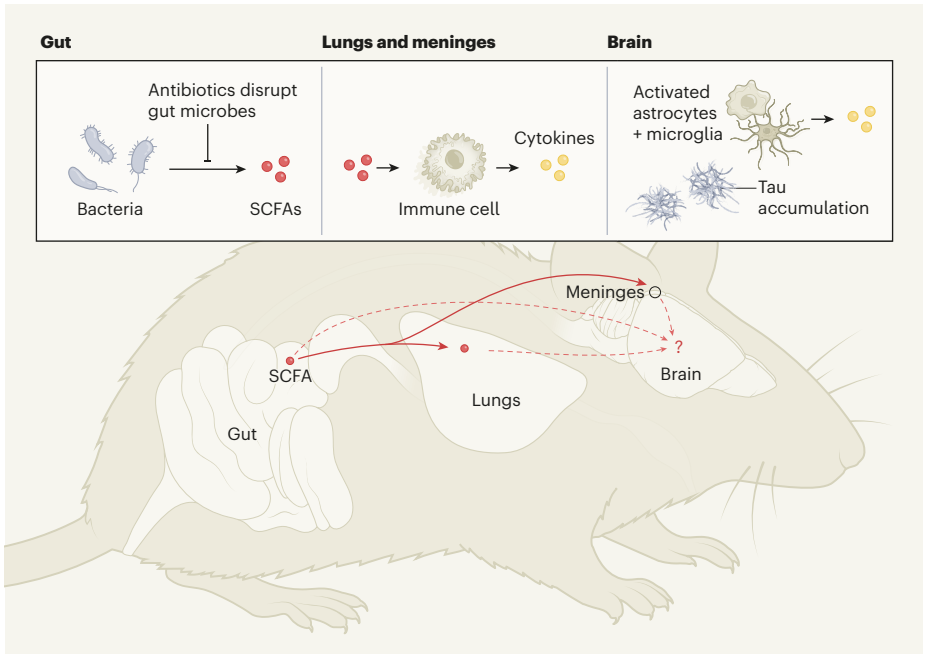

Figure 1 | Messages from the gut influence neurodegeneration. The gene *APOE* exists in four variants, of which *APOE3* and *APOE4* confer a higher-than-average risk of a person developing late-onset Alzheimer's disease (*APOE* variants not shown). Seo *et al.*[2] showed that, in mice carrying *APOE3* or *APOE4*, bacteria in the gut release short-chain fatty acids (SCFAs), which can signal to immune cells in the lungs and in tissues called the meninges that border the brain. These cells release cytokines that might signal to the brain (broken arrows), where immune cells called astrocytes and microglia become activated and release more cytokines, and tau protein accumulates (a key hallmark of Alzheimer's disease). It is possible that SCFAs also signal directly to the brain, although this was not tested. The authors show that treating mice with antibiotics to disrupt the gut microbes blocks this pathway and reduces neurodegeneration; this treatment is most effective in mice carrying *APOE3*.

terms of orchestrating this process.

It is not entirely surprising that raising mice germ-free or treating them with antibiotics can reduce tau aggregation, because another research group has shown that gut microbes are required for another type of aggregate, mutant α-synuclein, to have harmful effects[12]. But the finding that these effects involve *APOE*-dependent changes in the peripheral immune system is remarkable. *APOE* status is typically thought of as affecting mainly the brain, but Seo and colleagues' evidence for a mechanistic link between this genetic risk factor and the gut microbiota suggests that the gene might be a target of environmental triggers that promote neurodegeneration[11]. However, in what tissue and cell type *APOE* might be affected by these triggers remains to be seen.

It will be interesting to further investigate the sex differences that the authors observed in their experiments involving antibiotics. In a human study, higher SCFA levels in the stools of men who had Parkinson's disease correlated with later onset of motor symptoms[13], the implication being that SCFAs might be protective (although clinical trials would be needed to establish this link). Researchers should also aim to better understand the effect of the gut microbiota on brain-border immune cells, especially in light of work demonstrating that gut immune cells called B cells make immunoglobulin proteins that affect brain physiology[14].

