

Adult Crohn's Disease Exhibits Transcriptional Activation of the Atherosclerosis Signaling Pathway

Ahmed Hafez, BS¹, Kevin Le, BS¹, Stephanie Sutton, MPH¹, Sarah Voskamp, MD¹, Jennifer Nelson, MD, MS^{1,2,*}, Pablo Palomo, MD³

¹University of Central Florida College of Medicine, Orlando, FL, USA

²Department of Cardiovascular Services, Nemours Children's Hospital, Florida. Orlando, FL, USA 32827

³Nemours Children's Health, Division of Gastroenterology, Orlando, FL, USA

*Correspondence should be addressed to Jennifer Nelson, Jennifer.Nelson@nemours.org

Received date: July 10, 2026, **Accepted date:** September 18, 2026

Citation: Hafez A, Le K, Sutton S, Voskamp S, Nelson J, Palomo P. Adult Crohn's Disease Exhibits Transcriptional Activation of the Atherosclerosis Signaling Pathway. J Clin Cardiol. 2026;7(2):171-176.

Copyright: © 2026 Hafez A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Background and aim: Crohn's Disease (CD) is associated epidemiologically with Atherosclerotic Cardiovascular Disease (ASCVD), but the molecular overlap remains incompletely defined. We asked whether adult CD intestinal biopsy transcriptomes are enriched for genes assigned to the Ingenuity Pathway Analysis (IPA) atherosclerosis signaling pathway. This analysis was designed to evaluate transcriptional overlap, not clinical ASCVD or vascular disease.

Methods: An integrative gene-expression meta-analysis of adult CD versus non-IBD intestinal controls was conducted using the Search Tag Analyze Resource for Gene Expression Omnibus (STARGEO). Seventeen GEO series (594 CD samples, 355 controls) were pooled using inverse-variance fixed- and random-effects models; random-effects estimates were used for the primary analysis. Genes meeting the analysis threshold (false-discovery-rate-adjusted $p < 0.05$ and $|\text{experimental log ratio}| > 0.1$) were evaluated with IPA. Clinical metadata were insufficiently harmonized for reliable subgroup analyses.

Results: A total of 464 genes were differentially expressed in CD versus controls (344 upregulated, 120 downregulated). IPA ranked atherosclerosis signaling seventh among enriched canonical pathways (right-tailed Fisher exact $p = 1.3 \times 10^{-7}$; activation z-score = 3.357). Twenty-four upregulated genes mapped to this knowledge-base pathway, including inflammatory cytokines (IL1B, IL6, IFNG, CXCL8), adhesion molecules (SELE, SELP), chemokines (CCL2, CCL11), and inflammatory effectors (S100A8, MMP1, MMP3, LYZ). IPA also predicted inhibition of IL-10 signaling (z-score = -0.728).

Conclusions: Adult CD intestinal tissue shows an inflammatory transcriptional pattern enriched for genes contained in the IPA atherosclerosis signaling pathway. This finding demonstrates molecular overlap only; it does not establish accelerated atherosclerosis, increased ASCVD events, or a need to alter cardiovascular prophylaxis. Because no independent transcriptomic, qPCR, protein, serum, blood, or vascular-tissue validation was performed, the results should be considered hypothesis-generating.

Keywords: Atherosclerosis, Crohn's disease, IBD, Genetics, Molecular

Introduction

Crohn's Disease (CD) is a chronic Inflammatory Bowel Disease (IBD) characterized by relapsing Gastrointestinal (GI) inflammation. Although patients primarily suffer from GI symptoms, the impact of CD extends beyond the gut. Systemic inflammation in IBD has been increasingly associated with extraintestinal complications, including an

elevated risk of atherosclerotic cardiovascular disease [1,2]. Epidemiologic studies indicate higher rates of myocardial infarction and stroke in individuals with IBD, even after adjusting for traditional cardiovascular risk factors [3]. This risk appears most pronounced during periods of active disease flare, when inflammation is highest [2]. Such observations suggest that chronic inflammation is a critical driver linking CD to atherosclerosis development.

