Journal of Crohn's and Colitis, 2022, **16**, 1306–1320
https://doi.org/10.1093/ecco-jcc/jjac027
Advance access publication 12 February 2022
Review Article

Tailoring Multi-omics to Inflammatory Bowel Diseases: All for One and One for All

Padhmanand Sudhakar,[a] Dahham Alsoud,[a] Judith Wellens,[a] Sare Verstockt,[a] Kaline Arnauts,[a] Bram Verstockt,[a,b] Severine Vermeire[a,b]

[a]KU Leuven Department of Chronic Diseases, Metabolism and Ageing, Translational Research Center for Gastrointestinal Disorders [TARGID], Leuven, Belgium
[b]Department of Gastroenterology and Hepatology, University Hospitals Leuven, KU Leuven, Leuven, Belgium

Corresponding author: Padhmanand Sudhakar, Translational Research in Gastrointestinal Disorders [TARGID], ON I, Herestraat 49, box 701, 3000 Leuven, Belgium. Tel.: 0032 [0]16 19 49 40; email: padhmanand.sudhakar@kuleuven.be

Abstract Inflammatory bowel disease [IBD] has a multifactorial origin and originates from a complex interplay of environmental factors with the innate immune system at the intestinal epithelial interface in a genetically susceptible individual. All these factors make its aetiology intricate and largely unknown. Multi-omic datasets obtained from IBD patients are required to gain further insights into IBD biology. We here review the landscape of multi-omic data availability in IBD and identify barriers and gaps for future research. We also outline the various technical and non-technical factors that influence the utility and interpretability of multi-omic datasets and thereby the study design of any research project generating such datasets. Coordinated generation of multi-omic datasets and their systemic integration with clinical phenotypes and environmental exposures will not only enhance understanding of the fundamental mechanisms of IBD but also improve therapeutic strategies. Finally, we provide recommendations to enable and facilitate generation of multi-omic datasets.

Key Words: IBD complexity; -omic datasets; sparsity; exposome; validation; strategies tailored to IBD; coordinated funding; collaborative research

1. Introduction

Inflammatory bowel disease [IBD] is a disorder of the gut characterised by prolonged periods of relapsing and remitting inflammation. IBD incidence has risen mainly in Asia and in the Middle East over the past four to five decades and has now become a global disease. The increasing incidence rates of IBD also translates into increased years lived with disability [YLD] which has risen from 0.56 million to 1.02 million.[1] It has been estimated that IBD already poses a considerable threat to economic productivity as well as influencing early retirement.[2] As with most modern lifestyle-related diseases, the aetiology and pathogenesis of IBD is complex, and driven by a range of extrinsic and intrinsic factors. These include but are not limited to drug exposures, antibiotic treatments, smoking, lifestyle, stress, dietary patterns, genetics, immune responses, and the gut microbiome.[3–10] The complexity of IBD is also manifested by the heterogeneity of disease presentation and behaviour.[11–14] Heterogeneity in IBD is not only attributed to the complex phenotypes and the aetiological drivers, but also to the plethora of diverse molecules,[15–22] microbes, and cell types[23–25] and the interactions[26] among them.

The complexity and heterogeneity of IBD has resulted in a difficult endeavour for the scientific community to attribute causality as well as to raise treatment efficacies above the current therapeutic ceiling marked by endoscopic remission rates of around 30%,[27,28] although real-world observations recorded slightly higher efficacy rates.[29–31] The current standard of care, which includes histological examination [used currently only for diagnosis and not for therapeutic endpoints] following endoscopy, does not capture the fine print of disease biology such as cellular heterogeneities, cell type-specific expression, interactions among cell-types, and microbial influences. In response new approaches, like systems biology which integrates high-throughput datasets profiling different layers of biological organisation and the network of molecular interactions, have often been suggested as the 'holy grail' for understanding and treating IBD as well as other complex diseases.[32–36] However, despite the advances made in experimental [sequencing, high-throughput genotyping, etc] and computational [machine learning, data analysis tools, etc] methodologies, the -omic datasets in the IBD field are sparse and scattered, leading to incomplete coverage across -omic layers [genome, transcriptome, proteome, metabolome, etc]. The sparsity of -omic datasets is still a major bottleneck if the IBD research community were to harness the power of systems biology.[22,37–39]

In this review, we highlight the need for coordinated sampling to achieve horizontal [to increase sample size] and vertical [to cover different data types] -omic data coverage and the various challenges but also opportunities it presents. Motivated by improving disease knowledge and identifying therapeutic targets, we recommend how harmonised sampling strategies can be implemented at the level of decision making in devising science policy, grant funding, creating collaborative consortiums, and planning core infrastructure facilities such as biobanks and data generation centres. This calls not only for regional or national but above all international cooperation among IBD researchers and funding agencies spanning geopolitical divides.

Received: November 12, 2021. **Revised:** February 2, 2022. **Accepted:** February 10, 2022

2. The Need for Systems Biology and -Omic Datasets

2.1. Implications for experimental design and -omic data generation

The potential and utility of systems biology [Figure 1] in understanding and treating emerging diseases including IBD has been recognised over the past two to three decades. Various approaches employing a combination of top-down data-driven or bottom-up hypothesis-driven methodologies have been suggested and used to discover fundamental knowledge and uncover potential therapeutic pathways with clinical relevance.[37] For any systems biology endeavour, the availability of -omic datasets is an indispensable requirement and IBD research is no exception to it. The more complex the disease, the more the number of required data types [corresponding to -omic layers] to build models to better understand its pathogenesis, devise treatment options, and even suggest preventive strategies. This heavy need for -omic datasets has various implications at the level of experimental design and sampling, thus resulting in many challenges with multiple trade-offs and opportunities [Figure 2].

Broadly, before -omic datasets are generated, in the interest of best practice, considerations at the following levels [although not exhaustive; summarised further in Figure 3] need to be made subject to operational and financial constraints.

[a] Given the need for validation and discovery cohorts, and the requirement to achieve statistical power, how many samples will be required?
[b] What needs to be the split between the discovery and validation cohorts?
[c] Are the publicly available datasets annotated sufficiently with appropriate metadata so as to make them useable?
[d] How many types of -omic datasets need to be generated?
[e] Is sufficient level of interaction information available so as to interpret the -omic datasets using a network/systems approach, if the aim is to infer mechanisms at a systemic level?
[f] Is there an established standardised protocol to enable the harmonised collection of clinical phenotypes and subject information [diet, lifestyle, habits, etc]?
[g] Have the cohorts included a diversity of individuals representative of the population and the biological question[s] being investigated?
[h] Has the collection of subject information been customised to the expected cultural diversity in the population of interest?
[i] Has the gender aspect been addressed by an equitable inclusion of different gender categories?
[j] Have comorbidities been accounted for and recorded as part of the metadata?
[k] Have effect sizes been considered in the study design?

2.2. Omic dataset availability in IBD: identifying the gaps

Based on the information available in literature, the current gaps in IBD-omics include: [a] lack of more granular information [such as functional studies/effects on mutations]; [b] lack of IBD context-specific physical [mechanistic] interaction datasets; [c] small number of multi -omic studies [ie, studies profiling more than one -omic dataset]; and [d] an even smaller number of such studies using bimodal -omic data with independent validation. Each of the four above gaps are discussed more in detail below.

In order to gain an understanding of where IBD research stands in terms of [publicly] available -omic datasets, we compared it [Table 1] with a data-rich disease such as cancer and rheumatoid arthritis [RA]. Compared with IBD and RA, cancer is well represented not only in terms of the number of -omic-oriented public databases, but also with respect to the heterogeneity of datasets [ie, number of different -omic datatypes]. Furthermore, cancer datasets stand out in terms of their granularity or resolution such as, for example, the functional annotation of cancer-associated mutations provided by resources such as MutationAligner,[40] intOGen,[41] Cancer3D[42] or the collection of variants corresponding to specific genes such as BRCA1 and BRCA2 in breast cancer. Other examples of specialised public databases in cancer include Lnc2Cancer[43] which stores curated information on the expression of long non-coding RNAs associated with different types of cancers. These are in addition to the databases such as TCGA[44] which store different types of -omic datasets such as transcriptomics, genetics, methylomics, etc Despite the availability of -omic integration methods and datasets, examples of multi-omic data integration resulting in direct clinical translation are very few. Translation to the clinic depends on so many other aspects [regulatory, financial, ethical, time required] independent of the -omic datasets and the integration strategies, but there are several examples which reveal the potential of multi-omic data integration in disease sub-typing, biomarker discovery, and patient-specific treatments.

The WINTHER trial[45] stands out in particular and relies on the use of two -omic datasets [namely fresh biopsy-derived DNA sequencing and transcriptomics] along with other clinical data to make informed treatment decisions and recommendations for individual patients based on the -omic signatures. Several examples of individualised treatments or combinations of treatments resulting in disease remission based on the -omic signatures are provided. As a sequel to the WINTHER trial, the SPRING trial[46] was constituted following the same integration and predictive modelling strategy as the WINTHER trial with the added criteria that cancer patients were included earlier on in their disease course. Vitali *et al.* used network interaction data as a template to integrate genetic and gene expression datasets from triple negative breast cancer patients to prioritise drug targets which were subsequently verified using in vitro experiments.[47]

Other studies, such as the one by Buffa *et al.*,[48] identified new therapeutic targets for breast cancer by integrating miRNA and mRNA profiles. In addition, based on the integrative signatures, the authors identified pathways which could be involved in disease progression. Besides individual trials and studies, various consortium-level collaborative efforts [such as ACGT,[49] ContraCancrum,[50] p-medicine,[51] and CHIC[52]] have also been undertaken to create a translational framework in cancer by building on the power of -omic data integration on top of other methodologies such as multiscale modelling, cytology, toxigenomics, and pharmacology. However despite these advances, the question of whether the methodologies and approaches developed in the domain of cancer research can be applied to the more complex disease of IBD is still an open one.

