https://doi.org/10.1038/s41592-024-02307-5

Realization of cellomics to dive into the whole-body or whole-organ cell cloud

Tomoki T. Mitani, Etsuo A. Susaki, Katsuhiko Matsumoto & Hiroki R. Ueda

Tissues, organs and organ systems are composed of interacting cells (the cellome). We discuss the emergence of an omics approach that we refer to as cellomics. It enables cellome-wide analysis in whole-organ or whole-body specimens, based on advanced three-dimensional imaging and image analysis technology. We think that cellomics will pave the way for the incorporation of cellular, intercellular and spatial information across millions of cells in our body.

Since Robert Hooke's discovery of cells in cork, biological research has focused on the idea that living organisms are composed of cells as their primary unit of tissues and organs. Cellular networks consist of tissues, organs and organ systems, with homeostatic maintenance and disease traits attributed to multicellular systems and their constituents. To encompass spatial information on all cells in multicellular systems, a distinct -ome representation, called the cellome, is desirable. This Comment introduces the concept, methods and applications of cellomics, an omics workflow for analyzing the cellome.

Concept and workflow of cellomics

We defined cellomics as a part of omics technology, as it targets the cellome — the entirety of cells in an organ or body. Established omics approaches, such as genomics, examine the entire genome at the level of its discrete units (in the case of genomics, individual nucleotides). Similarly, cellomics investigates the entire cellome at the level of individual cells to link information on a specific trait to the cellular and intercellular information (Figs. 1a and 2a), which has already been practiced in *Caenorhabditis elegans*[1]. The cellome unit can be represented as each cell coordinate (the centroid of the cell body or nucleus) in the microscopy image covering a whole organ or a body. The coordinates may be associated with molecular information such as cell type, activity marker, protein expression (for example, of neuronal or immune cell type-related molecules, neuronal c-Fos, proliferation, cell death, inflammation or senescence-related molecules). In molecular-level omics such as proteomics or transcriptomics, the primary units (protein or RNA molecules) are perceived as discrete (that is, segmented and defined), but various modifications, such as phosphorylation, acetylation and methylation, make it clear that not all modified protein or RNA molecules are fully known. In this regard, cellomics also follows the principle of established omics in that cells are discrete entities, yet their properties are not fully known (for example, the definition of cell type).

We also propose cellomics as an approach to analyzing spatial context information, such as cell position, morphology, number, density, distributions and cell-to-cell interactions, across an organ or body. This approach differs from conventional omics technologies that focus on molecular contents within cells (Fig. 1b). Cellomics comprehensively analyzes spatial context information with adequate speed and coverage, focusing on minimal molecular content information. In this context, we view cellomics as a counterpart to single-cell sequencing technology, which is dedicated to the rapid and thorough analysis of molecular contents. While recent developments in spatial transcriptomics[2] have expanded molecular-targeting omics to include spatial context information, current spatial omics (for example, Visium) are based on processes for limited area of 2D sections (they collect molecular information with limited spatial information) and, as a trade-off, cannot cover the whole cellome across the three-dimensional (3D) organ. We assume that cellomics is used for population-level studies, which is in contrast to some current spatial omics projects covering a volume of a single sample at large cost (for example, MERFISH or Visium for a mouse brain hemisphere). Because of such complementary features, we believe that the significance of cellomics will increase as a result of the trend toward integrating the omics axes.

A standard omics workflow involves high-throughput collection and analysis of unit information, preparation and maintenance of reference data, and visualization and sharing of data. We illustrate corresponding cellomics procedures (Fig. 1c). Cells are labeled using genetic and histological techniques based on cell type or activity marker gene expression, or cell-to-cell interaction. Optical microscopy is used for data collection, with a micrometer-to-submicrometer resolution suitable for detecting cell-size objects. The position of labeled cells is calculated and aligned with reference data for multisample comparison. The analyzed data can be shared by integrating them into a single-cell-resolution cellome reference database.

