INNOVATION

The IBD interactome: an integrated view of aetiology, pathogenesis and therapy

Heitor S. P. de Souza, Claudio Fiocchi and Dimitrios Iliopoulos

Abstract | Crohn's disease and ulcerative colitis are prototypical complex diseases characterized by chronic and heterogeneous manifestations, induced by interacting environmental, genomic, microbial and immunological factors. These interactions result in an overwhelming complexity that cannot be tackled by studying the totality of each pathological component (an '–ome') in isolation without consideration of the interaction among all relevant –omes that yield an overall 'network effect'. The outcome of this effect is the 'IBD interactome', defined as a disease network in which dysregulation of individual –omes causes intestinal inflammation mediated by dysfunctional molecular modules. To define the IBD interactome, new concepts and tools are needed to implement a systems approach; an unbiased data-driven integration strategy that reveals key players of the system, pinpoints the central drivers of inflammation and enables development of targeted therapies. Powerful bioinformatics tools able to query and integrate multiple –omes are available, enabling the integration of genomic, epigenomic, transcriptomic, proteomic, metabolomic and microbiome information to build a comprehensive molecular map of IBD. This approach will enable identification of IBD molecular subtypes, correlations with clinical phenotypes and elucidation of the central hubs of the IBD interactome that will aid discovery of compounds that can specifically target the hubs that control the disease.

Over the past half century, medicine has scored impressive gains on multiple fronts: a better grasp of causes; a clearer understanding of mechanisms; a more precise identification of cells and molecules mediating disease; a more accurate targeting of mediators of tissue damage; and a substantial improvement in drug effectiveness and therapeutic outcomes. However, despite these successes, a large group of non-neoplastic diseases continue to pose formidable challenges in regard to understanding their pathogenesis and reaching a cure. Prominent examples are cardiovascular diseases, obesity, diabetes, rheumatoid arthritis, psoriasis, multiple sclerosis, asthma, as well as both forms of IBD (Crohn's disease and ulcerative colitis).

Several reasons explain why these conditions impose numerous challenges: all of them are chronic in nature and share common inflammatory pathways that might vary in time, distribution and intensity; they display both common and unique mechanisms; and some currently have no cure[1,2]. Active research is being undertaken for all of them, but even when advances are achieved they come with the frustrating realization of endless complexity at the pathophysiological level[3]. Hence, even though chronic inflammatory diseases represent distinct clinical entities, their commonalities have historically prevailed and they have been given the broad term of 'complex diseases' (REFS 3,4).

IBD is also a prototypical complex disease, terminology that goes beyond the recognized heterogeneity of clinical manifestations to indicate extreme complexity of pathogenic mechanisms, particularly when Crohn's disease and ulcerative colitis are examined under discrete perspectives[5]. The collective challenges coming from complex diseases are extended and magnified by the particular intricacies of each disorder, as is the case for both Crohn's disease and ulcerative colitis. The goal of this Perspectives article is to introduce the concept of the 'IBD interactome', a global disease network in which individual pathogenic components ('–omes') are integrated to avoid the pitfalls created by the study of each component in isolation. The merits of studying the IBD interactome will be highlighted, and its use justified by starting with a discussion of the current challenges intrinsic to IBD.

The challenges of IBD

Environmental challenges. Indirect but substantial evidence indicates that many complex diseases are 'modern diseases', a term indicating the emergence of several chronic conditions during the past century, including IBD, which have inflammation as a unifying background. The recent time of emergence, the progressive worldwide spreading of both Crohn's disease and ulcerative colitis, and the association with rapid and drastic changes in industrialization and social behaviour lend strong credence to the key contribution of environmental factors in the appearance of IBD[6]. The overall improved sanitary conditions in developed countries and the subsequent decreased incidence of infectious diseases have fueled the hygiene hypothesis as a probable explanation[7]. Concurrently, we now come into contact with man-made products and chemicals to which we can react in largely unknown ways. Thus, a myriad of major changes have created a brand new global environment of exposures — a new 'exposome' — that parallels the appearance of IBD worldwide. This understanding creates a number of IBD-related questions to which we have no answers. Have microorganisms been lost that might prevent IBD, for example,

Helicobacter pylori, helminths, soil or food microbiota, plant bacteria or viruses? Have we introduced factors into the exposome that might promote IBD, such as antibiotics, chemicals, pollutants, food additives or preservatives? Are there components missing from our diets that protected us from IBD, such as essential nutrients, vitamins or antioxidants? What component is in excess in modern diets that favours IBD? And what is the role of a daily stressful lifestyle on the body neuroendocrine–immune response? Currently, it is impossible to analyse each of these elements and factor all of them into a sensible pathogenic picture of cause and effect.

Genomic challenges. In the past two decades, technological advances and the creation of large consortia assembling thousands of patients with IBD have launched the era of IBD genetics, and so far ~200 genetic variants have been selectively associated with Crohn's disease, ulcerative colitis or both[8,9]. However, these variants combined only explain 20–25% of all IBD cases, indicating that genetics alone cannot explain IBD. Additionally, subgroups of patients with IBD carrying the same variants might be too numerous to evaluate and/or other subgroups might exist that are too rare or too small to reach statistical power for detection. Furthermore, individual IBD-associated genetic variants are concomitantly present in patients with IBD as well as healthy individuals[10], and the vast majority of variants only exert small effects[8]. These quantitative and qualitative realities prevent identifying genetics alone as a major culprit for IBD. Functional studies have suggested plausible pathways to disease mechanisms, as in the case for the nucleotide-binding oligomerization domain-containing protein 2 (*NOD2*) and autophagy-related 16 like 1 (*ATG16L1*) gene variants, which are thought responsible for an inappropriate response to microorganisms or their defective removal[11]. At the same time, 15 years after their discovery, the role of the very first Crohn's disease-associated *NOD2* variants has yet to be fully elucidated. How much longer will this investigation take, and will other variants ever be functionally investigated to a full extent? Even though we know that no single genetic variant can explain Crohn's disease or ulcerative colitis, will we examine variants one by one? How can we comprehend the contribution of single and combined multiple variants carried by the same patient? These lines of reasoning make it clear that IBD genetics might be reaching the limits of their potential and require integration with other pathogenic components to reveal their actual role in disease development.

Microbial challenges. The realization that microorganisms are an intrinsic part of humans, affecting all aspects of life, has led to an explosion of interest in the role of the gut microbiota in IBD[12]. The general consensus is that the gut microbiota is the target of an inappropriate immune response in genetically susceptible individuals and that this event is central in IBD pathogenesis[5]. This inappropriate response has been blamed on a 'loss of tolerance' towards the commensal microbiota, but it has been difficult to determine whether this process is secondary to an altered microbiota, a defective immune response or some other factors. Agreement on a loss of microbial diversity in IBD also exists, but it is still unclear whether this event is primary or secondary to chronic inflammation[13]. Reactivity against selected components of the gut microbiota is common, even in healthy individuals, and some Crohn's-disease-associated serological markers are present years before clinical manifestations in patients with Crohn's disease as well as healthy individuals[14]. Additionally, various microorganisms have been identified that are claimed to exert aggressive or protective functions relevant to Crohn's disease, such as adherent-invasive *Escherichia coli* and *Faecalibacterium prausnitzii*, respectively[15,16], suggesting that their relative abundances or balance might influence whether inflammation ensues. Finally, experimental evidence suggests that genes influence the microbiota and vice versa[17]. The number of reports implicating the gut microbiota in IBD pathogenesis continues to grow exponentially but, because of a lack of long-term prospective studies, there is still insufficient information to address the central question of whether dysbiosis in IBD is a primary or secondary phenomenon. The lack of a response to this crucial question is one of the most important challenges facing investigators, and one not likely to be answered by simply refining qualitative or quantitative screenings of gut bacteria, fungi or viruses.

Immunological challenges. Immunology dominated the investigation of IBD pathogenesis from the 1970s until a decade ago, when genetics and the microbiota started to receive more attention. This domination was justified considering that the inflammatory processes of Crohn's disease and ulcerative colitis are executed primarily by immune cells. Initially focused on antibody responses, then T cells and then cytokines, IBD research is currently more focused on innate immunity largely due to the reactivity against microbial agents[5]. This evolution has generated plenty of information and important lessons. One important discovery was that gut inflammation persists despite high concentrations of anti-inflammatory cytokines (such as IL-10 and transforming growth factor (TGF)-β) in IBD-affected mucosa[18,19], an observation compatible with a lack of immune regulation[5]. Based on evidence from animal models, Crohn's disease and ulcerative colitis were classified into conceptually acceptable but biologically dubious conditions based on type 1 T helper (T_H1), type 2 T helper (T_H2) and type 17 T helper (T_H17) cells, as well as cytokine profiles[20]. This categorization prompted therapeutic trials aimed at blocking the central mediator of each profile, with outcomes that were as disappointing as they were revealing: neutralizing IFNγ failed to improve Crohn's disease[21], neutralizing IL-17 worsened Crohn's disease[22], and neutralizing IL-13 failed to improve ulcerative colitis[23]. The simplistic assumption that each type of IBD is mediated by a particular type of immune response has proven to be inaccurate by experimental and clinical studies showing fluctuation of T_H1 and T_H2 responses during the disease course[24,25]. All these observations are fundamental because they demonstrate that what is measured by isolated immune parameters does not necessarily mediate IBD. Hundreds of immune cell subpopulations exist with highly complex gene signatures and associated functions[26], making interpretation of immune responses far more problematic. Thus, even though the immune system is the executioner of gut inflammation, its role is overemphasized and can only be fully understood in the context of signals received by other components of IBD pathogenesis that must be integrated.

