wilcoxon signed-rank test; *P*, *P* value.

[a]For the comparison of two groups, Student's t-test or Mann-Whitney U test was used for unmatched groups, and paired t-test or Wilcoxon signed-rank test was used for paired groups, depending on the statistical distribution of data.

[b]ANOVA or Kruskal-Wallis test was used to compare more than two groups, depending on the statistical distribution of data.

Bold indicates a statistically significant difference with a p-value <0.05, and bold, italic text indicates a significant difference with a Bonferroni-adjusted p-value <0.0015.

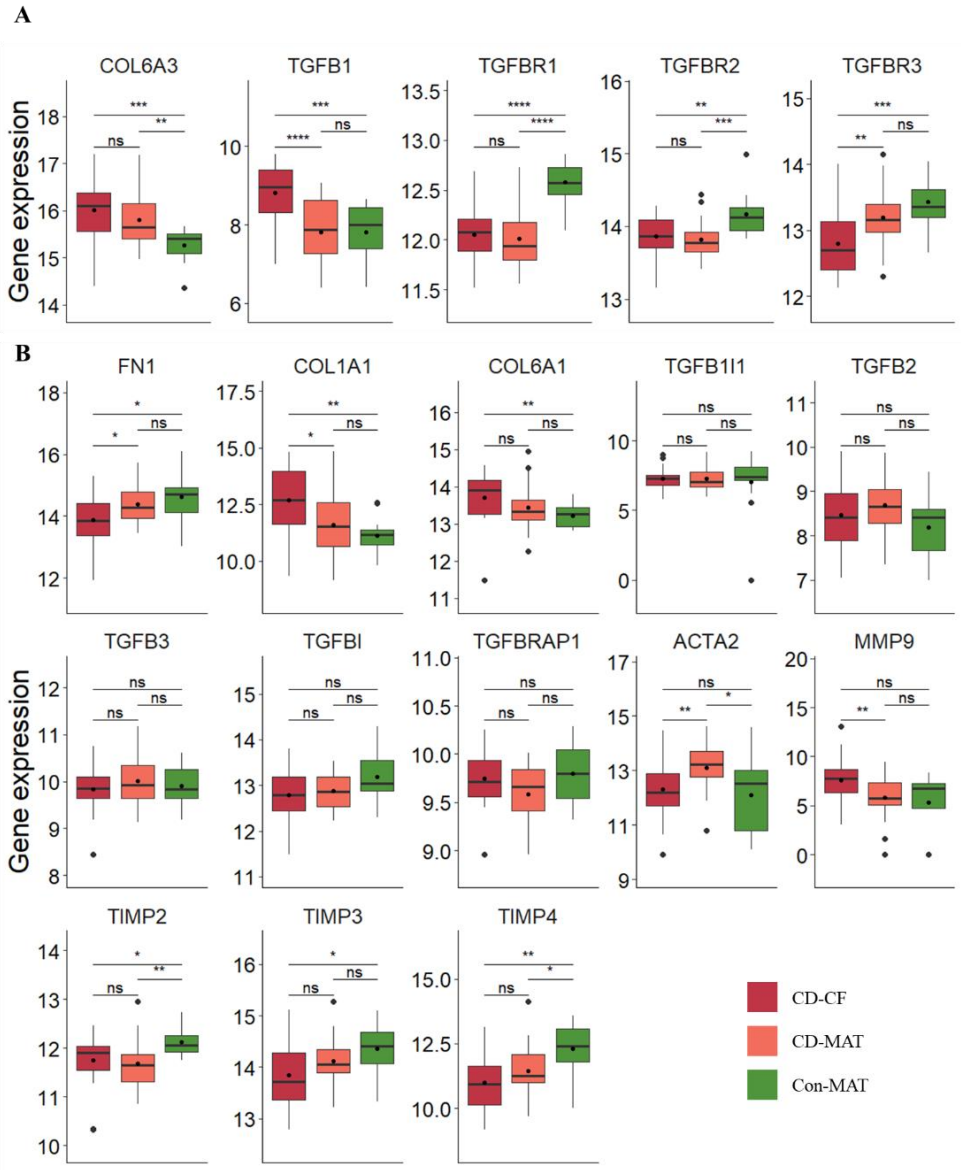

Figure 12. Box plots for gene ECM-related genes among groups.