Rapid imaging with sufficient resolution is crucial for collecting whole-cellome data. Light-sheet microscopy in combination with tissue clearing technologies can image an entire organ's cellome in a few hours[3]. Serial sectioning tomography systems, which combine physical serial sectioning with line-scan or spinning disc confocal microscopy, can achieve high speed and resolution[4,5]. Nevertheless, limitations in axial resolution pose a challenge, particularly in regions with high cell densities. Combining light-sheet microscopy with expansion microscopy could be a promising approach, achieving a practical axial resolution of ~5 µm. The detection of cellular nuclei in the granule cell layer of the cerebellum with sufficient accuracy has been demonstrated using this approach[3,6]. This technical specification supports precise cell detection in most organs, as most have cell densities similar to or less than that of the cerebellum. Efforts to form thinner light sheets for illumination, along with the combination with expansion microscopy protocols, push axial resolutions close to around 3 µm (refs. 7–9). This axial resolution could potentially approach or even surpass the

resolutions required for imaging exceptionally dense organs such as lymph nodes, which have two to four times the cell density of the cerebellar granule cell layer[10,11]. We anticipate that these next-generation 3D imaging technologies, which will have satisfactory imaging speed and resolution, will meet the requirement for omics-scale multisample data acquisition.

For subsequent data analysis, cellomics requires whole-organ or whole-body image processing, wherein cell positions are computationally digitized and aligned to corresponding cell coordinates in a whole-cellome reference (reference representing all cell coordinates within a specific organ or body) at both the global (the organ shape) and local (each cell coordinate) level. We see this process as analogous to genome-wide sequence data analysis with the corresponding reference genome, at both the global (the whole genome) and the local (the single base) levels[12]. The cell coordinate calculation (cell digitization) also requires substantial throughput to handle the entire cellome. Methods employing filter processing and point detection are fast and scalable[3], but may be less adept at intricate cell shapes. Learning-based algorithms are more versatile in detecting diverse cell morphologies, although they can be prone to overfitting and require extensive training[13]. Despite these challenges, recent software tool developments have advanced cell digitization, propelling cellomics forward.

Multisample comparison in cellomics can be achieved using image-based anatomical references with volumetric boundaries[14] or cellome references composed of cell coordinates[6]. The CUBIC single-cell resolution atlas (CUBIC-Atlas) for the whole mouse brain was an initial attempt to create the latter reference[6]. Digitized cell coordinates in the sample data are mapped to the CUBIC-Atlas globally (the step to register and align the entire organ image onto the reference image) and locally (the step to match each cell coordinate in the sample with a corresponding reference cell). This single-cell-resolution atlas has enabled the identification of cellular subgroups independent of human-defined anatomical regions[6]. This type of atlas, together with associated software, can also offer a database feature, allowing researchers to transfer and incorporate individual collected cellular information on cell type, cell activity and cell-to-cell interactions into the single-cell-resolution white map[15], a blank 3D point cloud that initially contains only coordinates of all cells without any annotations and can accommodate additional information related to each cell. We consider this function analogous to a whole-genome database such as the UCSC genome browser (https://genome.ucsc.edu/). Such databases integrate and distribute data among the research community to facilitate sharing, viewing, reanalysis and mining of the data collected by the corresponding omics scheme.