In contrast, IBD is characterised by a far lower number of databases, albeit containing the usual suspects [genetics, gut

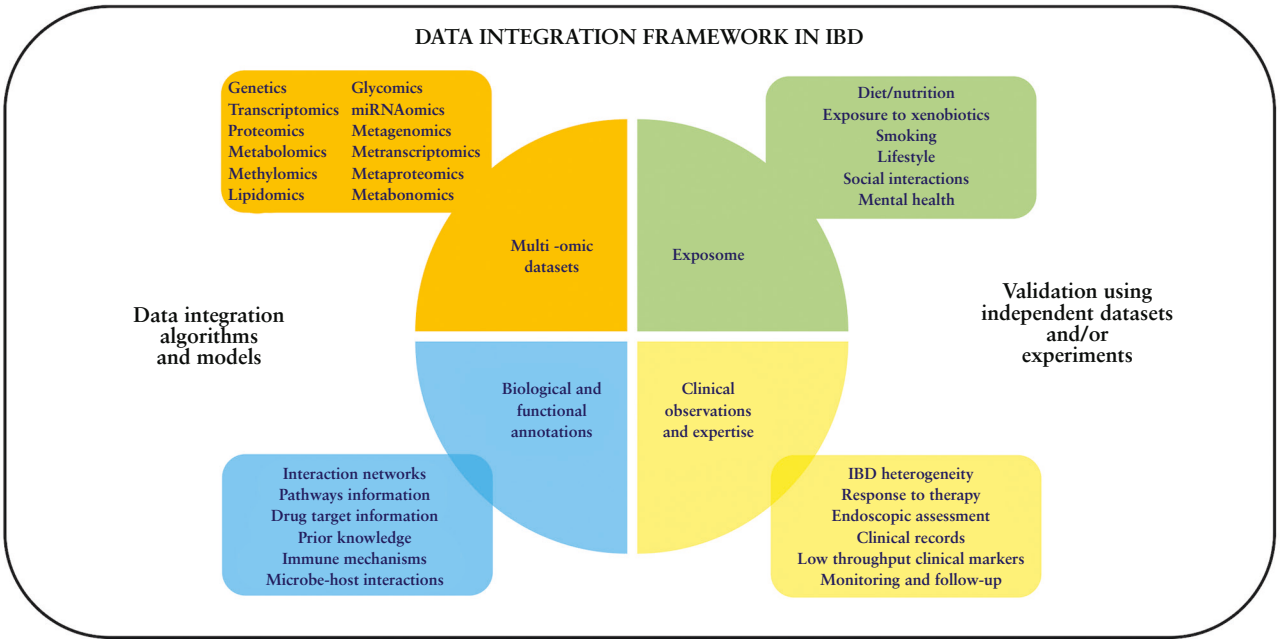

Figure 1. Data-integration framework in inflammatory bowel disease [IBD] characterised by the combination of heterogeneous information including multi-omic datasets, environment influences: exposome, clinical data, and network/functional annotations.

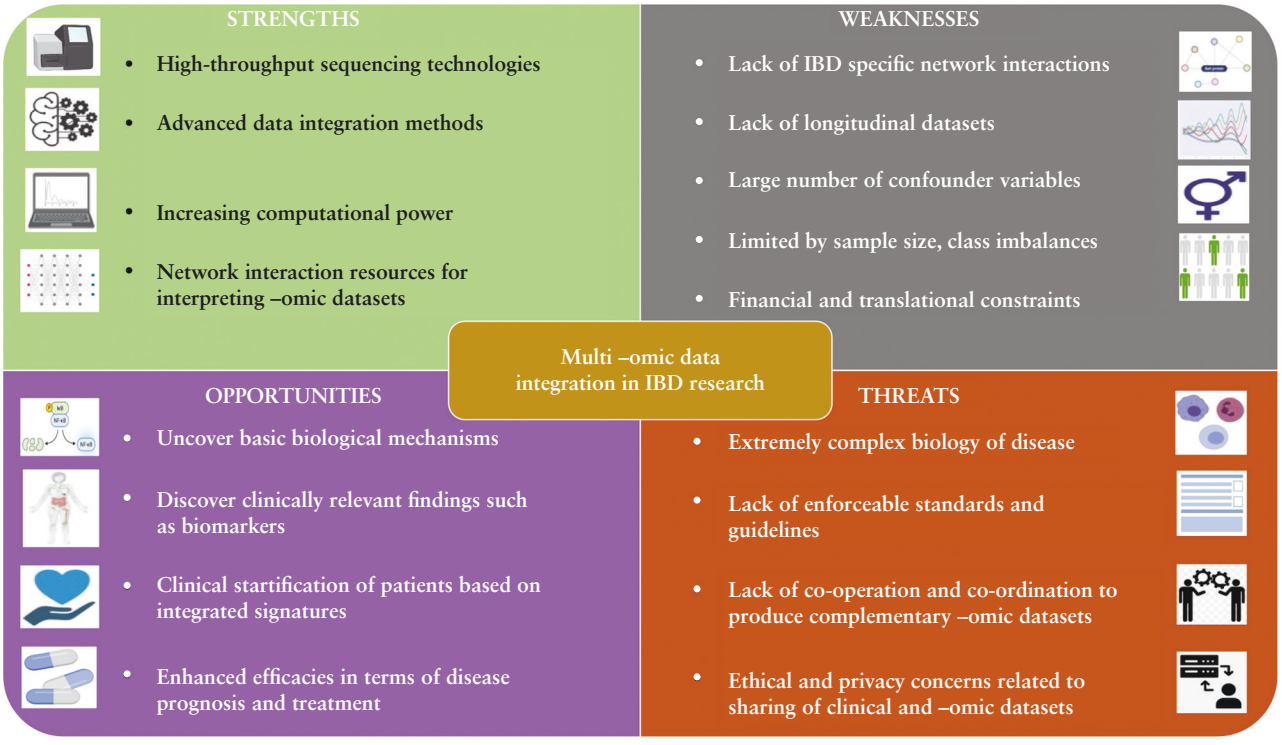

Figure 2. A SWOT [Strength, Weakness, Opportunity, and Threat] analysis of multi-omic data integration in inflammatory bowel disease [IBD] research.

microbiome profiles, proteomics, metabolomics, etc] in terms of -omic datatypes [Supplementary Table 1]. However, not all the publicly available IBD databases are populated with a wide range of data types. For example, whereas the IBDMDB database[4] affiliated to the HMP2/iHMP project is a repository with proteomics, host tissue transcriptomics, 16S faecal microbiome profiling, metagenomes, viromics, metabolomics, serology, and metatranscriptomics, others such as UK Biobank,[53] NIHR IBD bioresource,[54] and the International Inflammatory Bowel Disease Genetics Consortium [IIBDGC], are confined to particular data types such as genetics. Also noteworthy is the fact that granular information, such as functional characterisations of mutations, are not contained within any of the existing IBD resources. Nevertheless, IBD has a better public -omic data availability than RA as evidenced by the absence of any devoted RA -omic repositories [Table 1]. Furthermore, existing IBD patient-derived -omic repositories are also lacking in terms of read-outs from specific regulatory

