cell development, or are more-complex patterns observed, such as multifurcations and continuous transitions? Mittnenzweig *et al*. suggest the answer to be 'all of the above'.

For example, the developmental trajectory of cells in the primitive streak bifurcates sharply, such that these cells become either mesodermal or endodermal cells (Fig. 1). By contrast, the differentiation of cells in the nascent mesoderm is inferred to be gradual and continuous, and with more than two destinations. The model also enables the inference of flows that change with time; for instance, before E7.1, epiblast cells overwhelmingly transition to acquiring primitive-streak fates, but shortly after that point, they mostly transition to acquiring ectodermal fates.

Finally, what are the molecular factors that underlie differentiation, and do individual factors act alone or in combination? The authors claim that, with the exception of some lineages (notably, the node, cardiomyocyte and haemato-endothelial lineages), the landscape of gastrulation is predominantly characterized by a dependence on overlapping combinations of factors, as well as on a gradual unfolding of commitment. For example, although cells of the nascent mesoderm progress into a spectrum of fates, these fates are not sharply separated from one another, and there is no clear delineation between the sets of transcription factors that seem to specify each fate. The authors propose that, rather than a series of specific factors governing a stepwise, hierarchical progression of specification, combinations of molecular factors regulate diverse mesodermal fates in a 'fuzzy' and almost probabilistic manner. To highlight the delicacy of this program, the authors carried out experiments in which inferred key regulators were genetically disrupted, which led to delayed differentiation of affected lineages.

Of course, all models have limitations, and this model has its own. First, its resolution is limited by the underlying data, although simply processing more embryos would address this. Second, its metacells and flows are inferred solely from the similarity of the transcriptional profiles of the cells, and so there is a risk of missing or misinterpreting certain bona fide relationships[5]. Particularly rapid changes in the nematode worm *Caenorhabditis elegans,* for instance, elude efforts to reconstruct lineages in 'pseudotime' – that is, ordering cells by their developmental stage rather than their age in real time[6]. Third, the model ignores cells' spatial coordinates within embryos as well as their actual lineage relationships, two crucial aspects of development that are increasingly amenable to measuring and recording, respectively[7,8].

Notwithstanding these limitations, the model of mouse gastrulation developed by Mittnenzweig *et al*. is impressive, and shows how continuous maps of complex differentiation landscapes might be recovered despite discrete sampling. Together with other work published in the past few years[3,6,9], it represents a substantial step forward on the path to a complete understanding of cells' inner lives during this most important of times in an animal's life[10].

Chengxiang Qiu and **Jay Shendure** are in the Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA. **J.S.** is also at the Brotman Baty Institute for Precision Medicine, the Allen Discovery Center for Cell Lineage Tracing, and the Howard Hughes Medical Institute, Seattle.
e-mail: shendure@uw.edu

1. Mittnenzweig, M. *et al. Cell* https://dx.doi.org/10.1016/j.cell.2021.04.004 (2021).
2. McDole, K. M. *et al. Cell* **175**, 859–876 (2018).
3. Pijuan-Sala, B. *et al. Nature* **566**, 490–495 (2019).
4. Cao, J. *et al. Nature* **566**, 496–502 (2019).
5. Wagner, D. E. & Klein, A. M. *Nature Rev. Genet.* **21**, 410–427 (2020).
6. Packer, J. S. *et al. Science* **365**, eaax1971 (2019).
7. McKenna, A. *et al. Science* **353**, aaf7907 (2016).
8. Lohoff, T. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.11.20.391896 (2020).
9. Argelaguet, R. *et al. Nature* **576**, 487–491 (2019).
10. Vicente, C. *Development* **142**, 2547–2548 (2015).

This article was published online on 30 April 2021.

Medical research

A molecular connection to Crohn's disease risk

Scott Plevy

Mutations of the *NOD2* gene are risk factors for Crohn's disease. Many aspects of how they contribute to the condition are unknown. The discovery of cell populations that are involved suggests new therapeutic options. **See p.275**

Crohn's disease, a chronic inflammatory bowel disease, affects many people. For example, more than 0.3% of the populations of Canada and Germany have the condition, and its incidence is increasing worldwide[1]. Better therapies are needed, but progress in treating Crohn's disease has been hampered by the lack of understanding of how it arises. On page 275, Nayar *et al*.[2] shed light on a long-standing mystery about one risk factor for Crohn's disease, and their findings have important clinical implications.