Turning to humans, is it possible that antibiotic treatments to combat chronic infections might influence the course of neurodegenerative diseases? If so, would this be the case only in men, or for everyone? The answers to these questions will require clinical studies. Clinical trials could also begin to address whether manipulating the gut microenvironment — using pre- or probiotics, SCFA supplementation or faecal microbiota transplants — has the potential to protect some subsets of people from neurodegeneration.

Alfonso Martín-Peña and **Malú Gámez Tansey** are in the Department of Neuroscience, College of Medicine, Center for Translational Research for Neurodegenerative Disease, University of Florida, Gainesville, Florida 32610, USA.
e-mail: mgtansey@ufl.edu

1. Prinz, M., Jung, S. & Priller, J. *Cell* **179**, 292–311 (2019).
2. Seo, D.-O. *et al. Science* **379**, eadd1236 (2023).
3. Yamazaki, Y., Zhao, N., Caulfield, T. R., Liu, C.-C. & Bu, G. *Nature Rev. Neurol.* **15**, 501–518 (2019).
4. Cryan, J. F. *et al. Physiol. Rev.* **99**, 1877–2013 (2019).
5. MahmoudianDehkordi, S. *et al. Alzheimers Dement.* **15**, 76–92 (2019).
6. Dodiya, H. B. *et al. J. Exp. Med.* **216**, 1542–1560 (2019).
7. Lanfranco, M. F., Sepulveda, J., Kopetsky, G. & Rebeck, G. W. *Glia* **69**, 1478–1493 (2021).
8. TCW, J. *et al. Cell* **185**, 2213–2233 (2022).
9. Erny, D. *et al. Nature Neurosci.* **18**, 965–977 (2015).
10. Spichak, S. *et al. Brain Behav. Immun. Health* **16**, 100318 (2021).
11. Bettcher, B. M., Tansey, M. G., Dorothée, G. & Heneka, M. T. *Nature Rev. Neurol.* **17**, 689–701 (2021).
12. Sampson, T. R. *et al. Cell* **167**, 1469–1480 (2016).
13. Aho, V. T. E. *et al. Mol. Neurodegen.* **16**, 6 (2021).
14. Fitzpatrick, Z. *et al. Nature* **587**, 472–476 (2020).

The authors declare no competing interests.
This article was published online on 8 February 2023.

Developmental biology

An oracle predicts regulators of cell identity

Jeffrey A. Farrell

A computational tool called CellOracle can predict how networks of genes interact to program cell identity during embryonic development. The tool should help to hone efforts to understand how development is regulated. **See p.742**

As an animal develops, each of its thousands or even trillions of cells must be programmed to adopt one of many possible cell identities. This programming is controlled by a group of proteins and the genes that encode them, which are collectively known as developmental regulators. Kamimoto *et al.*[1] present a computational approach on page 742 to predict the shifts in cell identity that will occur if levels of individual developmental regulators are altered. Their systematic approach promises to help researchers determine the regulators most worthy of further study, saving valuable resources and suggesting biological phenomena that might previously have been missed.

The key process underlying development is differential expression of the genome. Although present in all cells, many genes are

expressed in only one or a few cell types, where they confer identity or specific functionality. This differential expression is driven mostly by transcription factors, which bind to specific DNA sequences to promote or suppress expression of target genes. Many of these target genes themselves encode transcription factors, forming a complex, interacting regulatory network.

Decades of work have provided insights into the gene regulatory networks that control development, mainly through painstaking experimental alteration of the expression of single genes[2,3]. Identifying new facets of this network would improve our ability to understand disease-causing mutations and to develop approaches to regenerative medicine involving cells of particular identities. With this aim in mind, Kamimoto and colleagues developed a computational tool called CellOracle that predicts potential regulatory interactions and the effects of disrupting regulators.