Tackling the challenges of IBD

Facing reality: overwhelming complexity. With the exception of the inherited genome, no other IBD-related –ome is fixed or predictable. This reality creates a whole universe of diverse and interacting variables that result in the disease complexity observed in practice. Biological diversity is intrinsic to human physiology, and the

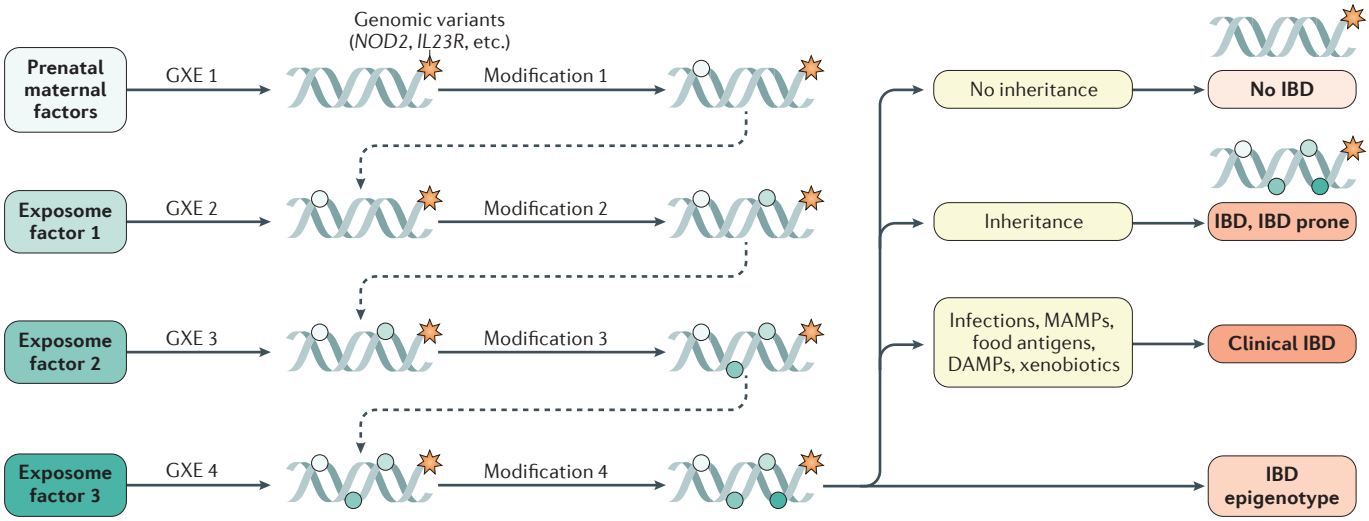

Figure 1 | **Role of epigenetic modifications in the development of IBD.** A fetus who inherited an IBD susceptibility genetic variant, or variants, is initially exposed to prenatal maternal factors, such as stress, infections, poor diet and obesity. Any of these factors might first induce a gene–environment (GXE) interaction (between fetal genes and maternal factors), resulting in an epigenetic modification of chromatin that alters the expression of a specific gene. After birth the individual encounters an exposome-derived factor that induces a second GXE interaction resulting in an epigenetic modification of chromatin that modifies the expression of another specific gene. Subsequent exposures lead to further GXE interactions, resulting in the accumulation of multiple epigenetic modifications and the establishment of an IBD epigenotype. Individuals carrying such an epigenotype might encounter exogenous (infections, microorganism-associated molecular patterns (MAMPs), dietary antigens, xenobiotics) or endogenous (damage-associated molecular pattern; DAMPs) signals that activate the clinical IBD phenotype. Owing to the potential inheritance of epigenetic modifications, if the IBD epigenotype is transmitted to the next generation, the progeny might also manifest IBD or be IBD-prone and develop IBD upon encounter of the same or other exogenous or endogenous signals. If the IBD epigenotype is not transmitted, the next generation will not develop IBD or be IBD-prone.

components of this diversity and why they might lead to different clinical expressions will be discussed. Although unpredictable, at the core of complex diseases are closely intertwined mechanisms, such as the immune response, fibrosis and thrombosis, which might shape specific endophenotypes relevant to IBD, such as ileal, ileocolonic or colonic inflammation[4].

Environmental complexity. The host, and future patient with IBD, does not live in 'splendid isolation' but rather in two incredibly complex milieus represented by the external environment (the general exposome) and the internal microbial environment (the gut microbiome), both of which influence each other as well as all other –omes in reciprocal fashion. The essential role of gene–environment interactions in inducing and modifying most diseases is now appreciated, including in IBD[6]. Here, the difficulty resides in understanding which and how many environmental factors are relevant, when they act on the host, and which host –omes are affected. Although a single environmental factor might have an effect, it is far more likely that multiple sequential gene–environment hits are needed to initiate and establish disease[27–31] (FIG. 1).

Currently, we have no tools at our disposal that enable us to identify all potentially IBD-relevant components of the exposome, but a few have emerged for which there is substantial experimental support, such as the diet and artificial sweeteners altering the gut microbiota[32], xenobiotics altering the immune response[33], and cohabitation[34].

Genomic complexity. Genome-wide association studies (GWAS) have been helpful in identifying hundreds of genetic variants in Crohn's disease, ulcerative colitis or IBD, and more will be discovered as deep re-sequencing continues to be performed[35]. Mendelian-like mutations can also be identified, such as those affecting the IL-10 receptor[36], but these mutations are quite rare and, unfortunately, "genomics is not enough" to solve the IBD puzzle[37]. Certain genetic variations, such as *NOD2* variants in Crohn's disease, induce specific phenotypes such as fibrostenotic ileal disease[38], but most IBD variants do not induce discrete phenotypes. This finding might be because different mutations in the same gene or locus lead to different interactions and result in distinct clinical presentations[39]. In addition to individual susceptibility genes defined by DNA variations, susceptibility loci alter gene networks that in turn lead to disease[40]. Moreover, all genes and gene networks are affected by epigenetic modifications that account for genotype-to-phenotype transition[41]. Despite their limitations, genomics can be useful in assessing risk of disease[42], and genomic classifications have been reported in lung cancer, pancreatic cancer and acute myeloid leukemia[43–45]. However, GWAS are more powerful when integrated with other platforms such as DNA copy number, DNA methylation, mRNA expression, microRNA (miRNA) expression and protein expression[46]. These comprehensive analyses capture disease biological features better than clinical or histopathological data, and enable precise molecular classification — the ultimate goal in IBD.

Epigenomic complexity. The majority of disease-associated gene variants reside in non-coding regions[47], making these regions vital regulators of gene expression. This understanding highlights the importance of nongenetic modifications of the genome. Epigenetic modifications, among other components such as enhancer bindings and transcription factors, have the potential to determine how, where and when genes are expressed[48]. Various enzymatic processes, such as methylation or histone

modifications, alter and vary the structure of chromatin from active to repressed[49], and these modifications are critical in determining ultimate gene expression. By profiling histone modifications, DNA accessibility, DNA methylation and RNA expression, the NIH Roadmap Epigenomics Mapping Consortium has generated a collection of human epigenomes[50] and shown that they are particularly enriched for disease-associated gene variants, providing a whole new insight into the molecular basis of human disease[50]. This finding has critical implications for complex diseases like IBD, in which epigenetic modifications occur from before clinical manifestations and throughout the disease course[27]. The emerging role of epigenetics in IBD is being recognized[51], and some studies show correlations between DNA methylation or specific miRNAs (non-coding RNA molecules regulating gene expression) with the status of patients with Crohn's disease or ulcerative colitis[52,53]. The fundamental importance of epigenetics goes beyond events in the human host, as gut epithelial-cell-derived miRNAs have also been shown to alter the gut microbiota[54]. When all these observations are viewed in the context of IBD, the multiplicity of reciprocal regulatory circuits involved in disease pathogenesis becomes obvious and studying genetics or epigenetics separately cannot explain Crohn's disease or ulcerative colitis.