Applications of cellomics in large-scale comparative studies

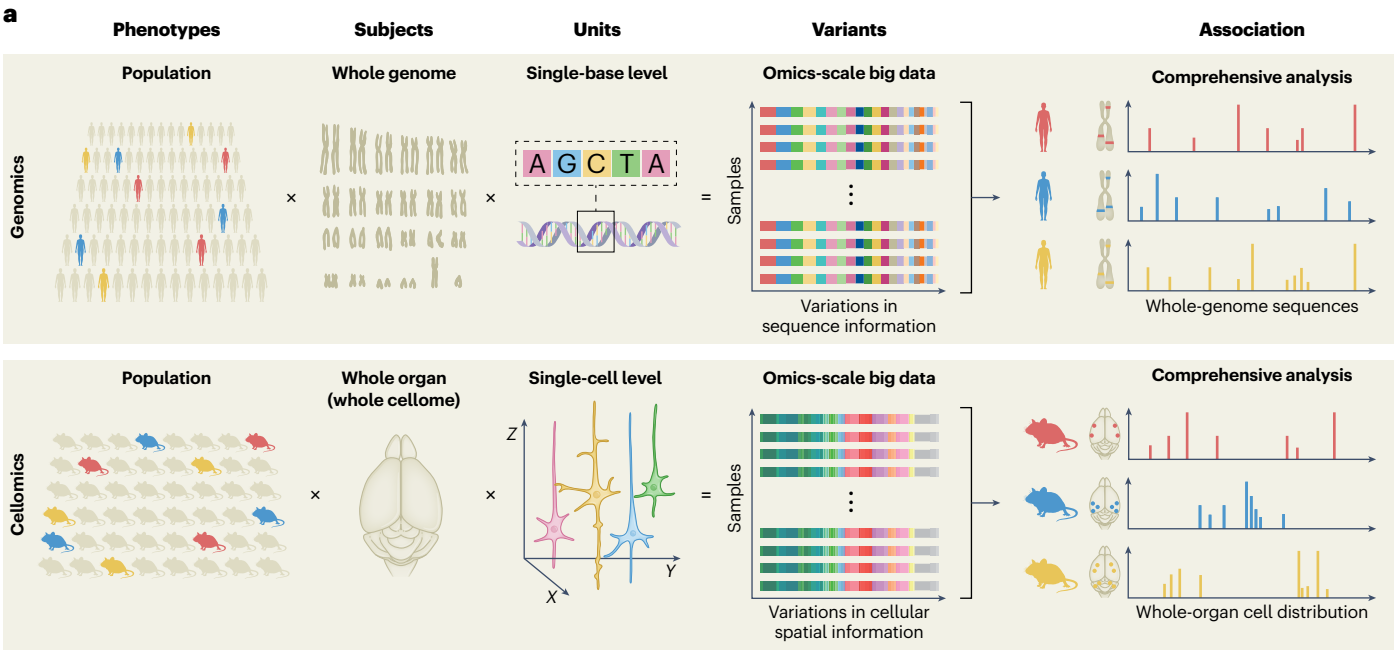
We foresee cellomics across an organ or a body enabling cellome-wide comparison and classification from various samples, eventually at a population scale (Fig. 2a). Here we consider examples for population-level cellome analysis in drug response[16–18], 3D pathology[19] and cross-species comparative studies.

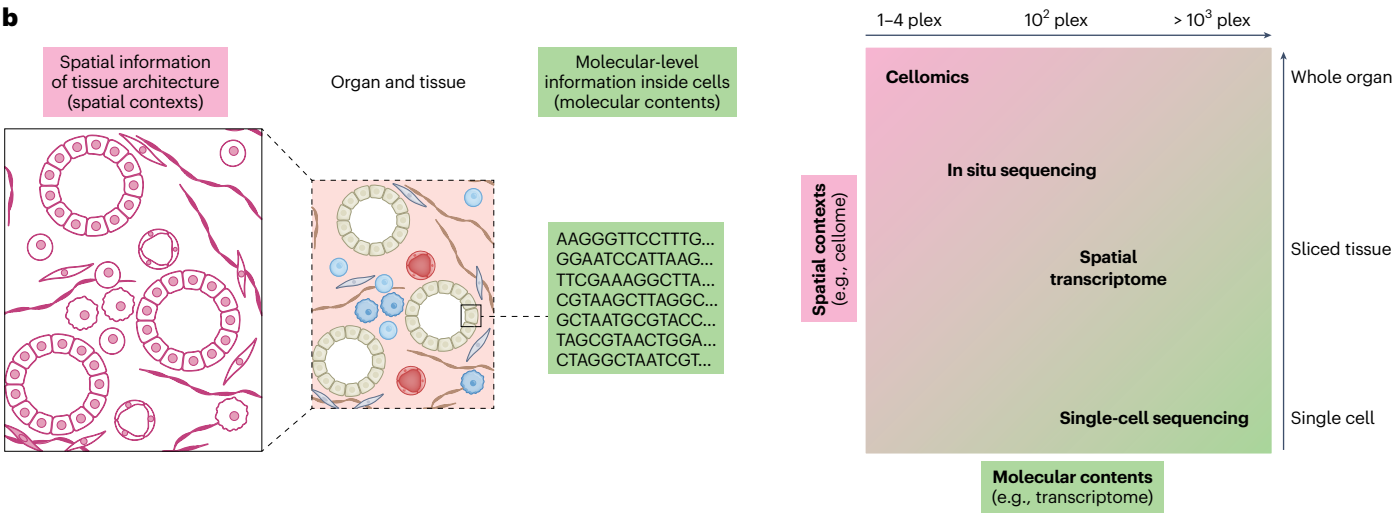
Integrating a drug–phenotype correlation panel (a collection of phenotypes induced by each drug) and a cellome–drug correlation panel (a collection of cell groups reacting to a specific drug administration; that is, a drug response cellome) from multiple drug administration experiments could yield information about a mutual cell population involved in the expression of a specific phenotype (a drug phenotype-specific cellome) (Fig. 2b). When conducting this type of study at a population scale, cellomics can support data-driven drug discovery to predict the drug's main and side effects and add biological meaning to the induced phenotype. The cellomics-based sensitive drug evaluation might be particularly helpful to select a more promising drug development pipeline during preclinical stages.