The authors' approach requires wild-type gene-expression data from thousands of single cells. These data can be optionally supplemented with information about the 'accessibility' of genomic regions, which can indicate whether a region is available for transcription factors to bind to it. CellOracle first uses these data to determine which transcription factors might regulate each gene's expression, by looking for transcription factors' preferred binding sequences in accessible DNA associated with a gene. The tool then prunes these sets of potential regulatory relationships to select ones that are active in each cell type, on the basis of the correlated expression of transcription factors and the genes that they might regulate.

A user of CellOracle can then investigate the effects of perturbations in the gene regulatory network, by shifting expression of a chosen transcription factor in the model. The expression of that factor's identified target genes shifts accordingly, and the process is then repeated (because target genes often encode transcription factors that target other genes). CellOracle compares the gene-regulatory network in its final, shifted state to states that are present in the original data, to predict the consequences of the change. For example, progress through a developmental transition might be abnormally fast, slow, or blocked, or an unusual change in cell identity might occur.

CellOracle joins a cast of computational tools for regulatory network inference[4]. Some of these require gene-expression data from cells carrying genetic mutations, but CellOracle needs only wild-type data, reducing the experimental effort required to get started. Furthermore, the purpose of many of these tools is mainly to identify a gene regulatory network, but CellOracle focuses on predicting the outcome of disrupting regulators in that network.

The authors first confirmed that CellOracle works by using published single-cell genomic data sets that describe the formation of blood and immune cells in mice and humans (a developmental setting in which many key regulators are already known). Most of the genes that CellOracle predicted as potential regulators of blood or immune identity have functions that have been reported previously, and that align with the tool's predictions.

The authors next applied CellOracle systematically to all embryonic transcription factors in zebrafish (*Danio rerio*), again using published data. The aim was to predict regulators of all cell-identity decisions in this species' development. The team then focused on the role of genetic regulators in forming a tissue called the axial mesoderm, which runs along the trunk of the embryo. In this tissue, an axial mesoderm progenitor cell type gives rise to a pair of descendant cell types — notochord and prechordal plate cells (Fig. 1a). These two cell types together send signals that pattern other tissues (such as the spinal cord and the brain) and provide a signalling barrier that divides the embryo into left and right halves.

"The value of this method is to sharpen hypotheses and prioritize or direct future experiments."

Genetic screens have identified several regulators that confer notochord identity, including the gene *noto*, deletion of which transforms notochord progenitors into muscle progenitors[5]. CellOracle predicted that loss of *noto* would also shift axial mesoderm progenitors towards a prechordal plate identity (Fig. 1b), and Kamimoto *et al.* confirmed this experimentally. The tool also predicted that loss of the gene *lhx1a* would inhibit early axial mesoderm differentiation (Fig. 1c). And indeed, the researchers found that zebrafish embryos lacking *lhx1a* had fewer mature derivatives of axial mesoderm, and expressed genes that were associated more strongly with early, immature states.

CellOracle is not without limitations. For instance, it did not predict some known effects of regulator disruption (such as that loss of the gene *pu.1* leads to depletion of red blood cell progenitors[6,7]), and did not identify some known regulators (such as the gene *foxa3* in the axial mesoderm[8,9]). Moreover, the tool predicts only shifts to a different 'normal' cell identity; it cannot predict whether perturbations would create a scrambled or non-normal cell state. Lastly, it cannot currently model more-complex outcomes, as might occur from mutations that do not solely increase or decrease the expression of transcription factors (such alterations include 'missense' mutations, which substitute one amino acid of a protein for another).

However, it would be unhelpful to demand perfect accuracy from an approach such as this. As stressed by Kamimoto and colleagues, the value of this method is to sharpen hypotheses and prioritize or direct future experiments. Screens in which random mutations are introduced to the genome have identified many developmental regulators (and have taught us fundamentally how developmental regulators work), but some regulators are hidden from those approaches — such as when two regulators must both be lost before observable changes can occur in development.

