CellPress

What Is Your Conceptual Definition of "Cell Type" in the Context of a Mature Organism?

What Is an Adult Cell Type, Really?

Hans Clevers
Hubrecht Institute

The human body is home to hundreds of cell types. Some are rather unobtrusive; others, like the cerebellar Purkinje cells, are outright spectacular. My personal favorites are the intestinal Paneth cell and the gastric chief cell. Classically, all cell types are considered to be hard-wired.

Each human cell represents the unique endpoint of a magical mystery tour that started from a single fertilized oocyte. This kaleidoscope of developmental trajectories of individual cells results somehow (yet with remarkably reproducibility) in a human body. Many decisions that cells take along their developmental journey are irreversible. But decisions taken toward the end of the journey are less definitive: cells in established tissues can exist in different maturation and activation states; they can even convert into other cell types of the same tissue.

My definition of "cell type" involves a description of hard-wired characteristics (e.g., "This cell is fated to be stomach epithelium"), combined with softer features such as actual morphology, gene expression pattern, location, function, and plasticity (e.g., "It has the defining features of a chief cell, located at the base of stomach glands. It secretes pepsinogen, yet it can become a stem cell upon damage"). Thus, the human body can be dissected into a limited number of hard-wired cellular "clouds," often coinciding with histologically defined tissues. At different time points, individual cells may occupy different positions within their respective clouds.

Defining Cell Type Space

Susanne Rafelski
Allen Institute for Cell Science

Canonical cell types, e.g., muscle and nerve, were originally defined by the functions of the tissues in which they reside, their unique morphologies, and their cytoplasmic architecture. Gene expression patterns provided molecular footing and permitted fine tuning into different cellular subtypes. However, considerable cell-to-cell variation in gene expression within a single cell type reflects the changing cellular microenvironment, various cell states (mitotic, migratory, etc.), and stochastic gene expression. This begins to blur the lines among cell types and introduces the notion of cell state.

However, some key features missing from this emerging view of cell types are the regulated, transient, and localized activities that determine cell behavior, and the structures that produce them. A useful way to classify cells might thus be a multiscale and multiparameter cell-type space that includes vectors for key intracellular organizational, dynamic, and functional features as well as tissue location, gene expression, etc. outlined above. This could result in a cell type classification system with some subpopulations of cells continuing to cluster into clearly separable groups while other cells might lie on a continuum due to small differences but distinct behaviors. It is intriguing to mull on how a discrete versus continuous nature of cell types relates to the fact that cells within a multi-cellular organism are tied to their tissue location and its environmental history. Clearly the definition of a cell type needs to be addressed with additional data that integrates cell organization with cell function and gene expression programs.

Cellular Demographies, Recorded

Michael Elowitz
Caltech

It seems to me that we are at the beginning of a paradigm shift on the issue of cell type. Once one starts looking at cells individually, in high dimensions and over time, previous notions of cell type don't always make sense. Cellular properties vary continuously as well as discretely, may not follow rigid hierarchies, and are highly dynamic. For this reason, I expect we will come to classify cells more like demographers classify people, without any singular, all-encompassing definition of "person type."

And as with people, history matters with cells. Where did the cell come from in the organism, and who is it related to? What external signals and molecular events did it experience? What might it become in the future? Knowing the answers to these questions would give us a better understanding of how the cell got to its current state, and how we might control it.

To access this information, several groups, including ours, are working on synthetic systems that enable cells to record their own histories in their genome. For example, together with Long Cai, we created a synthetic system called MEMOIR that enables cells to record lineage, signals, and other molecular events by modifying addressable, genomically integrated memory elements designed to enable subsequent in situ readout in single cells. These systems, together with other emerging single cell methods, allow us to envision a future in which we understand cell types not just as static molecular snapshots of individual cells but more broadly as cellular life histories unfolding in time.

Farewell, "Cell Type."

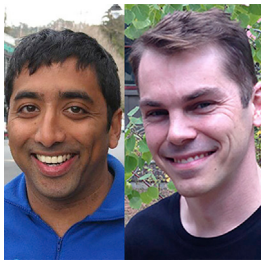
Allon Klein
Harvard Medical School

The concept of "cell type" is poorly defined and incredibly useful. We should keep it so and instead focus on defining coordinates underlying "cell states."

