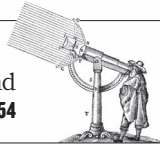


COMMENT

HISTORY The link between science and the Reformation **p.454**

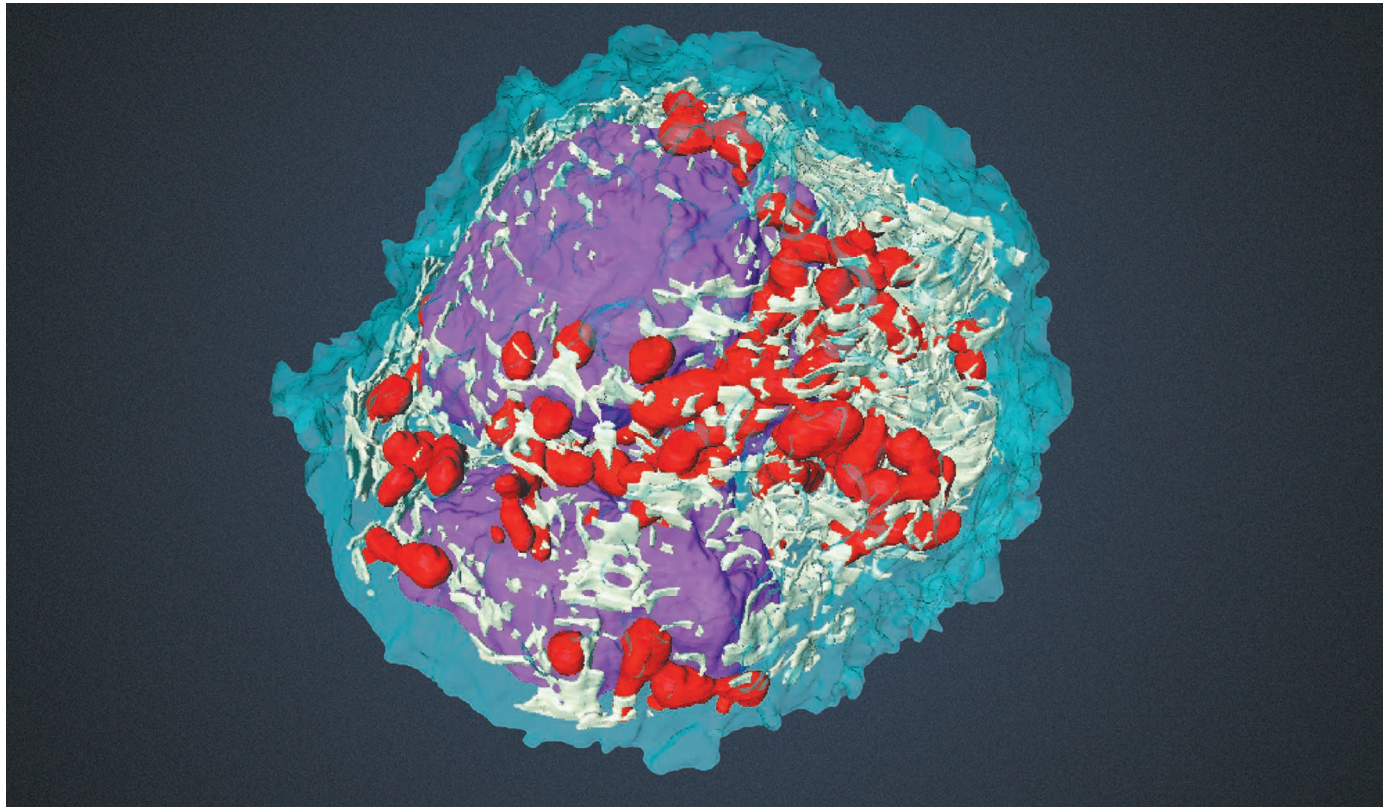


DISASTERS Art and interactives probe our preparedness for cataclysms **p.456**

CONSERVATION Optimistic tally could endanger snow leopards **p.457**

OBITUARY Pioneer in optics, lasers and NMR, remembered **p.458**

VILLANI, A.-C. ET AL. SCIENCE 356, EAAH453 (2017); IMAGE KATHRYN WHITE; RECONSTRUCTION JAMES FLETCHER



A new type of human dendritic cell recently discovered using single-cell RNA sequencing.

The Human Cell Atlas: from vision to reality

As an ambitious project to map all the cells in the human body gets officially under way, **Aviv Regev**, **Sarah Teichmann** and colleagues outline some key challenges.