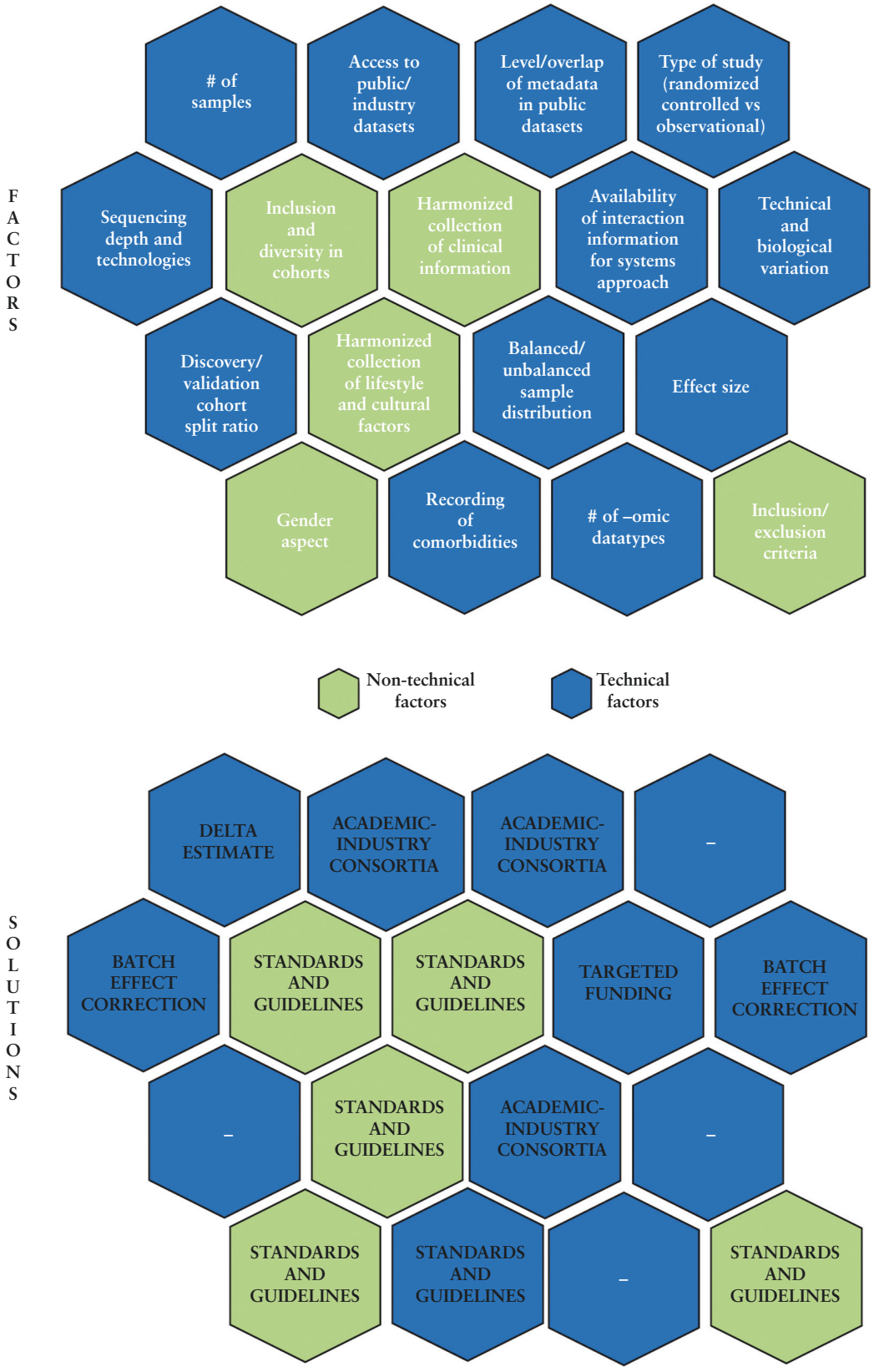

Figure 3. Some of the technical and non-technical factors [and possible solutions] which could potentially influence the outcome and statistical power of a multi-omic data integration project.

molecules such as long non-coding RNAs [lncRNAs] and circular RNAs [circRNAs]. None of the studies reported in Table 2 or the IBD-specific resources in Table 1 encompassed the profiles of lncRNA and circRNA expression. In other words, lncRNAs and circRNAs are neither profiled with other -omic datasets, which then results in disregarding the mechanisms

Table 1. Comparison of omic data repositories for three different diseases [cancer, inflammatory bowel disease, and rheumatoid arthritis]

Disease	Database	URL	Omic layers available	Downloadability?	Data format [raw/ processed]
IBD	HMP2/IBDMDB	https://ibdmdb.org/	Proteomics, host tissue transcriptomics, 16S gut microbiome profiling, metagenomes, viromics, metabolites, serology, metatranscriptomes	Yes	Both raw and processed
	Dutch IBD biobank	https://1000ibd.org/; https://ega-archive.org/studies/EGAS00001002702	Genetics, 16S gut microbiome profiling [faeces and intestinal biopsies], faecal metagenomics	Yes	Raw only
	International Inflammatory Bowel Disease Genetics Consortium [IIBDGC]	https://www.ibdgenetics.org/downloadshtml	Genetics	No	-
	NIHR IBD bioresource	https://www.ibdbioresource.nihr.ac.uk/index.php/resources/applying-for-access-to-the-ibd-bioresource-panel-2/	Genetics	No	-
	PRISM	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA400072/	WGS gut microbiome profiling, LC-MS metabolomics	Yes	Raw only
	UK Biobank	https://www.ukbiobank.ac.uk/	Genetics	No	-
Cancer	ArrayMap	http://www.arraymap.org	Copy number data	No	-
	BCNTBbp	https://www.breastcancertissuebank.org/bioinformatics	Genomics, methylomics, transcriptomics, proteomics and microRNA	No	-
	BRCA Share	http://www.umd.be/BRCA2/	Genomic variants on the BRCA1 and BRCA2 genes	Yes	Processed only
	BreCAN-DB	http://14.139.32.56/	Somatic DNA breakpoint profiles mapped using whole genome sequencing data	Yes	Raw only
	Cancer PPD	http://crdd.osdd.net/raghava/cancerppd/	Lists of proteins and peptides with anti-cancer activities	Yes	Processed only
	Cancer RNA-Seq Nexus	http://syslab4.nchu.edu.tw/	Expression of long non-coding RNAs and miRNAs	No	-
	Cancer3D	http://www.cancer3d.org/search	Map of cancer missense mutations on protein structures	Yes	Processed only
	cBioPortal	https://www.cbioportal.org/	Mutations, copy numbers, mRNA and protein expression	Yes	Both raw and processed
	COSMIC	https://cancer.sanger.ac.uk/cosmic/	Mutations and their annotations	Yes	Both raw and processed
	Database of Germline p53 Mutations	http://kolweb.lf2.cuni.cz//projects/germline_mut_p53.htm	Detailed annotations of p53 mutations	Yes	Processed only
	intOGen	https://www.intogen.org/search	Mutational cancer driver genes	Yes	Processed only
	Lnc2Cancer	http://www.bio-bigdata.com/lnc2cancer/home.jsp	Expression of cancer associated long non-coding RNAs	No	-
	MethHC	http://methhc.mbc.nctu.edu.tw/php/index.php	DNA methylation, gene expression, microRNA methylation, microRNA expression	Yes	Both raw and processed
	MOKCa	http://strubiol.icr.ac.uk/extra/mokca/	Map of cancer mutations and their phenotypic effects	No	-
	MutationAligner	http://www.mutationaligner.org/	Mutation 'hotspots' identified in protein domains	Yes	Processed only
	Network of Cancer Genes	http://ncg.kcl.ac.uk/index.php	Duplicated loci and expression of protein-coding cancer genes	Yes	Both raw and processed
	TCGA	https://portal.gdc.cancer.gov/	RNA-seq, methylation, genotyping, miRNA-seq, whole-exome high-throughput DNA sequencing	Yes	Both raw and processed
	TCGA Splice Seq	http://projects.insilico.us.com/TCGASpliceSeq/	Cross-tumour and tumour-normal alterations in mRNA splicing patterns in cancer	Yes	Processed only

Table 1. Continued

Rheumatoid arthritis	Biogps	http://biogps.org/dataset/tag/rheumatoid%20arthritis/	Gene expression	Yes	Both raw and processed
	Gene Expression Omnibus	https://www.ncbi.nlm.nih.gov/geo/	Gene expression	Yes	Both raw and processed
	Array Express	https://www.ebi.ac.uk/arrayexpress/	Gene expression	Yes	Raw only
	PMID: 23143596	https://pubmed.ncbi.nlm.nih.gov/23143596/	Genetics	Yes	Both raw and processed
	PMID: 22446963	https://pubmed.ncbi.nlm.nih.gov/22446963/	Genetics	Yes	Both raw and processed
	RADB	http://www.bioapp.org/RADB/	RA-related genetic polymorphisms extracted from published papers	Yes	Processed only

IBD, inflammatory bowel disease; LC-MS, Liquid Chromatography-Mass Spectrometry; RA, rheumatoid arthritis; WGS , whole genome sequencing.

and roles mediated by these novel regulatory molecules. This is conspicuous since lncRNAs and circRNAs are known to be involved in mediating disease pathogenesis and modulating the expression of various genes and proteins, as well as predicting response to therapeutic outcomes in IBD.[55–59] Although not quantitatively proven and not exclusively limited to this reason, the exclusion of lncRNAs and circRNAs, while identifying and prioritising therapeutic biomarkers using high-throughput profiling of other -omic datasets, could be a contributing factor to the middling therapeutic efficacy rates in IBD treatments.

Despite the availability of -omic data repositories in IBD research, there is still a relatively small number [n = 14] of studies integrating at least two different -omic datasets from IBD patients [Table 2]. Furthermore, none of these studies [with the exception of Franzosa *et al*.[60]] include independent validation, with four of the studies relying on internal validation [ie, assigning a specific proportion of the discovery/test cohort for validation] or experimental validation/mouse models. An additional drawback of many existing IBD multi-omic datasets is the lack of data from control/non-IBD samples. This could be an impeding factor in translational studies which measure the effects of therapeutic regimens in a particular population.

2.3. Omic datasets: just the tip of the iceberg

The past two decades have witnessed a relative increase in the amount of -omic datasets generated directly from IBD patients or mouse models or from in-vitro models [such as organoids and cell cultures] derived from IBD patients or mouse models. However, despite the context-specific nature of the -omic datasets, the underlying network information comprising different types of interactions [protein-protein, microRNA-mRNA, miRNA-lncRNA, etc] are generic. In other words the interaction information, even that derived entirely from experimental studies, is measured using experiments under conditions which may not be related to IBD. This poses a major bias while using network interaction information to interpret and/or integrate IBD-specific -omic datasets and to glean mechanisms. To draw an analogy, this is akin to having traffic and vehicular flow data of a city but using the wrong map of the city.

Whereas the non-availability of context-specific network interaction information is not confined to IBD research, it is a limitation which the researchers need to have in mind while drafting research proposals using systems approaches. It could also mean that when it comes to experimentally validating key regulators such as transcription factors or miRNAs, interactions mediated by such regulators need to be determined using targeted contextual approaches which profile the interactions from relevant IBD samples. Such targeted contextual approaches could include methodologies such as ChIP-seq[61]/ChIP-chip[62] to validate transcriptional regulatory interactions or HITS-CLIP[63]/PAR-CLIP[64] to validate miRNA-mRNA interactions.