Transcriptomic, proteomic and metabolic complexity. The complexity of interactions in IBD extends beyond the genome and the epigenome, and involves other components downstream of gene regulation, including the transcriptome, the proteome and the metabolome. Gene transcription is regulated by transcription factors, cofactor and chromatin regulators, and transcriptional dysregulation is associated with multiple diseases[55], including IBD, in which careful analysis of tissue mRNA expression can provide information on disease activity, disease course and response to therapy[56]. Interpretation of gene transcription, however, must be done with caution because the human transcriptome displays widespread RNA–DNA differences, and individual RNA variations result in individualized proteomes[57]. This understanding might explain the considerable variability in protein levels between individuals, populations and sex, and the usually poor correlations between mRNAs and protein levels[58]. Protein levels could be involved in the unpredictable responses to IBD therapies observed in the clinic. For example, a hypothetical patient with IBD that produces little TNF might display a delayed and mild disease course with an excellent response to anti-TNF agents; conversely, if the patient produces large amounts of TNF, he or she might display an aggressive and severe disease course, and be less responsive to the same agents. Partial IBD proteomes have been reported, mostly focused on detection of biomarkers and differentiating Crohn's disease from ulcerative colitis[59], but not on their pathophysiological implications. Similarly, in early metabolomic profiling studies of serum, plasma and urine, quantitative differences are found between patients with IBD and healthy individuals, but metabolic differences between Crohn's disease and ulcerative colitis are less pronounced, perhaps owing to commonalities imposed by chronic inflammation[60]. Metabolomics is a powerful tool to reveal key biological mechanisms[61], and more information on IBD should be forthcoming.

Microbial complexity. Studies of the gut microbiota have started to characterize quantity, quality and function, revealing an overwhelming diversity of interactions among microorganisms and with the host. The dynamic early-life microbiome becomes more stable with age[62] and during this evolution it exerts innumerous activities essential to health and disease, such as immune system development, establishing tolerance, metabolizing nutrients, and many others. All these processes occur in the context of interactions in which the gut microbiota are either influencing or being influenced by other –omes. For instance, genetics shape the composition of the gut microbiota[17], as do other intrinsic and extrinsic host factors, such as diet, drugs, xenobiotics and disease[63]. These interactions are common but certainly not identical among individuals, and their outcome is the creation of personal microbiomes; unique microbial fingerprints that distinguish one individual from all others[64]. Personal microbiomes are present in health and disease, which raises the intriguing question of whether a particular disease can only occur in a given person because of the personal microbiome. This situation would create a highly 'personalized disease', mediated by unique sets of pathogenic factors. Is this the case of IBD, in which interactions between specific gene variants and a unique microbiota contribute to disease pathogenesis[65]? This situation is not impossible considering the extreme heterogeneity of both Crohn's disease and ulcerative colitis in regard to age of onset, symptoms, clinical manifestations, phenotype, evolution, complications, response to therapy and need for surgery[66]. Such a situation would also explain why it is so difficult to correctly classify patients, select the right medications and predict clinical and therapeutic outcome with the tools we currently use. The logical conclusion is that we need new tools to better analyse and integrate the complexity of IBD to a much greater degree.

Solving complexity: the IBD interactome. The many challenges that remain in IBD cannot be addressed by the approaches taken so far; alternative approaches are needed to advance understanding and treatment. To start, it should be noted that, thus far, investigation of each of the –omes discussed earlier — exposome, genome, microbiome and 'immunome' — has been essentially researched in isolation without taking into account the concomitant role of and input of other –omes. Although overall knowledge of IBD continues to expand, new information generated by individual pathogenic components is not being integrated[5]. The result is increased, but isolated, sets of information relevant to single specific –omes without the benefit of data harmonization (FIG. 2a). As multiple pathogenic factors must come together to trigger IBD, a lack of harmonization hinders progress because accumulating data from different techniques remains separated and incapable of explaining the disease, regardless of the volume collected. On the other hand, if the growing body of data derived from each –ome is funneled together, complementary and deeper insights should be revealed, greatly improving our understanding of IBD (FIG. 2b). Scientific knowledge doubles every decade or so[67], making data integration increasingly urgent, but how can data integration be accomplished in IBD? Human diseases represent nonlinear systems that cannot be understood by analysing their components individually[68]. Moreover, nonlinear systems do not hold proportionality, and small alterations produce unpredictable outcomes[68]. Chronic diseases, including IBD, are the result of multiple biological events mediated by dynamic molecular networks with intricate, nonstationary, intermittent and nonlinear behaviours that depend

on the number and type of interactions among the network nodes (hubs), yielding an overall 'network effect' (the disease)[69]. Evaluation of the network effect in IBD requires state-of-the-art systems biology and bioinformatics tools that bring together all relevant –omes under a single umbrella capable of incorporating vast amounts of information. Based on the original definition of an interactome as a biological network with subunits that are physically and functionally linked into a whole[70], we believe this umbrella can be created by characterizing the IBD interactome, defined as a disease network in which dysregulation of individual –omes causes intestinal inflammation mediated by dysfunctional molecular modules controlling all biological responses. Thus, use of systems biology and bioinformatics is not only advantageous but also indispensable. How these tools can be utilized is the focus of the next section.

Systems biology

A system is defined as a set of interacting components forming an integrated whole. The first step in a systems approach is the identification of the components of the system, the second step is the functional characterization of each component, the third step is the integration of all of the components through mathematical models that describe the structure of the system, and the last step is the evaluation of the response of the system to external or internal perturbations[71]. The field of systems biology was born ~15 years ago and the initial definition included: the integration of molecular information into molecular networks and evaluation of the response of these networks to different stimuli[72]. The ability to generate genome-wide measurements on a system is arguably the single greatest force driving the popularity of the systems biology field. Moreover, systems biology is not only about high-throughput –omics measurements, but also a different philosophy of performing biomedical research. Compared with the classic hypothesis-driven approach, a systems approach is an unbiased data-driven strategy in which data integration reveals the key players of the system. Importantly, systems biology approaches have been used to characterize disease interactomes, identify key drivers of pathogenesis and develop specifically targeted therapies.

Systems biology and the IBD interactome. On the basis of the IBD interactome definition and the powerful bioinformatics tools able to query and integrate multiple –omes, IBD systems biology is ready for prime time[73]. IBD is an ideal condition for systems biology studies considering the easy access to all samples needed for a comprehensive evaluation of all

