# News & views

## Forum: Neuroscience

# Cellular atlases of the entire mouse brain

In a huge collaborative effort, millions of cells in the mouse brain have been mapped in detail. Two scientists examine the resulting wealth of insights into gene regulation in brain cells, neuronal connections and how our own brains evolved. **See pp.317–424**

## Maria Antonietta Tosches
### Maps to explore uncharted territory in the brain

For centuries, explorers have drawn and updated geographical maps, and used them to organize information and venture into new parts of the world. Similarly, brain atlases have been instrumental in neuroscience research. A series of papers from the BRAIN Initiative Cell Census Network (BICCN) present cellular maps of the mouse brain (Fig. 1), which will enable neuroscientists to explore previously uncharted territories with unprecedented depth and resolution.

In the past decade, single-cell RNA sequencing — a method of analysing the complete set of genes expressed in an individual cell (its transcriptome) — has fuelled the generation of cell-type atlases. So far, multimodal, high-resolution atlases have been limited mostly to particular parts of the brain, such as the motor cortex[1]. The latest BICCN papers by Yao et al.[2], Zhang et al.[3], Langlieb et al.[4] and Shi et al.[5] bring cell-type atlases to the next level. First, they present comprehensive resources that encompass not just specific regions, but the entire brain of the mouse (*Mus musculus*). Second, they combine single-cell RNA sequencing with spatial transcriptomics, which allows cell types to be placed in their native tissue context.

These brain-wide atlases allow for unbiased comparisons between brain regions. A key insight emerging from these data is that the number of neuron types in an area does not scale with the size of that area or the total number of neurons in it. Certain regions of the brain — such as the hypothalamus, midbrain and hindbrain — contain many types of

neuron, which are relatively similar to one another. By contrast, other regions of the brain, including the cerebral cortex and cerebellum, contain fewer types of neuron, but they are more divergent in their identities. These findings indicate that different rules govern neuronal diversity across the brain, perhaps because each region evolved under different constraints. The BICCN data sets will prompt further investigations of poorly studied brain regions by providing open access to the transcriptomic profiles and spatial distributions of all of their neuronal types.

After sequencing the RNA transcripts of millions of cells in the mouse brain, Yao et al.[2] and Langlieb et al.[4] report that cells can be grouped into more than 5,000 distinct clusters, most of which are neuronal. However, one set of genes — those that encode transcription factors — alone have the same classification power as complete transcriptomes, indicating that the immense diversity of brain cell types can be defined simply by the combinations of transcription factors that the cells express[2].
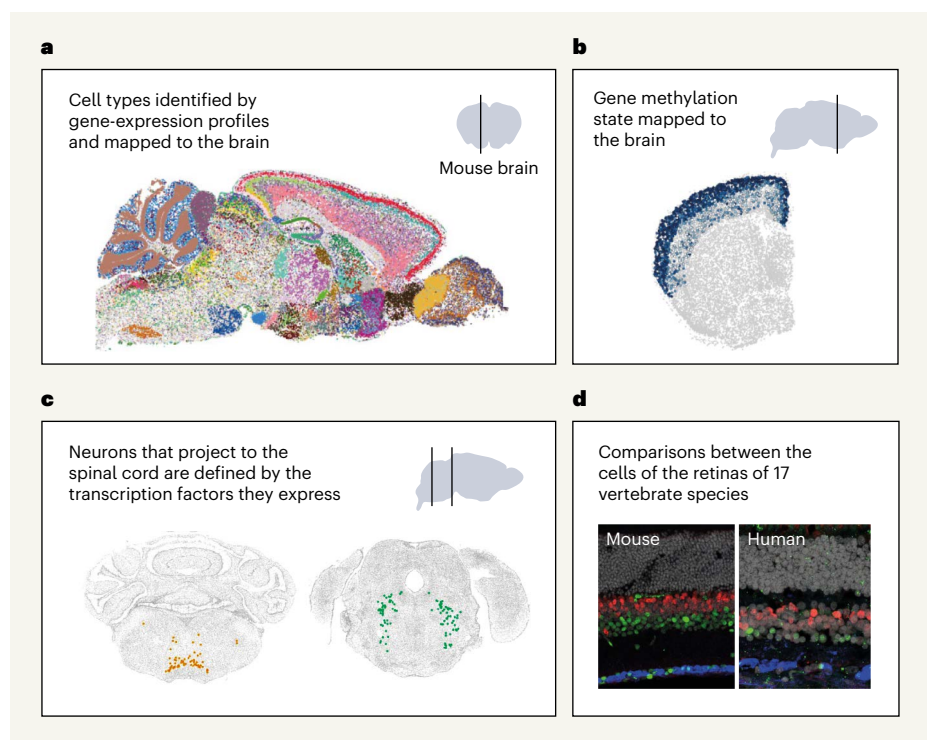