2.4. Exposomics: capturing the environmental stimuli

The term ‘exposomics’ was first introduced by the cancer epidemiologist Christopher Wild in 2005 to describe the totality of environmental exposures to which an individual is subjected over the course of a lifetime.[65] Another way to put it is that it is an omic-scale characterisation of the non-genetic drivers of health and disease.[66] The importance of these environmental exposures on human pathology becomes clear when considering the rapid increase in non-communicable diseases since the latter half of the 20th century, exceeding the speed at which genes are thought to evolve significantly.[67] Other examples include the worldwide increase in IBD prevalence [especially in regions that were historically spared from the disease, such as Asia] and the observation that IBD incidence in immigrants’ children resembles that of the new country rather than that of their parent’s former country.[1,68] Furthermore, associations of environmental factors including diet, smoking status, antibiotic intake, and early life exposures, have been repeatedly reported to impact on IBD onset and/or disease course.[69,70]

Another mediator to bridge the gap between the [meta] genome and environment is the gut microbiome, as alterations in this ‘other’ genome are increasingly linked to various pathological conditions.[71] This is especially the case in IBD, where an increasing number of studies suggests a crucial role for dysbiosis in disease pathogenesis, and where modulation of the microbiota through dietary interventions or faecal microbiome transplantation might alter disease course through the production of metabolites and interactions with the immune system.[72–76] This new angle was only possible because of the advent of novel technologies such as 16S rRNA sequencing and shotgun sequencing in combination with metabolomics, which can provide useful information on the exact composition of the microbiota and their functionality.[16] Using high-resolution mass spectrometry, metabolomics can be-

Table 2. A non-exhaustive list of multi -omic IBD studies. .

Study	Disease context	Study context	Omic layers	Source of validation dataset[a]	Control/ non-IBD samples?
[Lloyd-Price *et al.*, 2019][4]	CD, UC	Identification of multi-omic signatures associated with IBD patients	Faecal proteomics Host mucosal transcriptomics 16S faecal microbiome profiling WGS faecal metagenomics Faecal viromics Faecal metabolomics Serology Faecal metatranscriptomics	No validation	Yes
[Borren *et al.*, 2020][111]	CD, UC	Prediction of biomarkers associated with disease relapse	Faecal proteomics Faecal metabolomics WGS faecal metagenomics	Internal dataset[a]	No
[Suskind *et al.*, 2020][120]	CD	Investigating the effect of different diets on disease symptoms and inflammatory burden	Faecal metabolomics WGS faecal metagenomics Faecal proteomics	No validation	No
[Le *et al.*, 2020][121]	CD, UC	Prediction of metabolite abundances from microbial abundances	Faecal metabolomics WGS faecal metagenomics	Internal dataset[a]	No
[Dai *et al.*, 2019][122]	CD	Identification and characterisation of important drivers of CD pathogenesis	Host genetics TWAS Host mucosal transcriptomics Methylomics	No validation	Yes
[Liu *et al.*, 2021][123]	CD, UC	Role of microbiota in oxalate metabolism in IBD patients	WGS faecal metagenomics Faecal metatranscriptomics	Experimental validation	No
[Revilla *et al.*, 2021][124]	CD	Interdependent host genes and microbial genera in CD	Host mucosal transcriptomics 16S gut microbiome profiling	No validation	No
[Jin *et al.*, 2019][125]	CD, UC	Dysregulated genes and pathways in CD/ UC pathogenesis	Host mucosal transcriptomics Host mucosal proteomics	No validation	No
[Sudhakar *et al.*, 2020][22]	CD	Drivers of clinical heterogeneity in CD	PBMC gene expression CD4 gene expression Host genetics	No validation	No
[Nusbaum *et al.*, 2018][126]	UC	Influence of FMT on gut microbial and metabolic activity in paediatric UC patients	16S faecal microbiome profiling WGS faecal metagenomics Faecal viromics Faecal metabolomics	No validation	No
[Metwaly *et al.*, 2020][127]	CD	Integrative analysis of metabolic and microbial profiles in CD	16S faecal microbiome profiling WGS faecal metagenomics faecal metabolomics	Validation in mouse model	No
[Douglas *et al.*, 2018][113]	CD	Prediction of treatment response	16S gut microbiome profiling WGS gut metagenomics	No validation	Yes
1000IBD dataset	CD, UC, IBDU	Discover molecular sub-types of IBD	Host genetics 16S faecal microbiome profiling 16S gut microbiome profiling WGS faecal metagenomics Single cell RNA sequencing from biopsies	NA[a]	No
[Franzosa *et al.*, 2019][60]	CD, UC	Investigation of microbiome and metabolic activity in IBD	Faecal metabolomics WGS faecal metagenomics	Independent validation cohort	Yes

IBD, inflammatory bowel disease; UC, ulcerative colitis; CD, Crohn's disease; FMT, faecal microbiota transplanation; NA, not available; PMBC, peripheral blood mononuclear cells; TWAS, transcriptome-wide association study; WGS, whole genome sequencing;
[a]Internal independent dataset: defined as a dataset which is derived by ring-fencing a particular proportion of the test cohort for validation. Only published studies related to IBD and which integrate at least two different -omic datatypes were included. Publications based on original research were retrieved from PubMed using the co-occurrence of the search term 'multi -omics' with 'IBD', 'Inflammatory Bowel Disease', 'Ulcerative colitis', 'Crohns disease', or 'Crohn's disease'

come an increasingly interesting tool, since the endogenous compounds from various matrices [such as stool, blood, urine, which can give an idea of the biological responses] can be combined with exogeneous-derived small molecules [such as pesticides, herbicides, pharmaceuticals, and flame retardants].[77,78]

Of course, metabolomics [and other single -omic layers] alone will most likely not be sensitive and specific enough to adequately capture all human exposures, but can give a glimpse of the idea of designing a way to integrate various components with a single method. This method could then be implemented like DNA-sequencing and could complement genomic with environmental information.

We will illustrate the latter with the example of titanium dioxide [TiO_2]. TiO_2 is a white pigment and brightening agent that is widely used in various day-to-day products such as toothpaste, but also as a food additive [E171] in confectionery items, white sauces, and icing.[79] It can be categorised as a nanoparticle because of its size in the nanoparticle range and presents unique physical and chemical properties.[79] In-vitro studies of TiO_2 were able to link this nanoparticle to alterations in intestinal immunity, with uptake of the particle by macrophages and epithelial cells and the potential to induce inflammation through activation of the inflammasome and IL-1b secretion.[80,81] Studies focusing on the microbiota found that TiO_2 might be able to induce dysbiosis with long exposure times to human colon microbiome in vitro [5 days] at environmentally relevant concentrations, leading to significant changes in bacterial metabolites [including in short chain fatty acid production].[79,82] This might be of particular importance since bioavailability studies suggest that 99% of ingested TiO_2 accumulates in the gut lumen with a persistent contact of the particles with the commensals.[79] A randomised controlled trial that investigated the effect of a diet reduced in microparticles [TiO_2 and particulate silicates] in Crohn's disease failed to show any effect on remission.[83]

As illustrated, this classical 'one exposure—one disease' hypothesis-driven approach can be very informative for the specific exposure studied and provide mechanistic insights, but gives an incomplete picture as only one particle is being assessed. Importantly, the exposome implies a cumulative exposure over time and this includes [cumulative] dose, but also the possibility of critical windows when a certain exposure or dose might be more impactful.[84,85] Next, the exposome concept implies a multitude of various exposures that might interact with each other and with the individual genome, thus making a more holistic approach, like the untargeted and unbiased genome-wide association studies [GWAS], necessary to fully grasp the effect of the exposome.[84]

First steps towards studying the exposome in chronic diseases are already there, like the use of EWAS [environment-wide association study] as a tool in diabetes research[86] and the Groningen IBD Environmental Questionnaire to map the exposome.[87] These initiatives have been followed by the establishment of large collaborations such as the European Exposome Network launched by the European Commission in 2020[https://www.humanexposome.eu] [see Supplementary Table 2] and American HERCULES [the Human Exposome Research Center: Understanding Lifetime ExposureS] [https://emoryhercules.com], an NIH-funded project that started in May 2013 and has now grown to include 20 centres across the USA, which prove feasibility of a more structural and collaborative approach.

Unfortunately, proving causality and translating this information into meaningful interventions for prevention and treatment of IBD has been shown difficult. The main reason is that the impact on and interaction with other -omic layers remain poorly understood. That being said, there are large-scale IBD-including initiatives like ImmUniverse which try to understand the role of cross-talk of tissue and immune cells in a longitudinal fashion, making use of multiple -omic layers [https://immuniverse.eu]. Unfortunately, even with metabolomics and the gut microbiome included, capturing data on exposures like nutrition seems to be overlooked. Interestingly, a comprehensive multi-omics [including the exposome] project on the pathogenesis and outcomes in primary sclerosing cholangitis [PSC: a disease associated with IBD] will be started shortly and aims to combine large clinical databases and biorepositories, as well as expertise in PSC and related conditions, exposomics, metabolomics, methylomics, transcriptomics, metagenomics, genomics, and data analytics to better understand the role and interplay of the genome, exposome, and microbiome [https://mayoclinic.pure.elsevier.com/en/projects/dissecting-the-pathogenesis-and-outcomes-of-psc-using-multi-omics]. A similar but IBD-oriented endeavour with international collaborations and skills to map exposures and biological responses together with genetic information will likely be critical to move the field forwards [Figure 4A]. Recent initiatives, such as the one using dental matrices to evaluate early life exposures and their associations with IBD development, are worth highlighting [https://www.hospitaldaluz.pt/en/media/news/inflammatory-bowel-disease-joana-torres-ioibd-fellowship] [Supplementary Table 2].

3. Policy Recommendations for Coordinated and Effective -Omic Data Generation and Analysis

Multi-omic studies holds great promise to fill knowledge gaps in IBD pathogenesis, to advance therapeutic development, and to attain the yet unfulfilled goals of modifying the natural course of the disease. Unfortunately, the aforementioned challenges and risks hinder the widespread adoption of multi-omics in IBD studies thus far. However IBD researchers, research consortia, and other professional organisations can foster collaborations and facilitate the successful application of multi-omics in IBD. We list many [Table 3] and highlight a few relevant recommendations which could guide various stakeholders in IBD research to maximise coherence and efficiency.