Atherosclerosis is well recognized as an immune-inflammatory disease of the arteries, driven by maladaptive inflammation in the vessel wall, and CD similarly reflects an aberrant immune response in the gut. It is therefore plausible that common pathophysiological mechanisms underlie both conditions [4]. Both conditions involve innate and adaptive immune responses, cytokine release, recruitment of leukocytes, and endothelial dysfunction. For example, Tumor Necrosis Factor- α (TNF- α), a pivotal cytokine in CD, can induce endothelial activation and reduce nitric oxide availability, contributing to vascular dysfunction [2]. Interleukin-6 (IL-6), another cytokine elevated in active IBD, is linked to endothelial dysfunction and early atherogenesis [2]. Likewise, interleukin-1 β (IL-1 β), a master pro-inflammatory cytokine found in CD lesions, plays multiple roles in atherothrombotic plaque development including promoting adhesion molecule expression and inducing IL-6 among other chemokines [5]. Elevated systemic markers of inflammation in IBD patients such as C-reactive protein, TNF- α , IL-1, IL-6, and IL-8 may directly contribute to accelerated endothelial injury and plaque formation [1,6]. Chronic cytokine-mediated inflammation is a shared pathogenic feature of both CD and atherosclerosis.

Despite this biological overlap, direct molecular evidence connecting intestinal CD transcription to atherosclerosis-related signaling remains incomplete. The aim of this study was to determine whether the pooled gene-expression profile of adult CD intestinal tissue is enriched for genes assigned to an atherosclerosis-related canonical pathway and to describe the shared inflammatory mediators. The study did not measure vascular gene expression, atherosclerotic plaque, subclinical vascular imaging, serum biomarkers, or cardiovascular events; accordingly, its clinical interpretation is hypothesis-generating.

Methods

Data source and eligibility

STARGEO, a curated platform for integrated analysis of GEO gene-expression data, was used to compare adult CD and non-IBD intestinal tissue. The analysis included 17 GEO series and 949 samples: 594 adult CD samples and 355 controls derived from ileal or colonic biopsies. Series were eligible when they contained both adult CD intestinal samples and non-IBD intestinal controls. Samples labeled as pediatric, ulcerative colitis, active infection, experimental genetic manipulation, or a treatment-intervention time point were excluded when those attributes were identifiable in the GEO annotations. A series without an eligible internal control group was excluded. However, disease activity, exact biopsy location, age, sex, medication exposure, smoking status, and disease duration were variably or incompletely reported and could not be

harmonized across all series; treatment-naïve status therefore cannot be assumed for every sample.

All source data were publicly available and de-identified, and no new participants were recruited. Data use followed GEO requirements. Dataset-level clinical covariates were not available with sufficient completeness or consistency for adjusted models or reliable prespecified subgroup analyses.

STARGEO differential-expression meta-analysis

For each gene within each eligible series, STARGEO estimated the case-control difference in normalized expression and its standard error. Series-level estimates were combined using inverse-variance fixed-effects and DerSimonian-Laird random-effects models. Fixed-effects weights are proportional to $1/SE^2$, whereas random-effects weights are proportional to $1/(SE^2 + \tau^2)$, where τ^2 is the estimated between-series variance. Both model outputs were reviewed, and the random-effects pooled estimate was used for gene selection. Benjamini-Hochberg false-discovery-rate correction was applied across genes. STARGEO reports the pooled estimate as an experimental log ratio [15].

Heterogeneity, weighting, and robustness

STARGEO weighting and heterogeneity are gene-specific: the number of contributing series, study weights, Cochran Q , I^2 , and τ^2 may differ from one gene to another. Consequently, there is no single study weight or heterogeneity value for the IPA atherosclerosis signaling pathway. The analysis file retained for IPA contained pooled gene-level effects and p-values but not the series-specific estimates needed to reconstruct gene-level forest plots, funnel plots, or leave-one-series-out analyses. Formal small-study-effect or publication-bias testing was therefore not performed. Because clinical covariates were not harmonized, subgroup analyses by biopsy site, activity, treatment, demographic factors, smoking, or disease duration were also not performed. These constraints limit robustness assessment.

