The Human Biomolecular Atlas Program: mapping human structure in more than 37 trillion dimensions

Graham McLeod MD*

Max Rady College of Medicine, University of Manitoba
727 McDermot Avenue, Winnipeg, R3E 3P5

Abstract

Biologists wielding transformative technologies can describe the human body in increasingly complex ways. Single-cell gene expression and metabolic profiles, and three-dimensional maps of molecular distribution in tissue, are increasingly feasible. It is hoped that such high-definition analyses could better illuminate disease states to bring new insights about pathophysiology and natural history of disease processes. New insights, in turn, could help refine clinical descriptions of disease (e.g. revise diagnostic criteria) and contribute to generating new clinical therapies. But despite advancing biotechnology and accumulating knowledge of human biology, the enhanced characterization of disease states, especially from a clinically-salient perspective that must beware multiple aspects of disease, has been limited by lack of multi-disciplinary, interoperable framework to capture, integrate, and disseminate knowledge. The Fund Human Biomolecular Atlas Program (HuBMAP) is an ambitious effort to create such a framework. HubMAP Is an NIH-funded initiative with a nine-figure budget and several-year timeline to produce a three-dimensional "atlas" that will characterize the human body at cellular and molecular levels of detail, in health and in many disease states. Here I provide an overview of the HubMAP Consortium and describe its potential impacts on clinical medicine.

What is the precise physical nature of the human body? For centuries, science has revealed the body bit by bit, unconcealing cells that comprise tissue, molecules that form cells, and quarks composing atoms and molecules. With human biology increasingly revealed to be highly complex, distinct disciplines of analysis have emerged, each using distinct schemas and methodologies (e.g. structural biochemistry versus developmental neuroscience) such that expertise is increasingly siloed. With hypotheses generated and tested by highly specialized experts, there has been a lack of a unified, interoperable framework to integrate new knowledge, and no easy means to study transdisciplinary phenomena such as large-scale biological networks spanning multiple orders of magnitude and with effects in multiple research disciplines.[1,2] Now an exciting new initiative seeks to address these issues by building an open-access, global framework to create an exquisitely detailed, multifaceted "map" of the human body, answering the question "What's going on inside the human body?" in unprecedented scope, depth, and detail.

The NIH Common Fund Human Biomolecular Atlas Program (HuBMAP) seeks to develop "a comprehensive, accessible three-dimensional molecular and cellular atlas of the human body, in health and under various disease conditions."[3] Supported by an estimated $200 million USD over eight years,[4] the HuBMAP Consortium will assemble biologists of all kinds (molecular, cellular, developmental, computational), clinicians, software engineers, data scientists, and other experts to pursue this goal of characterizing new intricacies of physical humanity. HuBMAP is not the first effort to map the human body; other ongoing initiatives include the LifeTime initiative,[5] the Human Protein Atlas,[6] and the Human Cell Atlas[7] (which already involves 1500 researchers in 65 countries).[4] Notably, HuBMAP will collaborate with each of these contemporary mapping initiatives, and will also work alongside existent NIH-funded consortia dedicated to mapping specific organs, including LungMAP,[8] the Kidney Precision Medicine Project (KPMP),[9] and the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative.[10] Findings from these targeted efforts will be integrated into HuBMAP's interoperable and comprehensive framework,[11] alongside results of new data that HuBMap will generate. As it progresses, HuBMAP intends to release its findings iteratively into open-source online portals, with the first outputs expected in late 2020.[1]

How will HuBMAP work towards its goal? The inaugural publication in Nature[3] outlines a simultaneous three-part approach. The first facet specifically focuses on the development and implementation of transformative new technologies, including technologies to obtain

*correspondence to: graham.mcleod@ucalgary.ca

high-throughput single cell data and to capture three-dimensional spatial data of the distribution of important biomolecules within cells and amidst tissues. In a second and simultaneous facet, HuBMAP's "Tissue Mapping Centers" will generate single cell "omic" data (e.g. genomic, epigenomic, transcriptomic, proteomic, or metabolomic) and obtain three-dimensional spatial data of biomolecular distribution. These efforts will rely on existing technologies (see Stuart & Satija for an excellent recent review of novel single cell analytic techniques[12]) and new forthcoming technologies generated by HuBMAP contributors. In its third facet, HuBMAP will depend on its "HIVE" teams — HuBMAP integrations, visualization, and engagement (HIVE) — to organize, collate, and centralize data, and render accumulated information useable and accessible via friendly interfaces. It is intended that each of the three facets will work closely together, refining and releasing HuBMAP in improving, successive iterations.