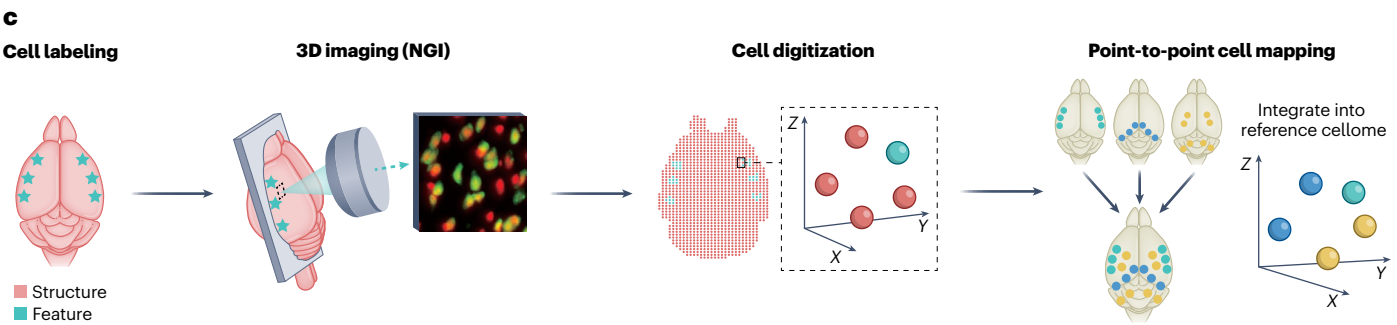
In digital pathology, the utility of 3D pathology in AI-based diagnosis has been investigated[20]. We also anticipate a future population-scale approach incorporating cellomics in pathology. This data-driven pathology can potentially contribute to revealing the association between a particular disease phenotype (for example, symptom and stage) and a particular cell group by integrating a patient–phenotype panel and a cellome–patient panel (Fig. 2c). This approach should be particularly beneficial for identifying cells and their functional alterations that lead to refractory diseases, such as neuropsychiatric and systemic autoimmune disorders. The cellomics pathology may thus provide evidence for developing new therapeutics that target disease-related cell populations.

Comparative genomics is a field of study that compares the genomic sequences of various species, supporting the discovery of evolutionarily relevant genes in the primate brain, for instance[21]. Similarly, comparative cellomics could be used to investigate cells and cell circuits that diverged during evolution in various species. In contrast to molecular networks, the structure and function of multicellular networks can directly implement mechanisms of organ- and organism-level phenotypes, such as sensing and cognition processes in the brain. As in the preceding examples, combining the species–phenotype and cellome–species panels facilitates population-scale cellomics. This approach can lead to the discovery of evolutionary associations between interspecies differences in a specific biological function and a unique multicellular circuit represented in the cellome (Fig. 2d).

