Contents lists available at ScienceDirect

Immunology Letters

Review

Multi-omics analysis of inflammatory bowel disease

Hu Huang [a], Pajau Vangay [a], Christopher E. McKinlay [a,b], Dan Knights [b,c,*]

[a] Biomedical Informatics and Computational Biology, University of Minnesota, Minneapolis, MN 55455, United States
[b] Biotechnology Institute, University of Minnesota, Saint Paul, MN 55108, United States
[c] Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, United States

ARTICLE INFO

ABSTRACT

Crohn's disease and ulcerative colitis, known together as inflammatory bowel disease (IBD), are severe autoimmune disorders now causing gut inflammation and ulceration, among other symptoms, in up to 1 in 250 people worldwide. Incidence and prevalence of IBD have been increasing dramatically over the past several decades, although the causes for this increase are still unknown. IBD has both a complex genotype and a complex phenotype, and although it has received substantial attention from the medical research community over recent years, much of the etiology remains unexplained. Genome-wide association studies have identified a rich genetic signature of disease risk in patients with IBD, consisting of at least 163 genetic loci. Many of these loci contain genes directly involved in microbial handling, indicating that the genetic architecture of the disease has been driven by host–microbe interactions. In addition, systematic shifts in gut microbiome structure (enterotype) and function have been observed in patients with IBD. Furthermore, both the host genotype and enterotype are associated with aspects of the disease phenotype, including location of the disease. This provides strong evidence of interactions between host genotype and enterotype; however, there is a lack of published multi-omics data from IBD patients, and a lack of bioinformatics tools for modeling such systems. In this article we discuss, from a computational biologist's point of view, the potential benefits of and the challenges involved in designing and analyzing such multi-omics studies of IBD.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Crohn's disease (CD) and ulcerative colitis (UC) are the two most common forms of inflammatory bowel diseases (IBD). CD involves intestinal inflammation mostly affecting the terminal ileum and colon, although it can also affect any other part of the gastrointestinal tract; UC is limited to the mucosa and submucosa of the colon. IBD is thus a complex chronic inflammatory disorder, likely caused by an series of interactions among multiple pathogenic factors, including genetics, environments and mucosal immune disorders [1–3]. The precise etiology of IBD in many cases is still unclear, and much ongoing research is based on defining causal mechanisms for known genetic susceptibility [1] and on understanding microbial risk factors [4].

The complexity of the IBD disease phenotype is driven by gene-environment interactions, possibly mediated by dysbiosis in the gut microbiota. Traditional bottom-up studies focus on individual candidate factors, such as candidate genes or microorganisms, and largely employ hypothesis-directed experimentation rather than exploratory research. Despite the large number of candidate factor studies, they do not readily provide holistic insight into the complexity of IBD etiology. IBD etiology is known to consist of intricate interacting networks formed by genetic and environmental factors, for instance diet, smoking status, age, medication history, and family history. Recent advances in high-throughput experimental "omics" data generation techniques along with necessary bioinformatics tools have resulted in new top-down research strategies to overcome these shortcomings in traditional approaches. Here "omics" refers to large-scale high-dimensional biological data generation aiming to comprehensively analyze one type of biochemical molecular species or interactions of molecules in a cellular system [5,6]. Experimental omics approaches, such as genomics, transcriptomics, proteomics and metabolomics, share several features when compared to traditional procedures [6,7]: (i) they are high-throughput, data-driven, holistic and top-down methodologies, in the sense that massive data are collected first with no prior hypothesis, and then the meaningful results are searched and explained within the obtained dataset; (ii) they consider the targets, such as cellular metabolism, as an "integrated system" by incorporating

the relationships between different measurements; (iii) the analysis of these high-throughput data are statistically complicated and computationally expensive.

So, multi-omics: what's in it for IBD? Combining data collection across the host–microbiome interface allows the potential for modeling complex interactions between host immune response, host environmental factors including diet, and the gut microbiome. Large-scale genome-wide association studies (GWAS) [8–11] and extensive research in mouse models [12–14] have provided us with a substantial base of knowledge covering IBD and signaling pathways in the host. Similarly, we now have a substantial representation of gut microbiome biodiversity in publicly available functionally annotated bacterial reference genomes [8,15–18]. Yet there are limited systematic methods for connecting these two bodies of knowledge. Collecting and integrating multi-omics data has begun to receive IBD researchers' attention, but because of technical and financial reasons multi-omics research on IBD is still in its infancy [3]. IBD researchers wishing to collect multi-omics data to measure simultaneous behavior of the host immune system and the gut microbiome, for example, via dual RNA-Seq [19], will be forced to use ad-hoc bioinformatics methods or to develop novel bioinformatics tools.

Discovery of novel host–microbiome disease pathways in IBD will benefit from multi-omics study designs including paired host genomics or transcriptomics, bacterial metagenomics or metatranscriptomics, metaproteomics, immunomics, and other data types. However, this type of study design requires careful consideration of a number of analytical challenges, and substantial bioinformatics development is needed to enable rapid progress in the field.

2. Conventional IBD 'omes

Here we provide a brief overview of several important data types in IBD 'omics research. For each IBD "ome" we discuss important applications as well as potential pitfalls and challenges expected in modeling each data type. These data types have a history of established application in IBD research. Based on systematic mining of published abstracts from IBD research literature on PubMed, we found that these now common 'omics data types have been increasingly utilized to study IBD over the last ten years (Fig. 1). We found a noted lack, however, of published multi-omics IBD research. With a small number of exceptions (for example, Erickson et al. [20]), the following IBD 'omes have been studied independently of one another.