As it is completed over the coming years, HuBMAP will bring many benefits to the biomedical community at large. One major benefit will be access to a detailed baseline framework for healthy human systems and select disease states. HuBMAP intends to resolve the human body at the single-cell level across dimensions such as genomics, epigenomics, transcriptomics, proteomics, and metabolomics.[1] In reference to this multi-omic framework, the precise molecular and cellular phenotypic deviations of pathology can be better characterized, leading to transformative progress in pre-clinical research efforts. This level of detail is crucial because there can be considerable heterogeneity between two cells nominally of the same type, even within the same tissue.[13] Two such cells, if in different functional activation states, may play vastly different roles in the pathophysiology of disease states, and may behave entirely differently in response to endogenous homeostatic events and exogenous therapeutics. HuBMAP's multi-omic single-cell analyses will allow disease-to-health comparison at the single-cell level, leading to better understanding of disease, better characterization of specific therapeutic targets, and exquisite frameworks within which to validate putative therapeutics.

Other early benefits of HuBMAP will be towards biomedical research. In generating tissue-wide, three-dimensional molecular maps, HuBMAP will allow novel analyses of cell-cell networks, intracellular localizations, behaviour of the extracellular matrix (ECM), and characterization of cell-ECM networks. The ECM plays a significant role in tissues; for example, the brain is 10-20% ECM by weight.[14] Moreover, the ECM is a crucial modulator of cell behaviour.[15] However, the ECM is historically challenging to study due to its broad spatial distribution and intricate molecular networks.[13] The tools and specific efforts of HuBMAP will help overcome this, possibly with profound contributions to expert understandings of biology. Further, HuBMAP is well-poised to characterize the role of motile effectors, such as immunologic cells, which may exert important yet difficult-to-characterize effects in regional tissue areas.[13] HuBMAP also aims to develop transformative new technologies that will be open for use by other researchers, such as new techniques for single-cell "omic" analyses, and spatially mapping molecular networks in three-dimensional tissue space.[3]

What will be the specific clinical benefits of HuBMAP? As discussed, HuBMAP may give rise to new waves of precision medicine treatments. However, it may be years before these new therapies are established. In the meantime, clinicians will have access to HuBMAP's centralized atlas framework, which may be a useful reference to consult, especially as disease states are increasingly mapped. The HuBMAP atlas may be especially useful for medical students, residents, and other learners; for established clinicians expanding the scope of their practice; for clinicians at all stages embarking on a new research interest; for clinicians facing licensure examination due to geographical relocation; or for any clinician wishing to update biological knowledge. Additionally, clinicians will also have the ability to refer patients to HuBMAP as a resource to learn more about their body in health (this may be especially of interest in primary care, or for preventive medicine) or disease (for any curious patient). It is conceivable that easy availability of biological information in HuBMAP, accessible in open-source portals,1 may raise public consciousness of the body as a machine, directly and indirectly educate patients about the positive and negative consequences of certain modifiable risk factors, and overall encourage health-promoting behaviours in the public or some subset thereof. Altogether, as HuBMAP grows, it will be an excellent clinical resource for healthcare professionals and patients alike.

Although its benefits are noteworthy, to succeed, the HuBMAP Consortium must overcome manifold challenges. HuBMAP must integrate vast amounts of data in various modalities, collected from various donor tissues, then processed and analyzed in numerous different laboratories in various locations. This multi-dimensional variability will present significant technical challenges. Beyond this, there are further technical challenges of creating new technologies and analytic techniques. HuBMAP is intending to pioneer new single-cell analytic techniques, three-dimensional molecular mapping techniques, and integration and visualization tools, to transform raw data into a comprehensive, three-dimensional, open-access, user-friendly atlas. Achieving these ambitious aims, especially under finite budgets and within anticipated time scales (NIH funding of HuBMAP currently scheduled until 2026), will require significant ingenuity on the part of the HuBMAP contributors, who will have to work efficiently, cost-effectively, and strategically.