Our knowledge of the cells that make up the human body, and how they vary from person to person, or throughout development and in health or disease, is still very limited. This week, a year after project planning began, more than 130 biologists, computational scientists, technologists and clinicians are reconvening in Rehovot, Israel, to kick the Human Cell Atlas initiative¹ into full gear. This international collaboration between hundreds of

scientists from dozens of universities and institutes — including the UK Wellcome Trust Sanger Institute, RIKEN in Japan, the Karolinska Institute in Stockholm and the Broad Institute of MIT and Harvard in Cambridge, Massachusetts — aims to create comprehensive reference maps of all human cells as a basis for research, diagnosis, monitoring and treatment.

On behalf of the Human Cell Atlas organizing committee, we outline here some

of the key challenges faced in building such an atlas — and our proposed strategies. For more details on how the atlas will be built as an open global resource, see the white paper² posted on the Human Cell Atlas website.

Cells have been characterized and classified with increasing precision since Robert Hooke first identified them under the microscope in the seventeenth century. But biologists have not yet determined all the molecular constituents of cells, nor have they established how ▶

▶ all these constituents are associated with each other in tissues, systems and organs. As a result, there are many cell types we don't know about. We also don't know how all the cells in the body change from one state to another, which other cells they interact with or how they are altered during development.

TECHNOLOGY REVOLUTION

New technologies offer an opportunity to build a systematic atlas at unprecedented resolution. These tools range from single-cell RNA sequencing to techniques for assessing a cell's protein molecules and profiling the accessibility of the chromatin. For example, we can now determine the RNA profiles for millions of individual cells in parallel (see 'From one to millions'). Protein composition and chromatin features can be studied in hundreds or thousands of individual cells, and mutations or other markers tracked to reconstruct cell lineages. We can also profile multiple variants of RNA and proteins *in situ* to map cells and their molecules to their locations in tissues.

We anticipate that the atlas will help researchers to answer key questions in diverse biological fields. In cellular taxonomy, it might enable the discovery and identification of cell types and molecular markers or signatures (a collection of genes, say, that characterize a specific cell type). In histology, it should enable researchers to relate tissue structure to the position of cells and molecules. Developmental biologists will be able to use it to track cell fate and lineage. Physiologists could characterize dynamic states, such as the cell cycle, and transient responses such as a T cell's reaction to a pathogen.

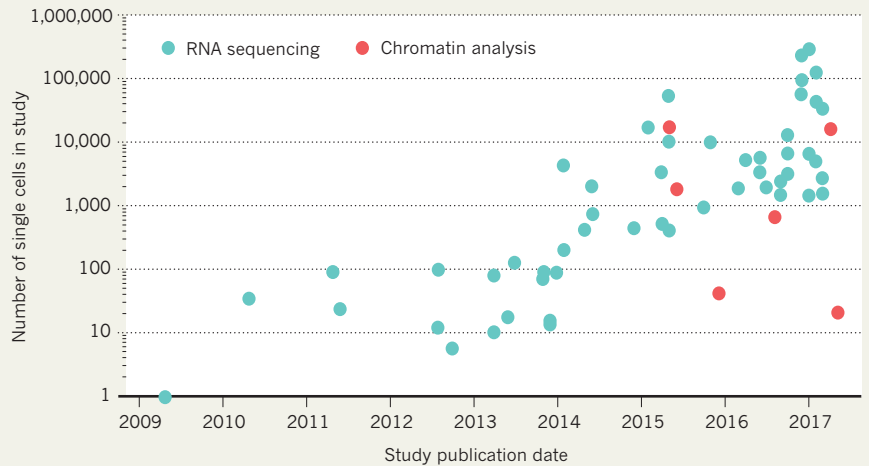
The atlas could also facilitate research on the molecular mechanisms of communication within and between cells. And it should allow biologists to compare cell types across species to better understand human evolution, and to determine to what extent animal model systems and organoids reflect human biology.

Crucially, the atlas should help researchers to compare healthy reference cells to diseased ones in the relevant tissues — and so facilitate the development of better drugs and more accurate predictions of unintended toxicity. The atlas could also aid regenerative medicine — the process of replacing, engineering or regenerating human cells, tissues or organs to establish normal function. Key diagnostic tests, such as the complete blood count — a routine blood screen that provides crude counts of white blood cells, red blood cells and so on — would become vastly more informative if cell types and states could be identified with much finer granularity. Such information could, for example, help to diagnose blood cancer, autoimmunity or infection before clinical symptoms appear.