3.1. Devise guidelines to design multi-omic studies tailored to the field of IBD

Before multi-omic studies can be widely implemented in IBD, clear guidelines are needed pertaining to which objectives can be pursued [disease prediction, outcomes prediction, drug target discovery, etc] and for each of these, the most suitable research design [cross-sectional, longitudinal, etc] and sampling schemes [time of sampling, number of omic layers, frequency of sampling, etc] [Table 4]. IBD consortia may take the lead in bringing experts and stakeholders together to develop these guidelines that are direly needed to execute suitable operational decisions. Taking such guideline-based decisions will have a huge impact on not only research discoveries, but also on financial and logistic aspects to ensure the efficient durability of multi-omic projects at low costs and labour efforts, a the same time providing the needed answers to fulfill the unmet needs. In addition to guidelines, we also recommend open-sourcing good practice recommendations generated within the IBD community as well as making -omic datasets

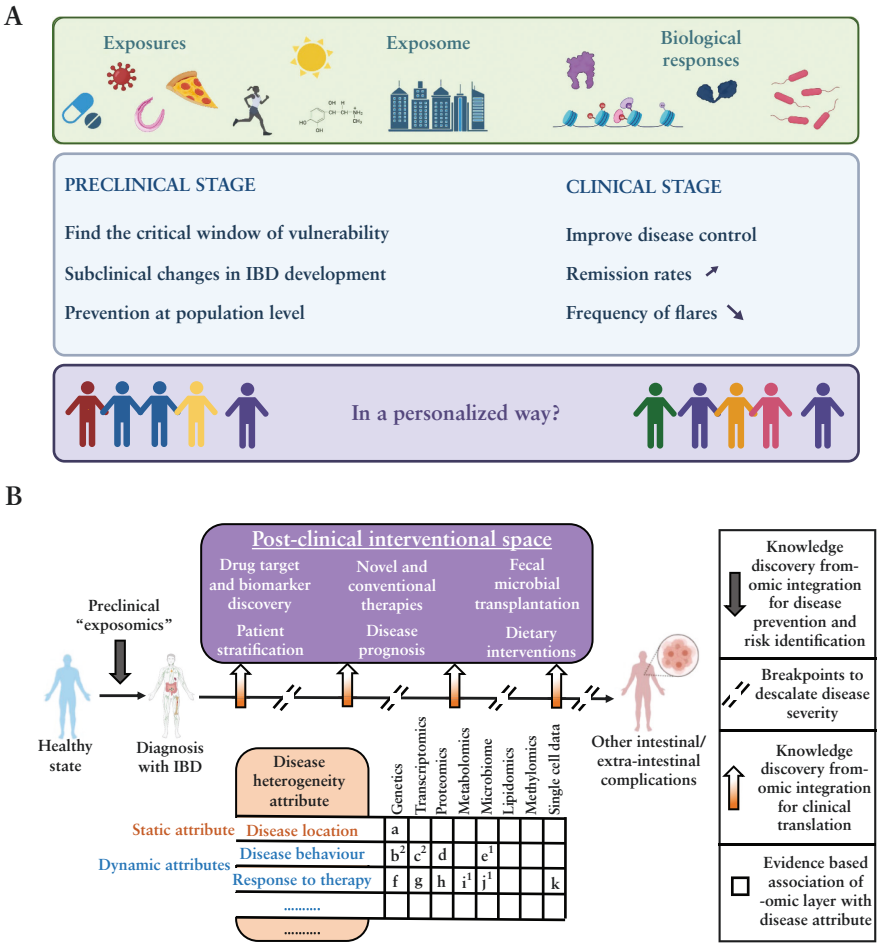

Figure 4. [A] Preclinical and clinical opportunities in inflammatory bowel disease [IBD] exposome research. [B] Omic data generation tailored to the wide range of IBD disease attributes which can be either static [like disease location] or dynamic [disease behaviour, response to therapy, perianal manifestations for example]. a, [88]; b, [89]; c, [22,90,91]; d, [92,93]; e, [90]; f, [89,94–97]; g, [98–107]; h, [108–110]; I, [111,112]; j, [111–115]; k, [116,117]. Numbers in superscript indicate the studies which have integrated at least two omic layers to study the corresponding disease attribute.

Table 3. Recommendations for the harmonious generation of omic datasets towards the goals of advancing mechanistic knowledge and clinical translation in IBD research.

Policy domain	Stakeholders	Recommendations
Funding and resource allocation	• Funding agencies • IBD consortia • Principal investigators	• Prolonged funding for generating omic datasets from IBD patients in a longitudinal manner • Targeted financial support for promoting international collaboration with research stakeholders in developing countries • 'Follow-up' grants to generate missing particular omic datasets to complement pre-existing omic datasets from the same samples • Ring-fencing funding for training skilled omic scientists
Collaboration and harmonisation	• Funding agencies • Principal investigators • Non-profit medical organisations in the realm of IBD • IBD consortia • Principal investigators	• Compilation of commonly agreed standards for IBD omic datasets • Coordinated strategies between international and national funders to avoid huge redundancies in terms of omic datasets • Creation of an easily shareable IBD omic dataset catalog between multiple funders and agencies in the IBD domain • Encouraging exchange programmes for interdisciplinary research
Legal and organisational	• Universities • Principal investigators • Industry	• Creating robust and flexible inter-institutional confidentiality agreements to promote exchange of knowledge and information while fulfilling GDPR requirements • Intra- and inter-project alignment for achieving multiple objectives using the same omic datasets

GDPR, general data protection regulation; IBD, inflammatory bowel disease.

Table 4. Best practices for IBD omic data generation and interpretation. Many decision points are influenced by clinical feasibility and translational application potential.

Phase of research	Decision points	Recommendations	Examples
Cohort design and sample collection		• Fine tune based on research question and record the clinical data	• Include or exclude inflamed/non-inflamed samples, treatment-naïve patients, postoperative setting etc
	Inclusion / exclusion criterion	• Think about the control population/samples.	• Should it be healthy controls, patients with other IMIDs, or non-diseased sites from the same IBD patients?
		• Consider enriching clinical data with environmental data.	• Smoking status, dietary factors, BMI, exercise etc
		• Using low-throughput biomarkers, determine the relationship between flares and patterns in high-throughout -omic datasets measured from longitudinal cohorts.	• Markers like fecal calprotectin and CRP
	Sampling dynamics	• Generate the same set of -omic datasets at multiple time-points from the corresponding samples (subject to technical limitations, available biomaterial and any possible time-lags between the -omic datasets) to enable sample-to-sample comparability.	• For example, host transcriptomics + microbial metatranscriptomics from biopsies at different time-points over treatment course or disease course.
		• Incorporate additional measurements after starting a new treatment	• For example, pharmacodynamics
	Sampling site	• Sample from the primary disease site but consider paired samples from other sites.	• Primary disease site (ileum, colon etc), other sites (such as PBMCs)
	Sampling for |microbiota	• Microbiota profiles from primary disease site generally preferred over faecal samples. Nonetheless, the latter might prove to be valuable given the ease of sampling.	• Primary disease site - luminal contents (from biopsies directly)
		• Additional host and microbial community based omics datasets can bridge the environmental and host aspects.	• Metabolomics, metabonomics, metatranscriptomics, metaproteomics, viromics, mycobiomics etc
	Sampling mass	• Pre-determine the number of samples (or biopsies) and/or amount of biomass required to generate the different -omic datasets from the same sample.	• Depends on the specification of the kit/ in-house protocols used for the extraction of the different molecular fractions
	Sample treatment	• Follow standardized protocols for sampling, processing and storage specific for each -omic dataset where applicable.	• Kit-based or in-house protocols
	Sample storage and documentation	• Use of registered biobanks recommended • Samples to be systematically indexed and barcoded during storage.	• UK Biobank • -
Data generation	Cellular resolution (bulk or single cell or purified cell-types/ fractions)	• Measurements from single cell technologies highly recommended.	• -
	Sequencing strategy (short or long read sequencing)	• Sequencing type. Budget, biological question(s) etc dictate the choice	• Long-read or short-read sequencing?
	-Omic data type (genomics, transcriptomics, proteomics etc)	• Dependent on biological question and various other factors like budget, research question, clinical and translational feasibility.	• For example, bulk transcriptomics provide a relatively cheap option for high-throughput genome-wide profiling of biological state. However, proteomics is closer to phenotype than transcriptomics but lags behind on coverage.
		• Include novel molecules for -omic measurements.	• Circular RNAs, long non-coding RNA
	Microbiota specific data resolution (16S or WGS)	• 16S for preliminary studies • WGS preferred for making in depth assessments due to its advantage of being able to make strain-level inferences	• - • -

IMID, Immune Mediated Inflammatory Disease; PBMC, Peripheral Blood Mononuclear Cells; BMI , Body Mass Index; CRP, C-reactive protein; WGS, Whole Genome Sequencing.

publicly available. Furthermore, IBD-specific metadata standards need to be compiled so that datasets can be made compatible for comparison with others and hence can be potentially harnessed for validation by the community.

3.2. Prevent redundant multi-omic IBD studies

Even though a certain degree of redundancy is required to ensure validation of findings, a large number of duplicated studies are a very much debated problem in all fields of med-

ical research, as it squanders public funds and delays the achievement of breakthroughs. In a resources-consuming domain as multi-omic studies, this redundancy should be minimised if not entirely avoided. Although this issue started to get attention from health care regulators [as in EU Clinical Trials Regulation 536/2014], IBD organizations and consortia should tackle this issue in the early stages of the multi-omic era. This can be achieved by establishing collaborative efforts involving IBD investigators for timely and transparent communication of study ideas and preliminary results to harmonise efforts and minimise redundancies.