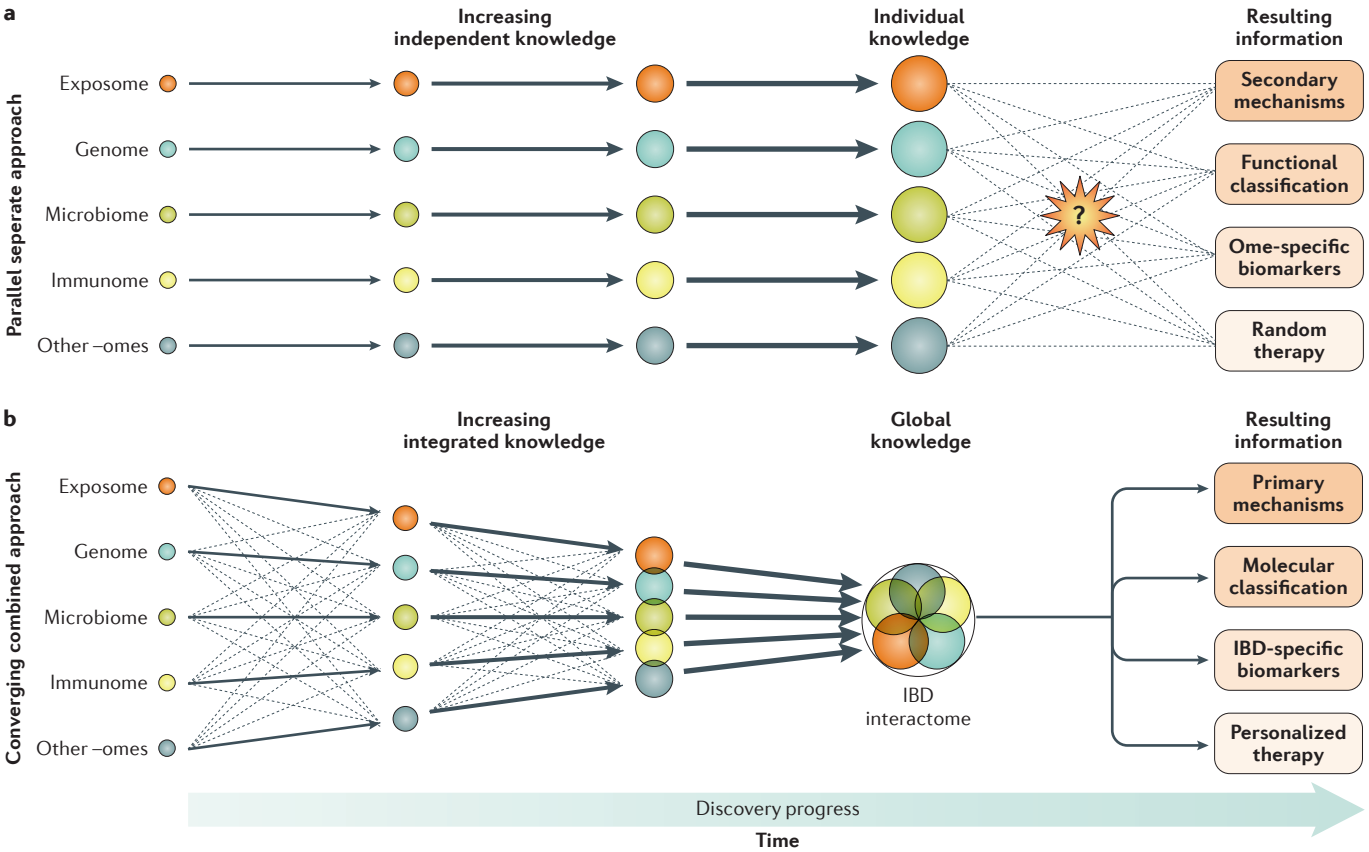

Figure 2 | **Knowledge expansion and building of the IBD interactome.** The figure illustrates two different ways to utilize the progressively accumulating knowledge of IBD pathogenesis. **a** | This panel shows the parallel separate approach, the most commonly utilized approach based on routine scientific methodologies. With this approach, increasing information on all IBD-relevant –omes develops over time but separately from each other, resulting in large bodies of distinct data that are not integrated. This approach results mainly in new information on secondary mechanisms, functional classifications, biomarkers specific to a single –ome and random therapeutic targets. **b** | This panel shows the converging combined approach, which is based on the utilization of state-of-the-art systems biology and bioinformatics tools. With this approach, increasing information on all IBD-relevant –omes also develops over time, but the data are progressively integrated among all –omes, resulting in large bodies of integrated information that converge to form the IBD interactome. This approach results in new information on primary mechanisms, molecular classifications, IBD-specific biomarkers and precise therapeutic targets (personalized therapy).

Table 1 | **Systems biology software programmes**

Name	Function of software	Weblink	Refs
iCluster	Identification of disease molecular subtypes	https://cran.r-project.org/web/packages/iCluster/index.html	98
PARADIGM	Multiomics data integration	https://sbenz.github.io/Paradigm/	101
TieDIE	Gene–protein interaction networks	https://github.com/epaull/TieDIE	102
iBAG	Identify genes linked to clinical outcomes	http://odin.mdacc.tmc.edu/~vbaladan	103
IPA	Multiomics data integration	http://www.ingenuity.com/products/ipa	104
Cytoscape	Visualization of –omics data	http://www.cytoscape.org/	105
BiNGO	Gene ontology of biological networks	http://www.psb.ugent.be/cbd/papers/BiNGO/Home.html	107
GenePro	Integration and visualization of networks	http://wodaklab.org/genepro/	108
OmicsViz	Visualization of large scale –omics datasets	http://metnet.vrac.iastate.edu/MetNet_fcmodeler.htm	109
Metscape	Visualization of metabolomics data	http://metscape.ncibi.org	110
CytoCom	Visualization of disease comorbidity networks	http://www.cl.cam.ac.uk/~mam211/	111
VisANT	Visualization of gene networks	http://visant.bu.edu/	112
Pathway Studio	Customizable network display	https://www.elsevier.com/solutions/pathway-studio-biological-research	113
ProViz	Visualization of protein–protein networks	http://www.cbib.u-bordeaux2.fr/eng/proviz.htm	114
Connectivity Map	Identification of drug–gene signature interactions	https://www.broadinstitute.org/cmap/	117
iFad	Pathway-based drug discovery	http://bioinformatics.oxfordjournals.org/content/28/14/1911.long	119

relevant –omes. Such samples include blood (genome, epigenome, immunome), serum (proteome, metabolome), stools (microbiome, proteome, metabolome) and mucosal tissue (transcriptome, proteome, metabolome, microbiome, immunome). Examples of how IBD-relevant –omes can be analysed and characterized to create the IBD interactome are given in the following sections, and a list of systems biology software programs is shown in TABLE 1.

IBD genome. GWAS and meta-analyses have identified ~200 genetic loci associated with IBD, but only a few single nucleotide polymorphisms (SNPs) correspond to nonsynonymous coding variation with a clear effect on protein function; the majority of these SNPs involve non-coding variations[74,75]. Additionally, expression quantitative trait loci (eQTL) studies revealed that gene expression is regulated by non-coding genetic variations and, interestingly, the SNPs identified from IBD GWAS are highly enriched within active enhancer marks[9]. These findings suggest that epigenomic alterations are prominently involved in IBD pathogenesis and, therefore, the characterization of the IBD epigenomic landscape should be performed at the level of DNA methylation and histone modifications.

IBD DNA methylome and hydroxymethylome. DNA methyltransferase 1 (DNMT1) and DNA methyltransferase 3B (DNMT3B) have been found to be upregulated in tissues from patients with IBD[76]. Furthermore, increased DNMT1 levels correlate with the abundance of $CD68^+$ macrophages[77], whereas DNMT3A has been found to regulate T-cell polarization through IL-4 and IFNγ promoter methylation upon ligation of T-cell receptors[78]. Establishing the methylation status of ~27,000 CpG sites showed 50 genes to be differentially methylated between healthy individuals and patients with Crohn's disease[52]. The application of DNA methylation sequencing technologies would enable identification of genome-wide DNA methylation alterations in IBD. Reduced representation bisulfite sequencing is highly applicable to evaluating DNA methylation changes in IBD clinical samples, as it requires very low amounts of genetic material[79]. In addition to DNA methylation, evaluating changes in 5′-hydroxymethylation, which are catalysed by teneleven translocation (TET) proteins[80], would also be of great interest in patients with IBD. In 2015, TET2 was identified as an active repressor of IL-6 transcription in innate myeloid cells[81]. Thus, it would be valuable to characterize similar genome-wide hydroxymethylation changes at other loci in patients with IBD.

IBD chromatin code. The presence of IBD-associated SNPs in enhancer areas suggests the importance of histone modifications in IBD[82,83]. Seven histone modifications (H3K4me3, H3K4me1, H3K36me3, H3K27me3, H3K9me3, H3K27ac, H3K9ac) evaluated by the NIH Roadmap Epigenomics Mapping Consortium[50] can accurately map and characterize the human epigenome. Thus, it would be important to characterize these chromatin modifications by using chromatin immunoprecipitation (ChIP)-sequencing technology, and integration of all seven histone marks could create a chromatin state model for healthy individuals and patients with IBD. This chromatin state model would consist of 15 different states (eight active states and seven repressed states)[50] and could provide new insights into gene expression and regulation in IBD.

IBD transcriptome. Several studies have identified genome-wide alterations at the level of mRNA and miRNAs in patients with IBD[53,56,84–86]. However, the potential role of long non-coding RNAs (lncRNAs) and circular RNAs (cRNAs) in IBD pathogenesis has yet to be investigated. LncRNAs have been found to regulate chromatin complexes[87], whereas cRNAs have been identified as negative regulators of miRNA expression[88], and both types of non-coding RNA could be involved in IBD pathobiology. Future studies should, therefore, perform integrated analysis of coding and non-coding RNA expression in patients with IBD by using RNA-sequencing technologies.

IBD proteome and metabolome. Proteomic studies, using mass spectrometry and liquid chromatography, have been performed in IBD biopsy and serum samples, revealing proteomic signatures that correlate with IBD disease activity and response to

anti-TNF therapeutics[89,90]. However, novel high-throughput technologies are required to analyse in depth the phosphoproteomic alterations in patients with IBD. EasyPhos is a scalable phosphoproteomics platform that could rapidly quantify hundreds of IBD phosphoproteomes at a depth of >10,000 sites[91]. In addition, metabolomics approaches, such as gas and high-performance or ultra-performance liquid chromatography, have been applied in human IBD biomaterials, revealing distinct metabolomic profiles[92,93]. Thus, it would be extremely valuable to integrate IBD proteomic and metabolomic data to build a comprehensive biochemical map of IBD pathogenic events.

IBD microbiome. The number of studies evaluating alterations in gut microbiota in IBD has exploded, but IBD-associated microbial alterations have yet to be interpreted in light of the other IBD –omes. The abundance of Enterobacteriaceae and Fusobacteria has been found to be increased in both experimental and clinical IBD[94,95], but it is still unclear whether this finding, as well as other reported microbial abnormalities, represent a primary or secondary event. A commonly cited example of a link between IBD genetics and the microbiome is through *NOD2*, which triggers an innate immune response on recognizing muramyl dipeptide, a cell wall peptidoglycan constituent of bacteria[96]. One report proposed that polymorphisms in IBD susceptibility genes promote disease through defects in sensing protective signals from the microbiome[65]. Despite intriguing correlations among genetic variants, shifts in microbial composition and IBD phenotypes, to date, no evidence exists for a causative role of the microbiome. Thus, it is essential to integrate microbiome data with other –omics data to evaluate the true importance of the gut microbiota in IBD pathogenesis.