2.1. Genomics (host)

Host genetics has received the most attention of any of the "omics" data types we survey here. A single gene, nucleotide-binding oligomerization domain-containing protein 2 (NOD2), holds the largest component of disease risk for Crohn's disease (CD); NOD2 was the first known IBD-related gene [21–23]. Now thanks to very large-scale genotyping efforts the IBD genotype is known to contain at least 163 genetic loci [8], with some loci specific to CD and some to ulcerative colitis (UC). Only approximately 30% of the IBD loci are coding variants [24], and most of the loci have unknown function. Furthermore, most of the SNPs identified through GWAS are merely representative of a signal somewhere within a broader genetic locus. Fine-mapping studies of the exomes and regulatory landscape surrounding these putative IBD SNPs ("exomics") in specific cell types and tissues is expected to yield important insights into the functional roles of the loci in IBD pathogenesis [24]. Although knowledge of the genetics of IBD has grown exponentially for the last decade, the contribution of specific genes and the impact
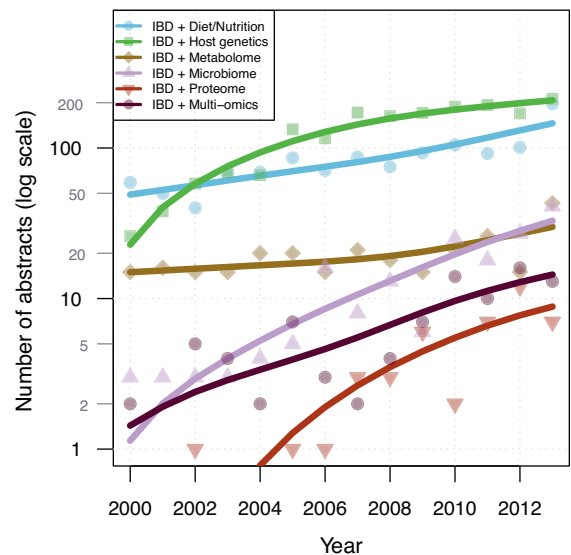

Fig. 1. Historical trends of 'omics terms in IBD-related research articles. The number of abstracts per year (log scale) for research articles mentioning both inflammatory bowel diseases and several kinds of high-throughput 'omics data since the year 2000. Trend data was mined using pattern matching within the results from a search of the PubMed database [85] for inflammatory bowel disease(s), Crohn's disease, or ulcerative colitis. We also mined abstracts for combinations of two or more of these "omes" (data types), but there were too few to show any pair individually on this plot. Instead we included a single series including IBD-related publications with any two or more different omics terms. This indicates that although high-throughput 'omics data are increasingly being collected for studying IBD, very little multi-omics IBD research has been published to date. This is likely due to a combination of the increased challenges associated with both multi-omics data collection and analysis.

of their associated variation is largely still unclear, indicating that that genomics alone is not sufficient to uncover the causes of IBD.

2.2. Transcriptomics (host)

Transcriptomics refers to the study of the complete set of RNA molecules, or "transcripts", present in a population of cells. While genomics examines static DNA information, transcriptomics measures the dynamic expression of RNA molecules and their variation under different circumstances at the genome scale, hence reflecting the genes that are actively expressed at any given time [7], with the exception of mRNA degradation phenomena [25,26]. Techniques that allow assessment of RNA profiles on a genome-wide scale are microarrays [27], serial analysis of gene expression (SAGE) [28], and RNA-seq [29]. More recently, a "dual RNA-seq" technique has been proposed, which focuses on both host and microbiome transcriptome data [19].

Measuring the host transcriptomes in, for example, tissue from a mucosal biopsy may shed light on the phenotypic effects of the host genotype of immune system response under various exposures to bacterial products or other bioactive compounds. Transcriptomics via microarrays and RNA-Seq has been an essential component of IBD research for mapping immune regulatory and signaling networks and for understanding the etiology of IBD, often in controlled ex vivo experiments [29]. Investigation using high-density cDNA microarrays have identified several important regulatory molecules and genes associated with a disrupted immune response that may be involved in the pathogenesis of IBD [30]. Similar research on pediatric-onset IBD patients [31] confirmed the association of previously published IBD genes, but also expanded the number of immune-related genes differentially regulated in IBD. Moreover, transcriptomics has been applied to the study of the effect of anti-inflammatory diet interventions in Crohn's disease patients [32], and biomarker discovery using non-coding RNAs,

such as MicroRNAs (miRNAs), in diagnosis and medical therapy of UC and CD patients [33–36]. One of the major challenges in transcriptomics is the management of expression profile variability in diverse host cell types. Isolation of individual host cells from mucosal biopsies via new developments in single-cell sequencing can alleviate this problem, although obtaining matched biopsies from healthy subjects remains an additional challenge.

2.3. Proteomics and metabolomics

In addition to analysis on the DNA level and the RNA transcript level, studies at the protein level can bring direct insight into cellular behavior and metabolic pathway utilization. Proteomics is the high-throughput measurement of constituent protein structures in populations of host or bacterial cells to provide insight into how cell metabolism and cell networks are involved in responses to changes in environment. Proteomics plays an important role in biomarker discovery due to its ability to measure directly the enzymatic activity of cells. However, there is no exact linear association between genes and their respective protein products [37], and most proteins are further altered via a variety of post-translational modifications (PTMs) [38]. PTMs dramatically increase the diversity of protein structures and functions, which cannot be directly studied from genomes and transcriptomes alone. Hence, identifying PTMs and quantifying the diverse array of proteins in a cellular system is important for biomarker discovery [39]. The most commonly used analytical proteomic techniques are mass spectrometry (MS)-based approaches [40,41] and nuclear magnetic resonance (NMR) spectroscopy, although others exist, including quantitative/comparative proteomics (2D-PAGE, ICAT, SILAC and iTRAQ) [42], array-based technologies (antibody arrays, protein lysate arrays, peptide arrays, aptamer arrays and bead-based arrays) [43] and the multi-epitope-ligand cartography (MELC) technology [42].