A putative comprehensive atlas of the human body must also contend with the challenge of innate biological variability from one person to the next. To the extent that human beings are biologically heterogeneous, even in healthy states, a single atlas is unlikely to be ad-

equate to represent the human population as a whole. This problem has already been encountered in efforts to produce a standard human reference genome. It is now known that "thousands of DNA sequences of various lengths" that are not found in extant human reference genomes are present in various populations worldwide.[16] Thus, there cannot be a single "reference" human genome. Likewise, a unique single cell-resolved "reference" human atlas would likely be inadequate. Time will tell how HuBMAP plans to overcome the challenge of innate human heterogeneity.

In his famous 1910 address "Man's Redemption of Man" at the University of Edinburgh, celebrated physician Sir William Osler rejoiced that "the leaves of the tree of science have availed for the healing of the nations" and that he considered this the "greatest glory" of humankind.[17] To the extent that HuBMAP succeeds, the tree of biological sciences will grow considerably; its leaves will be more numerous and better understood, and clinicians, scientists, and patients alike may avail themselves of its fruits. It may take years for the work of HuBMAP to pay direct clinical dividends in forms of improved diagnostic categories and targeted new therapeutics, but if HuBMAP can help decipher some of the complexity of the human body, there is grounds for a specific optimism that the future of medicine can be better than the past. Moreover, humanity will be moved leaps and bounds closer towards realizing one of our longest-lived dreams: to know what's going on inside our bodies, and understand this magnificent structure with which we live our lives.

References

[1] Sara Green, Maria Serban, Raphael Scholl, Nicholaos Jones, Ingo Brigandt, and William Bechtel. Network analyses in systems biology: new strategies for dealing with biological complexity. *Synthese*, 195:1751–1777, 2018.

[2] Peter Schuster. Networks in biology: Handling biological complexity requires novel inputs into network theory. *Complexity*, 16:6–9, 2011.

[3] The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature*, 574:187–192, 2019.

[4] Elizabeth Pennisi. An ambitious effort to map the human body's individual cells gets backing from nih. *Sci. Mag.*, 2019.

[5] LifeTime Consortium. The lifetime initiative, 2020.

[6] Peter J Thul and Cecilia Lindskog. The human protein atlas: A spatial map of the human proteome. *Protein Sci.*, 27, 2018.

[7] Aviv Regev and et al. The human cell atlas. *Elife*, 6, 2017.

[8] Maryanne E Ardini-Polekse and et. al. Lungmap: The molecular atlas of lung development program. *Am. J. Physiol. Cell. Mol. Physiol.*, 313:L733–L740, 2017.

[9] Jenna M Norton, Christian J Ketchum, Andrew S Narva, Robert A Star, and Griffin P Rodgers. Complementary initiatives from the niddk to advance kidney health. *Clin J Am Soc Nephrol.*, 12, 2017.

[10] Meghan C Mott, Joshua A Gordon, and Walter J Koroshetz. The nih brain initiative: Advancing neurotechnologies, integrating disciplines. *PLOS Biol.*, 16, 2018.

[11] HuBMAP - General Frequently Asked Questions - NIH Common Fund.

[12] Tim Stuart and Rahul Satija. Integrative single-cell analysis. *Nature Reviews Genetics*, 20:257–272, 2019.

[13] Ananda L Roy and Richard S Conroy. Toward mapping the human body at a cellular resolution. *Molecular Biology of the Cell*, 29:722–729, 2013.

[14] Lorraine W Lau, Rowena Cua, Michael B Keough, Sarah Haylock-Jacobs, and V Wee Yong. Pathophysiology of the brain extracellular matrix: A new target for remyelination. *Nature Science Reviews*, 14:722–729, 2013.

[15] Dirk Hubmacher and Suneel S Apte. The biology of the extracellular matrix: Novel insights. *Current Opinion in Rheumatology*, 25:65–70, 2013.

[16] Rachel M Sherma and Steven L Salzberg. Pan-genomics in the human genome era. *Nat. Rev. Genet.*, pages 1–12, 2020.

[17] Wi Osler. Man's redemption of man: A lay sermon. 1910.