Computational integration of –omics data. During the past decade, there has been an explosion in the generation of high-throughput –omics data, and this has created several computational and statistical obstacles related to their storage, analysis, integration and, above all, interpretation[97]. Thus, extensive efforts have been undertaken by the scientific community to integrate –omics data into disease clusters and networks, aiming to comprehensively understand disease mechanisms, identify the key drivers of disease, and to develop novel therapeutics targeting specific molecular networks. To accomplish these goals, a number of biology software programmes have been developed and made available to the scientific community (TABLE 1).

A fundamentally important issue in the biomedical field is the identification of patient clusters (subtypes) based on multiomic high-throughput data. An integrative clustering method called iCluster was developed based on a joint latent variable model able to simultaneously integrate multiomic data sets and generate a single integrated cluster assignment[98]. An advanced version of iCluster, called iCluster+, is able to perform pattern discovery and has been extensively used by the NIH-sponsored consortia, The Cancer Genome Atlas, to identify molecular subtypes for multiple cancer types[99,100]. These studies have shown that evaluation of multiple (3–5) –omics datasets have a better predictive value and statistical significance than single –omics to identify human disease subtypes, emphasizing the importance of evaluating the combined –omic dimensions forming the disease interactome.

Another tool for integrating molecular patterns at the level of gene expression, copy number variation and epigenetic changes is PARADIGM, a pathway recognition algorithm[101]. This algorithm contributes to the identification of patient subtypes that correlate with specific clinicopathological parameters. To determine whether specific genotypes (SNPs) are significantly associated with the network hubs identified with the PARADIGM software, the user can utilize the integrative approach named Tied Diffusion Through Interacting Events (TieDIE)[102], which enables searching for statistically significant interconnections between genotype perturbations and molecular network alterations. In addition to the identification of disease subtypes based on multiomic datasets, it is also essential to discover specific genes or gene signatures related to clinical conditions. A model developed in 2013, called iBAG (integrative Bayesian Analysis of Genomic data), provides the framework for identifying key genes associated with clinical outcome[103]. Furthermore, the Ingenuity Pathway Analysis (IPA) tool is a commercially available web-based system that transforms –omics data into a set of relevant networks based on extensive records maintained in the Ingenuity Pathways Knowledge Base (IPKB)[104]. This knowledge base is currently a database that includes updated information from current literature, as well as public and third-party databases (for example, the Online Mendelian Inheritance in Man catalogue (OMIM), GWASdb, Human Metabolome Database, ClinicalTrials.gov, KEGG, TargetScan).

Visualization of molecular interactomes. Novel software packages have been developed to assist with modelling of molecular interactomes, providing an intuitive interface and visual display of the –omics data. The visualization of the –omics data and the interactome is essential to our understanding of the perturbations that exist during human disease initiation and progression. One of the most popular software tools used for visualization of biological networks is called Cytoscape[105], a free computational tool that enables data import and export, data integration, and visualization of the interactome (see Supplementary information S1 (figure)). Cytoscape features VizMapper, which enables control of the shape, colour and size of the nodes and edges of the networks, and is appropriate for overlaying multiple –omics datasets into a single network context. Importantly, Cytoscape's architecture permits the development of

Box 1 | **A systematic stepwise approach to the identification of the IBD interactome**

- Assemble bioinformatics, computational biology, high-throughput data analysis tools and expertise
- Select and follow-up a phenotypically homogeneous group of patients with IBD
- Prospectively collect patient biosamples during the evolution of IBD, including blood, serum, endoscopic biopsy samples and stool
- Using appropriate biosamples, analyse the genome, transcriptome, epigenome, proteome, microbiome and other –omes relevant to IBD pathogenesis
- Using a systems approach, integrate the data derived from individual –omes and identify the various IBD molecular subtypes
- Each molecular subtype will consist of a unique IBD interactome with specific central regulators whose biological outcome is gut inflammation
- Screen the regulators against drug databases to discover compounds targeting the key regulators responsible for IBD in individual patients or group of patients

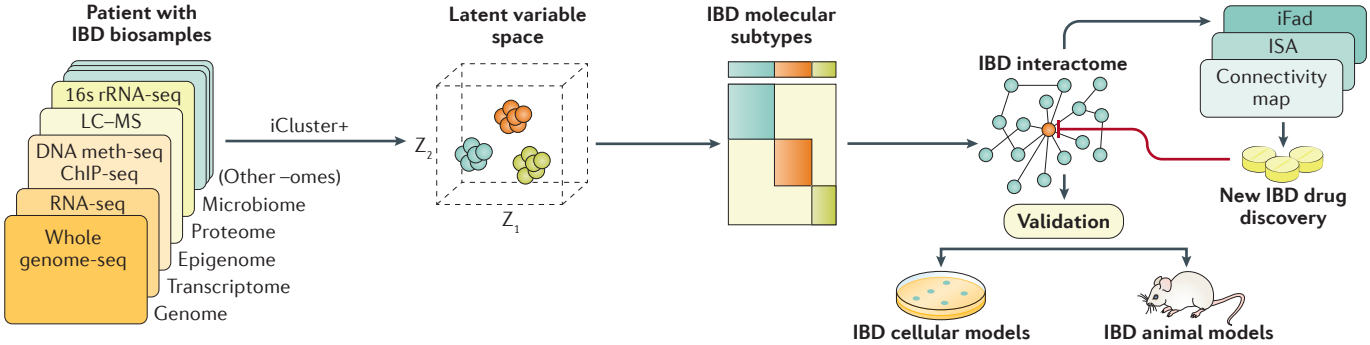

Figure 3 | **Building and therapeutic targeting of the IBD interactome.** Using a variety of patient-derived biosamples (blood, serum, endoscopic biopsy samples, stool), high-throughput –omics technologies can be applied to characterize the genome, epigenome, transcriptome, proteome, metabolome and microbiome (as well as additional –omes) of individual patients with IBD. The resulting –omes can then be integrated with different computational procedures, such as the iCluster algorithm, resulting in the identification of IBD molecular subtypes, and bioinformatics tools such as Cytoscape can be used to build and visualize the IBD interactome. Characterizing the IBD interactome enables the identification of the central regulators (hubs) of the network. In this example, the central hub (orange circle) could be a gene that is highly interconnected with other genes and regulates the whole network of biological mechanisms underlying disease pathogenesis. The biological significance and pathogenic relevance of these hubs can be tested and validated in *in vitro* (cellular models, organoids, tissues) and *in vivo* (normal and genetically manipulated animals) IBD models. Specific targeting of the regulatory hub or hubs can be accomplished using computational tools for drug discovery such as Connectivity Map, Iterative Signature Algorithm (ISA), or Integrative Factor Analysis for Drug Pathway Associated Inference (iFad), which could lead to new drug discovery for personalized therapy. ChIP-seq, chromatin immunoprecipitation followed by sequencing; LC–MS, liquid chromatography–mass spectroscopy; meth-seq, methylation sequencing; RNA-seq, RNA sequencing; rRNA-seq, ribosomal RNA sequencing.

plugins by programmers. To date, there are 152 available plugins compatible with the latest Cytoscape release[106]. The most important available plugins are: BiNGO, a plugin to assess overrepresentation of gene ontology categories in biological networks[107]; GenePro, a plugin for advanced visualization and analysis of interactomes[108]; OmicsViz, a plugin for visualizing –omics data across species[109]; Metscape, a plugin for visualizing and interpreting metabolomics data in the context of human metabolic networks[110]; and CytoCom, a plugin application for the visualization of disease comorbidity networks[111].

In addition to Cytoscape, another freely available software tool that can be used to visualize the interactome is VisANT[112], an application for integrating –omics data into a cohesive and graphical interface. Pathway Studio is a Windows-based application that includes customizable network display styles for assigning visual attributes such as node size and shape, and subcellular localization[113]. ProViz[114] is another software that leverages the power of the graph drawing package Tulip[115] for handling graphs containing millions of nodes and edges, while maintaining a predefined length of time. Tulip also supports plugins enabling third-party programmers to extend the system. In summary, multiple software visualization tools are currently available to graphically present molecular networks in basic biological systems and human diseases. The appropriate choice of visualization tool depends on the size of the imported information, the number of –omes to be included, and the aim of the study. Optimized versions for all these tools are continuously being developed, enabling scientists to quickly identify and analyse the network central hubs at the experimental level.

Systems biology and drug discovery

System biology has created high-hopes for expediting the drug identification and development process in comparison to the traditional 'one target–one drug' perspective that ignores real-life complex molecular interactions. Furthermore, systems biology is more effective for drug discovery because it is now obvious that most complex diseases, including IBD, result from system-level abnormalities rather than defects in individual genes[116].