The first published proteomics study in IBD investigated protein changes in intestinal epithelial cells of IBD patients induced by multiple cytokines in vitro [44]. With the fast development of proteomics technologies and bioinformatics tools, proteomic analysis of IBD has gained popularity with different study targets, such as a proteomic profile of intestinal mucosa from UC patients [45], label-free comparative proteomics biomarker discovery for UC dysplasia detection in UC progressors (cancer) from rectal samples [46] and proteomic serum profiling for biomarker discovery in UC and CD patients [47]. Predictive utility of proteomic features has been investigated by employing a machine learning technique, support vector machines (SVMs), to differentiate CD and UC, with moderate success for UC classification [48]. More systematic reviews on applications of proteomic study in IBD can be found in [39,49].

Colonic inflammations are often characterized by dysregulation of gut microbiota and their metabolic activities. It is also possible to measure the overall proteome in microbial communities via *metaproteomics*, although a number of challenges remain. Metaproteomics refers to the measurement of protein production in diverse populations of organisms, for example the gut microbiome. Major challenges in analyzing and collecting proteomics and metaproteomics data include the identification of PTMs, such as phosphorylation and ubiquitination. In the case of metaproteomics it is also challenging to identify an appropriate reference database for identifying observed peptide sequences [20].

In contrast to proteins, the primary functional unit of cell physiology, metabolites are small molecules resulting from biochemical activity in metabolism, and hence "provide a functional readout of cellular state" [50]. Metabolomics refers to profiling of metabolites, that is, the quantitation of small molecules produce by cellular metabolism. Metabolomics via nuclear magnetic resonance (NMR) spectroscopy, for example, was used to identify non-invasively

characteristic metabolomics profiles in fecal samples for different types of IBD [51]. In this study, the authors observed more marked differences in the fecal metabolite profiles in the CD patients than the UC patients when compared to the healthy control group, suggesting that metabolic disturbances are more severe in CD patients than in UC patients [51]. In addition to metabolomics of fecal extracts [52–55], other IBD researchers have studied metabolomics of serum and plasma [56–59], colon tissues [60–62] and urinary metabolomics [60,63,64], both in mouse models and human patients. Metabolomics has obvious applications as a complement to other 'omics techniques [65]. For example, metabolomics can be combined with metagenomics or metatranscriptomics of the gut microbiota to strengthen discovery of relevant pathways in microbial metabolism.

2.4. Metagenomics and metatranscriptomics (gut microbes)

Metagenomics, or *whole-genome shotgun metagenomics*, refers to the unbiased sequencing of whole-community DNA in an environment [66]. Shotgun metagenomics sequences are typically compared to annotated genes from bacterial reference genomes to obtain an estimate of the functional profile of a bacterial community. The functional profile is usually summarized using a gene functional categorization scheme such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) [67] functional pathways or modules. One of the main findings of the Human Microbiome Project (HMP) [66] was that host-associated microbiomes tend to have highly conserved function, even while the taxonomic profiles are quite variable. This indicates that functional profiles may be expected to produce more stable disease associations than taxonomic profiles. IBD associations with functional biomarkers in the gut microbiome may also lead more directly to an understanding of the mechanisms of the diseases, because the functional annotations describe the underlying microbial metabolic and signaling pathways involved. A major challenge with analyzing shotgun metagenomics data is that we generally have to rely on reference databases containing annotated microbial genomes, and a large portion of observed sequences fail to match anything in the current databases. It is also possible to amplify only a specific taxonomic marker gene, typically the 16S ribosomal RNA gene. Such "marker gene" sequencing requires much less sequencing effort, and is therefore less costly, but only provides information at the taxonomic level. A recent bioinformatics pipeline, PICRUSt, can be used to predict community-wide functional profiles from marker gene profiles with reasonable accuracy in the human gut [68]. Metatranscriptomics in the gut microbiome is similar to transcriptomics in host cells, but here it measures total microbiome-wide gene expression levels.

3. Future IBD 'omes

3.1. Enviromics

The host "envirome", including dietary history and relevant clinical metadata such as patient age, race, gender, antibiotic usage, and history of immunosuppressants, plays an important role in the etiology of IBD, and has a profound effect on the microbiome phenotype exhibited by the patient [69,70]. Thus it is essential to include such metadata in the design and analysis of any multi-omics IBD study involving the gut microbiome. We list the envirome as a future IBD "ome" because we believe that measurement and analysis of available and relevant environmental factors should be given just as much care and importance as the other "omes" being measured, and yet the processing and modeling of environmental and clinical variables is often treated in an ad hoc fashion. Table 1

Table 1

Essential envirome factors in IBD microbiome research. These clinical, dietary, demographic, and environmental factors have been found to influence strongly the taxonomic and functional composition of the gut microbiome. Any IBD study including measurements of the gut microbiome profile should take into account at least these factors, either during study design or during analysis.

Host/environmental factor	Important aspects
Antibiotics history	Broad vs. narrow spectrum, time since last dose
Immunosuppressants/ immunomodulators history	Type; time since last dose
Smoking history	
Diet	Long- and short-term diet
Age	
Gender	
Inflammation status	Should be determined locally for mucosal microbiota samples
Ileal involvement	
Race	Subjects of mixed race may need to be considered distinct from those with homogeneous racial background
Geography	Country and city of residence

lists major host and environmental factors known to affect the gut microbiome in patients with IBD and in other populations [66,70].