**Figure 1 | Whole mouse-brain atlases from the BRAIN Initiative Cell Census Network provide a wealth of spatial transcriptomic and epigenomic data. a**, Thousands of cell types (shown in different colours) can be identified on the basis of their transcriptomic profiles, and mapped to the anatomical regions of the mouse brain. (Adapted from Fig. 1 of ref. 2.) **b**, Epigenetic information, such as the methylation state of a gene (shown in shades of blue), can also be spatially mapped to allow comparisons between cell types and brain regions. (Adapted from Fig. 3 of ref. 13.) **c**, Neurons that project to the spinal cord can be investigated using spatial transcriptomic information. This reveals, for example, that the combination of transcription factors that spinal-projecting neurons express (shown by different colours) defines where they are in the brain. (Adapted from Fig. 4 of ref. 7.) **d**, Transcriptomic and epigenomic data sets enable comparisons between species. For example, comparing gene-expression patterns in neurons of the retinas of the eyes of different vertebrate species (including mice and humans) reveals characteristics of cells that have either been conserved or diverged throughout evolution. (Adapted from Fig. 1 of ref. 10.)

Further to these observations, Zhou et al.[6] and Winter et al.[7] investigated the relationships between the molecular identity of neurons and the region of the central nervous system to which they extend their neuronal projections (axons). By focusing on a specific population of neurons that project to the spinal cord, Winter and colleagues show that patterns of expression of a particular family of transcription factors distinguish five spatially segregated subpopulations. These results support the hypothesis that multiple facets of neuronal identity, including its transcriptome and projection specificity, are controlled in concert by cell-type-specific combinations of transcription factors[8].

If this is the case, it is tempting to speculate that neurons that are evolutionarily conserved across species should express the same transcription factors, and that they might have some similar anatomical and functional features[9]. This is precisely what Hahn et al.[10] found when they studied eye evolution by comparing single-cell RNA sequencing data from the retinas of 17 vertebrate species. The retina is an excellent choice to study cell-type evolution: it has a conserved organization, with six major cell classes, but it nevertheless serves the visual needs of animals inhabiting different environments, suggesting that there have been evolutionary adaptations at the cellular level.

Using comparative transcriptomics, Hahn and colleagues show that neurons that were thought to be specific to primates, known as midget retinal ganglion cells, actually have counterparts in rodents. Furthermore, the granularity of the cell-type data allows scientists to test whether evolutionary adaptations affect certain cell classes more than others. Indeed, the authors show that cell-type diversification is highly pronounced in retinal ganglion cells, which are the neurons that transmit visual information from the retina to the brain.

Cell-type atlases are important not only for understanding brain architecture at the cellular scale, but also for making precise inferences about how brains evolved. As the studies by Hahn et al.[10] and others[11] show, each class of neuron evolves at its own pace. The kinship between gene expression (including that of transcription factors) and axonal projections suggests that the molecular evolution of neurons is tied to the evolution of their function. In the future, these hypotheses can be tested rigorously by comparing comprehensive maps of gene expression and neural connections between multiple species.

**Maria Antonietta Tosches** is in the Department of Biological Science, Columbia University, New York, New York 10027, USA.
e-mail: mt3353@columbia.edu

## Heather J. Lee
### Buried treasure in the epigenomic landscape

Whether you got to work this morning by bike, car or boat, your choice was probably dictated by your environment. In the same way that our daily lives are shaped by the landscapes in which we live, the specialized functions of our cells are constrained by the 'epigenomic landscape' — the pattern of the molecular features across a genome that control access to the genetic code. Such features include biochemical modifications to DNA (such as the addition of a methyl group), the compaction of chromatin (the packaged form of DNA in complex with proteins) and the conformation of chromosomes, all of which are dynamically rearranged during development. This leads to distinct epigenetic profiles that maintain expression of the appropriate genes at the appropriate levels in each specialized cell type. The BICCN presents an incredibly detailed survey of the epigenomic landscapes across thousands of specialized cell types from the adult mouse brain.

The authors used advanced experimental and analytical approaches to map multiple epigenetic features in the nuclei of individual

> **"Epigenomic landscapes can shape interactions between cell types in different brain regions."**

cells and to connect information about gene transcription with information about the anatomical location of cells in the brain. This collaborative effort in cellular cartography has revealed insights into the rules that govern how the genetic code is read and interpreted in different cell types and species.