Furthermore, observing these changes requires looking in the right way, in the right place, at the right time. CellOracle and related computational approaches (including

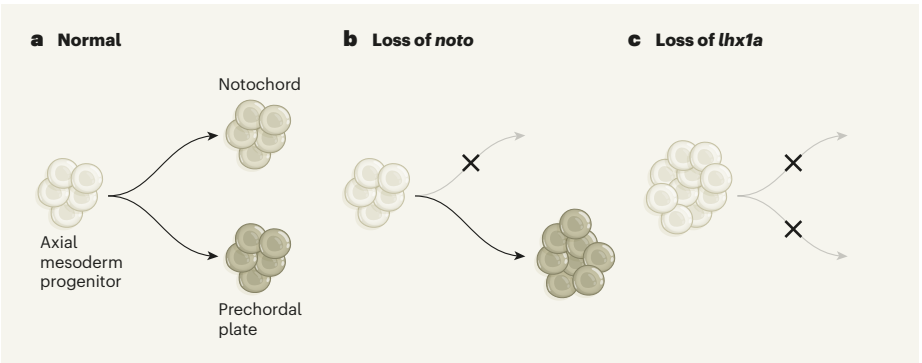

Figure 1 | Predicting developmental regulators of axial mesoderm tissue. a, In zebrafish embryos, cells called axial mesoderm progenitors give rise to two types of descendant — notochord and prechordal plate cells. Kamimoto *et al.*[1] used their newly developed computational tool, CellOracle, to predict how disruption of particular genes in axial mesoderm progenitors would alter the identities of their descendants. **b**, CellOracle accurately predicted that loss of the gene *noto* would prevent cells from becoming notochord (a change that has previously been demonstrated experimentally[5]), and identified, for the first time, that some cells would be funnelled into prechordal plate identities. **c**, The tool also correctly predicted that loss of the gene *lhx1a* would prevent axial mesoderm progenitor cells from maturing.

SCENIC+)[10] can suggest regulators and trait changes of interest that might have been missed, from data that are often available already. Moreover, testing the tool's predictions is easier than ever, thanks to gene-editing tools such as CRISPR–Cas9 and ever-cheaper single-cell genomic assays.

Even in the well-studied axial mesoderm, CellOracle identified previously unknown regulators and predicted new roles for previously identified regulators. The authors' predictions can be browsed online, and their tool is freely available (www.celloracle.org). It has already been used by the authors to identify regulators of cell reprogramming in a tissue called the endoderm[11]. Other groups have used it to predict regulators of immune-cell identity[12]; to further explore known regulators of the formation of cell types in the thymus[13], immune system[14], cartilage and bone[15]; and to study progenitors of an embryonic tissue called the neuromesoderm[16].

My own prediction is that approaches such as CellOracle will hasten our understanding of the regulatory networks that determine cell identity. Let us hope that, in doing so, they will accelerate development of medical interventions that manipulate these networks. But future users beware: just as, in Greek mythology, Apollo had to slay Python to establish his oracle at Delphi, you, too, will have to conquer Python (in this case, the scripting language) before you can use CellOracle.

Jeffrey A. Farrell is in the Division of Developmental Biology, NICHD Intramural Research Program, Bethesda, Maryland 20814, USA.
e-mail: jeffrey.farrell@nih.gov

1. Kamimoto, K. *et al. Nature* **614**, 742–751 (2023).
2. Davidson, E. H. *et al. Science* **296**, 1669–1678 (2002).
3. Simões-Costa, M. & Bronner, M. E. *Development* **142**, 242–257 (2015).
4. Nguyen, H., Tran, D., Tran, B., Pehlivan, B. & Nguyen, T. *Brief. Bioinform.* **22**, bbaa190 (2021).
5. Halpern, M. E. *et al. Development* **121**, 4257–4264 (1995).
6. Wontakal, S. N. *et al. PLoS Genet.* **7**, e1001392 (2011).
7. Moreau-Gachelin, F. M., Tavitian, A. & Tambourin, P. *Nature* **331**, 277–280 (1988).
8. Dal-Pra, S., Thisse, C. & Thisse, B. *Dev. Biol.* **350**, 484–495 (2011).
9. Seiliez, I., Thisse, B. & Thisse, C. *Dev. Biol.* **290**, 152–163 (2006).
10. Bravo, C. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2022.08.19.504505 (2022).
11. Kamimoto, K. *et al. Stem Cell Rep.* **18**, 97–112 (2023).
12. Chopp, L. B. *et al. Immunity* **53**, 1182–1201 (2020).
13. Magaletta, M. E. *et al. Nature Commun.* **13**, 457 (2022).
14. Ravi, V. M. *et al. Nature Commun.* **13**, 925 (2022).
15. Herpelinck, T. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2022.03.14.484345 (2022).
16. Argelaguet, R. *et al.* Preprint at biorxiv https://doi.org/10.1101/2022.06.15.496239 (2022).