In a survey of several hundred healthy adults living in the USA, the HMP revealed that pathway and taxon abundances were significantly associated with subject ethnicity, suggesting that variation in the gut microbiome may be associated with host genetic factors [66]. These results have major implications for distinguishing between taxon differences attributable to IBD phenotype or differences that are a result of other host characteristics. Another critical factor to consider is antibiotic exposure; antibiotics can cause major alterations in microbiome community structure for weeks or longer after the course of antibiotics ends. During and immediately after a course of antibiotics, the diversity and abundance of the microbial community is drastically reduced [71,72], with recovery to pre-administration biodiversity in most subjects taking up to four weeks [72]. Despite the return of many taxa and a replenished bacterial biomass, long-term antibiotic-related shifts in the composition of the gut microbiome have been shown to persist for as long as two years [71]. Although decreased *Firmicutes* abundance may be associated with IBD status [73], knowledge of recent antibiotic exposures may reveal otherwise; for example, a patient who had recently taken a high dose of clarithromycin may also show a decreased proportion of *Firmicutes* [74]. Morgan et al. provide an IBD analysis combining the microbiome and the envirome that incorporates several major host factors, including age, antibiotic use, anti-inflammatory drug use, and location of the sample [70]. The researchers found that after adjusting for all known factors, IBD status was associated with a shift in the microbiome but with large covariation between the disease state and various host factors; they concluded that the sample source and subject age were the largest independent factors associated with changes in the microbiome [70]. Diet is also known to be a strong determinant of microbiome structure and function on both long- and short-term time scales [75,76], and IBD studies involving the gut microbiome should control for diet as much as possible either through study design or during data collection and analysis.

3.2. "Knowmics"

The study of an autoimmune disorder like IBD requires understanding the function and topology of numerous host signaling and metabolic pathways, as well as the complex interplay between host immune response and gut-resident microorganisms. Although there are extensive public databases representing known host genetic interactions and host functional pathways, as well as extensive functional annotations of bacterial genomes, these resources may not always be current, and may have differing levels of noise due to incorrect annotations. Therefore we suggest treating these available data as its own "ome", with appropriate filtering, quality control, and data integration methods. When treated as a proper data source in conjunction with other 'omics data in a multi-omics study, we refer to this knowledge-base resource as "knowmics", a combination of the word "knowledge" and the neologism "omics". We use this term *knowmics* to describe the total sum of knowledge existing prior to the analysis of experimental data; the data themselves are called the "knowme" (pronounced like "gnome"). *Knowmics* includes both publicly available knowledge extracted from published annotation databases or directly from the scientific literature via text mining and knowledge supplied manually by an expert to refine or constrain the space of models under consideration. Given the complexity of IBD genotypes and phenotypes, and the vast library of IBD-related 'omics research published to date, IBD is an excellent candidate for mining relevant knowledge from the scientific literature.

The main challenges associated with utilizing *knowmics* data are the integration of disparate public annotation resources, and the manual input and representation of human expert knowledge. An immunologist designing a multi-omics experiment or analyzing multi-omics experimental data may have specific knowledge about host or microbial pathways from preliminary findings that have not yet been published. Therefore a knowledge-representation system that allows experts to easily add simple constraints or facts to the existing knowledge base may play a useful role in bringing new insight into the field. Knowledge-based expert systems, a subfield of artificial intelligence (AI), are designed to incorporate human knowledge when designing expert systems to solve problems that would normally require human intelligence. In order to interact with the expert systems and benefit from them, they need to at least have the following capabilities: knowledge representation, automated reasoning and the ability to learn from new data [77]. Knowledge representation often employs symbolic programming, allowing computers to store and manipulate human knowledge. Automated reasoning is the process of manipulating the stored knowledge to solve problems and to derive representations of new knowledge. An important aspect in AI systems is that computers "learn" patterns from known data, and can generalize those patterns to future data. Modeling human expert knowledge has been explored widely in AI research and has been applied to diverse fields including biology [78], medicine [79,80] and business [81], and should also provide benefit in multi-omics IBD research.

4. Combining the IBD 'omes

IBD exhibits complex disease phenotypes with likely contributions from impaired immune response, environmental microbial exposures, diet, and medication history, as shown in Fig. 2 [82]. Deleterious genetic mutations can lead to impaired immune function along several pathways [82]. Impaired immune response to commensal bacteria can alter the gut microbiome structure, in turn causing an imbalance or dysbiosis in bacterial metabolism in the gut. This dysbiosis can then exacerbate any imbalances in immune response, for example if an increase in opportunistic pathogens causes increased invasion of the intestinal lining. Dietary intake can alter microbiome profile [75], and administration of antimicrobials and immunosuppressants also have comprehensive effects on immune system behavior and microbiome profile. Medication history is expected to be confounded with genetics since IBD symptom severity can in part be predicted by host genetics [83,84]. Therefore there are complex multi-way interactions between the host immune system, the envirome, and the gut microbiome (Fig. 2). Certain aspects of each of these components can be observed in a
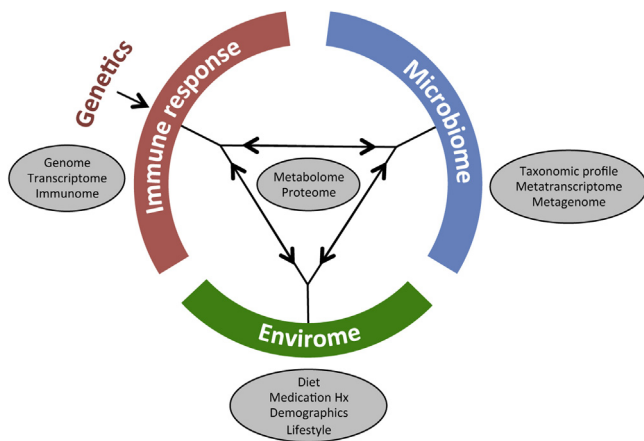
Fig. 2. Causal interactions between IBD 'omes. Inflammatory bowel diseases likely involve complex interactions between impaired host immunity, diet and other environmental exposures, and the gut microbiome structure and function. Impaired immune response to commensal bacteria may cause shifts in the taxonomic profile of the gut microbiome; this in turn may lead to a shift in the metabolic function of the microbiome [24]. Diet strongly affects gut microbiome structure and function [75], as does history of antibiotic and immunosuppressant usage [70,86]. Various high-throughput "omics" technologies may be used to interrogate these major components of IBD pathogenesis; metabolomics and proteomes or metaproteomics can provide simultaneous information about immune system and gut microbiome behavior, as well as diet.