Box plots of gene expression level of ECM-related genes in CD-CF, CD-MAT, and Con-MAT. (A) Although the expression patterns were inconsistent, the expression levels of *COL6A3*, *TGFB1*, *TGFBR1*, *TGFBR2*, *TGFBR3*, *FN1*, *COL1A1*, *TGFBRAP1*, *ACTA2*, *MMP9*, *TIMP2*, *TIMP3*, and *TIMP4* significantly differed among CD-CF, CD-MAT, and Con-MAT (with the Kruskal-Wallis test or ANOVA P-value <0.05). (B) In contrast, other ECM-related genes demonstrated neither a specific expression pattern nor statistically different expression level among CD-CF, CD-MAT, and Con-MAT. P-values are depicted on each boxplot (blue: Kruskal-Wallis test or ANOVA test, black: two-tailed t-test or Mann-Whitney U test between unmatched group, depending on the data distribution, paired t-test between CD-CF and CD-MAT, *P-value <0.05; **P-value <0.01; ***P-value <0.001).

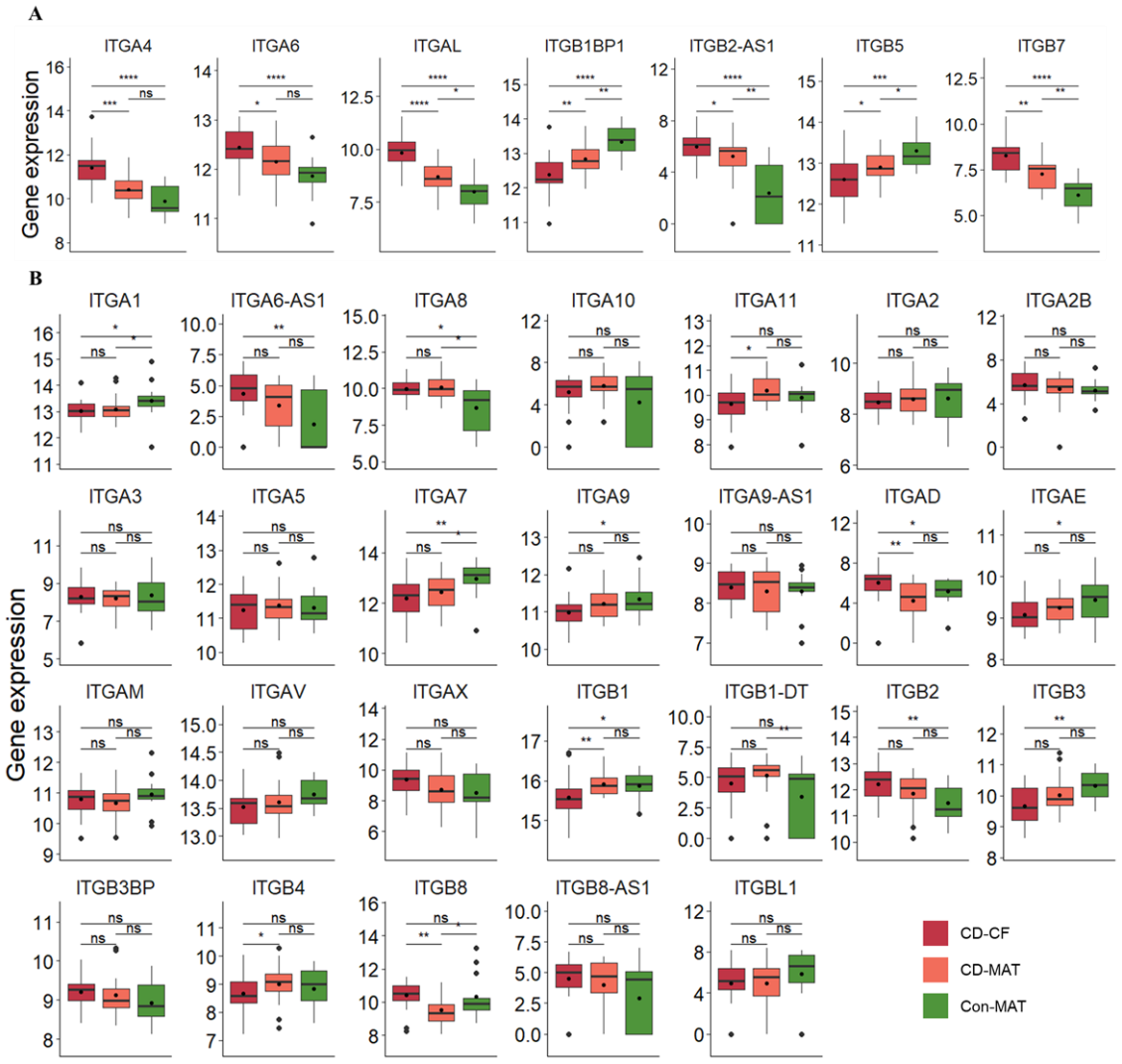

Figure 13. Box plots for gene expression of integrin among groups.

Box plots were used to visualize the gene expression levels of integrin in CD-CF, CD-MAT, and Con-MAT. In (A), although the expression patterns were inconsistent, the expression levels of 19 integrin genes were found to be significantly different among the groups (with the Kruskal-Wallis test or ANOVA P-value <0.05). Notably, *ITGA4*, *ITGA6*, and *ITGB7* showed an incremental pattern, while *ITGB5* showed a decremental pattern. In (B), however, other integrin genes demonstrated neither a specific expression pattern nor statistically different expression levels among CD-CF, CD-MAT, and Con-MAT. P-values were depicted on each boxplot (blue: Kruskal-Wallis test or ANOVA test, black: two-tailed *t*-test or Mann-Whitney U test between unmatched group, depending on the data distribution, paired *t*-test or Wilcoxon signed-rank test between CD-CF and CD-MAT, depending on the data distribution, *P-value <0.05; **P-value <0.01; ***P-value <0.001).