Key features of epigenomic landscapes are segments of non-coding DNA called regulatory elements, which control gene activity. Zu et al.[12], Liu et al.[13] and Zemke et al.[14] have mapped these regulatory elements in individual cell types of the mouse brain. Zu and colleagues also developed deep-learning models that can identify potential cell-type-specific regulatory elements on the basis of the underlying DNA sequence. When tested on human DNA sequences, these models predicted regulatory elements in the corresponding cell types from the mouse brain, illustrating that the links between the genetic sequence and cell-type-specific epigenomic landscapes are conserved across species.

A detailed comparison between four mammalian species found that highly conserved regulatory elements tend to be located near

the starts of genes that they regulate, whereas species-specific regulatory elements were at sites farther away. Another interesting observation is that regulatory elements specific to either mouse or human cells often overlap with transposable elements — genetic sequences that moved around the genome during mammalian evolution[14]. This suggests that transposable elements could have been co-opted to fine-tune gene activation in specialized cell types and in different species.

Another feature of the epigenomic landscape is the segregation of active and inactive segments of the genetic code into distinct chromatin compartments. Liu and colleagues mapped these features in each cell type and found that DNA segments that switched between active and inactive compartments contained genes that are important for specialized neuronal functions. They also identified physical interactions between different parts of the genetic sequence and, using machine learning, demonstrated that these interactions could be used to predict which version of a gene is active in each cell type.

Next, the same group mapped places in the genome that limit interactions between genes and regulatory elements — known as genomic boundaries — across brain cells. Interestingly, boundaries that were specific to a particular cell type were often found at the start or end of long neuronal genes that were active in that population of cells. When comparing between species, Zemke and colleagues found that evolutionarily conserved genomic boundaries were associated with genes that were uniformly expressed in all the species that they tested. Similar to regulatory elements, boundary sites that differed between species contained many transposable elements.

Zhou et al.[6] explored links between epigenomic landscapes and synaptic connections between cell types in different parts of the brain. In a region of the cerebrum (the forebrain) called the isocortex, for example, cells that project to other regions in the cerebrum had epigenetic profiles that were distinct from those that project to regions outside the cerebrum, such as the hindbrain. This suggests that epigenomic landscapes can shape interactions between cell types in different brain regions, as well as genetic activity in individual cells.

Will this survey of epigenomic landscapes point future studies to any buried treasure? The directory of cell-type-specific gene regulatory networks — which describe the interactions between hundreds of transcription factors, regulatory elements and genes — will certainly be an excellent resource for researchers investigating atypical cell differentiation in neurodevelopmental disorders. In other neurological conditions, epigenomic landscapes will guide the interpretation of

genetic contributions to disease risk. Indeed, Zemke and colleagues demonstrate that genetic variants linked to multiple sclerosis, anorexia nervosa and tobacco-use disorder often overlap with regulatory elements that are conserved across species[14]. Thus, the work presented by the BICCN has laid a firm foundation for many further discoveries relating to neurobiology and neurological disorders.

Heather J. Lee is in the School of Biomedical Sciences and Pharmacy, The University of Newcastle, Callaghan 2308, Australia and the Precision Medicine Research Program, Hunter Medical Research Institute, New Lambton Heights, Australia.
e-mail: heather.lee@newcastle.edu.au

1. BRAIN Initiative Cell Census Network (BICCN), *Nature* **598**, 86–102 (2021).
2. Yao, Z. *et al. Nature* **624**, 317–332 (2023).
3. Zhang, M. *et al. Nature* **624**, 343–354 (2023).
4. Langlieb, J. *et al. Nature* **624**, 333–342 (2023).
5. Shi, H. *et al. Nature* **622**, 552–561 (2023).
6. Zhou, J. *et al. Nature* **624**, 355–365 (2023).
7. Winter, C. C. *et al. Nature* **624**, 403–414 (2023).
8. Hobert, O. *Annu. Rev. Cell Dev. Biol.* **27**, 681–696 (2011).
9. Arendt, D. *et al. Nature Rev. Genet.* **17**, 744–757 (2016).
10. Hahn, J. *et al. Nature* **624**, 415–424 (2023).
11. Hain, D. *et al. Science* **377**, eabp8202 (2022).
12. Zu, S. *et al. Nature* **624**, 378–389 (2023).
13. Liu, H. *et al. Nature* **624**, 366–377 (2023).
14. Zemke, N. R. *et al. Nature* **624**, 390–402 (2023).