high-throughput way using various "omics" technologies. The host genome and immune response can be observed using genomics, transcriptomics, and immunomics (immune-related proteomics, T- and B-cell repertoires, and targeted immune-related gene expression arrays). The gut microbiome can be measured with taxonomic

marker gene sequencing, shotgun metagenomics, and metatranscriptomics. Proteomics, metaproteomics, and metabolomics can be used to measure both host and microbial signaling and co-metabolism. Dual RNA-Seq, a relatively new approach, is designed to quantify simultaneously the transcriptomes of both hosts and microbial symbionts [29]. Finally, the envirome must be measured through food diaries, food frequency questionnaires, and carefully collected, annotated, and standardized clinical records.

Each IBD "ome" allows us to interrogate only one aspect of this multi-faceted disease model. We believe that multi-omics IBD research has the potential to lead more quickly to novel insights about the mechanisms of these interactions, although the integration and modeling of multi-omics data remains challenging. The process of modeling such data will require us to integrate vast quantities of published knowledge and manually entered expert knowledge (collectively, "knowmics" data) with multiple IBD "omes" (Fig. 3). The goal will be to use all available prior knowledge about the functional genomics of the host and the microbes, metabolic reaction networks, and disease pathways, as a foundation on which to analyze a smaller quantity of newly collected experimental multi-omics data. These disparate sources of prior knowledge will support the modeling and inference of novel multi-omics interactions (the "meta-interactome") with a much smaller sample size than would be required to infer all interactions de novo strictly from new experimental data. The complexity of the multi-omics modeling process can be reduced further when performing hypothesis-driven research instead of exploratory research.

Another primary challenge in designing multi-omics studies involving IBD patients is the determination of appropriate sample sizes. For this purpose we must rely on past experiments to estimate the relative size of the effect that we are investigating. This process is in fact not straightforward because of the many host and
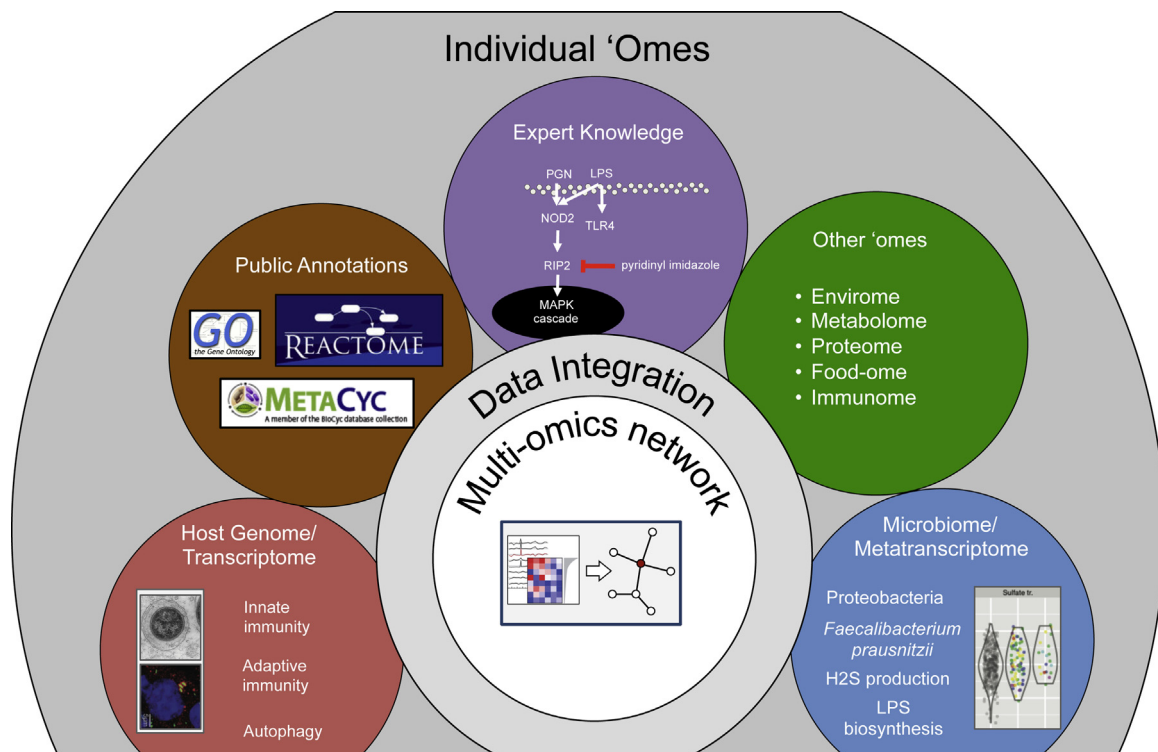

Fig. 3. Ideal IBD multi-omics model, integrating all available data types. The ideal model for analyzing multi-omics IBD data would link prior knowledge about host and microbiome functional genomics and metabolic networks to the new experimental data. This would in theory allow the model to share statistical strength across related features, and to reduce the number of tests performed by constraining the set of possible reactions. Multi-omics analysis tools could potentially utilize tools from knowledge representation theory for inferring semantic relations between gene products (both human and microbial) given both manually entered expert knowledge and predicted interactions derived from published annotations. The ideal multi-omics data set would include measurements of host genetics and host immune system function, microbiome taxonomy and function, gut-wide metabolism via proteomics and metabolomics, and dietary, clinical, and other environmental exposures.

environmental factors that influence host cell behavior and microbiome structure in patients, often with overlapping effects, and because of the heterogeneity of the disease phenotypes. As suggested by Fig. 1, there is an increasing number of large cohort studies that are beginning to allow us to measure these effect sizes in the individual 'omes. However, with respect to measuring multi-omics interactions, more exploratory research is necessary to establish effect sizes. The simplest approach in lieu of having published multi-omics data is to simply power one's study for the analysis of the individual 'omes being collected, and to treat the analysis of multi-omics interactions as exploratory.