Crohn's disease can affect any part of the gut. Most commonly, it affects the ileum region, causing inflammation that frequently results in fibrosis (the deposition of fibrous connective tissue as an injury response). This leads to the narrowing (or stricture) of the lumen of the ileum, which often requires surgical intervention[3]. Crohn's disease provides a useful model of illnesses that are mediated by genes and environmental interactions. In this case, genetic susceptibility underpins the disease-causing inflammatory responses to gut microorganisms.

Genetic variations, called polymorphisms, of the *NOD2* gene are the strongest genetic risk association for Crohn's disease; approximately 20% of all such risk of developing the disease is related to three single nucleotide polymorphisms of this gene[4]. Furthermore, *NOD2* mutations are strong predictive factors for the development of ileum strictures and for the need for surgery in Crohn's disease, which is a widely validated association between the genetic underpinnings of this condition and manifestations of the disease[3].

However, connecting the *NOD2* gene to disease susceptibility presented a paradox. NOD2 is an intracellular receptor (Fig. 1) that recognizes the molecule muramyl dipeptide (MDP) – a ubiquitous component of bacterial cell walls. Before *NOD2* was described as a risk gene for Crohn's disease, NOD2 function was best understood in immune cells that aid the innate branch of immune defences. NOD2 activation in these cells leads to the expression of inflammatory molecules called cytokines, and an abnormally intense inflammatory response can mediate intestinal damage in Crohn's disease[5,6]. One might therefore have expected that *NOD2* mutations known as loss-of-function mutations, which do not generate a fully functional version of the encoded protein, would protect against Crohn's disease. Yet such loss-of-function mutations of *NOD2* were identified as risk factors for the disease. Subsequent research therefore pivoted to focus on a different aspect of NOD2 biology in the intestine, investigating how functional NOD2 maintains homeostasis in the intestine, where the body's largest biomass of immunologically active cells is constantly exposed to MDP from gut microbes, and how *NOD2* mutations perturb this balance and lead to disease[5].