The development of novel computational methods focusing on drug–network interactions is revolutionizing the drug discovery field. Specifically, drug targets can be identified by comparing RNA profiling data in human cellular models treated with a library of chemical compounds. The Connectivity Map software uses a nonparametric rank-based pattern-matching strategy based on the Kolmogorov-Smirnov statistic[117]. Another bi-clustering method, called Iterative Signature Algorithm (ISA), searches for 'co-modules' representing gene–drug associations[118]. A Bayesian sparse factor analysis model, called integrative Factor Analysis for Drug–pathway pathway association inference (iFad), is able to integrate two distinct data types in a unified framework to identify drug–target pathways[119]. For example, the iFad model has been applied to the joint analysis of gene expression and drug sensitivity profiles of the NCI-60 cell lines, representing a comprehensive resource for various –omics data from 60 human cancer cell lines, as well as treatment response to >100,000 chemical compounds[119,120]. This analysis has revealed novel drugs potentially effective against melanoma and endometrial cancers[119].

Another important issue related to drug discovery is the identification of the mechanism of action of small-molecule compounds. One study introduced a molecular network-based approach that is able to elucidate genome-wide mechanism of action proteins based on the assessment of the global dysregulation of their molecular interactions following compound perturbation[121]. This innovative methodology revealed altretamine as an inhibitor of glutathione peroxidase 4 lipid repair activity, which was experimentally confirmed, validating its strong predictive ability. Similarly, a novel version of the VisANT (v4.0) network platform has the ability to perform an integrated search and navigation of disease and drug hierarchies, it can annotate gene–drug interactions using disease therapy information and it can also predict therapeutics for a given set of genes using enrichment analysis[122].

All these computational tools can identify novel chemical compounds targeting large molecular networks, helping us to reposition current FDA-approved therapeutics and identify the mechanism of actions of novel and old drugs. One of the first examples of the power of this strategy was the discovery that metformin, a drug for diabetes, has potent anticancer properties[123]. This finding was validated experimentally[124] and metformin is currently being used in phase II and phase III clinical trials for breast cancer[125]. The approach to discovering drugs that could benefit other unrelated conditions is called computational drug repositioning, which has also been accomplished in IBD[126]. Using publicly available molecular data on gene expression in IBD samples and small-molecule drugs, the anticonvulsant topiramate was found to be effective in improving colitis in mice, suggesting that topiramate might represent a therapeutic option for human IBD[126].

Together, these tools can identify the central hubs of the IBD interactome that can then be confirmed experimentally in cellular models. This systematic, unbiased and comprehensive strategy can yield an unprecedented wealth of information related to IBD pathobiology and open entirely new therapeutic opportunities for IBD patients (BOX 1).

Conclusions

The challenges imposed by complex diseases such as IBD cannot be solved by traditional investigational, clinical and therapeutic tools because each one of these tools can only examine few facets of a much bigger problem[127]. Strong evidence already shows that the understanding of complex diseases benefits from personal –omic profiling studied in a prospective longitudinal fashion, which enables defining molecular and clinical subtypes that become the targets of personalized medicine[128,129]. Jameson and Longo define personalized medicine as, "treatments targeted to the needs of individual patients on the basis of genetic, biomarker, phenotypic or psychosocial characteristics that distinguish a given patient from other patients with similar clinical presentations" (REF. 130). They also point out that, "the most daunting challenge for precision medicine is to manage the complexity associated with the progressively refined nosology of disease". Today, this challenge can be met using the systems biology and bioinformatics tools discussed in this Perspectives; identifying the molecular networks forming the IBD interactome in patients with Crohn's disease and ulcerative colitis[131] (FIG. 3; also see Supplementary information S2 (figure)). Disease-mediating interactomes will vary from patient to patient, posing the intimidating prospect of customizing therapy on a one-by-one basis. However, most patients will probably fall into discrete molecular categories that will benefit from similar therapies. The success of these therapies will fail if based on single-target approaches, no matter how sophisticated and precise the method is. For instance, genetic editing with unprecedented accuracy is feasible using CRISPR–Cas9 technology and one could envision replacing an IBD-associated variant with a wild-type one[132]. However, we already know that gene variants alone do not cause IBD and genome editing would have to be accompanied by complementary therapies. These therapies, however, will be undoubtedly complex and will require multiple and concomitant molecular targeting to achieve complete and permanent remission, if not a real cure[2].

Heitor S. P. de Souza is at the Department of Gastroenterology & Multidisciplinary Research Laboratory, Federal University of Rio de Janeiro, Rio de Janeiro 21941–913, Brazil.

Claudio Fiocchi is at the Department of Pathobiology, Lerner Research Institute, Department of Gastroenterology and Hepatology, Digestive Disease and Surgery Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA.

Dimitrios Iliopoulos is at the Center for Systems Biomedicine, Vatche and Tamar Manoukian Division of Digestive Diseases, Department of Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California 90095, USA.

Correspondence to C.F.
fiocchc@ccf.org

doi:10.1038/nrgastro.2017.110
Published online 23 Aug 2017

1. Busse, R., Blümel, M., Scheller-Kreinsen, D. & Zentner, A. *Tackling Chronic Diseases in Europe — Strategies, Interventions and Challenges.* (European Observatory on Health Systems and Polices & World Health Organization, 2010).
2. Iyengar, R. Complex diseases require complex therapies. *EMBO Rep.* **14**, 1039–1042 (2013).
This insightful commentary highlights the molecular complexity of chronic diseases and suggests that current therapeutic approaches may have reached their limits.
3. Sivakumaran, S. *et al.* Abundant pleiotropy in human complex diseases and traits. *Am. J. Hum. Genet.* **89**, 607–618 (2011).
4. Ghiassian, S. D. *et al.* Endophenotype network models: common core of complex diseases. *Sci. Rep.* **6**, 27414 (2016).
This article presents an alternative view emphasizing that different diseases often have common underlying mechanisms and shared endophenotypes. The authors construct endophenotype network models and explore their relation to different diseases, identifying critical modules within the subnetworks of the human interactome.
5. de Souza, H. S. & Fiocchi, C. Immunopathogenesis of IBD: current state of the art. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 13–27 (2016).
The authors present the current knowledge of the immunity and expand the concept of IBD immunopathogenesis to include various novel components that affect the intestinal inflammatory process.
6. Rogler, G. & Vavricka, S. Exposome in IBD: recent insights in environmental factors that influence the onset and course of IBD. *Inflamm. Bowel Dis.* **21**, 400–408 (2015).
7. Saidel-Odes, L. & Odes, S. Hygiene hypothesis in inflammatory bowel disease. *Ann. Gastroenterol.* **27**, 189–190 (2014).
8. Cleynen, I. & Vermeire, S. The genetic architecture of inflammatory bowel disease: past, present and future. *Curr. Opin. Gastroenterol.* **31**, 456–463 (2015).
9. McGovern, D. P., Kugathasan, S. & Cho, J. H. Genetics of inflammatory bowel diseases. *Gastroenterology* **149**, 1163–1176.e2 (2015).
10. Wang, M. H. *et al.* A novel approach to detect cumulative genetic effects and genetic interactions in Crohn's disease. *Inflamm. Bowel Dis.* **19**, 1799–1808 (2013).
11. Stappenbeck, T. S. *et al.* Crohn disease: a current perspective on genetics, autophagy and immunity. *Autophagy* **7**, 355–374 (2011).
12. Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **146**, 1489–1499 (2014).
13. Manichanh, C., Borruel, N., Casellas, F. & Guarner, F. The gut microbiota in IBD. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 599–608 (2012).
14. Choung, R. S. *et al.* Serologic microbial associated markers can predict Crohn's disease behaviour years before disease diagnosis. *Aliment. Pharmacol. Ther.* **43**, 1300–1310 (2016).
15. Darfeuille-Michaud, A. *et al.* High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology* **127**, 412–421 (2004).
16. Machiels, K. *et al.* A decrease of the butyrate-producing species *Roseburia hominis* and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. *Gut* **63**, 1275–1283 (2013).
17. Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789–799 (2014).
18. Braunstein, J., Qiao, L., Autschbach, F., Schurmann, G. & Meuer, S. C. T cells of the human intestinal lamina propria are high producers of interleukin 10. *Gut* **53**, 215–220 (1997).
19. Monteleone, G. *et al.* Blocking Smad7 restores TGF-β1 signaling in chronic inflammatory bowel disease. *J. Clin. Invest.* **108**, 601–609 (2001).
20. Neurath, M. F. Cytokines in inflammatory bowel disease. *Nat. Rev. Immunol.* **14**, 329–342 (2014).
21. Reinisch, W. *et al.* A dose escalating, placebo controlled, double blind, single dose and multidose, safety and tolerability study of fontolizumab, a humanised anti-interferon gamma antibody, in patients with moderate to severe Crohn's disease. *Gut* **55**, 1138–1144 (2006).
22. Hueber, W. *et al.* Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. *Gut* **61**, 1693–1700 (2012).
23. Danese, S. *et al.* Tralokinumab for moderate-to-severe UC: a randomised, double-blind, placebo-controlled, phase IIa study. *Gut* **64**, 243–249 (2015).
24. Spencer, D. M., Veldman, G. M., Banerjee, S., Willis, J. & Levine, A. D. Distinct inflammatory mechanisms mediate early versus late colitis in mice. *Gastroenterology* **122**, 94–105 (2002).
25. Kugathasan, S. *et al.* Mucosal T-cell immunoregulation varies in early and late Crohn's disease. *Gut* **56**, 1696–1705 (2007).
26. Wang, P., Han, W. & Ma, D. Electronic sorting of immune cell subpopulations based on highly plastic genes. *J. Immunol.* **197**, 665–673 (2016).
27. Berg, J. Genes-environment interplay. *Science* **354**, 15 (2016).
28. Fan, S., Hansen, M. E., Lo, Y. & Tishkoff, S. A. Going global by adapting local: A review of recent human adaptation. *Science* **354**, 54–59 (2016).
29. Miska, E. A. & Ferguson-Smith, A. C. Transgenerational inheritance: Models and mechanisms of non-DNA sequence-based inheritance. *Science* **354**, 59–63 (2016).
30. Yosef, N. & Regev, A. Writ large: Genomic dissection of the effect of cellular environment on immune response. *Science* **354**, 64–68 (2016).