5. Concluding remarks

IBD has complex etiology involving strong host, gut microbial, and environmental components. Although relatively new high-throughput "omics" technologies such as genomics, metagenomics, transcriptomics, metatranscriptomics, metaproteomics, and metabolomics allow us to interrogate different aspects of these components contributing to IBD pathogenesis, there have been relatively few "multi-omics" studies incorporating multiple data types from the same subjects. We suggest that additional focus be placed on formalizing approaches to modeling "enviromics", the total sum of relevant environmental and clinical data, and "knowmics", the total sum of prior available knowledge from annotation databases, published experimental data, and unpublished expert knowledge. These relatively under-represented IBD "omes" are of comparable importance to the other high-throughput "omes". There remains substantial need for bioinformatics development specifically for multi-omics data integration. Furthermore, data acquisition rates are growing faster than computer processing power, and it takes a relatively long time to train new bioinformaticians, especially given that they need to understand several sub-fields of bioinformatics in order to specialize in multi-omics modeling. Thus there are a number of modeling challenges involved in unifying and processing these disparate multi-omics data types; however, we expect such integrative multi-omics studies to play an essential role in mapping the complex etiology of inflammatory bowel diseases.

References

[1] Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. Nature 2007;448:427–34.
[2] Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature 2011;474:307–17.
[3] Fiocchi C. Towards a 'cure' for IBD. Dig Dis 2012;30:428–33.
[4] Weinstock JV, Elliott DE. Helminths and the IBD hygiene hypothesis. Inflamm Bowel Dis 2009;15:128–33.
[5] Yadav SP. The wholeness in suffix -omics, -omes, and the Word Om. J Biomol Tech 2007;18:277.
[6] Ishii N, Tomita M. Multi-omics data-driven systems biology of E. coli. In: Lee SY, editor. Systems biology and biotechnology of Escherichia coli. Dordrecht, Netherlands: Springer; 2009. p. 41–57.
[7] Zhang W, Li F, Nie L. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. Microbiol 2010;156:287–301.
[8] Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–24.
[9] Thompson AI, Lees CW. Genetics of ulcerative colitis. Inflamm Bowel Dis 2011;17:831–48.
[10] Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 2011;43:246–52.
[11] Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 2010;42:1118–25.
[12] Baur P, Gruber L, Bosco N, Brahmbhatt V, Collino S, Guy P, et al. Metabolic phenotyping of the Crohn' s disease-like IBD etiopathology in the TNF ($\triangle$ARE/WT) mouse model. J Proteome Res 2011;10:5523–35.
[13] Wirtz S, Neurath MF. Mouse models of inflammatory bowel disease. Adv Drug Deliv Rev 2007;59:1073–83.
[14] Sugimoto K, Ogawa A, Mizoguchi E, Shimomura Y, Andoh A, Bhan AK, et al. IL-22 ameliorates intestinal inflammation in a mouse model of ulcerative colitis. J Clin Invest 2008;118:534–44.
[15] Gophna U, Sommerfeld K, Gophna S, Doolittle WF, Veldhuyzen van Zanten SJO. Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. J Clin Microbiol 2006;44:4136–41.
[16] Bibiloni R, Mangold M, Madsen KL, Fedorak RN, Tannock GW. The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients. J Med Microbiol 2006;55:1141–9.
[17] Peterson DA, Frank DN, Pace NR, Gordon JI. Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. Cell Host Microbe 2008;3:417–27.
[18] Lepage P, Häsler R, Spehlmann ME, Rehman A, Zvirbliene A, Begun A, et al. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. Gastroenterol 2011;141:227–36.
[19] Westermann A, Gorski S, Vogel J. Dual RNA-seq of pathogen and host. Nat Rev Microbiol 2012;10:618–30.
[20] Erickson AR, Cantarel BL, Lamendella R, Darzi Y, Mongodin EF, Pan C, et al. Integrated metagenomics/metaproteomics reveals human host–microbiota signatures of Crohn's disease. PLOS ONE 2012;7:e49138.
[21] Hugot J, Chamaillard M, Zouali H, Lesage S, Ce J, Macry J. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 2001;411:599–603.
[22] Ogura Y, Bonen D, Inohara N, Nicolae D. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 2001;411:603–6.
[23] Ogura Y, Inohara N, Benito A, Chen FF, Yamaoka S, Nunez G. NOD2, a NOD1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. J Biol Chem 2001;276:4812–8.
[24] Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. Gut 2013;62(10):1505–10.
[25] Horak CE, Snyder M. Global analysis of gene expression in yeast. Funct Integr Genomics 2002;2:171–80.
[26] Ye RW, Wang T, Bedzyk L, Croker KM. Applications of DNA microarrays in microbial systems. J Microbiol Methods 2001;47:257–72.
[27] Kok K, Stokkers P, Reitsma P. Genomics and proteomics. Implications for inflammatory bowel diseases. Inflamm Bowel Dis 2004;10:1–6.
[28] Hermeking H. Serial analysis of gene expression and cancer. Curr Opin Oncol 2003;15:44–9.
[29] Wang Z, Gerstein M, Snyder M. RNA-seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.
[30] Costello CM, Mah N, Häsler R, Rosenstiel P, Waetzig GH, Hahn A, et al. Dissection of the inflammatory bowel disease transcriptome using genome-wide cDNA microarrays. PLoS Med 2005;2:e199.
[31] Carey R, Jurickova I, Ballard E, Bonkowski E, Han X, Xu H, et al. Activation of an IL-6:STAT3-dependent transcriptome in pediatric-onset inflammatory bowel disease. Inflamm Bowel Dis 2008;14:446–57.
[32] Marlow G, Ellett S, Ferguson IR, Zhu S, Karunasinghe N, Jesuthasan AC, et al. Transcriptomics to study the effect of a Mediterranean-inspired diet on inflammation in Crohn's disease patients. Hum Genomics 2013;7:24.
[33] Lin J, Welker NC, Zhao Z, Li Y, Zhang J, Reuss SA, et al. Novel specific microRNA biomarkers in idiopathic inflammatory bowel disease unrelated to disease activity. Mod Pathol 2014;27:602–8.
[34] de Bruyn M, Machiels K, Vandooren J, Lemmens B, Van Lommel L, Breynaert C, et al. Infliximab restores the dysfunctional matrix remodeling protein and growth factor gene expression in patients with inflammatory bowel disease. Inflamm Bowel Dis 2014;20:339–52.
[35] Clark PM, Dawany N, Dampier W, Byers SW, Pestell RG, Tozeren A. Bioinformatics analysis reveals transcriptome and microRNA signatures and drug repositioning targets for IBD and other autoimmune diseases. Inflamm Bowel Dis 2012;18:2315–33.
[36] Wu F, Guo N, Tian H. Peripheral blood microRNAs distinguish active ulcerative colitis and Crohn's disease. Inflamm Bowel Dis 2011;17:241–50.
[37] Pandey A, Mann M. Proteomics to study genes and genomes. Nature 2000;405:837–46.
[38] Mann M, Jensen O. Proteomic analysis of post-translational modifications. Nat Biotechnol 2003;21:255–61.
[39] Bennike T, Birkelund S, Stensballe A, Andersen V. Biomarkers in inflammatory bowel diseases: current status and proteomics identification strategies. World J Gastroenterol 2014;20:3231–44.
[40] Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol 2005;1:252–62.
[41] Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 2003;422:198–207.
[42] Alex P, Gucek M, Li X. Applications of proteomics in the study of inflammatory bowel diseases. Inflamm Bowel Dis 2009;15:616–29.
[43] Chaerkady R, Pandey A. Applications of proteomics to lab diagnosis. Annu Rev Pathol 2008;3:485–98.
[44] Barceló-Batllori S, André M, Servis C. Proteomic analysis of cytokine induced proteins in human intestinal epithelial cells: implications for inflammatory bowel diseases. Proteomics 2002;2:551–60.
[45] Fogt F, Jian B, Krieg RC, Wellmann A. Proteomic analysis of mucosal preparations from patients with ulcerative colitis. Mol Med Rep 2008;1:51–4.
[46] May D, Pan S, Crispin D. Investigating neoplastic progression of ulcerative colitis with label-free comparative proteomics. J Proteome Res 2010;10:200–9.