Fig. 1 | The concept and possible workflow of cellomics. a, Conceptual analogies of genomics and cellomics. Both omics approaches examine omics-scale information at the unit level (single base or single cell). They can identify the relationship between the characteristics of omics units (such as a single-base mutation or single-cell functional change) and a particular trait (such as a disease phenotype) when the omics-scale data are obtained from a large-scale population. **b**, Left: a graphical representation of the biological information contained within tissues and organs. Biological systems include both spatial contexts and molecular contents. Cellomics targets the former information while established omics focus primarily on the latter category of information. Right: a diagram depicting two axes of biological information and their respective omics approaches. **c**, A possible cellomics workflow. Cell labeling using genetic and histological methods enables the visualization of specific cells. Next-generation 3D imaging (NGI) enables data collection from entire specimens with single-cell resolution. High-throughput serial sectioning and advanced light-sheet fluorescence microscopy with clearing and expansion microscopy techniques meet this requirement. Cell digitization is the process of converting image data into a 3D point cloud representing the coordinates of each cell, via image segmentation using filters or machine learning. Point-to-point cell mapping entails the registration and alignment of the entire organ image onto a reference, followed by the transfer of the sample cell characteristics (for example, genetically or histologically labeled cell type, cell activity and cell-to-cell interactions) to the corresponding reference cell.


a
Phenotypes
Subjects
Units
Variants
Association
Genomics
Population
Whole genome
Single-base level
Omics-scale big data
Comprehensive analysis
A G C T A
Samples
Variations in sequence information
Whole-genome sequences
Cellomics
Population
Whole organ (whole cellome)
Single-cell level
Omics-scale big data
Comprehensive analysis
Z
Y
X
Samples
Variations in cellular spatial information
Whole-organ cell distribution


b
Spatial information of tissue architecture (spatial contexts)
Organ and tissue
Molecular-level information inside cells (molecular contents)
AAGGGTTCCTTTG...
GGAATCCATTAAG...
TTCGAAAGGCTTA...
CGTAAGCTTAGGC...
GCTAATGCGTACC...
TAGCGTAACTGGA...
CTAGGCTAATCGT...
1–4 plex
10² plex
> 10³ plex
Cellomics
In situ sequencing
Spatial transcriptome
Single-cell sequencing
Spatial contexts (e.g., cellome)
Whole organ
Sliced tissue
Single cell
Molecular contents (e.g., transcriptome)


c
Cell labeling
3D imaging (NGI)
Cell digitization
Point-to-point cell mapping
Structure
Feature
Z
Y
X
Integrate into reference cellome