31. Franks, P. W. & McCarthy, M. I. Exposing the exposures responsible for type 2 diabetes and obesity. *Science* **354**, 69–73 (2016).
32. Suez, J. *et al.* Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* **514**, 181–186 (2014).
33. Kreitinger, J. M., Beamer, C. A. & Shepherd, D. M. Environmental immunology: lessons learned from exposure to a select panel of immunotoxicants. *J. Immunol.* **196**, 3217–3225 (2016).
34. Carr, E. J. *et al.* The cellular composition of the human immune system is shaped by age and cohabitation. *Nat. Immunol.* **17**, 461–468 (2016).
Using a systems-level approach, the authors establish a resource of cellular immune profiles of healthy individuals.
35. Ye, B. D. & McGovern, D. P. Genetic variation in IBD: progress, clues to pathogenesis and possible clinical utility. *Expert Rev. Clin. Immunol.* **12**, 1091–1107 (2016).
36. Glocker, E. O. *et al.* Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N. Engl. J. Med.* **361**, 2033–2045 (2009).
37. Chakravarti, A. Genomics is not enough. *Science* **334**, 15 (2011).
38. Abreu, M. T. *et al.* Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. *Gastroenterology* **123**, 679–688 (2002).
39. Sahni, N. *et al.* Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* **161**, 647–660 (2015).
40. Chen, Y. *et al.* Variations in DNA elucidate molecular networks that cause disease. *Nature* **452**, 429–435 (2008).
41. Benfey, P. N. & Mitchell-Olds, T. From genotype to phenotype: systems biology meets natural variation. *Science* **320**, 495–497 (2008).
The authors propose that systems biology approaches can be applied to decipher gene's function and related abnormalities, generating genotype-phenotype maps.
42. Manolio, T. A. Genomewide association studies and assessment of the risk of disease. *N. Engl. J. Med.* **363**, 166–176 (2010).
43. Silvestri, G. A. *et al.* A bronchial genomic classifier for the diagnostic evaluation of lung cancer. *N. Engl. J. Med.* **373**, 243–251 (2015).
44. Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47–52 (2016).
45. Papaemmanuil, E. *et al.* Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).
46. Ellison, D. W. Multiple molecular data sets and the classification of adult diffuse gliomas. *N. Engl. J. Med.* **372**, 2555–2557 (2015).
47. Edwards, S. L., Beesley, J., French, J. D. & Dunning, A. M. Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.* **93**, 779–797 (2013).
48. [No authors listed.] Editorial: beyond the genome. *Nature* **518**, 273 (2015).
49. Kasowski, M. *et al.* Extensive variation in chromatin states across humans. *Science* **342**, 750–752 (2013).
50. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
In this study, investigators establish global maps of regulatory elements and define regulatory modules, revealing biologically relevant cell types for diverse human traits, and demonstrating the central role of epigenomic events for the development of disease.
51. Ventham, N. T., Kennedy, N. A., Nimmo, E. R. & Satsangi, J. Beyond gene discovery in inflammatory bowel disease: the emerging role of epigenetics. *Gastroenterology* **145**, 293–308 (2013).
52. Nimmo, E. R. *et al.* Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflamm. Bowel Dis.* **18**, 889–899 (2012).
53. Polytarchou, C. *et al.* MicroRNA214 is associated with progression of ulcerative colitis, and inhibition reduces development of colitis and colitis-associated cancer in mice. *Gastroenterology* **149**, 981–992.e11 (2015).
54. Liu, S. *et al.* The host shapes the gut microbiota via fecal microRNA. *Cell Host Microbe* **19**, 32–43 (2016).
55. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
56. Planell, N. *et al.* Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* **62**, 967–976 (2013).
57. Li, M. *et al.* Widespread RNA and DNA sequence differences in the human transcriptome. *Science* **333**, 53–58 (2011).
58. Wu, L. *et al.* Variation and genetic control of protein abundance in humans. *Nature* **499**, 79–82 (2013).
59. Starr, A. E. *et al.* Proteomic analysis of ascending colon biopsies from a paediatric inflammatory bowel disease inception cohort identifies protein biomarkers that differentiate Crohn's disease from UC. *Gut* http://dx.doi.org/10.1136/gutjnl-2015-310705 (2016).
60. Schicho, 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.* **11**, 3344–3357 (2012).
61. Patti, G. J., Yanes, O. & Siuzdak, G. Innovation: metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **13**, 263–269 (2012).
62. Lim, E. S. *et al.* Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat. Med.* **21**, 1228–1234 (2015).
63. Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).
64. Franzosa, E. A. *et al.* Identifying personal microbiomes using metagenomic codes. *Proc. Natl Acad. Sci. USA* **112**, E2930–E2938 (2015).
65. Chu, H. *et al.* Gene–microbiota interactions contribute to the pathogenesis of inflammatory bowel disease. *Science* **352**, 1116–1120 (2016).
66. Imhann, F. *et al.* Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut* http://dx.doi.org/10.1136/gutjnl-2016-312135 (2016).
67. Van Noorden, R. Global scientific output doubles every nine years. *Nature News Blog* http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html (2014).
68. Goldberger, A. L. Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *Lancet* **347**, 1312–1314 (1996).
69. West, B. J. A mathematics for medicine: the Network Effect. *Front. Physiol.* **5**, 456 (2014).
70. Sanchez, C. *et al.* Grasping at molecular interactions and genetic networks in *Drosophila melanogaster* using FlyNets, an Internet database. *Nucleic Acids Res.* **27**, 89–94 (1999).
71. Barabasi, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
72. Ideker, T., Galitski, T. & Hood, L. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* **2**, 343–372 (2001).
A very comprehensive review of the concept and application of systesm biology in biomedicine.
73. Polytarchou, C., Koukos, G. & Iliopoulos, D. Systems biology in inflammatory bowel diseases: ready for prime time. *Curr. Opin. Gastroenterol.* **30**, 339–346 (2014).
A systems biology approach, including novel computational methodologies capable of integrating high-throughput molecular data, is proposed for the study of pathogenic mechanisms in IBD.
74. Kabakchiev, B. & Silverberg, M. S. Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. *Gastroenterology* **144**, 1488–1496.e3 (2013).
75. Adams, A. T. *et al.* Two-stage genome-wide methylation profiling in childhood-onset Crohn's Disease implicates epigenetic alterations at the VMP1/MIR21 and HLA loci. *Inflamm. Bowel Dis.* **20**, 1784–1793 (2014).
76. Saito, S. *et al.* DNA methylation of colon mucosa in ulcerative colitis patients: correlation with inflammatory status. *Inflamm. Bowel Dis.* **17**, 1955–1965 (2011).
77. Foran, E. *et al.* Upregulation of DNA methyltransferase-mediated gene silencing, anchorage-independent growth, and migration of colon cancer cells by interleukin-6. *Mol. Cancer Res.* **8**, 471–481 (2010).
78. Jones, B. & Chen, J. Inhibition of IFN-gamma transcription by site-specific methylation during T helper cell development. *EMBO J.* **25**, 2443–2452 (2006).
79. Gu, H. *et al.* Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **6**, 468–481 (2011).
80. Chen, T. & Dent, S. Y. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat. Rev. Genet.* **15**, 93–106 (2014).
81. Zhang, Q. *et al.* Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature* **525**, 389–393 (2015).
82. Meddens, C. A. *et al.* Systematic analysis of chromatin interactions at disease associated loci links novel candidate genes to inflammatory bowel disease. *Genome Biol.* **17**, 247 (2016).
83. Ellinghaus, D. *et al.* Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat. Genet.* **48**, 510–518 (2016).
84. Koukos, G. *et al.* MicroRNA-124 regulates STAT3 expression and is down-regulated in colon tissues of pediatric patients with ulcerative colitis. *Gastroenterology* **145**, 842–852.e2 (2013).
85. Koukos, G. *et al.* A microRNA signature in pediatric ulcerative colitis: deregulation of the miR-4284/CXCL5 pathway in the intestinal epithelium. *Inflamm. Bowel Dis.* **21**, 996–1005 (2015).
86. Wu, F. *et al.* MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2 alpha. *Gastroenterology* **135**, 1624–1635.e24 (2008).
87. Davidovich, C. & Cech, T. R. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. *RNA* **21**, 2007–2022 (2015).
88. Lasda, E. & Parker, R. Circular RNAs: diversity of form and function. *RNA* **20**, 1829–1842 (2014).
89. M'Koma, A. E. *et al.* Proteomic profiling of mucosal and submucosal colonic tissues yields protein signatures that differentiate the inflammatory colitides. *Inflamm. Bowel Dis.* **17**, 875–883 (2011).
90. Meuwis, M. A. *et al.* Proteomics for prediction and characterization of response to infliximab in Crohn's disease: a pilot study. *Clin. Biochem.* **41**, 960–967 (2008).
91. Humphrey, S. J., Azimifar, S. B. & Mann, M. High-throughput phosphoproteomics reveals *in vivo* insulin signaling dynamics. *Nat. Biotechnol.* **33**, 990–995 (2015).
92. Le Gall, G. *et al.* Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. *J. Proteome Res.* **10**, 4208–4218 (2011).
93. Stephens, N. S. *et al.* Urinary NMR metabolomic profiles discriminate inflammatory bowel disease from healthy. *J. Crohns Colitis* **7**, e42–e48 (2013).
94. Lupp, C. *et al.* Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe* **2**, 119–129 (2007).
95. Ohkusa, T. *et al. Fusobacterium varium* localized in the colonic mucosa of patients with ulcerative colitis stimulates species-specific antibody. *J. Gastroenterol. Hepatol.* **17**, 849–853 (2002).
96. Ogura, Y. *et al.* A frameshift mutation in Nod2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
97. Gedela, S. Integration, warehousing, and analysis strategies of Omics data. *Methods Mol. Biol.* **719**, 399–414 (2011).
98. Shen, R., Olshen, A. B. & Ladanyi, M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics* **25**, 2906–2912 (2009).
99. Shen, R. *et al.* Integrative subtype discovery in glioblastoma using iCluster. *PLoS ONE* **7**, e35236 (2012).
100. Mo, Q. *et al.* Pattern discovery and cancer gene identification in integrated cancer genomic data. *Proc. Natl Acad. Sci. USA* **110**, 4245–4250 (2013).
101. Vaske, C. J. *et al.* Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* **26**, i237–245 (2010).
102. Paull, E. O. *et al.* Discovering causal pathways linking genomic events to transcriptional states using Tied Diffusion Through Interacting Events (TieDIE). *Bioinformatics* **29**, 2757–2764 (2013).
103. Wang, W. *et al.* iBAG: integrative Bayesian analysis of high-dimensional multiplatform genomics data. *Bioinformatics* **29**, 149–159 (2013).