change in CF, due to the limitation of bulk RNA sequencing data. To overcome this, we adopted public scRNA sequencing data to perform deconvolution analysis, but it was limited to discriminate different cell subtypes, such as B-cell vs plasma cell, and M1 vs M2 macrophages. Second, small sample size limited the generalizability of the analysis, and we could not investigate the relationship of CF between clinical outcomes, such as predicting postoperative recurrence. Third, we did not assess transcriptomes of ileal mucosa. Previous studies[14, 40] have already reported that the inflamed ileal mucosa contained profuse immune cell infiltration, dominated by proinflammatory M1 macrophages, whereas the CF was infiltrated with anti-inflammatory M2 macrophages. Given this cell compositional difference, molecular crosstalk between ileal mucosa and CF needs to be further investigated. Finally, this study only investigated transcriptomic data and we could not identify the proteomic or metabolomic aspects, which can function differently from expectation. Further studies with larger sample sizes including a multi-omics approach can play a greater role in uncovering the mechanism of CF formation and its role in CD.

In summary, our study shows that CD-CF and CD-MAT have a distinct gene expression profile, characterized by dynamic immune responses in the progression from Con-MAT to CD-MAT and CD-CF status. These transcriptomic alterations, mostly with an incremental expression of genes enriched with immune response and a decremental expression of genes related to lipid metabolism, seem to mainly arise from changes in cellular composition during the shift from Con-MAT to CD-MAT and CD-CF.