Cell types have historically been defined by morphology, ontogeny, function, or molecular composition. No single attribute has served for cell type classification. Yet "we know it when we see it." We are left with a functional but flawed taxonomy: functional, because it provides a language to describe biology; yet flawed, because it lacks consistency. We agree that neutrophils differ from basophils, but argue about how they overlap with myeloid derived suppressor cells. Moreover, differences within cell types can be as large as differences between cell types, as seen in comparing fibroblasts between tissues.

On cue, enter the "cell state:" as a community, we have embarked on describing each individual cell by its molecular composition. We are discovering coherent patterns in this data that serve to define "cell states." The hard work now begins in relating cell states to historically defined "cell types." We will witness arguments about whether novel cell states are in fact distinct "cell types," and whether different cell types in fact represent points in a continuum of states. We will also find cells that associate with two or more classical "cell types." This situation may be familiar: as we try to carve biology into manageable chunks (neurobiology, immunology, development), we find that there are no clean dividing lines. The organism is a continuum of interactions among different cells over time, and cells rest on a continuum manifold of cell states. Cell type divisions are useful, but we must remember that they are artificial, imposed for our convenience and not because biology needs them.

C. elegans Is a Test-Bed for Ideas



Jay Shendure and Cole Trapnell
University of Washington

As our field considers embarking on the task of comprehensively delineating human cell types, it is crucial that we identify a useful working definition for the slippery concept of "cell type." It is tempting to define cell types based on clustering of global molecular profiles of cells recovered from human tissues (e.g., by single-cell RNA-seq), but such an approach is unlikely to cleanly distinguish differences in cell type versus state (wherein a cell type is a collection of states among which cells can reversibly transition). However, if we are able to develop methods for querying the lineage history of each cell in conjunction with molecular profiling, we would at once know its developmental relationships to other cells, such that we can quantify how often transitions between putative states/types occur.

Because its complete cell lineage history is known and extensively annotated with functional information, *C. elegans* affords an outstanding opportunity to work through some of these issues. In particular, the worm would help us assess how much developmental data (i.e., lineage and state transitions) are needed to meaningfully classify cell types at the adult stage. If we cannot develop methods and agree on principles that support a cellular taxonomy for *C. elegans* and other simple model organisms, we are unlikely to be able to do it in humans!

Consider the Brain's Circuits

Ed Lein
Allen Institute for Brain Science

The neuroscience perspective on cell types is unique in that the brain is both an organ and a circuit. The concept of "cell type" is necessary to reduce the complexity of its >80 billion post-mitotic neurons and the functional circuits they form, but its precise definition has remained elusive. However, traditional approaches to neuronal classification rely on single-cell anatomy and physiology, which are typically qualitative and under-sampled. Transcriptomics has recently offered an unbiased, quantitative, and high-throughput alternative. It reveals a high degree of cellular diversity: cells whose relationships mirror developmental origin and whose molecular variation reflects their connectivity. As in immunology, lineage, and molecular classification are powerful organizing principles for brain cells.

Combining these concepts, a meaningful neuronal type could be considered a genetically encoded circuit element, while an elegant taxonomy should reflect cells' developmental origin, connectivity, and function. A testable definition of "cell type" is a set of cells with a common transcriptomic signature and low variation in other phenotypes (including connectivity). Its fundamental identity is a product of developmental lineage and defined by a progressive process of gene regulation. This includes intrinsic programs of differentiation and paracrine influences with the environment as the cells migrate, extend axonal and dendritic processes, and refine their synaptic connections. It is an exciting time where molecular tools are now available to test this idea at scale.

CellPress

Mapping as a Key First Step

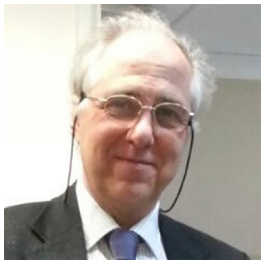
Emma Lundberg and Matthias Uhlen
KTH Royal Institute of Technology, Science for Life Laboratory

The definition of a cell type should be attributed to its specialized cellular function. In the context of a mature organism, this is both a difficult and ambiguous task. In order to classify phenotypically different cells, single cell assays to measure parameters that set them apart are needed.

