

Evolution at the cellular level



Capitalizing on developments in genome sequencing technology, the Biodiversity Cell Atlas is a multinational project that uses single-cell transcriptomics to map cell types of whole organisms across the tree of life.

As laid out at the Biodiversity Cell Atlas (BCA)'s [launch meeting](#) last month in Barcelona's Centre for Genomic Regulation, the project is a joint effort of three communities that develop and apply single-cell sequencing technology to build cell atlases of model organisms, sequence genomes of a diversity of species and compare single-cell genomics datasets.

Cell atlases are maps of cell types of a given organism inferred from gene expression profiles of single cells. This resource not only includes general characteristics of cells and their function but also contains information about the location of cells within tissues or within the body. If cell atlases are available for multiple developmental stages, we can trace cell trajectories during development; and if cell atlases are available for multiple species, we can study the evolutionary origins and diversification of cell types¹.

For those who are interested in developmental gene regulatory networks and the evolution of gene regulation, the BCA will include valuable single-cell transcriptomic data that enable integration of other types of data such as genome sequences, bulk RNA sequencing (RNA-seq), whole-genome chromatin accessibility (assay for transposase-accessible chromatin with sequencing (ATAC-seq)) or 3D genome structure inferred from chromatin interactions across the genome (chromosome conformation capture (Hi-C)). For example, comparing single-cell RNA-seq (scRNA-seq) data of key developmental stages between closely related species that differ in a morphological trait of interest could resolve the timing of cell fate specification that underlies that trait and reveal associated regulatory factors.

Cells are the units of life but they can also be seen as evolutionary units². Each cell type is characterized by unique genomic

information, and changes in the regulatory signature of a given cell type may lead to the evolution of new cell entities. From this perspective, the BCA will be a tremendous resource for comparative studies that aim to elucidate how cell types are related by descent as well as the molecular mechanisms that underlie changes in the gene regulatory networks that determine cell identity. As with many evolutionary questions, understanding the origin of cell types requires comparisons within and across species. For example, on the basis of whole-organism RNA-seq data, Musser et al.³ defined 18 cell types in a freshwater sponge and reconstructed their evolutionary relationships. In a study that shows the importance of comparing scRNA-seq data across the different life stages of a single species, Steger et al.⁴ identified the developmental origins of neural cell types in a cnidarian (the sister group to bilaterians) and offered insights into the evolutionary origin and diversification of bilaterian neural cell types. Li et al.⁵ compared brain cell atlases of different castes of a social ant to identify neural mechanisms linked to behavioural specialization. On the basis of whole-organism scRNA-seq data, Seb e-Pedr os et al.⁶ compared cell type diversity across three early-branching animal lineages (sponges, comb jellies and Placozoa) and showed different architectures of the gene regulatory networks that determine cell type in these animals. In another cross-species study, Shafer et al.⁷ compared cell atlases of the hypothalamus of zebrafish and Mexican tetra (surface and cave morphs) to determine conservation and diversification of hypothalamic cell types in teleosts, as well as their genetic mechanisms of evolution.

Although some of these studies include cross-species comparison of single-cell datasets, identifying homologous cell types is particularly difficult between distantly related species. It is fair to say that, as more species are added to the BCA, comparing cell atlases across the tree of life will be one of the major challenges for those interested in studying cell type evolution. High-quality genome assemblies and annotations as well as high-quality scRNA-seq data are paramount for informative comparative studies. Different approaches are emerging: for example,

SAMap⁸ identified similar cell types with shared expression programs across distantly related species such as a sponge and zebrafish, and the demonstration that cell transcriptomes have tree structure⁹ opens up the possibility of phylogenetic comparative approaches to cell biology. A phylogenetic tree of cells is a promising framework to address exciting topics such as the inference of ancestral cell types, reconstruction of the evolutionary history of cell divergence and whether cell types with similar morphology and/or function, in different organisms, represent homology or convergent evolution.

Building whole-organism cell atlases across the tree of life is no small feat but there are similarly ambitious projects, such as the [Earth Biogenome Project](#) or the [Human Cell Atlas](#) and the [Fly Cell Atlas](#), that can provide information on operational and technical issues. From a technical perspective, the community must agree on standards for sample collection and processing, data generation and analysis, including detailed reporting of metadata that allows for reproduction and interpretation of results. This is particularly important given that the BCA will include non-model organisms collected in the field, which will require development of new protocols for specimen dissociation and RNA preservation.

The launch of the BCA reflects advances in genomics and comparative evo-devo but is also an exciting, ambitious project that opens the prospect of building the cellular tree of life. Beyond that, understanding the mechanisms behind the tremendous cell diversity that we see in nature will inform biomedical research and synthetic biology. As with other biodiversity-related projects, the BCA should consider issues of justice, equity, diversity and inclusion from its inception. This project requires networking between different scientific communities and collection of species across the globe, thus offering opportunities for meaningful collaborations with researchers in the Global South and the involvement of local communities, and continual consideration of how benefits of the outcomes of the research may be equitably shared.

Published online: 3 July 2023

References

1. Tanay, A. & Sebé-Pedrós, A. *Trends Genet.* **37**, 919–932 (2021).
2. Arendt, D. et al. *Nat. Rev. Genet.* **17**, 744–757 (2016).
3. Musser, J. M. et al. *Science* **374**, 717–723 (2021).
4. Steger, J. et al. *Cell Rep.* **40**, 111370 (2022).
5. Li, Q. et al. *Nat. Ecol. Evol.* **6**, 1191–1204 (2022).
6. Sebé-Pedrós, A. et al. *Nat. Ecol. Evol.* **2**, 1176–1188 (2018).
7. Shafer, M. E. R., Sawh, A. N. & Schier, A. F. *Nat. Ecol. Evol.* **6**, 63–76 (2022).
8. Tarashansky, A. J. et al. *eLife* **10**, e66747 (2021).
9. Liang, C., FANTOM Consortium, Forrest, A. R. R. & Wagner, G. P. *Nat. Commun.* **6**, 6066 (2015).