The role of *NOD2* mutations in the emergence

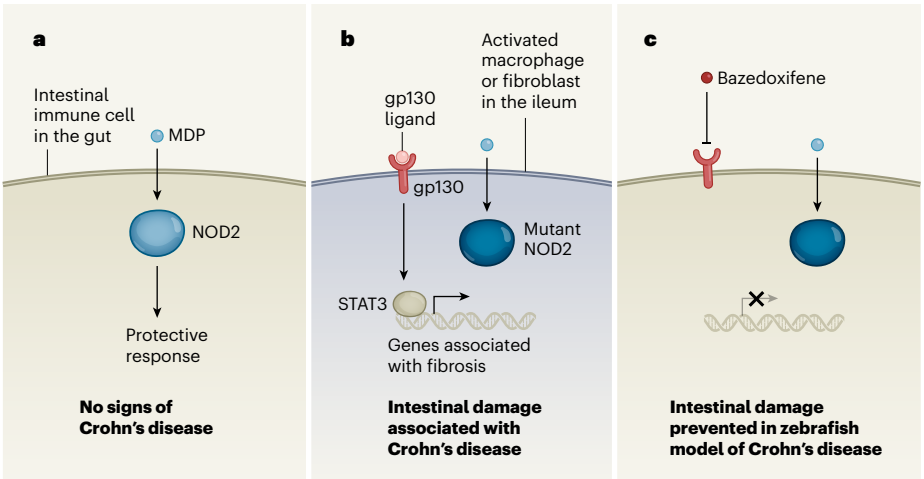

Figure 1 | Molecular underpinnings of an inflammatory bowel disease. The causes of Crohn's disease are not fully understood. **a**, NOD2 is an intracellular protein that can recognize the bacterial molecule muramyl dipeptide (MDP), which is commonly found in the gut. This is part of a normal protective response by immune cells to maintain gut homeostasis. **b**, Mutations that generate a non-functional version of NOD2 are a risk factor for Crohn's disease[4]. Studying zebrafish and clinical samples, Nayar *et al.*[2] reveal details of processes underlying the disease that are associated with mutant NOD2. The authors identify populations of activated immune cells called macrophages and fibroblasts as being responsible for fibrosis, a tissue abnormality in the ileum region of the bowel that occurs in Crohn's disease. Ligand binding to the gp130 receptor of these immune cells triggers a gene-expression pathway mediated by the STAT3 transcription factor. **c**, The authors report that a gp130 inhibitor molecule called bazedoxifene prevented gene expression that leads to damage in a zebrafish model of Crohn's disease.

of fibrosis of the ileum was unknown before the present study. The authors sought to understand what drives inflammation and fibrosis in Crohn's disease, and linked these biological insights to NOD2 through research using human cells, human intestinal tissue and a zebrafish model.

First, the authors used single-cell sequencing of RNA from the inflamed tissue of ileum samples removed during surgery from people with Crohn's disease. These cells revealed a gene-expression signature associated with activated macrophage and fibroblast cells. The authors also identified a key cell type that expresses markers of both myeloid and fibroblast cellular lineages. These discoveries suggest that a population of inflammatory macrophages in the ileum differentiates to become activated fibroblasts during the course of disease.

Strikingly, the authors demonstrate the evolutionary conservation of these cellular populations in an experimental model of intestinal inflammation — zebrafish treated with the molecule dextran sodium sulfate (DSS). This molecule has long been used to induce intestinal damage and inflammation in a standard rodent model. The *in vivo* modelling of human inflammatory bowel diseases has been dominated by mouse models. However, as Nayar and colleagues demonstrate, zebrafish offer a useful alternative for relatively high-throughput investigations and rapid assessment of correlations with human disease. Indeed, zebrafish and mammalian intestines have a similar form (morphology). Moreover, like humans, zebrafish have innate and adaptive branches of their immune-defence responses, and intestinal inflammation of zebrafish is also dependent on the community of gut microorganisms[7]. Gene-editing tools, such as CRISPR, aid the rapid modification of genes of interest in zebrafish.

The authors studied intestinal inflammation in zebrafish engineered to have *nod2* deficiency. These fish, treated with DSS, had increased numbers of leukocyte immune cells in their intestines, a hallmark of inflammation, compared with zebrafish with normal *nod2*. But the zebrafish model is relevant only if a human correlation can be established. Accordingly, using data from children newly diagnosed with Crohn's disease, the authors show that an increase in the number of copies of a *NOD2* mutation (associated with the risk of Crohn's disease) indeed correlated with an activated macrophage and fibroblast gene-expression signature in ileum tissue.

To understand NOD2 function in human cells that can differentiate *in vitro*, the authors used peripheral blood monocytes from healthy volunteers, and determined whether the cells had one, two or no copies of *NOD2* mutations linked to susceptibility to Crohn's disease. The cells were then differentiated *in vitro* with and without MDP. The authors observed a higher number of activated fibroblasts for cells with two copies of *NOD2* mutations compared with cells with wild-type *NOD2*. Furthermore, an increase in the number of *NOD2* mutations was associated with a corresponding enrichment in the number of fibroblasts with a gene-expression signature characteristic of activated cells. Interestingly, zebrafish with *nod2* deficiencies, which were given MDP, had a gene-expression signature characteristic of activated fibroblasts that persisted even during recovery from injury mediated by DSS, compared with zebrafish that have wild-type *nod2*. These data suggest that *nod2* deficits inhibit efficient recovery (resolution) from fibrosis and inflammation.

To further elucidate the molecular basis of the fibrosis-linked gene-expression signature associated with *NOD2* risk mutations, the authors searched for upstream transcriptional regulators of this pathway. They identified the gene encoding STAT3 as being markedly upregulated in activated fibroblasts and macrophages. STAT3 is a transcriptional regulator of key components of inflammatory and fibrotic responses in inflammatory bowel diseases, and acts through the cytokine receptor gp130. Analyses of clinical data revealed upregulated expression of gp130-regulated genes encoding the proteins IL-6, oncostatin M and IL-11 in people with Crohn's disease who did not respond to therapy targeting the tumour-necrosis factor (TNF) protein (anti-TNF antibodies are a common treatment for Crohn's disease). The discovery supports a role for gp130 signalling in this group of therapy-resistant individuals.