The cells in the body differentiate with numerous intermediate steps. Large-scale studies have shown that the majority of genes are expressed in all cell types. Although there are cell types for which specific markers for recognition exist, such cells also show heterogeneous gene expression patterns. In addition, the proteome of a single cell is highly regulated in space and time to enable the cell to react to intrinsic or extrinsic factors. For example, we have demonstrated that the spatial distribution of proteins varies between cell types and that half of the human proteome localizes to multiple subcellular compartments. This increases the functionality of the proteome and the complexity of the cell from a systems perspective. Thus, we believe that a robust classification of a cell type, including variable cell states, needs to encompass a systems level view of the cell's function; comprising not only gene expression patterns but also information on the protein components and their activity, localization, post-translational modifications, interactions, and the wiring of the pathway activities in the cell.

Although the cell is more than its molecular building blocks, mapping the protein constituents of every cell type is a key first step for a comprehensive understanding of the human cell and its complex molecular machinery, as well as a major step toward modeling the mature organism.

In Search of Definitive Concepts



Alfonso Martinez-Arias
University of Cambridge

A mature organism is made up of many different kinds of cells whose individual physiology and molecular activities are tailored to their function, e.g., long-term stem cells that take care of repair upon need and short-lived cells that do exhausting work in the skin or the intestine. Your blood, for example, needs to produce 200 billion cells per day and contrasts with the central nervous system where many neurons last a life time. A cell should be defined for what it does, by its phenotype. However, biologists are suckers for techniques that we exploit to death. These days the arrival of single-cell transcriptomics has created a fad and, unconsciously, the thought that a cell can be defined in terms of the genes it expresses and by what people call their "epigenomes" is spreading. However, to define a cell in this manner is like defining a person by their looks; neither enough nor accurate and potentially misleading. The proteins that make up a cell and their relative locations might be a better descriptor but at the moment, we are far from reaching the level of sensitivity associated with nucleic acids. More importantly, we do not have ways to capture this information in a satisfactory manner. What we really need is some conceptual breakthrough that allows us to link the molecular make up of a cell to its physiology. Sometimes one feels that we are in a situation like the early days of physics, looking for the equivalents of kinetic and potential energy, work, free energy and how to link these to the variables which are genes and proteins. The answer then? We don't have the concepts, yet.

Moving Forward Despite Quarrels

Joshua R. Sanes
Harvard University

Although cell types would ideally be defined by function, that is still an unattainable goal for most neurons, so classification must be based on morphological, physiological, or molecular properties. Unfortunately, no neuronal group is truly homogeneous in any of these respects, and a few of my colleagues have seized on this fact to question whether neurons can really be divided into discrete types. To them, I offer two encouraging thoughts. First, work from my lab and many others makes a strong case that we can comprehensively classify neurons in one part of the brain, the mouse retina. So far, all neurons fit into discrete types with little evidence for intermediate types. Moreover, divisions based on molecular, morphological, and physiological criteria all lead to the same categories. I see no reason to think that other brain areas will obey different rules. Second, there is a related field from which we can learn: taxonomy. Like my quarrelsome colleagues, systematists continue to debate about how to define species and even whether they exist, but this has not stood in the way of their managing to preside over a successful enterprise. To this end, they avoid classification based on single features, use hierarchical rather than flat classification schemes, and accept that all groups are heterogeneous, focusing on discontinuous variation among groups rather than continuous variation within groups. The problems of defining species and neuronal cell types are similar in many ways, so perhaps we can adopt some of their common-sense rules.

Dynamic Cellular Personalities

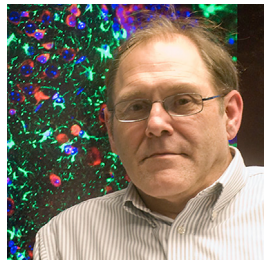
Paul Blainey
Broad Institute and Massachusetts Institute of Technology

I consider cells of a type to be those that are functionally equivalent. My particular view is that cells with different dynamic responses to the same stimulus are functionally distinct and therefore not the same type.

Cells are dynamic, with many context-dependent functional properties that influence both health and disease. Characterizing these cellular functions comprehensively requires a seemingly endless set of assays, many of them mutually incompatible. While the research community has developed new multi-omic methods to capture the molecular correlates of many cellular properties, it remains imperative that we make principled choices about which cellular functions are most relevant for each cellular genre and insist on high-quality assessments of these functions.

To delineate human cell types, we must account for the "business-as-usual" *homeostatic dynamic rhythms* in the functional properties of cells, such as how they interact with neighboring cells and surrounding matrix, their diurnal cycles, and their response to the changing nutritional and activity states of the organisms in which they reside. We must also account for the *dynamic functional responses* of cells to transient phenomena such as migration, replication, differentiation, and stresses including infection and inflammation. In so doing, we will need new *time-resolved* assays of single cells to capture cells' dynamic functionality in relevant biological contexts—and ultimately to advance our understanding of human health. When cells are functionally equivalent by these measures, we can describe them as the same cell type.