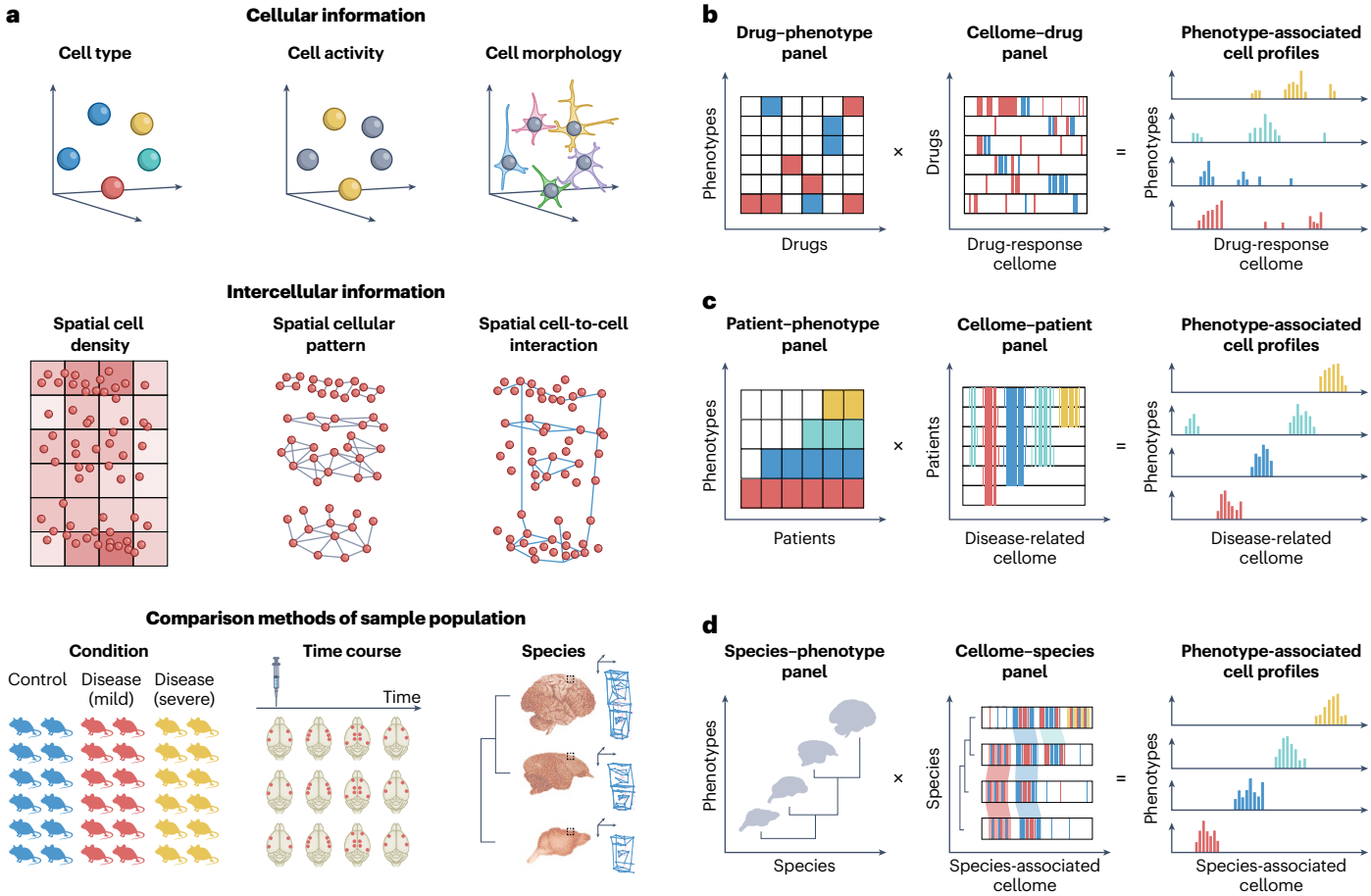


Fig. 2 | Application of cellomics. a, Various forms of information associated with cellomics data and their usage in a large-scale cellomics study. Cellomics enables the addition or deciphering of cellular and intercellular information to and from the cell coordinates across the specimen's space. Cell type and activity can typically be visualized by a marker gene or protein expression. Cell activity markers include molecules involved in activation, proliferation, inflammation, stress or cellular senescence. Spatial cell-to-cell interactions range from direct contact (for example, synaptic connections or cell–cell adhesion) and ligand–receptor binding to signaling via secreted proteins. Cellomics can be expanded to population-scale samples to compare and classify the cellome-associated information. **b**, A cellomics application in drug discovery and development. A group of systemic phenotypes for each drug administration form a drug–phenotype correlation panel. A drug–cell response correlation panel is created by combining various drug classes with whole-organ cell populations that react to various drugs (the drug-related cellome), each labeled with a cell activity marker. Combining these correlation panels from various drug administration studies can reveal a set of cell profiles associated with a particular phenotype. **c**, A cellomics application in 3D pathology. Each patient's phenotypes (for example, symptoms and stage) are gathered to create a patient–phenotype correlation panel. The cell populations of each patient marked by disease-related molecules, such as inflammation or cell death markers, are combined to create a patient–disease cell correlation panel. An integrated panel can offer phenotype-associated cell profiles for a particular disease symptom and progression stage. **d**, A cellomics application in a cross-species comparative study. A group of phenotypes from each species assembles to form a species–phenotype correlation panel. Gathering cellomes from various species (species-associated cellomes) produces a cross-species cellome correlation panel. Phenotype-associated cell profiles that express interspecies consistency or variation regarding a particular physiological function and its relevance to the cellome can be created by putting these correlation panels together.