[47] Meuwis MA, Fillet M, Geurts P, de Seny D, Lutteri L, Chapelle JP, et al. Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling. Biochem Parmacol 2007;73:1422–33.

[48] Seeley EH, Washington MK, Caprioli RM, M'Koma AE. Proteomic patterns of colonic mucosal tissues delineate Crohn's colitis and ulcerative colitis. Proteomics Clin Appl 2013;7:542–9.

[49] Roda A, Caponi A, Sartini A, Cevenini M, Colliva C, Roda A. Update on current applications of proteomic in the study of inflammatory bowel disease. Ann Gastroenterol 2012;25:303–8.

[50] Patti GJ, Yanes O, Siuzdak G. Innovation Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol 2012;13:263–9.

[51] Marchesi J, Holmes E. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. J Proteome Res 2007;6:546–51.

[52] Marchesi JR, Holmes E, Khan F, Kochhar S, Scanlan P, Shanahan F, et al. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. J Proteome Res 2007;6:546–51.

[53] Jansson J, Willing B, Lucio M, Fekete A, Dicksved J, Halfvarson J, et al. Metabolomics reveals metabolic biomarkers of Crohn's disease. PLoS ONE 2009;4:e6386.

[54] Baur P, Martin FP, Gruber L, Bosco N, Brahmbhatt V, Collino S, et al. Metabolic phenotyping of the Crohn's disease-like IBD etiopathology in the TNF(ΔARE/WT) mouse model. J Proteome Res 2011;10:5523–35.

[55] Le Gall G, Noor SO, Ridgway K, Scovell L, Jamieson C, Johnson IT, et al. Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. J Protemoe Res 2011;10:4208–18.

[56] Williams HRT, Willsmore JD, Cox IJ, Walker DG, Cobbold JFL, Taylor-Robinson SD, et al. Serum metabolic profiling in inflammatory bowel disease. Dig Dis Sci 2012;57:2157–65.

[57] Hisamatsu T, Okamoto S, Hashimoto M, Muramatsu T, Andou A, Uo M, et al. Novel, objective, multivariate biomarkers composed of plasma amino acid profiles for the diagnosis and assessment of inflammatory bowel disease. PLOS ONE 2012;7:e31131.

[58] Schicho R, Shaykhutdinov R, Ngo J, Nazyrova A, Schneider C, Panaccione R, et al. Quantitative metabolomic profiling of serum, plasma, and urine by ^{1}H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. J Proteome Res 2012;11:3344–57.

[59] Ooi M, Nishiumi S, Yoshie T, Shiomi Y, Kohashi M, Fukunaga K, et al. GC/MS-based profiling of amino acids and TCA cycle-related molecules in ulcerative colitis. Inflamm Res 2011;60:831–40.

[60] Bjerrum JT, Nielsen OH, Hao F, Tang H, Nicholson JK, Wang Y, et al. Metabonomics in ulcerative colitis: diagnostics, biomarker identification, and insight into the pathophysiology. J Proteome Res 2010;9:954–62.