Pathway and network analysis

To interpret pooled gene-level differences, we used IPA, a curated knowledge-base tool [7]. Genes meeting the prespecified threshold (false-discovery-rate-adjusted $p < 0.05$ and absolute experimental log ratio > 0.1) yielded 464 differentially expressed genes for the IPA core analysis. Canonical-pathway enrichment was assessed with a right-tailed Fisher exact test comparing observed with expected pathway overlap. IPA's activation z-score compares the direction of expression changes with directionally curated relationships and predicts relative pathway activation or inhibition. The z-score is not a measure of effect size, vascular pathology, or clinical ASCVD risk.

The IPA atherosclerosis signaling canonical pathway was prespecified as the pathway of interest. After enrichment was identified, differentially expressed genes mapping to that pathway were extracted with their pooled experimental log ratios and meta-analysis p-values. The pathway label reflects IPA knowledge-base membership and contains many mediators of general inflammation; enrichment therefore cannot by itself distinguish atherosclerosis-specific biology from active intestinal inflammation.

Upstream regulator analysis

IPA's Upstream Regulator Analysis identifies molecules whose curated downstream relationships are consistent with the observed expression pattern and assigns an activation z-score predicting relative activation or inhibition [7]. These results are computational predictions and were not validated experimentally.

Results

Differentially regulated genes in adult CD

Using the random-effects pooled estimate and the prespecified thresholds, 464 genes were differentially expressed in adult CD intestinal tissue relative to controls (344 upregulated and 120 downregulated; false-discovery-rate-adjusted $p < 0.05$). **Table 1** lists the 10 largest positive and 10 largest negative experimental log ratios. Highly upregulated genes included DUOX2, MMP1, LCN2, MMP3, S100A8, CXCL1, and CXCL6; genes with negative estimates included CLDN8, TMIGD1, CDHR1, GUCA2A, and GUCA2B. MMP1, MMP3, and S100A8 also mapped to the IPA atherosclerosis signaling pathway, demonstrating annotation overlap.

Dataset heterogeneity and analysis scope

The pooled contrast combined ileal and colonic biopsies profiled on different platforms. Disease activity, biopsy segment, age, sex, treatment exposure, smoking, and disease duration were not reported consistently enough to evaluate effect consistency across these strata. No covariate-adjusted, subgroup, or leave-one-series-out analysis was available; the pooled findings therefore represent an aggregate adult intestinal CD signal and should not be interpreted as uniform across clinical or tissue subgroups.

Enrichment of the atherosclerosis signaling pathway

In the IPA canonical pathway analysis, atherosclerosis signaling ranked seventh by $-\log(p\text{-value})$ among enriched pathways. The pathway-overlap p-value was 1.3×10^{-7} , and the IPA activation z-score was 3.357, indicating that the direction of the intestinal CD expression pattern was consistent with predicted activation of this curated pathway. Twenty-four differentially expressed genes mapped to the pathway, and all 24 had positive pooled estimates. These included inflammatory cytokines (IL1A, IL1B, IL6, IFNG), adhesion molecules (ICAM1, SELE, SELP), chemokines (CCL2, CCL11, CXCL8), matrix-remodeling enzymes (MMP1, MMP3), and other inflammatory or lipid-related mediators. **Table 2** presents the 24 genes with pooled experimental log ratios and meta-analysis p-values. This enrichment result does not measure arterial expression, plaque burden, or cardiovascular events.

Adult CD also showed enrichment of multiple innate and adaptive immune pathways (**Table 3**), including pathogen-induced cytokine storm, neutrophil degranulation, leukocyte extravasation signaling, IL-6 signaling, IL-1 signaling, and toll-

Table 1. Top differentially expressed genes in adult Crohn's disease.

Upregulated Genes	Experimental Log Ratio	Downregulated Genes	Experimental Log Ratio
DUOX2	0.649	CLDN8	-0.414
MMP1*	0.538	TMIGD1	-0.378
LCN2	0.523	CDHR1	-0.345
MMP3*	0.520	PAQR5	-0.291
S100A8*	0.520	GUCA2A	-0.283
CXCL1	0.469	LINC01389	-0.279
SLC6A14	0.469	GUCA2B	-0.277
LOC102724971/LOC642131	0.460	MS4A10	-0.259
REG1B	0.450	MT-ND3	-0.248
CXCL6	0.444	DPP10-AS1	-0.245

*Genes also present in atherosclerosis signaling pathway

Table 2. Crohn's disease-atherosclerosis overlap: differentially expressed genes in the atherosclerosis signaling pathway.