The author declares no competing interests.
This article was published online on 8 February 2023.

Forum: Astronomy

JWST opens a window on exoplanet skies

An unprecedented glimpse of a distant planet reveals clues about how it might have formed. Scientists explain why it's a win for atmospheric chemistry, and celebrate the technology that made it possible. **See p.649, p.653, p.659, p.664 & p.670**

The papers in brief

- **NASA's James Webb Space Telescope (JWST) was launched in December 2021 and now orbits the Sun, some 1.5 million kilometres from Earth.**
- **Early data released last year confirm that this is an ideal vantage point for investigating exoplanets — the distant worlds that orbit stars other than the Sun.**
- **Five papers in *Nature* report analyses profiling the atmospheric chemistry of WASP-39b, a hot exoplanet with a Saturn-like mass.**
- **The studies settle questions about this exoplanet's atmosphere, and showcase the power and versatility of JWST.**

Julia V. Seidel & Louise D. Nielsen

Composition and origins of far-flung worlds

Rustamkulov *et al.*[1] (page 659), Alderson *et al.*[2] (page 664), Feinstein *et al.*[3] (page 670) and Ahrer *et al.*[4,5] (pages 649 and 653) used three different instruments on board JWST — each with its own advantages and shortcomings — but reported largely complementary results (Fig. 1). In all five investigations, the teams found that elements heavier than hydrogen and helium are more abundant in the atmosphere of WASP-39b than they are in the Sun, whereas the ratio of carbon to oxygen is lower than that of the Sun and commensurate with that of Saturn. These findings offer crucial information about the planet's formation, the basic composition of its atmosphere and its potential to host life.

The carbon/oxygen ratio of an exoplanet's atmosphere is a telltale sign of where the planet formed[6]. This is particularly useful in the case of giant planets that are close to their host stars, because their formation mechanism has been an open question since the first exoplanet was found. The ratio measured for WASP-39b indicates that the planet might have formed at a location beyond the system's water-ice line — the distance from the host star at which it is cold enough for compounds such as water and carbon dioxide to condense into solid ice. At this location, the planet could have accreted the oxygen-rich solids measured by JWST, before migrating inwards to its current position.

The sulfur/oxygen ratio is another piece in the puzzle of planetary formation. But the sulfur content of an exoplanet's atmosphere is intriguing for another reason. Sulfur dioxide is much like the protective ozone in Earth's atmosphere: it is produced during chemical reactions that are triggered by ultraviolet radiation from the host star[7]. Rustamkulov *et al.* and Alderson *et al.* both detected sulfur dioxide in the atmosphere of WASP-39b. This observation marks the first direct evidence of light-induced (photochemical) reactions in an exoplanet atmosphere —a milestone in the quest for a truly habitable planet.

Much work remains to be done to probe the limits of this habitability. However, the finding is a step towards understanding how photochemistry protects exoplanetary surfaces from high-energy irradiation. It also tightens constraints on the parameters used in models of planetary formation. Both advances pave the way to future observations of planets that are similar to Earth.

Part of these efforts involve profiling the characteristics of the exoplanetary atmosphere itself. By comparing the measured chemical abundances with those of several cloud models, Feinstein *et al.* determined that the clouds of WASP-39b are broken up along the day–night terminator, the line that separates day and night on a planet. Such cloud structure has previously been associated with other hot exoplanets that have