[61] Sharma U, Singh RR, Ahuja V, Makharia GK, Jagannathan NR. Similarity in the metabolic profile in macroscopically involved and un-involved colonic mucosa in patients with inflammatory bowel disease: an in vitro proton ((1)H) MR spectroscopy study. Magn Reson Imaging 2010;28:1022–9.

[62] Balasubramanian K, Kumar S, Singh RR, Sharma U, Ahuja V, Makharia GK, et al. Metabolism of the colonic mucosa in patients with inflammatory bowel diseases: an in vitro proton magnetic resonance spectroscopy study. Magn Reson Imaging 2009;27:79–86.

[63] Williams HRT, Cox IJ, Walker DG, North BV, Patel VM, Marshall SE, et al. Characterization of inflammatory bowel disease with urinary metabolic profiling. Am J Gastroenterol 2009;104:1435–44.

[64] Stephens NS, Siffledeen J, Su X, Murdoch TB, Fedorak RN, Slupsky CM. Urinary NMR metabolomic profiles discriminate inflammatory bowel disease from healthy. J Crohns Colitis 2013;7:e42–8.

[65] Storr M, Vogel HJ, Schicho R. Metabolomics: is it useful for inflammatory bowel diseases. Curr Opin Gastroenterol 2013;29:378–83.

[66] Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, et al. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207–14.

[67] Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 2012;40:109–14.

[68] Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 2013;31:814–21.

[69] Betteridge JD, Armbruster SP, Maydonovitch C, Veerappan GR. Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population. Inflamm Bowel Dis 2013;19: 1421–7.

[70] Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol 2012;13:R79.

[71] Jernberg C, Löfmark S, Edlund C, Jansson JK. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. ISME J 2007;1:56–66.

[72] Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. PNAS 2011;108:4554–61.

[73] Sokol H, Seksik P. The intestinal microbiota in inflammatory bowel diseases: time to connect with the host. Curr Opin Gastroenterol 2010;26:327–31.

[74] Pérez-Cobas AE, Artacho A, Knecht H, Ferrús ML, Friedrichs A, Ott SJ, et al. Differential effects of antibiotic therapy on the structure and function of human gut microbiota. PLOS ONE 2013;8:e80201.

[75] David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature 2013;505:559–63.

[76] Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;334:105–8.

[77] Russell S, Norvig P, Canny J, Malik J, Edwards D. Artificial intelligence: a modern approach. Egnlewood Cliffs: Prentice-Hall; 1995.

[78] Higuera C, Pajares G, Tamames J, Morán F. Expert system for clustering prokaryotic species by their metabolic features. Expert Syst Appl 2013;40:6185–94.

[79] Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Developing healthcare rule-based expert systems: case study of a heart failure telemonitoring system. Int J Med Inform 2012;81:556–65.

[80] Muthukaruppan S, Er MJ. A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. Expert Syst Appl 2012;39:11657–65.

[81] Lee K, Lee S. A causal knowledge-based expert system for planning an Internet-based stock trading system. Expert Syst Appl 2012;39:8626–35.

[82] Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. Gut 2013;62:1505–10.

[83] Tyler AD, Milgrom R, Stempak JM, Xu W, Brumell JH, Muise AM, et al. The NOD2insC polymorphism is associated with worse outcome following ileal pouch-anal anastomosis for ulcerative colitis. Gut 2013;62:1433–9.

[84] Sehgal R, Berg A, Hegarty JP, Kelly AA, Lin Z, Poritz LS, et al. NOD2/CARD15 mutations correlate with severe pouchitis after ileal pouch-anal anastomosis. Dis Colon Rectum 2010;53:1487–94.

[85] NCBI PubMed database. Bethesda, MD: US National Library of Medicine; 2009. Available from: https://www.ncbi.nlm.nih.gov/pubmed [cited 07.02.14].

[86] Antonopoulos DA, Huse SM, Morrison HG, Schmidt TM, Sogin ML, Young VB. Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. Infect Immun 2009;77:2367–75.

Hu Huang is pursuing a Ph.D. in Biomedical Informatics and Computational Biology in the Knights Lab at the University of Minnesota, and is currently a fellow of that program. His research includes developing novel bioinformatics tools for modeling multi-omics studies of host-associated microbiomes. He received a B.S. from Jiangsu University and an M.S. from Carnegie Melon University, both in Biomedical Informatics.

Pajau Vangay is currently pursuing a Ph.D. in Biomedical Informatics and Computational Biology in the Knights Lab at the University of Minnesota. She is interested in modeling host–microbiome interactions in the developing gut microbiome, especially in response to environmental perturbations such as exposure to antimicrobials. She is currently a fellow of the Biomedical Informatics and Computational Biology program. She received her B.S. and M.S. in Math and Computer Science from Colorado School of Mines, followed by an M.S. in Food Microbiology from Cornell University.

Christopher E. McKinlay is a postdoctoral scholar in the Knights Lab at the University of Minnesota, where he is modeling fine-grained temporal dynamics of host-associated microbial communities. He received his Ph.D. in Applied Mathematics from the University of California, Los Angeles. Prior to pursuing his Ph.D., he obtained a B.A. in Chinese from Middlebury College. His past research has applied wavelet decomposition methods and other temporal dynamics modeling to the analysis of turbulence.

Dan Knights is an Assistant Professor in the Department of Computer Science and Engineering and the BioTechnology Institute at the University of Minnesota. His research includes developing theory and tools for modeling genome-scale host–microbiome interactions as well as predictive modeling of microbiome response to perturbation. He received his B.A. from Middlebury College, and his M.S. and Ph.D. from the University of Colorado, all in Computer Science, with a certificate in Interdisciplinary Quantitative Biology from the University of Colorado BioFrontiers Institute. He also completed a post-doctoral research fellowship at Harvard Medical School and the Broad Institute of MIT and Harvard.