The authors hypothesized that gp130 blockade might lessen the abnormalities that occur with *NOD2* mutation. They tested this idea by using bazedoxifene, a gp130 inhibitor, on MDP-treated human cells with *NOD2* mutations. Bazedoxifene indeed lessened the fibrosis-associated gene-expression signature and reversed the cellular shape changes that are characteristic of activated fibroblasts. This drug also reduced the intestinal damage found in *nod2* mutant zebrafish treated with DSS.

Starting with the clinical characteristics of fibrosis in Crohn's disease, this work describes a molecular pathway linked to *NOD2* mutations associated with the disease, and concludes with a potential therapeutic insight to address the pressing clinical problems of fibrosis and anti-TNF drug resistance. By underpinning the genetics and the clinical outcomes to this cellular and molecular pathway, the study provides a road map to understanding present and future therapeutic approaches.

Many interesting avenues of investigation remain. NOD2 is the only described recognition pathway for MDP, yet this paper demonstrates MDP-induced cellular and molecular changes in the absence of nod2 in zebrafish. This implies that there are MDP signalling pathways that have not yet been described. Bazedoxifene was initially characterized as a selective inhibitor of the oestrogen receptor[8], raising the concern that the drug might have adverse effects on other signalling pathways if used as a therapeutic for Crohn's disease. The gp130 receptor has

multiple ligand binding partners that influence a broad range of immune responses. Hence, understanding the specific gp130 ligands that orchestrate NOD2-mediated molecular events could lead to more selective, effective and safer therapeutic interventions than would globally inhibiting gp130 signalling, or targeting other clinically relevant signalling pathways, such as the Janus kinase enzymes (inhibitors for which are in late-stage clinical development for Crohn's disease).

Not everyone with Crohn's disease has *NOD2* mutations associated with disease risk. Indeed, in individuals of certain ethnic groups, such as people of Chinese, Malay or Indian heritage, disease of the ileum is a prominent clinical feature of Crohn's disease, yet *NOD2* is not associated with disease risk in this population[5,9]. Perhaps the molecular signature of activated macrophages and fibroblasts is the relevant unifying signature for individuals with Crohn's disease of the ileum. It is probable that different genetic landscapes might result in the same clinical and molecular outcomes. Hence, Nayar and colleagues' results move the field a step closer to a molecular classification of Crohn's disease that might clarify a complex condition that has approximately 200 genetic regions associated with disease risk[10], and diverse clinical manifestations.

Scott Plevy is at Senda Biosciences, Cambridge, Massachusetts 02140, USA.
e-mail: splevy@sendabiosciences.com

1. Ng, S. C. *et al. Lancet* **390**, 2769–2778 (2017).
2. Nayar, S. *et al. Nature* **593**, 275–281 (2021).
3. Rieder, F., Zimmermann, E. M., Remzi, F. H. & Sandborn, W. J. *Gut* **62**, 1072–1084 (2013).
4. Cuthbert, A. P. *et al. Gastroenterology* **122**, 867–874 (2002).
5. Sidiq, T., Yoshihama, S., Downs, I. & Kobayashi, K. S. *Front. Immunol.* **7**, 367 (2016).
6. Franchi, L., Warner, N., Viani, K. & Nuñez, G. *Immunol. Rev.* **227**, 106–128 (2009).
7. Brugman, S. *Dev. Comp. Immunol.* **64**, 82–92 (2016).
8. Duggan, S. T. & McKeage, K. *Drugs* **71**, 2193–2212 (2011).
9. Hilmi, I., Tan, Y. M. & Goh, K. L. *World J. Gastroenterol.* **12**, 1435–1438 (2006).
10. Mondal, K. & Kugathasan, S. *Nature Rev. Gastroenterol. Hepatol.* **14**, 266–268 (2017).

This article was published online on 19 April 2021.