References

1. Silverberg MS, Satsangi J, Ahmad T, *et al.* Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005;**19 Suppl A**:5A-36A.
2. Torres J, Mehandru S, Colombel JF, Peyrin-Biroulet L. Crohn's disease. Lancet 2017;389(10080):1741-55.
3. Peyrin-Biroulet L, Chamaillard M, Gonzalez F, *et al.* Mesenteric fat in Crohn's disease: a pathogenetic hallmark or an innocent bystander? *Gut* 2007;**56**:577-83.
4. Coffey JC, O'Leary DP. The mesentery: structure, function, and role in disease. *Lancet Gastroenterol Hepatol* 2016;**1**:238-47.
5. Ha CWY, Martin A, Sepich-Poore GD, *et al.* Translocation of viable gut microbiota to mesenteric adipose drives formation of creeping fat in humans. *Cell* 2020;**183**:666-83 e17.
6. Mao R, Doyon G, Gordon IO, *et al.* Activated intestinal muscle cells promote preadipocyte migration: a novel mechanism for creeping fat formation in Crohn's disease. *Gut* 2022;**71**:55-67.
7. Karaskova E, Velganova-Veghova M, Geryk M, Foltenova H, Kucerova V, Karasek D. Role of adipose tissue in inflammatory bowel disease. *Int J Mol Sci* 2021;**22**:4226.
8. Han SJ, Glatman Zaretsky A, Andrade-Oliveira V, *et al.* White adipose tissue is a reservoir for memory t cells and promotes protective memory responses to infection. *Immunity* 2017;**47**:1154-68 e6.
9. Crohn BB, Ginzburg L, Oppenheimer GD. Landmark article Oct 15, 1932. Regional ileitis. A pathological and clinical entity. By Burril B. Crohn, Leon Ginzburg, and Gordon D. Oppenheimer. *JAMA* 1984;**251**:73-9.
10. Mao R, Kurada S, Gordon IO, *et al.* The mesenteric fat and intestinal muscle interface: creeping fat influencing stricture formation in Crohn's disease. *Inflamm Bowel Dis* 2019;**25**:421-6.
11. Althoff P, Schmiegel W, Lang G, Nicolas V, Brechmann T. Creeping fat assessed by small bowel MRI is linked to bowel damage and abdominal surgery in Crohn's disease. *Dig Dis Sci* 2019;**64**:204-12.
12. Li Y, Zhu W, Gong J, *et al.* Visceral fat area is associated with a high risk for early postoperative recurrence in Crohn's disease. *Colorectal Dis* 2015;**17**:225-34.

13. Coffey CJ, Kiernan MG, Sahebally SM, *et al.* Inclusion of the mesentery in ileocolic resection for Crohn's disease is associated with reduced surgical recurrence. *J Crohns Colitis* 2018;**12**:1139-50.
14. Kredel LI, Jodicke LJ, Scheffold A, *et al.* T-cell composition in ileal and colonic creeping fat - separating ileal from colonic Crohn's disease. *J Crohns Colitis* 2019;**13**:79-91.
15. da Silva FAR, Pascoal LB, Dotti I, *et al.* Whole transcriptional analysis identifies markers of B, T and plasma cell signaling pathways in the mesenteric adipose tissue associated with Crohn's disease. *J Transl Med* 2020;**18**:44.
16. Zuo L, Li J, Zhang X, *et al.* Aberrant mesenteric adipose extracellular matrix remodelling is involved in adipocyte dysfunction in Crohn's disease: the role of tlr-4-mediated macrophages. *J Crohns Colitis* 2022;**16**:1762-76.
17. Coope A, Pascoal LB, da Silva FAR, *et al.* Transcriptional and molecular pathways activated in mesenteric adipose tissue and intestinal mucosa of Crohn's disease patients. *Int J Inflam* 2017;**2017**:7646859.
18. DeLuca DS, Levin JZ, Sivachenko A, *et al.* RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* 2012;**28**:1530-2.
19. Andrews S. FsatQC. A quality control tool for high throughput sequence data. Cambridge: Babraham Institute, 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc Accessed Feb 12, 2023.
20. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 2016;**32**:3047-8.
21. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;**17**:10-2.
22. Dobin A, Davis CA, Schlesinger F, *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15-21.
23. Frankish A, Diekhans M, Ferreira AM, *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 2019;**47**:D766-D73.
24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550.
25. Subramanian A, Tamayo P, Mootha VK, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;**102**:15545-50.
26. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The

Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;**1**:417-25.