Are All Cells Unicorns?



James Eberwine
University of Pennsylvania Perelman School of Medicine

When Robert Hooke published the first description of a cell in 1665, he did so with the phrase "these pores, or cells….." He called these structures pores before naming them as cells, therefore implying function from form. In the interim more sensitive single-cell resolvable descriptors of cells have been developed. For example, RNaseq permits analysis of the 1000s of transcribed genes in a single cell and shows that no two cells are transcriptionally the same. Such "transcriptional phenotypes" show a cell's "potential" for making proteins, but not that it necessarily does so. There is hope that simultaneous measurement of varied static pictures of intracellular chemistries such as the epigenome, transcriptome, proteome, and metabolome will provide hierarchical cell descriptors that will better classify cells.

However, in most cases, it is the functional cell type that is of most interest. Functioning cells are not static, as coordinated interactions of multiple cellular constituents give rise to function. Therefore, the best descriptor of cell type may be a dynamical measure of the rates of the cellular constituent changes giving rise to particular physiologies. Indeed, a dynamical cell type definition or theory would account for the intercellular variability and plasticity that exists for cellular systems for which static measures are insufficient. Importantly, such a dynamical definition would explain work from my lab and others, which shows that cells can trans-differentiate from one functional cell type to another, thus revealing that a continuum of possible functional states exist for any cell. This highlights the need for a dynamical definition of cell type as a stop or pause on the continuum. This idea correlates, in part, with the ecological definition put forth by Junhyong Kim.

Cell Types as Ecological Guilds

Junhyong Kim
University of Pennsylvania

The "theory of types" has an unsettled discourse in many domains (e.g., philosophy, math, linguistics, etc.). In biology, a classic "type" problem is "What is a species?" Darwin's *On the Origin of Species*, after some complexities (e.g., "doubtful species") lands on a provisional definition of "a set of individuals closely resembling each other." Much later, G.G. Simpson defined species as "a lineage evolving… separately," emphasizing both lineage and selective process as key characteristics. E. Mayr provided a functional definition with his species concept of reproductively isolated groups, allowing, for example, empirical tests. But, as in other domains, none of these and myriad other definitions provide a satisfactory resolution (e.g., what about asexual organisms?).

For cell types, with so many new measurements, applying "closely resembling cells" is appealing but difficult, especially given widely different notions of "closeness" available from machine learning methods. Overall similarity as a criterion for defining cell type has a mixed history in terms of *robustness* (i.e., stability of these definitions vis-a-vis new measurements). Similarly, lineages are clearly important in cells, but often inconsistent with well-recognized cell types. As found in systematics, a universally satisfying definition of types is not likely to be possible. But a definition that is testable and maximally *predictive* of a cell's properties would be desirable. James Eberwine's idea of "dynamical program"-based definition might be an example. But, what cellular properties are important? I would argue, properties relevant to the system-level roles of a cell. A cell provides structure and it dynamically processes input environmental materials into outputs. The totality of its activity characterizes an ecological guild within the ecosystem of the organism. That is, we need an ecological definition of the cell type.

CellPress

New Technology Up-Ends Histology

J. Christopher Love
Koch Institute and Massachusetts Institute of Technology

To define a cell type requires first considering what a cell is. From one vantage, it is just a collection of parts: its genomic DNA, epigenome, RNA, proteins, metabolites, and other categorical elements together represent the most atomic form of self-replicating, living organisms. From another, it is a dynamic open system that operates under non-steady-state conditions. It samples and senses its noisy environment, and (sometimes) responds to those inputs, modifying its "state" on a variety of timescales (from seconds to days to weeks to years).

The notion of a "cell type" has arisen as a matter of convenience in cataloging, and of historically limited techniques for measurements. Pathological classification based on the cellular morphology or simple staining of fixed (and most certainly no longer functional) cells mounted on a slide has shaped our conceptualization of "cell types." With new technologies to measure both the fixed and dynamic "states" of single cells today, it is unclear whether or not this modest classification remains appropriate. Rather, an opportunity now presents itself to redefine the meaning of cell type to encompass the dynamic landscapes on which cells reside, including their plausible trajectories of phenotypic evolution given time or perturbation.