In the same context, the conduct of clinical trials has been accompanied in recent years by the collection of numerous biospecimens with detailed clinical characteristics from included patients and the generation of corresponding omic datasets. Unfortunately, these resources are often out of reach for principal investigators/researchers since the accrued data are rarely published. The pharmaceutical industry should be encouraged to grant IBD investigators access to these piled-up data or at least share preliminary findings to avoid duplication, save resources and time, integrate knowledge, and guide future research.

3.3. Disseminate training and education for harnessing the power of analytical and computational approaches to analyse high-throughput multi-omic datasets

Once the guidelines outlining design of multi-omic studies are firmly established, IBD researchers need to be able to analyesthe resulting data using methods tailored to IBD. Currently, there is a lack of concise computational methodologies in this research domain in IBD. This can be overcome by enhancing familiarity of clinical investigators with analysing multi-omic data through fostering collaboration with skilled bioinformaticians and organising hands-on training workshops for interested IBD researchers.

3.4. Ensure sufficient funding and resources

To ensure flexibility in investigating the different facets of IBD and dissect its complex pathogenesis, adequate amounts of multi-omic data need to be generated in the first phases, and this requires a stable funding approach and not to be solely reliant on sporadic funding opportunities. Therefore, inter-institutional collaboration among multi-omic investigators is necessary to collectively convey their message to policy makers in national scientific funding agencies to secure sufficient funding for multi-omic projects and to illustrate the impact of such investments on the prospects of national health quality and expenditure.

3.5. Approach IBD in a global framework

As the incidence and prevalence of IBD are increasing worldwide, it should be tackled globally with the same sense of solidarity as communicable diseases are being and have been dealt with before.[118] Furthermore, the genome end exposome components are likely to differ between different ethnicities and geographical areas.[70,119] IBD consortia in developed countries should assist IBD clinicians and scientists in countries with limited resources in establishing an affordable framework to collect and generate multi-omic data that are critical to give answers for their IBD patient cohorts. In addition, valuable knowledge and best practices as to how IBD can be prevented and managed [despite being genetically susceptible] can be gathered from different cultural backgrounds. Such knowledge can be transferred across cultures, if found to be inter-culturally appropriate.

4. Conclusion

Investigating the causes of any complex disease, including IBD, requires not only a multidisciplinary effort but also an extensive multi-omic effort to propel a systems view of the disease. Concurrently, multi-omic research deals with sampling the different levels of biological complexity and organisation. Since multi-omic datasets are at the heart of both multi-omic research and multidisciplinarity, it is essential to holistically integrate various technical and non-technical aspects while compiling patient cohorts from whom multi-omic datasets will be generated. Acknowledging and apprising such factors have a huge influence not only on the study design but also on the overall impact of the study. In this review, we attempt to provide an overview of disease complexity, the systems biology-based research framework, and multi-omic datasets tailored to the heterogeneous disease attributes [Figure 4B] as well as the technical and non-technical factors which could influence sampling and study design. Although nowhere near exhaustive or comprehensive, outlining the above-discussed aspects will help aid the IBD research community in generating informative multi-omic datasets to investigate research questions relevant to IBD biology and/or clinical/translational importance.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

Funding

PS was supported by the European Research Council Advanced Grant [ERC-2015-AdG, 694679, CrUCCial]. BV is funded by the Clinical Research Fund [KOOR], University Hospitals, Leuven, Belgium.

Conflict of Interest

SV has received research support from AbbVie, Johnson & Johnson, Pfizer, and Takeda; lecture fees from AbbVie, Centocor, Ferring, Genentech/Roche, Hospira, Johnson & Johnson, Merck Sharp & Dohme, Pfizer, Takeda, and Tillotts; and consulting fees from AbbVie, Abivax, Celgene, Celltrion, Centocor, Ferring, Galapagos, Genentech/Roche, Gilead, GlaxoSmithKline, Hospira, Johnson & Johnson, Merck Sharp & Dohme, Mundipharma, Pfizer, ProDigest, Prometheus, Second Genome, Takeda, and Tillotts. BV reports research support from Pfizer; speaker's fees from Abbvie, Biogen, Bristol Myers Squibb, Chiesi, Falk, Ferring, Galapagos, Janssen, MondayNightIBD, MSD, Pfizer, R-Biopharm, Takeda, and Truvion; consultancy fees from Alimentiv, Applied Strategic, Atheneum, Bristol Myers Squibb, Galapagos, Guidepont, Ipsos, Janssen, Progenity, Sandoz, Sosei Heptares, and Takeda.

Author Contributions

PS, JW, and DA performed the survey of -omic databases and resources and contributed to writing the manuscript. SV and BV provided critical scientific feedback and discussions on the

manuscript. SV provided clinical suggestions and supervision. All authors read and approved the final version of the manuscript.

Data Availability

No analytical or methodological datasets were generated from the study. The auxillary datasets summariing the databases and resources are provided via Supplementary Tables 1 and 2 [submitted as part of the review process].

References

1. GBD 2017 Inflammatory Bowel Disease Collaborators. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2020;5:17–30.
2. Fróes R de SB, Carvalho ATP, V Carneiro AJ de, *et al*. The socio-economic impact of work disability due to inflammatory bowel disease in Brazil. *Eur J Health Econ* 2018;19:463–70.
3. Di Narzo AF, Peters LA, Argmann C, *et al*. Blood and intestine eQTLs from an anti-TNF-resistant Crohn's disease cohort inform IBD Genetic Association Loci. *Clin Transl Gastroenterol* 2016;7:e177.
4. Lloyd-Price J, Arze C, Ananthakrishnan AN, *et al*. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 2019;569:655–62.
5. Halfvarson J, Brislawn CJ, Lamendella R, *et al*. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* 2017;2:17004.
6. Vieira-Silva S, Sabino J, Valles-Colomer M, *et al*. Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat Microbiol* 2019;4:1826–31.
7. Pascal V, Pozuelo M, Borruel N, *et al*. A microbial signature for Crohn's disease. *Gut* 2017;66:813–22.
8. Gevers D, Kugathasan S, Denson LA, *et al*. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15:382–92.
9. Jostins L, Ripke S, Weersma RK, *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119–24.
10. Lee JC, Biasci D, Roberts R, *et al*. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat Genet* 2017;49:262–8.
11. Ananthakrishnan AN. Epidemiology and risk factors for IBD. *Nat Rev Gastroenterol Hepatol* 2015;12:205–17.
12. Van Kruiningen HJ, Joossens M, Vermeire S, *et al*. Environmental factors in familial Crohn's disease in Belgium. *Inflamm Bowel Dis* 2005;11:360–5.
13. Cleynen I, Vermeire S. The genetic architecture of inflammatory bowel disease: past, present and future. *Curr Opin Gastroenterol* 2015;31:456–63.
14. Verstockt B, Bressler B, Martinez-Lozano H, *et al*. Time to revisit disease classification in IBD: is the current classification of IBD good enough for optimal clinical management? *Gastroenterology* 2021. doi:10.1053/j.gastro.2021.12.246.
15. Schaefer JS, Attumi T, Opekun AR, *et al*. MicroRNA signatures differentiate Crohn's disease from ulcerative colitis. *BMC Immunol* 2015;16:5.
16. Kim H-Y, Kwon HY, Ha Thi HT, *et al*. MicroRNA-132 and microRNA-223 control positive feedback circuit by regulating FOXO3a in inflammatory bowel disease. *J Gastroenterol Hepatol* 2016;31:1727–35.
17. Wang H, Zhang S, Yu Q, *et al*. Circulating microRNA223 is a new biomarker for inflammatory bowel disease. *Medicine* 2016;95:e2703.
18. Nguyen HTT, Dalmasso G, Müller S, *et al*. Crohn's disease-associated adherent invasive Escherichia coli modulate levels of microRNAs in intestinal epithelial cells to reduce autophagy. *Gastroenterology* 2014;146:508–19.
19. Treveil A, Sudhakar P, Matthews ZJ, *et al*. Regulatory network analysis of Paneth cell and goblet cell enriched gut organoids using transcriptomics approaches. *Mol Omics* 2020;16:39–58.
20. Longhi MS, Kokkotou E. Lnc-ing RNA expression with disease pathogenesis: MALAT1 and ANRIL in ulcerative colitis. *Dig Dis Sci* 2020;65:3061–3.
21. Padua D, Mahurkar-Joshi S, Law IKM, *et al*. A long noncoding RNA signature for ulcerative colitis identifies IFNG-AS1 as an enhancer of inflammation. *Am J Physiol Gastrointest Liver Physiol* 2016;311:G446–57.
22. Sudhakar P, Verstockt B, Cremer J, *et al*. Understanding the molecular drivers of disease heterogeneity in Crohn's disease using multi-omic data integration and network analysis. *Inflamm Bowel Dis* 2021;27(6):870–86.
23. Smillie CS, Biton M, Ordovas-Montanes J, *et al*. Intra- and intercellular rewiring of the human colon during ulcerative colitis. *Cell* 2019;178:714–30.e22.
24. Drokhlyansky E, Smillie CS, Van Wittenberghe N, *et al*. The enteric nervous system of the human and mouse colon at a single-cell resolution. *Cell* 2020;182:1606–1622.e23.
25. Neumann C, Blume J, Roy U, *et al*. c-Maf-dependent Treg cell control of intestinal TH17 cells and IgA establishes host-microbiota homeostasis. *Nat Immunol* 2019;20:471–81.
26. Sudhakar P, Machiels K, Verstockt B, *et al*. Computational biology and machine learning approaches to understand mechanistic microbiome-host interactions. *Front Microbiol* 2021;12:618856.
27. Reinglas J, Gonczi L, Kurt Z, *et al*. Positioning of old and new biologicals and small molecules in the treatment of inflammatory bowel diseases. *World J Gastroenterol* 2018;24:3567–82.
28. Alsoud D, Verstockt B, Fiocchi C, *et al*. Breaking the therapeutic ceiling in drug development in ulcerative colitis. *Lancet Gastroenterol Hepatol* 2021;6:589–95.
29. Mühl L, Becker E, Müller TM, *et al*. Clinical experiences and predictors of success of treatment with vedolizumab in IBD patients: a cohort study. *BMC Gastroenterol* 2021;21:33.
30. Al-Bawardy B, Shivashankar R, Proctor DD. Novel and emerging therapies for inflammatory bowel disease. *Front Pharmacol* 2021;12:651415.
31. Sands BE, Sandborn WJ, Panaccione R, *et al*. Ustekinumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2019;381:1201–14.
32. Polytarchou C, Koukos G, Iliopoulos D. Systems biology in inflammatory bowel diseases: ready for prime time. *Curr Opin Gastroenterol* 2014;30:339–46.
33. Moco S, Candela M, Chuang E, *et al*. Systems biology approaches for inflammatory bowel disease: emphasis on gut microbial metabolism. *Inflamm Bowel Dis* 2014;20:2104–14.
34. Suzuki S, Takai-Igarashi T, Fukuoka Y, *et al*. Systems analysis of inflammatory bowel disease based on comprehensive gene information. *BMC Med Genet* 2012;13:25.
35. Huang H, Vangay P, McKinlay CE, *et al*. Multi-omics analysis of inflammatory bowel disease. *Immunol Lett* 2014;162:62–8.
36. Dovrolis N, Filidou E, Kolios G. Systems biology in inflammatory bowel diseases: on the way to precision medicine. *Ann Gastroenterol* 2019;32:233–46.
37. Seyed Tabib NS, Madgwick M, Sudhakar P, *et al*. Big data in IBD: big progress for clinical practice. *Gut* 2020;69:1520–32.
38. He FQ, Wang W, Zheng P, *et al*. Essential O_2-responsive genes of Pseudomonas aeruginosa and their network revealed by integrating dynamic data from inverted conditions. *Integr Biol* 2014;6:215–23.
39. Sudhakar P, Andrighetti T, Verstockt S, *et al*. Integrated analysis of microbe-host interactions in Crohn's disease reveals potential mechanistic effects of microbial proteins on host gene expression. *iScience* 2022. doi:10.1016/j.isci.2022.103963.
40. Gauthier NP, Reznik E, Gao J, *et al*. MutationAligner: a resource of recurrent mutation hotspots in protein domains in cancer. *Nucleic Acids Res* 2016;44:D986–91.

41. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, *et al*. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods* 2013;10:1081–2.
42. Porta-Pardo E, Hrabe T, Godzik A. Cancer3D: understanding cancer mutations through protein structures. *Nucleic Acids Res* 2015;43:D968–73.
43. Gao Y, Shang S, Guo S, *et al*. Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. *Nucleic Acids Res* 2021;49:D1251–8.
44. Wang Z, Jensen MA, Zenklusen JC. A practical guide to the cancer genome atlas [TCGA]. *Methods Mol Biol* 2016;1418:111–41.
45. Rodon J, Soria JC, Berger R, *et al*. Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med* 2019;25:751–8.
46. Solomon BM, Callejo A, Bar J, *et al*. SPRING: a Worldwide Innovative Network [WIN] Consortium phase I study of triple therapy [avelumab, axitinib, and palbociclib] in advanced non-small cell lung cancer [NSCLC] with genomic and transcriptomic correlates. *JCO* 2020;38:9581.
47. Vitali F, Cohen LD, Demartini A, *et al*. A network-based data integration approach to support drug repurposing and multi-target therapies in triple negative breast cancer. *PLoS One* 2016;11:e0162407.
48. Buffa FM, Camps C, Winchester L, *et al*. microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. *Cancer Res* 2011;71:5635–45.
49. Martin L, Anguita A, Graf N, *et al*. ACGT: advancing clinico-genomic trials on cancer – four years of experience. *Stud Health Technol Inform* 2011;169:734–8.
50. Marias K, Sakkalis V, Roniotis A, *et al*. Clinically oriented translational cancer multilevel modeling: the contracancrum project. In: Dössel O, Schlegel WC, editors. *World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany: Vol. 25/4 Image Processing, Biosignal Processing, Modelling and Simulation, Biomechanics*. Vol 25/4. IFMBE Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010; 2124–7.
51. Rossi S, Christ-Neumann M, Rüping S, *et al*. p-Medicine: from data sharing and integration via VPH models to personalized medicine. *Ecancermedicalscience* 2011;5:218.
52. Yang Y-J, Lee S-H, Kim BS, *et al*. Combination therapy of rosuvastatin and ezetimibe in patients with high cardiovascular risk. *Clin Ther* 2017;39:107–17.
53. Sudlow C, Gallacher J, Allen N, *et al*. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
54. Parkes M; IBD BioResource Investigators. IBD BioResource: an open-access platform of 25 000 patients to accelerate research in Crohn's and Colitis. *Gut* 2019;68:1537–40.
55. Ye Y-L, Yin J, Hu T, *et al*. Increased circulating circular RNA_103516 is a novel biomarker for inflammatory bowel disease in adult patients. *World J Gastroenterol* 2019;25:6273–88.
56. Nie J, Zhao Q. Lnc-ITSN1-2, derived from RNA sequencing, correlates with increased disease risk, activity and promotes CD4+ T cell activation, proliferation and Th1/Th17 cell differentiation by serving as a ceRNA for IL-23R via sponging miR-125a in inflammatory bowel disease. *Front Immunol* 2020;11:852.
57. Lu J-W, Rouzigu A, Teng L-H, *et al*. The construction and comprehensive analysis of inflammation-related ceRNA networks and tissue-infiltrating immune cells in ulcerative progression. *Biomed Res Int* 2021;2021:6633442.
58. Sun F, Liang W, Tang K, *et al*. Profiling the lncRNA-miRNA-mRNA ceRNA network to reveal potential crosstalk between inflammatory bowel disease and colorectal cancer. *PeerJ* 2019;7:e7451.
59. Yin J, Tong F, Ye Y, *et al*. Hsa_circRNA_103124 upregulation in Crohn's disease promotes cell proliferation and inhibits autophagy by regulating the Hsa-miR-650/AKT2 signaling pathway. *Front Genet* 2021;12:753161.
60. Franzosa EA, Sirota-Madi A, Avila-Pacheco J, *et al*. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 2019;4:293–305.
61. Furey TS. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet* 2012;13:840–52.
62. Zheng M, Barrera LO, Ren B, *et al*. ChIP-chip: data, model, and analysis. *Biometrics* 2007;63:787–96.
63. Neame E. HITS-CLIP hits the microRNA target. *Nat Rev Genet* 2009;10:510–1.
64. Danan C, Manickavel S, Hafner M. PAR-CLIP: a method for transcriptome-wide identification of RNA binding protein interaction sites. *Methods Mol Biol* 2016;1358:153–73.
65. Wild CP. Complementing the genome with an 'exposome': the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14:1847–50.
66. Niedzwiecki MM, Walker DI, Vermeulen R, *et al*. The exposome: molecules to populations. *Ann Rev Pharmacol Toxicol* 2019;59:107–27.
67. Bauer-Mehren A, Bundschus M, Rautschka M, *et al*. Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases. *PLoS One* 2011;6:e20284.
68. Ko Y, Butcher R, Leong RW. Epidemiological studies of migration and environmental risk factors in the inflammatory bowel diseases. *World J Gastroenterol* 2014;20:1238–47.
69. Agrawal M, Sabino J, Frias-Gomes C, *et al*. Early life exposures and the risk of inflammatory bowel disease: systematic review and meta-analyses. *EClinicalMedicine* 2021;36:100884.
70. Piovani D, Danese S, Peyrin-Biroulet L, *et al*. Environmental risk factors for inflammatory bowel diseases: an umbrella review of meta-analyses. *Gastroenterology* 2019;157:647–59.e4.
71. Claesson MJ, Clooney AG, O'Toole PW. A clinician's guide to microbiome analysis. *Nat Rev Gastroenterol Hepatol* 2017;14:585–95.
72. Lees CW. Environmental stimuli and gut inflammation via dysbiosis in mouse and man. *Nat Rev Gastroenterol Hepatol* 2020;17:715–6.
73. David LA, Maurice CF, Carmody RN, *et al*. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2014;505:559–63.
74. Schaubeck M, Clavel T, Calasan J, *et al*. Dysbiotic gut microbiota causes transmissible Crohn's disease-like ileitis independent of failure in antimicrobial defence. *Gut* 2016;65:225–37.
75. Shanahan F, Sinderen D van, O'Toole PW, *et al*. Feeding the microbiota: transducer of nutrient signals for the host. *Gut* 2017;66:1709–17.
76. Albenberg LG, Lewis JD, Wu GD. Food and the gut microbiota in inflammatory bowel diseases: a critical connection. *Curr Opin Gastroenterol* 2012;28:314–20.
77. Vermeulen R, Schymanski EL, Barabási A-L, *et al*. The exposome and health: where chemistry meets biology. *Science* 2020;367:392–6.
78. Walker DI, Valvi D, Rothman N, *et al*. The metabolome: a key measure for exposome research in epidemiology. *Curr Epidemiol Rep* 2019;6:93–103.
79. Lamas B, Martins Breyner N, Houdeau E. Impacts of foodborne inorganic nanoparticles on the gut microbiota-immune axis: potential consequences for host health. *Part Fibre Toxicol* 2020;17:19.
80. Yazdi AS, Guarda G, Riteau N, *et al*. Nanoparticles activate the NLR pyrin domain containing 3 [Nlrp3] inflammasome and cause pulmonary inflammation through release of IL-1α and IL-1β. *Proc Natl Acad Sci U S A* 2010;107:19449–54.
81. Becker HM, Bertschinger MM, Rogler G. Microparticles and their impact on intestinal immunity. *Dig Dis* 2012;30:47–54.