Gene	Random-effects Experimental Log ₂ Ratio	Meta-analysis p-value
APOL1	0.174	1.36E-13
CCL11	0.195	2.39E-11
CCL2	0.213	5.29E-08
COL18A1	0.104	4.92E-09
COL1A1	0.136	2.03E-11
COL1A2	0.151	1.20E-09
CXCL8	0.419	2.11E-10
F3	0.101	1.53E-07
ICAM1	0.173	2.35E-07
IFNG	0.131	1.27E-04
IL1A	0.161	2.13E-05
IL1B	0.442	1.02E-08
IL1RN	0.197	1.83E-08
IL6	0.203	1.36E-07
LPL	0.205	2.16E-07
LYZ	0.103	3.19E-05
MMP1	0.538	2.30E-07
MMP3	0.52	9.60E-08
PLA2G2A	0.117	1.21E-04
PLA2G7	0.114	5.87E-04
PLAAT3	0.11	6.93E-06
S100A8	0.52	3.63E-09
SELE	0.194	3.00E-07
SELP	0.212	1.08E-07

Note. Gene-level pooled values are shown. Membership in the IPA atherosclerosis signaling pathway indicates knowledge-base overlap and does not establish clinical or histologic atherosclerosis.

Table 3. Significant enriched canonical pathways in adult Crohn's disease.

Canonical Pathway	IPA Activation z-score (prediction)	# of Overlapped Genes
Pathogen induced cytokine storm	5.231	37
Neutrophil degranulation	4.617	33
Leukocyte extravasation signaling	4.112	27
IL-6 signaling	3.989	19
IL-1 signaling	3.478	16
Toll-like receptor signaling	3.675	22
Atherosclerosis signaling	3.357	24
IL-10 signaling	-0.728	19

Note. Positive or negative z-scores are IPA predictions based on curated directional relationships; they are not clinical effect sizes. Pathways are displayed as reported and are not all ordered strictly by z-score.

like receptor signaling. IPA predicted relative inhibition of IL-10 signaling (activation z-score = -0.728).

IPA upstream regulator analysis identified TNF, IFNG, bacterial lipopolysaccharide, and IL1B as predicted regulators whose known targets were consistent with the observed expression pattern. These predictions support a shared inflammatory transcriptional program but were not confirmed by regulator-specific experiments.

Discussion

This integrative analysis identified an intestinal inflammatory expression pattern in adult CD that overlaps with genes contained in a curated atherosclerosis-related pathway. The overlap includes IL1B, IL6, IFNG, CCL2, CCL11, CXCL8, ICAM1, selectins, and MMPs. However, many of these mediators are not specific to atherosclerosis and are expected in active mucosal inflammation. The result is therefore best interpreted as a molecular hypothesis that is biologically compatible with, but does not prove or quantify, the epidemiologic association between IBD and ASCVD.

Shared inflammatory mediators (IL-1, IL-6, IFN- γ)

Several cytokine transcripts implicated in both CD and atherosclerosis, including IL1B and IL6, had positive pooled estimates. IL-1 β promotes innate immune activation in CD and contributes to endothelial and leukocyte activation in experimental atherosclerosis [1,5]. IL-6 participates in active intestinal inflammation, endothelial dysfunction, and hepatic acute-phase signaling [1]. Epidemiologic studies report a modestly higher myocardial infarction risk in IBD populations [3,8]. The present intestinal transcriptomic result offers a possible shared-inflammatory explanation for that association, but it does not estimate the magnitude of clinical risk.

IFN- γ also had a positive pooled estimate, consistent with the Th1-skewed immune response characteristic of CD. IFN- γ is expressed in atherosclerotic plaques and can promote macrophage activation, inflammatory cytokine release, adhesion-molecule expression, and impaired collagen synthesis [2]. STAT1, a downstream transcription factor, also had a positive pooled estimate (experimental log ratio 0.18). Together, these observations indicate overlaps in IFN- γ -associated inflammatory signaling.