Immunology

# DNA is tethered together in immune cells

Zhi Liu & Ye Zheng

A transcription factor in immune cells forms an unexpectedly ladder-like complex with two DNA molecules, allowing the expression of genes that these cells need to suppress harmful immune responses. **See p.433**

The human immune system is meant to defend the body from pathogens such as bacteria and viruses. When the immune system goes awry, however, it can start to attack the body itself, resulting in autoimmune diseases such as type 1 diabetes and rheumatoid arthritis. To prevent this from happening, the immune system has a braking mechanism, in the form of regulatory T ($T_{reg}$) cells that rein in harmful immune reactions by suppressing the activity of other immune cells. The differentiation of $T_{reg}$ cells from precursor cells (a process referred to as $T_{reg}$ cell development) depends on a transcription factor called FOXP3, which turns genes on or off by binding to DNA. On page 433, Zhang *et al.*[1] describe the ladder-like structure of a complex of ten FOXP3 proteins and two DNA molecules, which could be responsible for bringing remote regions of DNA together to control gene expression.

Most $T_{reg}$ cells arise in the thymus gland along with other T cells; the remainder develop in the body's periphery, such as the small and large intestines. A key step in $T_{reg}$ cell development is the expression of FOXP3, which kick-starts differentiation. This leads to the expression of a suite of genes specific to $T_{reg}$ cells, called signature genes, that allows them to carry out their immune-suppressive function[2,3].

The *FOXP3* gene was originally identified by mapping mutations that cause a severe autoimmune disease in humans and a similar condition in a mutant mouse strain[4,5]. It belongs to a family of at least 40 'forkhead box' proteins, all of which contain a forkhead DNA-binding domain that recognizes a short DNA sequence known as a forkhead motif. When FOXP3 was first crystallized, it was thought to form a type

of dimer in which identical domains from each protein are swapped over[6]. However, a structural study last year revealed that FOXP3 proteins actually form head-to-head dimers when they bind to DNA[7].

Like other forkhead-box proteins, FOXP3 has been shown to bind to the forkhead motif *in vitro*[6,7]. However, when researchers sequenced the DNA motifs bound by FOXP3 *in vivo*, the forkhead motifs barely showed up[8–10]. This result raised the intriguing question of whether FOXP3 functions in $T_{reg}$ cells by binding to DNA sequences other than the forkhead motif.

In the current study, Zhang and colleagues started to tackle this question by looking for sequences in the mouse genome to which FOXP3 was bound. Surprisingly, they found that FOXP3 was strongly associated with DNA fragments containing repeat sequences of several thymine (T) nucleotides, followed by one guanine (G) nucleotide (denoted $T_nG$), instead of the expected forkhead motif. Notably, repeat sequences such as $T_nG$ are usually considered non-functional and are discarded as junk DNA (sequences with no apparent biological function) by sequence-analysis programs, explaining why previous studies ignored these sequences when looking for fragments of DNA that bind to FOXP3.

Using cryo-electron microscopy, the authors resolved the structure of FOXP3 proteins in a complex with DNA molecules that contained 18 $T_nG$ repeats. The structure of the complex is exquisitely ladder-like: two double-stranded DNA molecules form the 'side rails' of the ladder, which are pulled together by five pairs of FOXP3 proteins that serve as 'rungs'. This represents a new geometry of the FOXP3–DNA assembly, which is in contrast to the simple head-to-head dimer that FOXP3 proteins form on binding to a single DNA molecule (Fig. 1).
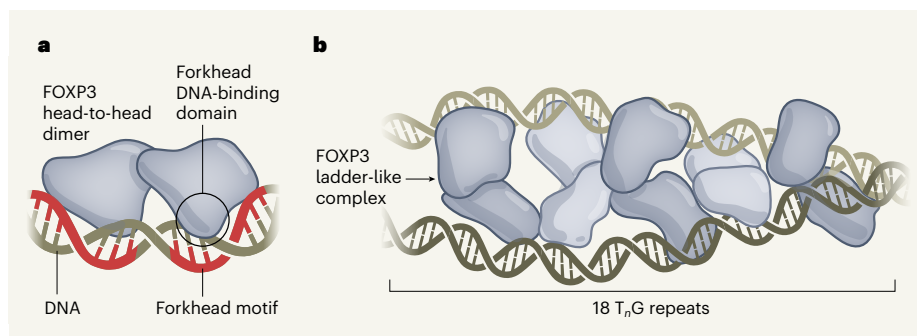
The ladder-like structure of the FOXP3–DNA



**Figure 1 | The transcription factor FOXP3 can control gene expression in regulatory T cells through two distinct mechanisms. a**, Two FOXP3 proteins form a head-to-head dimer that binds to a region of DNA called the forkhead motif, which directly regulates gene expression. **b**, Zhang *et al.*[1] report that multiple FOXP3 proteins can also bind to regions of DNA called $T_nG$ repeats. This interaction forms a ladder-like complex consisting of two double-stranded DNA molecules held together by five pairs of FOXP3 proteins, which brings together regions of DNA that would otherwise be far apart. Both mechanisms seem to be important for the transcription of genes that allow regulatory T cells to carry out their function of suppressing the activity of other immune cells.