Early studies are already showing

FROM ONE TO MILLIONS

Biologists can now analyse RNA transcripts or chromatin accessibility in thousands or even millions of individual cells in parallel.



tremendous potential in all these areas. New cell types have been found in the brain³⁻⁷, gut⁸, retina⁹ and immune system¹⁰, and these discoveries have yielded new insight — into how the immune system¹¹ functions, for example, and into the dynamics of tumour ecosystems¹². Yet, to take the next step — to build a human cell atlas that is truly useful — requires taking the long view and addressing various systemic and organizational challenges, as well as technical and scientific ones.

THE CHALLENGES

Agree on scope. In light of the enormous complexity of the human body, and the rapid evolution of technologies for probing cells and tissues, and for analysing the data, we plan to build this resource in phases and generate reference maps at increasing resolution as the project progresses.

The first draft of the atlas will profile cells' molecular and spatial characteristics, capturing only those cell types that occur above a pre-specified rarity — ones that make up more than 1% of a sample, say. These cells will be obtained from major tissues from healthy donors, taking into account the genetic diversity, geographical location and person's age. Although disease will not be a focus of the first draft of the atlas, we plan to look at some disease samples to compare them with healthy cell types.

The first draft will focus on tissues, not whole organs. Extremely rare cells may be missed, and sample sizes may be too small to fully reveal the links between cellular characteristics and human diversity. In later phases, the atlas could take on entire organs, include small cohorts of people (say, 50–60) with diseases of interest, gather bigger sample sizes and provide greater power to associate molecular variation with the underlying genetic diversity. A similar step-wise strategy

was deployed in the Human Genome Project; even a partially assembled genome proved immediately useful to researchers, and human genetic variation in health and disease was tackled over several years after the full genome was sequenced.

The atlas will provide an important starting point for functional studies — for instance, those aimed at establishing the mechanistic links between cell states and disease. But such studies are themselves beyond its scope. Again, this parallels what happened with the Human Genome Project: studies of functional elements in the genome, which are ongoing, have relied on the reference sequence obtained through the project.

The atlas will aim to provide a detailed representation of molecules, cells, tissues, organs and systems, allowing researchers to zoom in and out to identify patterns and interactions at various levels of resolution. To this end, those compiling the atlas must establish how many cells to sample, which types of molecular features to analyse, how to assign cells to different categories and how to subdivide those categories. At the spatial level, they must decide how to sample complex anatomies and histologies. Lastly, they need to establish ways of connecting the various layers of cellular and spatial information from different samples to a single anatomical reference by developing what is termed a common coordinate framework.

To ensure the best use of resources, those involved in the initiative must agree on the desired resolution for each phase of the atlas. Researchers could, of course, try to pursue ever-rarer cell types, but potentially at ever-greater expense. In this respect, the Human Cell Atlas will pursue similar approaches to those used in human genetic studies that focus on variants present at a

certain frequency. Here, geneticists have begun to tackle increasingly rare variants as technologies have advanced.

Be open and fair. To have maximum impact, the Human Cell Atlas must be an open resource, on many levels.

The project is already open to all interested participants who are committed to its values. Discussions about particular organs, tissues, technologies or computational approaches are running on more than a dozen Slack channels that anyone can join.

Wherever consent agreements allow, atlas data will be made publicly available in an open-source data-coordination platform as soon as possible, after they have been collected and have passed quality-control checks. All standards established to ensure the production of high-quality data, and any updates to those standards, will also be shared. The same goes for new technologies and computational methods resulting from the project.

Atlas data and analysis products will exist in multiple public clouds (currently, those hosted by Google, Amazon and Microsoft) to ensure that people with different preferred cloud environments can access them. Because computation will happen in the cloud, individual researchers will not need to download and store all the data or have access to their own high-performance computing power. Finally, in addition to the continual release of data and periodic formal data releases, publications interpreting the data will help to establish standardized approaches and disseminate the insights and value that can be gained from them.

As much as possible, the atlas must reflect the diversity of humans and human experience. The broad distribution of participating researchers, institutions and countries involved in the initiative will, in itself, help to ensure tissue diversity. The initiative currently includes members from 5 continents and more than 18 countries, including Japan, Israel, South Africa, China, India, Singapore, Canada and Australia.

Getting appropriate consent agreements and fostering public trust from the outset will also help efforts to obtain sufficient geographical, gender, age and genetic diversity in sampling. As part of the global initiative, an ethics working group will establish how best to obtain informed consent from sample donors, how the terms of that consent can be adhered to and how to protect the privacy of participants and donors appropriately. Various existing projects involving human samples, such as the public-research

project ENCODE (the Encyclopedia of DNA Elements), which aims to identify all the functional elements of the human genome, can provide guidance on this.