27. Ashburner M, Ball CA, Blake JA, *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;**25**:25-9.
28. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;**28**:27-30.
29. Chen EY, Tan CM, Kou Y, *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013;**14**:128.
30. Kuleshov MV, Jones MR, Rouillard AD, *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;**44**:W90-7.
31. Steen CB, Liu CL, Alizadeh AA, Newman AM. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. *Methods Mol Biol* 2020;**2117**:135-57.
32. Hildreth AD, Ma F, Wong YY, et al. Single-cell sequencing of human white adipose tissue identifies new cell states in health and obesity. *Nat Immunol* 2021;22(5):639-53.
33. de Lange KM, Moutsianas L, Lee JC, *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;**49**:256-61.
34. Bunker JJ, Flynn TM, Koval JC, *et al.* Innate and adaptive humoral responses coat distinct commensal bacteria with immunoglobulin A. *Immunity* 2015;**43**:541-53.
35. Wrackmeyer U, Hansen GH, Seya T, Danielsen EM. Intelectin: a novel lipid raft-associated protein in the enterocyte brush border. *Biochemistry* 2006;**45**:9188-97.
36. Nonnecke EB, Castillo PA, Dugan AE, *et al.* Human intelectin-1 (ITLN1) genetic variation and intestinal expression. *Sci Rep* 2021;**11**:12889.
37. Kanke M, Kennedy Ng MM, Connelly S, *et al.* Single-cell analysis reveals unexpected cellular changes and transposon expression signatures in the colonic epithelium of treatment-naive adult Crohn's disease patients. *Cell Mol Gastroenterol Hepatol* 2022;**13**:1717-40.
38. Liu J, DeYoung SM, Zhang M, Zhang M, Cheng A, Saltiel AR. Changes in integrin expression during adipocyte differentiation. *Cell Metab* 2005;**2**:165-77.
39. Caputa G, Matsushita M, Sanin DE, *et al.* Intracellular infection and immune system cues rewire adipocytes to acquire immune function. *Cell Metab* 2022;**34**:747-60 e6.
40. Lissner D, Schumann M, Batra A, *et al.* Monocyte and M1 macrophage-induced barrier defect contributes to chronic intestinal inflammation in IBD. *Inflamm Bowel Dis* 2015;**21**:1297-305.

국문요약

크론병은 위장관의 만성 재발성 염증 질환으로 아직 그 발생 기전은 불명확하다. 장간막 지방이 비후해져 장을 둘러싸는 현상은 크론병의 질병 징후학적 특징이다. 그러나 크론병 병인에서 장간막 비후의 역할은 규명되어 있지 않다. 본 연구는 전사체 분석을 통해 크론병에서 비대해진 장간막 지방의 유전자 발현의 특징을 확인하고 이들의 면역학적 역할에 대해 탐색하고자 한다. 크론병 환자 23명에서 염증부위와 비염증부위의 장간막 지방 조직 한쌍을 수집하였으며, 비크론병 환자 13명의 장간막 지방 조직을 수집하였다. 염증영역 장간막 지방, 비염증영역 장간막 지방, 및 비크론병 환자의 장간막 지방을 대상으로 차등유전자발현 분석을 각각 수행하여, 529개의 차등발현유전자를 식별하였다. 크론병 장간막 지방에서 B 세포/형질세포 및 T 세포 마커 발현의 증가로 면역 관련 유전자 발현의 증가가 확인되었으며, 세포 대사 및 지방생성 관련 유전자의 감소를 확인하였다. 또한, 529개의 차등발현유전자 중 323개의 유전자는 비크론병 장간막 지방에서 크론병 비염증영역 장간막 지방, 크론병 염증영역 장간막 지방 순으로 발현이 증가하는 양상을 보였으며, 105개의 유전자는 감소하는 양상을 보였다. 세포 유형 구성에서는 크론병 염증영역 장간막 지방에서 B 세포/형질세포 비율의 증가가 크론병 비염증영역 장간막 지방, 비크론병 장간막 지방에 비해 유의하게 증가($p=6.00\times10^{-8}$)한 것을 확인할 수 있으며, 지방전구세포는 유의하게 감소($p=6.20\times10^{-4}$)한 것을 확인할 수 있다. 이러한 결과는 크론병 염증영역 장간막 지방이 크론병 비염증영역 장간막 지방, 비크론병 장간막 지방 보다 T 세포, B 세포/형질세포 관련 유전자군이 상향 발현하며, 면역세포의 구성이 증가함을 알 수 있다. 이를 통하여 크론병 염증영역의 비대해진 장간막 지방의 전사체의 구조에 대한 통찰력을 제공하고 크론병에서의 면역학적 역할에 대한 추론을 제공한다.