Synthesis

Unprecedented reactions for molecular editing

William P. Unsworth & Alyssa-Jennifer Avestro

Many scientific fields and industries rely on the synthesis of small organic molecules. A chemical reagent has been developed that allows such molecules to be made by 'deleting' nitrogen atoms from readily accessible precursors. **See p.223**

On page 223, Kennedy *et al.*[1] report a strategy for molecular editing in which nitrogen atoms are 'deleted' from organic molecules. The idea of deleting, rather than adding, atoms to molecules runs counter to the way chemists usually think about making organic molecules (with a few notable exceptions; see ref. 2, for example) . But the authors' reactions could dramatically change the way in which such synthesis is planned.

Chemists attach great pride to the idea that, given sufficient time and resources, they can synthesize almost any small organic molecule. Such efforts are the basis of many technologies that have enormous societal value, such as medicines, polymers and agrochemicals. To make the range of molecules that is needed for these applications, chemists are armed with an array of methods that promote specific chemical changes, often with exquisite selectivity.

Moreover, countless chemical-synthesis methods are discovered and published daily. Most involve relatively small, practical changes to existing methods, or modest advances in the scope of known reaction types. These advances are important – incremental improvements are crucial to scientific progress. Nonetheless, methods occasionally emerge that have more far-reaching implications. Kennedy and colleagues' chemistry is one such example. To explain why, let's consider the way in which chemical syntheses are usually conceived, using a process known as retrosynthetic analysis[3,4].

In retrosynthetic analysis, the chemist starts by considering the chemical structure of the target molecule, and then works backwards by mentally 'disconnecting' individual bonds in the target molecule – the idea being to break it down into smaller and simpler chemical fragments. A synthetic route is then devised by working out a series of reactions that leads from the fragments back to the target, in the reverse sequence. Typically, there are multiple possible ways to disconnect any given target molecule, but a key consideration is that each step in the forward chemical synthesis must be a known type of chemistry, or a reaction that can be developed. Chemists therefore typically rely on tried-and-tested disconnections for common molecular motifs, because this usually ensures that the forward synthesis is productive.

Knowledge of which bonds can (or cannot) be disconnected using established chemistry, and the ability to apply this knowledge systematically, is crucial. But there is also a large creative aspect to synthesis; indeed, many of the best syntheses are said to be on the borderline between science and art[5]. Proposing a disconnection for which no synthetic methods exist for the equivalent forward reaction requires inspiration and creativity, and subsequent development of the requisite methods is hard. But chemists will always be drawn to such challenges[6–9], because they open up strategies for synthesis that would previously have been considered impossible.

This is precisely what Kennedy *et al.* have achieved. They report a reaction that enables challenging molecular targets to be made by excising single nitrogen atoms from easily accessible starting materials (Fig. 1). The authors developed a new, easily prepared chemical reagent to promote the reaction, the mechanism of which involves an unprecedented molecular rearrangement: a molecule of nitrogen is lost from a reaction intermediate, producing two highly reactive free radicals that combine to form a new carbon–carbon (C–C) bond (see Fig. 1b of the paper[1]).

The Oxford English Dictionary defines 'synthesis' as the "combination of components or elements to form a connected whole" – so isn't deleting atoms, rather than adding them, counterproductive to this goal? The value of Kennedy and colleagues' strategy lies in the fact that the nitrogen-containing starting materials are typically much easier to make, or to source commercially, than are the analogous molecules that don't contain nitrogen. Chemists can therefore simplify their syntheses by making intermediates that contain a nitrogen atom, and then removing it later. This is similar to the way in which scaffolding aids in the construction of a skyscraper, but is removed once the main structure has been built. Notably, the removal of nitrogen fundamentally alters the molecular skeleton of the molecule, because an internal atom is lost[10–13]; this contrasts with most other molecular-editing strategies, which focus on making less drastic changes on the molecule's periphery.

A practical advantage to Kennedy and colleagues' synthetic strategy is that it mitigates the costs and safety problems associated with many established C–C bond-forming methods, which usually require expensive or toxic metal reagents. The authors also demonstrate that their chemistry can delete nitrogen from commercially available drugs and natural