Current limitations of cellomics

Cellomics has allowed comprehensive comparisons of whole organs across multiple specimens. Yet cell populations such as immune cells and cancer cells may not have a standardized place in a cellome reference. However, alterations in the distribution of these cells within a particular organs, where a cellome atlas can be defined, are comparable in the same global space. This comparability is made possible by cellome-wide data analysis, which involves sequential registrations and alignments at two distinct scales: locally, at the coordinate level of each cell; and globally, at the anatomical structure level of the organ. The primary objective of global alignment is to minimize the computational workload. Nevertheless, the application of global alignment enables quantitative comparisons of cells that do not possess fixed positions, including cancer and immune cells, across different specimens. Despite the difficulties associated with achieving a precise local alignment of these cells in the coordinate system of an organ (that is, cellome atlas), it is possible to align these cell populations on a global scale and compare them across samples. This is due to discernible variations in cell populations within a given organ, enabling quantitative assessments of sample cellomes from normal

and diseased organs. In vertebrates, however, the variation in cell position and type between individuals and species is an obvious obstacle when considering a more precise cell-to-cell comparison. Therefore, practical cell-to-cell alignment methods that take variability into account should be pursued[5,22]. Examples include the use of global optimization approaches, in which the spatial coordinate information and molecular expression profile of each cell are integrated as Bayesian probability distributions. The objective is to minimize the cost function, which evaluates the discrepancies or differences in cell alignment, as exemplified by studies on *C. elegans*[22]. Variability can arise from biological diversity but also from a lack of technical reproducibility. A standardized platform is needed for source data, code, user instructions, parameter values and methods for data compatibility and software reproducibility evaluation[20].

How cellomics could be applied to soft organs (for example, intestine) or multicellular samples (for example, organoids) without a fixed atlas is another potential issue. A possible solution is to extend the scheme to the atlas-free approach by extracting multidimensional data on cellular spatial distribution features. This strategy facilitates cellome-based multisample clustering by using hidden patterns in 3D cell distribution, similarly to dimension reduction methods in single-cell transcriptomics. While this approach can even be adapted to a limited tissue area, careful consideration of the observation volume (that is, the coverage size of the cellome) to accurately reflect each sample's unique spatial characteristics and variance would be required. In particular, for heterogeneous samples such as tumor masses or organoids, analyzing the entire sample is crucial for standardized comparison. However, the large 3D point clouds from cellome-scale coordinates may be more complicated to interpret than a simple 1D analysis of DNA, RNA or protein sequences. Point cloud informatics can decipher these intricate patterns and extract feature information effectively[23]. The field is currently focused on point clouds on object surfaces in engineering, while the future holds the promise of expanded interdisciplinary collaborations. Cellomics may drive the development of new computational tools as research into mammalian cellomes advances.

Conclusion and perspective

We envision cellomics to resolve spatial context information from organ- or body-scale single-cell resolution data. We believe that this new omics has the potential to serve researchers examining the 100 million cells in the mouse brain[2], 200 billion cells in the human brain and 30 trillion cells in the human body as an integrated platform of spatial context information. The platform aspires to eventually support the integration of molecular content information from all cells in an organ and a body, potentially using large-scale resources such as the Human Cell Atlas[24]. Population-scale cellomics studies could assist in spatially exploring important cells in physiological and pathological conditions, potentially enabling broader investigations. The number of analyzed molecules is a current limitation; this limitation would be partially overcome by the application of multiplexing technologies[25,26]. We expect cellomics to substantially impact numerous biomedical disciplines, similarly to other established omics.

Tomoki T. Mitani[1,2,3,7], Etsuo A. Susaki[1,4,5,7], Katsuhiko Matsumoto[1,6] & Hiroki R. Ueda[1,6]

[1]Laboratory for Synthetic Biology, RIKEN Center for Biosystems Dynamics Research, Osaka, Japan. [2]Department of Systems Biology, Graduate School of Medicine, Osaka University, Osaka, Japan. [3]Department of Neurology, Graduate School of Medicine, Osaka University, Osaka, Japan. [4]Department of Biochemistry and Systems Biomedicine, Juntendo University Graduate School of Medicine, Tokyo, Japan. [5]Nakatani Biomedical Spatialomics Hub, Juntendo University Graduate School of Medicine, Tokyo, Japan. [6]Department of Systems Pharmacology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. [7]These authors contributed equally: Tomoki T. Mitani, Etsuo A. Susaki.