82. Taylor AA, Marcus IM, Guysi RL, *et al*. Metal oxide nanoparticles induce minimal phenotypic changes in a model colon gut microbiota. *Environ Eng Sci* 2015;32:602–12.
83. Lomer MCE, Grainger SL, Ede R, *et al*. Lack of efficacy of a reduced microparticle diet in a multi-centred trial of patients with active Crohn's disease. *Eur J Gastroenterol Hepatol* 2005;17:377–84.
84. Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;41:24–32.
85. Miller GW, Jones DP. The nature of nurture: refining the definition of the exposome. *Toxicol Sci* 2014;137:1–2.
86. Hall MA, Dudek SM, Goodloe R, *et al*. Environment-wide association study [EWAS] for type 2 diabetes in the Marshfield Personalized Medicine Research Project Biobank. *Pac Symp Biocomput* 2014:200–11.
87. Sloot KWJ van der, Weersma RK, Dijkstra G, *et al*. Development and validation of a web-based questionnaire to identify environmental risk factors for inflammatory bowel disease: the Groningen IBD Environmental Questionnaire [GIEQ]. *J Gastroenterol* 2019;54:238–48.
88. Cleynen I, Boucher G, Jostins L, *et al*. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016;387:156–67.
89. Visschedijk MC, Spekhorst LM, Cheng S-C, *et al*. Genomic and expression analyses identify a disease-modifying variant for fibrostenotic Crohn's disease. *J Crohns Colitis* 2018;12:582–8.
90. Kugathasan S, Denson LA, Walters TD, *et al*. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* 2017;389:1710–8.
91. Marigorta UM, Denson LA, Hyams JS, *et al*. Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn's disease. *Nat Genet* 2017;49:1517–21.
92. Townsend P, Zhang Q, Shapiro J, *et al*. Serum proteome profiles in stricturing crohn's disease: a pilot study. *Inflamm. Bowel Dis* 2015;21:1935–41.
93. Stidham RW, Wu J, Shi J, *et al*. Serum glycoproteome profiles for distinguishing intestinal fibrosis from inflammation in Crohn's disease. *PLoS One* 2017;12:e0170506.
94. Sazonovs A, Kennedy NA, Bewshea C, *et al*. OP013 HLA-DQA1 contributes to the development of antibodies to anti-TNF therapy in Crohn's disease. *J Crohns Colitis* 2018;12:S009–10.
95. Billiet T, Vande Casteele N, Van Stappen T, *et al*. Immunogenicity to infliximab is associated with HLA-DRB1. *Gut* 2015;64:1344–5.
96. Barber GE, Yajnik V, Khalili H, *et al*. Genetic markers predict primary non-response and durable response to anti-TNF biologic therapies in Crohn's disease. *Am J Gastroenterol* 2016;111:1816–22.
97. Dubinsky MC, Mei L, Friedman M, *et al*. Genome wide association [GWA] predictors of anti-TNFalpha therapeutic responsiveness in pediatric inflammatory bowel disease. *Inflamm Bowel Dis* 2010;16:1357–66.
98. Haberman Y, Karns R, Dexheimer PJ, *et al*. Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. *Nat Commun* 2019;10:38.
99. Verstockt B, Verstockt S, Creyns B, *et al*. Mucosal IL13RA2 expression predicts nonresponse to anti-TNF therapy in Crohn's disease. *Aliment Pharmacol Ther* 2019;49:572–81.
100. Prins MM, Verstockt B, Ferrante M, *et al*. Monocyte TREM-1 levels associate with anti-TNF responsiveness in IBD through autophagy and Fcγ-receptor signaling pathways. *Front Immunol* 2021;12:627535.
101. Verstockt B, Verstockt S, Dehairs J, *et al*. Low TREM1 expression in whole blood predicts anti-TNF response in inflammatory bowel disease. *EBioMedicine* 2019;40:733–42.
102. Tew GW, Hackney JA, Gibbons D, *et al*. Association between response to etrolizumab and expression of integrin αe and granzyme A in colon biopsies of patients with ulcerative colitis. *Gastroenterology* 2016;150:477–87.e9.
103. Gaujoux R, Starosvetsky E, Maimon N, *et al*. Cell-centred meta-analysis reveals baseline predictors of anti-TNFα non-response in biopsy and blood of patients with IBD. *Gut* 2019;68:604–14.
104. Schmitt H, Billmeier U, Dieterich W, *et al*. Expansion of IL-23 receptor bearing TNFR2+ T cells is associated with molecular resistance to anti-TNF therapy in Crohn's disease. *Gut* 2019;68:814–28.
105. West NR, Hegazy AN, Owens BMJ, *et al*. Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* 2017;23:579–89.
106. Telesco SE, Brodmerkel C, Zhang H, *et al*. Gene expression signature for prediction of golimumab response in a phase 2a open-label trial of patients with ulcerative colitis. *Gastroenterology* 2018;155:1008–11.e8.
107. Arijs I, Li K, Toedter G, *et al*. Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. *Gut* 2009;58:1612–9.
108. Gazouli M, Anagnostopoulos AK, Papadopoulou A, *et al*. Serum protein profile of Crohn's disease treated with infliximab. *J. Crohns Colitis* 2013;7:e461–70.
109. Meuwis M-A, Fillet M, Lutteri L, *et al*. Proteomics for prediction and characterization of response to infliximab in Crohn's disease: a pilot study. *Clin Biochem* 2008;41:960–7.
110. Heier CR, Fiorillo AA, Chaisson E, *et al*. Identification of pathway-specific serum biomarkers of response to glucocorticoid and infliximab treatment in children with inflammatory bowel disease. *Clin Transl Gastroenterol* 2016;7:e192.
111. Borren NZ, Plichta D, Joshi AD, *et al*. Multi-'-Omics' profiling in patients with quiescent inflammatory bowel disease identifies biomarkers predicting relapse. *Inflamm Bowel Dis* 2020;26:1524–32.
112. Ding NS, McDonald JAK, Perdones-Montero A, *et al*. Metabonomics and the gut microbiome associated with primary response to anti-TNF therapy in Crohn's disease. *J Crohns Colitis* 2020;14:1090–102.
113. Douglas GM, Hansen R, Jones CMA, *et al*. Multi-omics differentially classify disease state and treatment outcome in pediatric Crohn's disease. *Microbiome* 2018;6:13.
114. Doherty MK, Ding T, Koumpouras C, *et al*. Fecal microbiota signatures are associated with response to ustekinumab therapy among Crohn's disease patients. *MBio* 2018;9:e02120–17.
115. Ananthakrishnan AN, Luo C, Yajnik V, *et al*. Gut microbiome function predicts response to anti-integrin biologic therapy in inflammatory bowel diseases. *Cell Host Microbe* 2017;21:603–10.e3.
116. Nayar S, Morrison JK, Giri M, *et al*. A myeloid-stromal niche and gp130 rescue in NOD2-driven Crohn's disease. *Nature* 2021;593:275–81.
117. Martin JC, Chang C, Boschetti G, *et al*. Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy. *Cell* 2019;178:1493–508.e20.
118. Hurst JR, Dickhaus J, Maulik PK, *et al*. Global Alliance for Chronic Disease researchers' statement on multimorbidity. *Lancet Glob Health* 2018;6:e1270–1.
119. Cho JH, Brant SR. Recent insights into the genetics of inflammatory bowel disease. *Gastroenterology* 2011;140:1704–12.
120. Suskind DL, Lee D, Kim Y-M, *et al*. The specific carbohydrate diet and diet modification as induction therapy for pediatric crohn's disease: a randomized diet controlled trial. *Nutrients* 2020;12:3749.
121. Le V, Quinn TP, Tran T, *et al*. Deep in the bowel: highly interpretable neural encoder-decoder networks predict gut metabolites from gut microbiome. *BMC Genomics* 2020;21:256.

122. Dai Y, Pei G, Zhao Z, *et al*. A convergent study of genetic variants associated with Crohn's disease: evidence from GWAS, gene expression, methylation, eQTL and TWAS. *Front Genet* 2019;10:318.
123. Liu M, Devlin JC, Hu J, *et al*. Microbial genetic and transcriptional contributions to oxalate degradation by the gut microbiota in health and disease. *Elife* 2021;10:e63642.
124. Revilla L, Mayorgas A, Corraliza AM, *et al*. Multi-omic modelling of inflammatory bowel disease with regularized canonical correlation analysis. *PLoS One* 2021;16:e0246367.
125. Jin L, Li L, Hu C, *et al*. Integrative analysis of transcriptomic and proteomic profiling in inflammatory bowel disease colon biopsies. *Inflamm. Bowel Dis* 2019;**25**:1906–8.
126. Nusbaum DJ, Sun F, Ren J, *et al*. Gut microbial and metabolomic profiles after faecal microbiota transplantation in pediatric ulcerative colitis patients. *FEMS Microbiol Ecol* 2018;94fiy133.
127. Metwaly A, Dunkel A, Waldschmitt N, *et al*. Integrated microbiota and metabolite profiles link Crohn's disease to sulfur metabolism. *Nat Commun* 2020;11:4322.