Adhesion-related transcripts provided another area of overlap. SELE, SELP, and ICAM1 had positive pooled estimates. In inflamed gut, cytokine-induced adhesion molecules facilitate leukocyte recruitment; in early atherogenesis, related molecules support monocyte adhesion to arterial endothelium [2,9]. The shared gene annotation is mechanistically plausible, but intestinal expression cannot be used as a surrogate for systemic endothelial activation without blood- or vascular-tissue validation.

MMP1 and MMP3 were among the positively expressed genes mapping to the IPA pathway. MMPs participate in intestinal extracellular-matrix remodeling and, in vascular plaques, can contribute to matrix degradation [10,11].

Protective counter-regulatory pathways (IL-1Ra, IL-10)

The results also contained potential counter-regulatory signals. IL1RN, which encodes the interleukin-1 receptor antagonist, had a positive pooled estimate and may reflect feedback to local IL-1 activity. The CANTOS trial supports a causal role for IL-1 β signaling in recurrent cardiovascular events in a selected post-myocardial-infarction population, [5] but it does not establish that IL-1 blockade is indicated for ASCVD prevention in CD or that intestinal IL1RN expression is cardioprotective.

IPA predicted inhibition of IL-10 signaling (activation z-score = -0.728). IL-10 has anti-inflammatory functions in both intestinal and experimental vascular inflammation; IL-10 deficiency aggravates atherosclerosis in animal models and causes severe early-onset intestinal inflammation in humans and mice [12–14].

The concurrent positive inflammatory signals and predicted IL-10 pathway inhibition generate hypotheses about regulatory imbalance in CD. They do not provide evidence to recommend IL-1, IL-6, or IL-10-directed therapy for cardiovascular prevention.

Interpretation, inflammatory confounding, and tissue specificity

The observed expression pattern arose from intestinal biopsies and included cytokines, chemokines, adhesion molecules, and proteases that are common to many inflammatory states. IPA pathway enrichment is based on overlap with a curated gene set and directional knowledge; it does not demonstrate histologic atherosclerosis, accelerated plaque formation, or increased ASCVD events in the sampled individuals. The result may primarily reflect active intestinal inflammation rather than an atherosclerosis-specific program. Thus, the analysis supports biological overlap and hypothesis generation, not individual risk prediction.

Tissue specificity further limits inference. Intestinal expression cannot be assumed to represent peripheral blood, vascular endothelium, or atherosclerotic plaque. The clinical relationship between CD and cardiovascular disease is multifactorial and may be influenced by disease activity, systemic inflammatory burden, genetics, smoking, diet, metabolic risk factors, and anti-inflammatory treatment. These variables were not uniformly available in the pooled GEO metadata and were not adjusted for in this analysis.

Limitations

The principal limitation is the absence of independent validation. We did not reproduce the signature in a separate external cohort or confirm selected genes using RNA sequencing, qPCR, protein expression, serum biomarkers, peripheral blood, vascular tissue, or functional experiments. The IPA result therefore remains computational and hypothesis-generating, and causality or clinical cardiovascular relevance cannot be inferred.

The 17 GEO series also differed in platform, normalization, tissue segment, disease activity, participant characteristics, and exposure reporting. Age, sex, medication exposure, smoking, and disease duration were incompletely harmonized, precluding covariate adjustment and meaningful subgroup analyses. Gene-specific heterogeneity statistics, series-level weights, leave-one-series-out analyses, forest plots, funnel plots, and formal small-study-effect assessments were not available from the retained IPA input. Combining ileal and colonic tissue may obscure site-specific biology, and the modest effect-size cutoff may admit small signals. Finally, IPA may be affected by literature and knowledge-base bias, and the pathway includes nonspecific inflammatory genes. These limitations reduce confidence in the robustness, specificity, and generalizability of the findings.

Conclusions

In pooled adult intestinal transcriptomic data, CD was associated with an inflammatory expression pattern enriched for genes assigned to the IPA Atherosclerosis Signaling pathway. This finding indicates transcriptional overlap between intestinal CD inflammation and a curated atherosclerosis-related gene set; it does not demonstrate accelerated atherosclerosis, predict ASCVD events, or justify changes to cardiovascular screening, prophylaxis, or anti-cytokine therapy.