✉e-mail: uedah-tky@umin.ac.jp

Published online: 13 June 2024

References

1. Aoki, W. et al. *Sci. Rep.* **8**, 10380 (2018).
2. Moses, L. & Pachter, L. *Nat. Methods* **19**, 534–546 (2022).
3. Matsumoto, K. et al. *Nat. Protoc.* **14**, 3506–3537 (2019).
4. Seiriki, K. et al. *Neuron* **94**, 1085–1100.e6 (2017).
5. Economo, M. N. et al. *eLife* **5**, e10566 (2016).
6. Murakami, T. C. et al. *Nat. Neurosci.* **21**, 625–637 (2018).
7. Chakraborty, T. et al. *Nat. Methods* **16**, 1109–1113 (2019).
8. Glaser, A. K. et al. *Nat. Methods* **19**, 613–619 (2022).
9. Glaser, A. et al. *bioRxiv* 2023.06.08.544277 (2023). https://doi.org/10.1101/2023.06.08.544277
10. Keller, D., Erö, C. & Markram, H. *Front. Neuroanat.* **12**, 83 (2018).
11. Cupedo, T., Stroock, A. & Coles, M. *Front. Immunol.* **3**, 343 (2012).
12. Langmead, B. & Salzberg, S. L. *Nat. Methods* **9**, 357–359 (2012).
13. Laine, R. F., Arganda-Carreras, I., Henriques, R. & Jacquemet, G. *Nat. Methods* **18**, 1136–1144 (2021).
14. Wang, Q. et al. *Cell* **181**, 936–953.e20 (2020).
15. Rood, J. E. et al. *Cell* **179**, 1455–1467 (2019).
16. Mano, T. et al. *Cell Rep. Methods* **1**, 100038 (2021).
17. Renier, N. et al. *Cell* **165**, 1789–1802 (2016).
18. Pan, C. et al. *Cell* **179**, 1661–1676.e19 (2019).
19. Glaser, A. K. et al. *Nat. Biomed. Eng.* **1**, 0084 (2017).
20. Liu, J. T. C. et al. *Nat. Biomed. Eng.* **5**, 203–218 (2021).
21. Enard, W. et al. *Nature* **418**, 869–872 (2002).
22. Bubnis, G., Ban, S., DiFranco, M. D. & Kato, S. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1903.09227 (2019).
23. Guo, Y. et al. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 4338–4364 (2021).
24. Regev, A. et al. *eLife* **6**, e27041 (2017).
25. Kishi, J. Y. et al. *Nat. Methods* **16**, 533–544 (2019).
26. Goltsev, Y. et al. *Cell* **174**, 968–981.e15 (2018).

Acknowledgements

We thank T. Mano (OIST) and the Ueda lab members, particularly S. Harada, K. Yamaura and F. Kinoshita, for their advice in enhancing the manuscript. This study was supported by grants from the JST ERATO JPMJER2001 (to H.R.U.); JST CREST JPMJCR23B7 (to E.A.S.); MEXT QLEAP JPMXS0120330644 (to H.R.U.); JST Moonshot R&D JPMJMS2023 (to K.M.); AMED/MEXT JPMJER2001 (to H.R.U.); JP22ama221517, JP20gm6210027, JP21ak0101181 and JP21wm0425003 (to E.A.S.); the JSPS KAKENHI JP18H05270 (to H.R.U.), 20K06885 (to K.M.), 22H02824, 22H04926 and 23K20044 (to E.A.S.), and 20K16498 (to T.T.M.); grants-in-aid from the Human Frontier Science Program (to H.R.U.);Operating Costs Subsidies for Private Universities (to E.A.S.); Takeda Science Foundation, UTEC-UTokyo, Uehara Memorial Foundation and Nakatani Foundation (to E.A.S.).

Author contributions

E.A.S. and H.R.U. came up with the concept of cellomics. T.T.M. and E.A.S. prepared the manuscript and figures. K.M. and H.R.U. were involved in revising the content and manuscript.

Competing interests

CUBICStars Inc., which H.R.U. founded and where T.T.M., K.M. and E.A.S. are employed, has filed patents regarding CUBIC-HV and CUBIC-Cloud. The company also provides CUBIC-Cloud web services. H.R.U., E.A.S., T.T.M. and K.M. are co-inventors on patents and patent applications owned or filed by RIKEN covering the CUBIC reagents and/or MOVIE. E.A.S. receives collaboration funding from Kantum Ushikata Co., LTD., that is related to the development of devices for cellomics.

Additional information

Peer review information *Nature Methods* thanks Raju Tomer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.