Procure samples appropriately. Obtaining tissue samples using standardized procedures, with appropriate consent and in a way that enables other researchers to know exactly where the sample came from is a complex endeavour. To access the diversity of human tissues needed, researchers will work with both fresh tissue from live donors and specimens obtained postmortem or from transplant organ donors.

We plan to learn from, and build on, pre-existing reliable procurement processes. Examples include those used in the Genotype-Tissue Expression Project (GTEx, a database and tissue bank designed to help researchers to gain insight into the mechanisms of gene regulation in humans) and the Cambridge Biorepository for Translational Medicine, a resource for multidisciplinary research projects for which fresh tissue is required.

Organize effectively. The Human Cell Atlas consortium is built on four distinct and interconnected pillars. Collaborative biological networks involve experts in biological systems or organs as well as in genomics, computation and engineering, working together to build maps of each tissue, system or organ. Several biological-network pilot projects have been formulated through grass-roots efforts in the Human Cell Atlas community. As well as revealing new biology and helping to build a collaborative international network, these activities are informing the community about how to structure sampling and conduct analyses for a full-scale cell atlas.

A technical forum involving genomics experts, imaging specialists and biotechnologists, is developing new technologies, and testing, comparing and disseminating existing ones. A data-coordination platform is being designed to bring researchers to the data by developing the software to upload, store, process and serve data. The platform also provides an open environment in which computational methods and algorithms developed by any interested group can be shared. Finally, an analysis garden involves computational biologists working together to develop sophisticated techniques for data mining and interpretation.

Activities across all areas are currently governed by a scientific steering group, the Human Cell Atlas organizing committee. Co-chaired by two of us (A. R. and S. A. T.), this includes 27 scientists from 10 countries and diverse areas of expertise. The committee establishes working groups (about 5 so far, consisting of about 5 to 15 members each) that tackle specific key areas.

For instance, an analysis working group is crafting best practices for computational analysis through a community-wide process, including workshops and jamborees. The committee governs the data-coordination platform, including making all policy decisions and approving its overall plan.

JOIN THE EFFORT

Having a catalogue of genes at our fingertips has transformed research in human biology and disease. Similarly, we believe that the Human Cell Atlas will catalyse progress in biology and medicine. Descriptors such as 'cell type' and 'cell state' can be difficult to define at the moment. An integrative, systematic effort by many teams of scientists working together and bringing different expertise to the problem could dramatically sharpen our terminology, and revolutionize the way we see our cells, tissues and organs. We invite you to join the effort. ■

Orit Rozenblatt-Rosen is lead scientist for the Human Cell Atlas Initiative at the Klarman Cell Observatory at the Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. **Michael J. T. Stubbington** is lead scientist for the Human Cell Atlas Initiative at the Wellcome Trust Sanger Institute, Hinxton, UK. **Aviv Regev** is co-chair of the Human Cell Atlas organizing committee, and at the Klarman Cell Observatory at the Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA, and the Massachusetts Institute of Technology, Cambridge, Massachusetts. **Sarah A. Teichmann** is co-chair of the Human Cell Atlas organizing committee, and at the Wellcome Trust Sanger Institute, Hinxton, UK, and the Cavendish Laboratory, Cambridge. e-mails: orit@broadinstitute.org; ms31@sanger.ac.uk; aregev@broadinstitute.org; st9@sanger.ac.uk

On behalf of the Human Cell Atlas organizing committee, a list of whose members accompanies this Comment online (see go.nature.com/2yo17mx).

1. Regev, A. *et al.* *The Human Cell Atlas* Preprint available at bioRxiv at <http://dx.doi.org/10.1101/121202> (2017).
2. White Paper available at https://www.humancellatlas.org/files/HCA_WhitePaper_18Oct2017.pdf
3. Darmanis, S. *et al.* *Proc. Natl Acad. Sci. USA* **112**, 7285–7290 (2015).
4. Lake, B. B. *et al.* *Science* **352**, 1586–1590 (2016).
5. Pollen, A. A. *et al.* *Nature Biotechnol.* **32**, 1053–1058 (2014).
6. Tasic, B. *et al.* *Nature Neurosci.* **19**, 335–346 (2016).
7. Zeisel, A. *et al.* *Science* **347**, 1138–1142 (2015).
8. Grün, D. *et al.* *Nature* **525**, 251–255 (2015).
9. Shekhar, K. *et al.* *Cell* **166**, 1308–1323 (2016).
10. Villani, A. C. *et al.* *Science* **356**, eaah4573 (2017).
11. Lönnberg, T. *et al.* *Sci. Immunol.* **2**, eaal2192 (2017).
12. Tirosh, I. *et al.* *Science* **352**, 189–196 (2016).