Because the analysis lacks independent transcriptomic and orthogonal validation and is limited by heterogeneous, incompletely characterized intestinal datasets, the conclusions should be regarded as hypothesis-generating. Future work should validate the signal in independent cohorts, confirm key targets by qPCR or protein assays, evaluate blood and vascular compartments, and test whether validated molecular measures correlate with disease activity, vascular imaging, or prospective cardiovascular outcomes.

Specific Author Contributions

A.H.: Conceptualization, study design, data interpretation, and drafting of the introduction, abstract, discussion, and results sections.

S.V.: Data acquisition and organization of gene expression datasets.

S.S. and K.L.: Writing and editing of the results section and assistance with data interpretation.

J.N.: Supervision of laboratory activities, coordination of IPA software access in collaboration with P.P., and critical review and editing of the manuscript.

P.P.: Oversight of data analysis environment, facilitation of IPA availability, and substantive review and editing of the final manuscript.

Financial Support

The IPA Software license was supported by Nemours Central Florida Research.

Potential Competing Interests

The authors have no financial disclosures or conflicts of interest.

References

1. Łykowska-Szuber L, Rychter AM, Dudek M, Ratajczak AE, Szymczak-Tomczak A, Zawada A, et al. What Links an Increased Cardiovascular Risk and Inflammatory Bowel Disease? A Narrative Review. *Nutrients.* 2021 Jul 30;13(8):2661.
2. Gabbiadini R, Dal Buono A, Mastroiocco E, Solitano V, Repici A, Spinelli A, et al. Atherosclerotic cardiovascular diseases in inflammatory bowel diseases: to the heart of the issue. *Front Cardiovasc Med.* 2023 May 16;10:1143293.
3. Panhwar MS, Mansoor E, Al-Kindi SG, Sinh P, Katz J, Oliveira GH, et al. Risk of Myocardial Infarction in Inflammatory Bowel Disease: A Population-based National Study. *Inflamm Bowel Dis.* 2019 May 4;25(6):1080–7.
4. Masson W, Fernández-Villar G, Martínez-Elhelou S. Management of Atherosclerotic Cardiovascular Risk in Inflammatory Bowel Disease: Current Perspectives. *Adv Ther.* 2025 May;42(5):2118–34.
5. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, et al. Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med.* 2017 Sep 21;377(12):1119–31.
6. Principi M, Mastroiocco M, Scicchitano P, Gesualdo M, Sassara M, Guida P, et al. Endothelial function and cardiovascular risk in active inflammatory bowel diseases. *J Crohns Colitis.* 2013 Nov;7(10):e427–33.
7. Krämer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014 Feb 15;30(4):523–30.

8. Sun HH, Tian F. Inflammatory bowel disease and cardiovascular disease incidence and mortality: A meta-analysis. *Eur J Prev Cardiol.* 2018 Oct;25(15):1623–31.
9. Hashimoto M, Shingu M, Ezaki I, Nobunaga M, Minamihara M, Kato K, et al. Production of soluble ICAM-1 from human endothelial cells induced by IL-1 beta and TNF-alpha. *Inflammation.* 1994 Apr;18(2):163–73.
10. Kirkegaard T, Hansen A, Bruun E, Brynskov J. Expression and localisation of matrix metalloproteinases and their natural inhibitors in fistulae of patients with Crohn's disease. *Gut.* 2004 May;53(5):701–9.
11. Ait-Oufella H, Taleb S, Mallat Z, Tedgui A. Recent advances on the role of cytokines in atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2011 May;31(5):969–79.
12. Mallat Z, Besnard S, Duriez M, Deleuze V, Emmanuel F, Bureau MF, et al. Protective role of interleukin-10 in atherosclerosis. *Circ Res.* 1999 Oct 15;85(8):e17–24.
13. Glocker EO, Kotlarz D, Boztug K, Gertz EM, Schäffer AA, Noyan F, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med.* 2009 Nov 19;361(21):2033–45.
14. Kühn R, Löhler J, Rennick D, Rajewsky K, Müller W. Interleukin-10-deficient mice develop chronic enterocolitis. *Cell.* 1993 Oct 22;75(2):263–74.
15. Hadley D, Pan J, El-Sayed O, Aljabban J, Aljabban I, Azad TD, et al. Precision annotation of digital samples in NCBI's gene expression omnibus. *Sci Data.* 2017 Sep 19;4:170125.