104. Calvano, S. E. *et al.* A network-based analysis of systemic inflammation in humans. *Nature* **437**, 1032–1037 (2005).
105. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
106. Saito, R. *et al.* A travel guide to Cytoscape plugins. *Nat. Methods* **9**, 1069–1076 (2012).
107. Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449 (2005).
108. Vlasblom, J. *et al.* GenePro: a Cytoscape plug-in for advanced visualization and analysis of interaction networks. *Bioinformatics* **22**, 2178–2179 (2006).
109. Xia, T. & Dickerson, J. A. OmicsViz: Cytoscape plug-in for visualizing omics data across species. *Bioinformatics* **24**, 2557–2558 (2008).
110. Gao, J. *et al.* Metscape: a Cytoscape plug-in for visualizing and interpreting metabolomic data in the context of human metabolic networks. *Bioinformatics* **26**, 971–973 (2010).
111. Moni, M. A., Xu, H. & Lio, P. CytoCom: a Cytoscape app to visualize, query and analyse disease comorbidity networks. *Bioinformatics* **31**, 969–971 (2015).
112. Hu, Z., Mellor, J., Wu, J. & DeLisi, C. VisANT: an online visualization and analysis tool for biological interaction data. *BMC Bioinformatics* **5**, 17 (2004).
113. Nikitin, A., Egorov, S., Daraselia, N. & Mazo, I. Pathway studio — the analysis and navigation of molecular networks. *Bioinformatics* **19**, 2155–2157 (2003).
114. Iragne, F., Nikolski, M., Mathieu, B., Auber, D. & Sherman, D. ProViz: protein interaction visualization and exploration. *Bioinformatics* **21**, 272–274 (2005).
115. Agoston, V., Cemazar, M., Kajan, L. & Pongor, S. Graph-representation of oxidative folding pathways. *BMC Bioinformatics* **6**, 19 (2005).
116. Pujol, A., Mosca, R., Farres, J. & Aloy, P. Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol. Sci.* **31**, 115–123 (2010).
117. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
118. Kutalik, Z., Beckmann, J. S. & Bergmann, S. A modular approach for integrative analysis of large-scale gene-expression and drug-response data. *Nat. Biotechnol.* **26**, 531–539 (2008).
119. Ma, H. & Zhao, H. iFad: an integrative factor analysis model for drug-pathway association inference. *Bioinformatics* **28**, 1911–1918 (2012).
120. Shankavaram, U. T. *et al.* CellMiner: a relational database and query tool for the NCI-60 cancer cell lines. *BMC Genomics* **10**, 277 (2009).
121. Woo, J. H. *et al.* Elucidating compound mechanism of action by network perturbation analysis. *Cell* **162**, 441–451 (2015).
122. Hu, Z. *et al.* VisANT 4.0: integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Res.* **41**, W225–W231 (2013).
123. Hirsch, H. A., Iliopoulos, D., Tsichlis, P. N. & Struhl, K. Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission. *Cancer Res.* **69**, 7507–7511 (2009).
124. Iliopoulos, D., Hirsch, H. A. & Struhl, K. Metformin decreases the dose of chemotherapy for prolonging tumor remission in mouse xenografts involving multiple cancer cell types. *Cancer Res.* **71**, 3196–3201 (2011).
125. Niraula, S. *et al.* Metformin in early breast cancer: a prospective window of opportunity neoadjuvant study. *Breast Cancer Res. Treat.* **135**, 821–830 (2012).
126. Dudley, J. T. *et al.* Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci. Transl Med.* **3**, 96ra76 (2011).
127. Loscalzo, J., Barabasi, A. L. & Silverman, E. K. *Network Medicine: Complex Systems in Human Disease and Therapeutics.* (Harvard Univ. Press, 2017).
This book introduces and explains the concept of network medicine as the correct approach to tackle the molecular underpinnings of most complex diseases. Various chapters address molecular, translational and clinical aspects of network medicine in a fashion accessible to non-experts.
128. Li-Pook-Than, J. & Snyder, M. iPOP goes the world: integrated personalized Omics profiling and the road toward improved health care. *Chem. Biol.* **20**, 660–666 (2013).
129. Banchereau, R. *et al.* Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. *Cell* **165**, 551–565 (2016).
In this study, the molecular heterogeneity of systemic lupus erythematosus, a complex autoimmune disease, is revealed through a systems biology approach, offering an explanation for the failure of current clinical trials.
130. Jameson, J. L. & Longo, D. L. Precision medicine — personalized, problematic, and promising. *N. Engl. J. Med.* **372**, 2229–2234 (2015).
This article discusses the overall scope and potential applications of precision medicine, and analyses questions regarding its implementation in the practice of medicine.
131. Chan, A. C. & Behrens, T. W. Personalizing medicine for autoimmune and inflammatory diseases. *Nat. Immunol.* **14**, 106–109 (2013).
132. Cox, D. B., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat. Med.* **21**, 121–131 (2015).

Author contributions
All authors contributed equally to all aspects of this manuscript.

Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
NIH Roadmap Epigenomics Mapping Consortium: http://www.roadmapepigenomics.org/
Online Mendelian Inheritance in Man: https://www.omim.org/
GWASdb: http://jjwanglab.org/gwasdb
Human Metabolome Database: http://www.hmdb.ca/
ClinicalTrials.gov: https://clinicaltrials.gov/
KEGG: http://www.genome.jp/kegg/
TargetScan: http://www.targetscan.org/

SUPPLEMENTARY INFORMATION
See online article: S1 (figure) | S2 (figure)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF