

Research highlight

Open Access

## Towards a primate single-cell atlas

Biology in the 21<sup>st</sup> century is shifting substantially towards single-cell analysis. Mammalian tissues are highly heterogeneous and contain multiple cell types in different states whose tight coordination determines overall function. Bulk measurements, including RNA sequencing, provide average values, which may dilute cell-specific effects or overlook rare effects of functional importance. This greatly limits our understanding of mammalian physiology and disease. For example, the histological structures of the cortical and medullary regions of the kidney differ substantially, and their cells exhibit distinct characteristics related to environmental gradients such as oxygenation, glomerular filtrate-induced shear stress, and solute concentrations within the nephron. Furthermore, not all nephrons, mesenchymal cells, or vascular cells function in the same way, especially in the context of disease and aging. Therefore, analysis results are highly dependent on the region being sampled. This issue can be partially addressed by bulk profiling of isolated cell populations (e.g., flow cytometry sorting) (Krausgruber et al., 2020). Although relevant, this approach typically relies on surface markers, which suggests that cells lacking or displaying low levels of a specific marker will be absent from analysis. The generation of single-cell suspensions can also induce stress, which may modify gene expression and limit the ability to distinguish cell states (e.g., activation, damage, or cell fate transition). The rapid advances in high-throughput single-cell RNA sequencing (scRNA-seq) technologies underpin a scientific revolution, enabling researchers to study tissue cell heterogeneity with unprecedented resolution. This has led to the generation of increasingly refined single-cell taxonomic atlases (whole-body and tissue-specific) for a variety of species ranging from invertebrates to mammals, including rodents and humans (Han et al., 2018; Han et al., 2020; Tabula Muris et al., 2018). Recent efforts in human settings (Cao et al., 2020; Dominguez Conde et al., 2022; Eraslan et al., 2022; Tabula Sapiens et al., 2022) have greatly expanded the number of tissues tested and cells profiled. This

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2022 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

information can be used to understand cell-cell communication networks and predict disease susceptibility or drug response by mapping the expression of disease-causing genes and drug targets expressing certain cell types. However, realizing the full potential of a human single-cell atlas will require coverage of different ethnic groups, ages, sexes, environmental conditions, and diseases. This complexity can be overwhelming considering the difficulties associated with procuring human samples of optimal quality and samples representing different time points of a process rather than the end stage.

The use of animal models is a long-standing practice in biomedical research, providing valuable information regarding anatomy, histology, physiology, and disease pathophysiology. Research has greatly relied on small animal models, including worms (*Caenorhabditis elegans*), fruit flies (*Drosophila melanogaster*), zebrafish (*Danio rerio*), and rodents, as they offer the easiest and most cost-effective way to maintain high-breeding animals long-term. However, while these models have proven very useful for dissecting a large range of biological mechanisms, they often fail to mimic specific aspects of human disease (Seok et al., 2013). Although problematic for all diseases, this issue is exemplified in the study of central nervous system diseases, in which rodents often exhibit phenotypes that differ from humans. Non-human primates (NHPs) and humans diverged from a common ancestor as recently as 25–30 million years ago. Consequently, NHPs are physiologically and behaviorally very similar to humans, and thus represent a potentially advantageous model for biomedical research. NHPs also provide an important opportunity to bridge many of the current gaps in generating a comprehensive single-cell atlas relevant to understanding human biology and disease.

We recently selected the cynomolgus monkey (*Macaca fascicularis*) to generate an adult NHP cell atlas (Han et al., 2022) (Figure 1). While previous reports have described scRNA-seq analysis of specific monkey tissues, comprehensive mapping of a single NHP species using the same technical platform has not yet been performed. The

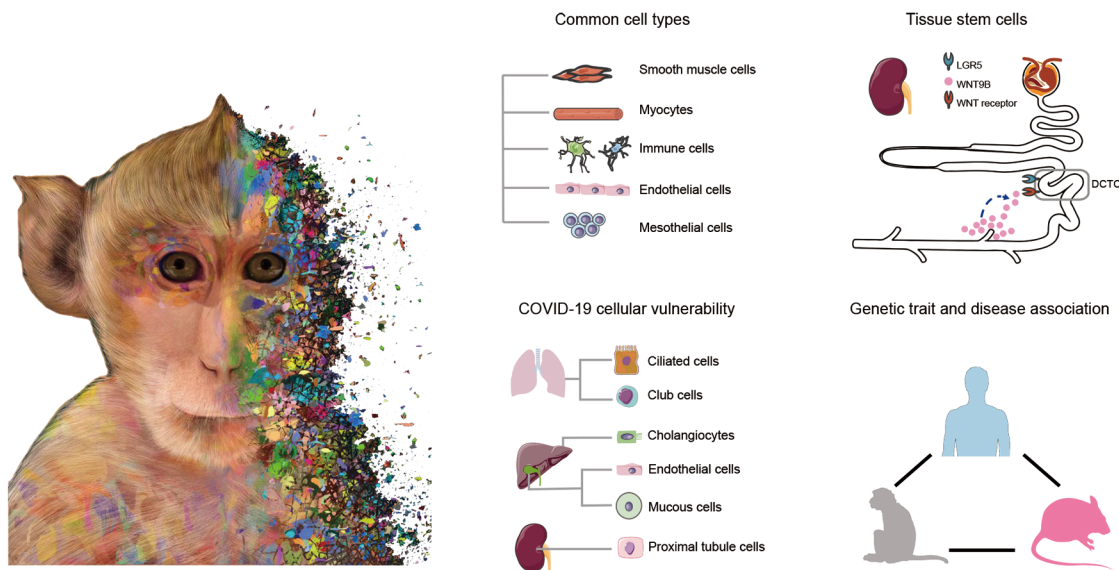
Received: 05 July 2022; Accepted: 15 July 2022; Online: 15 July 2022  
Foundation items: Work on this topic in the Esteban Laboratory at the Guangzhou Institutes of Biomedicine and Health was supported by the National Natural Science Foundation of China (92068106)

latter consideration is important because comparing datasets generated with different scRNA-seq methodologies can introduce bias. *Cynomolgus* monkeys are broadly distributed across Southeast Asia and are one of the most widely used NHPs in biomedical research (e.g., vaccine development, toxicological testing, disease modeling, developmental biology, and aging) (Liu et al., 2021). There are special ethical considerations for NHP research and China's regulatory approach is in line with international expectations. We isolated single nuclei and cells from 45 tissues from a cohort of eight monkeys (male and female) and applied droplet-based analysis, obtaining 1.14 million single-cell/single-nucleus transcriptomes. From this, we identified 113 major cell types belonging to all major systems, i.e., cardiovascular, adipose, digestive, endocrine, nervous, reproductive, respiratory, skeletal, and urinary. Most tissues were profiled as single nuclei rather than cells to avoid any potential bias in cell types captured and any potential changes in gene expression resulting from enzymatic tissue digestion. This was critical to achieve the correct representation of specific cell types in different tissues. For example, single-nucleus profiling of the monkey liver yielded almost 80% hepatocytes, recapitulating the correct proportion in mammalian livers. It also allowed the capture of otherwise difficult-to-profile cells such as adipocytes and skeletal muscle fibers.

To demonstrate the power of our monkey dataset in transforming our current understanding of primate tissue function, we explored the molecular signatures of common cell types shared by various tissues across the body, including stromal, endothelial, mesothelial smooth muscle and myeloid cells (macrophages and microglia), adipocytes, and skeletal muscle myonuclei. Consistent with previous studies

(Dominguez Conde et al., 2022; Kalucka et al., 2020), this approach highlighted tissue-specific differences and diversity within specific cell populations. For mesothelial cells, comprehensive annotation resulted in the identification of a subcluster of cells with immune properties, consistent with the observation that structural cells exhibit immune properties (Krausgruber et al., 2020). These results have biomedical implications as they suggest the presence of a cell population within visceral adipose tissue that may modulate responses triggered by gut bacteria (Ha et al., 2020).

Similarly, we performed a comprehensive investigation of cell-cell interactions at the body level, mapping the expression of key components of the Wnt pathway, which is essential for development as well as homeostasis and growth in several tissues in adulthood (Clevers, 2006). We focused on leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5), a Wnt signaling amplifier used to mark and isolate progenitor cells from mammalian organs (Barker, 2014). LGR5 mapping studies in mice have primarily relied on the use of genetically engineered reporters, as moderate expression levels and limited expressing cells make it difficult to study effectively with antibodies (Barker et al., 2013). Consequently, LGR5 distribution has not been comprehensively studied in larger mammals. Interestingly, among other unexpected locations, we found that LGR5 was enriched in distal convoluted tubule cells (DCTC) from the kidney and, to a lesser extent, in other parts of the nephron. Furthermore, we observed expression of the LGR5 ligand RSPO1 in kidney myofibroblasts and WNT9B in collecting duct cells. Notably, LGR5 itself is a Wnt target and WNT9B is implicated in nephrogenesis during fish and amphibian development and regeneration (Karner et al., 2009). In addition, Wnt factors act locally over short distances,



**Figure 1** Generation of an adult NHP cell atlas

NHP cell atlas composed of single-cell/single-nucleus transcriptomes of 1.14 million cells (nuclei), covering 113 major cell types across 45 tissues. Cross-tissue comparison of cell types and molecular features revealed tissue specificity in common cell types, cellular vulnerability to COVID-19, tissue stem cells, and cell-type association with complex human traits and monogenic heritable diseases. The monkey drawing was generated by Ming-Jiang Lu and Rui Zeng (BGI-Shenzhen). Part of this diagram was drawn using free-to-use artwork from Servier Medical Art (<https://smart.servier.com/>).

and the distal convoluted tubule is the anatomical site in the nephron closest to the collecting duct. To assess whether these findings are unique to monkeys, we performed comparative analysis using scRNA-seq datasets from other species, with *LGR5* and *WNT9B* being barely detected in the human and mouse kidney cells compared to monkey. However, *LGR5* is robustly expressed during mouse and human development and is instrumental in kidney formation (Barker et al., 2012). These findings suggest that epithelial cells with progenitor or homeostatic capacity may be present in the monkey distal convoluted tubule but have been lost in humans during evolutionary divergence. If true, this could have important biomedical implications, as mammalian kidneys have a lower ability to regenerate (in particular, there is no new nephrogenesis) when damaged compared to other vertebrates such as fish and amphibians (Chang-Panesso & Humphreys, 2017). By understanding the underlying mechanisms, it may be possible to re-awaken similar capacities in humans. Although our observations were validated through single-molecule fluorescence *in situ* hybridization, caution should be exercised when interpreting such findings, as further analysis is warranted to demonstrate the real nature of *LGR5*<sup>+</sup> DCTC. A potential approach to evaluate the functional relevance of these cells in monkeys may be scRNA-seq after controlled acute kidney injury.

Our large-scale monkey cell atlas is also important for advancing knowledge about disease pathogenesis. Given the global pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Wong et al., 2020), we used our dataset to map the expression of the virus receptor *ACE2* and its co-receptor *TMPRSS2* across the body of a species commonly used to study COVID-19 pathogenesis (Delorey et al., 2021). Comparison between NHPs and humans showed similarity in the potential target tissues in both species (e.g., lung, liver, gallbladder, and kidney), but also differences in the relative expression of constituent cells, which may influence tissue targeting and viral dissemination. We also integrated transcriptomic and epigenomic data for monkey kidney cells to obtain clues regarding the mechanisms of COVID-19. Interestingly, we observed that proximal tubule cells displaying the highest *ACE2* levels contained highly accessible chromatin regions in the promoter and enhancer regions at this locus, and these sites were enriched in DNA-binding motifs for several transcription factors, including *STAT* (*STAT1* and *3*), and interferon response factor (*IRF*) proteins. These transcription factors play essential immunomodulatory roles in the innate immune response; in particular, *STAT* proteins are induced by *IL6*. This is relevant because monoclonal antibodies for *IL6* blockade used in patients with rheumatoid arthritis have shown promising results in recent clinical trials for severe COVID-19 patients (Salama et al., 2021). This was confirmed by our observation that *ACE2*<sup>+</sup> cells in the monkey kidney also displayed enriched expression of the *IL6* receptor (*IL6R*), suggesting that a paracrine link between *IL6* and *ACE2* may facilitate viral entry. In support, tissue levels of *IL6* are enhanced in patients with chronic inflammatory disease and in the elderly, potentially explaining why these individuals develop more aggressive forms of COVID-19. Although

whether the kidney is indeed targeted by SARS-CoV-2 in humans is currently under debate (Delorey et al., 2021), the same mechanism may apply to other tissues. In addition to viral infections, our dataset can also be used to map the enrichment of complex human traits and monogenic heritable diseases across the monkey body and to assess potentially relevant similarities and differences with humans. In this regard, we observed that cell types expressing disease-causing genes showed better correlation between monkey and human than between either species and mice. Thus, our monkey atlas may provide insights into tissue vulnerability in human diseases, especially if additional datasets involving monkey tissue-specific perturbations are included. However, differences between monkeys and humans also exist, suggesting that caution should be taken when modeling human diseases using NHPs.

Our study has several limitations. Notably, as the monkey body contain trillions of cells, many more cells need to be profiled to understand true body complexity. With the rapid development of scRNA-seq technologies, we foresee that single-cell atlases of hundreds of millions of cells will be feasible for multiple mammalian species, including NHPs, and will be extended to include different forms of perturbation. Likewise, combining scRNA-seq with epigenetic information, such as single-cell analysis of transposase-accessible chromatin using sequencing, will help elucidate the gene regulatory networks and underlying epigenetic mechanisms. Furthermore, high-resolution spatially resolved transcriptomic technologies (Chen et al., 2022) will help clarify cellular organization and interactions at the tissue level. Altogether, this will bring us one step closer to generating a comprehensive primate single-cell atlas that can be used to both study NHPs and improve human health.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTIONS

All authors wrote, revised, read, and approved the final version of the manuscript.

Xiao Zhang<sup>1</sup>, Guang-Yao Lai<sup>2</sup>, Giacomo Volpe<sup>3</sup>, Lei Han<sup>4</sup>, Patrick H. Maxwell<sup>5</sup>, Long-Qi Liu<sup>4,\*</sup>, Miguel A. Esteban<sup>1,4,6,\*</sup>, Yi-Wei Lai<sup>4,6</sup>

<sup>1</sup> Jilin Provincial Key Laboratory of Animal Embryo Engineering, Institute of Zoonosis, College of Veterinary Medicine, Jilin University, Changchun, Jilin 130062, China

<sup>2</sup> Joint School of Life Sciences, Guangzhou Institutes of Biomedicine and Health and Guangzhou Medical University, Guangzhou, Guangdong 511436, China

<sup>3</sup> Hematology and Cell Therapy Unit, IRCCS-Istituto Tumori 'Giovanni Paolo II', Bari 70124, Italy

<sup>4</sup> BGI-Shenzhen, Shenzhen, Guangdong 518103, China

<sup>5</sup> Cambridge Institute for Medical Research, Department of Medicine, University of Cambridge, Cambridge CB2 0XY, UK

<sup>6</sup> Laboratory of Integrative Biology, Guangzhou Institutes of

## REFERENCES

- Barker N. 2014. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology*, **15**(1): 19–33.
- Barker N, Rookmaaker MB, Kujala P, Ng A, Leushacke M, Snippet H, et al. 2012. Lgr5<sup>ve</sup> stem/progenitor cells contribute to nephron formation during kidney development. *Cell Reports*, **2**(3): 540–552.
- Barker N, Tan S, Clevers H. 2013. Lgr proteins in epithelial stem cell biology. *Development*, **140**(12): 2484–2494.
- Cao JY, O'Day DR, Pliner HA, Kingsley PD, Deng M, Daza RM, et al. 2020. A human cell atlas of fetal gene expression. *Science*, **370**(6518): eaba7721.
- Chang-Panesso M, Humphreys BD. 2017. Cellular plasticity in kidney injury and repair. *Nature Reviews Nephrology*, **13**(1): 39–46.
- Chen A, Liao S, Cheng MN, Ma KL, Wu L, Lai YW, et al. 2022. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, **185**(10): 1777–1792.e21.
- Clevers H. 2006. Wnt/β-catenin signaling in development and disease. *Cell*, **127**(3): 469–480.
- Delorey TM, Ziegler CGK, Heimberg G, Normand R, Yang YM, Segerstolpe Å, et al. 2021. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature*, **595**(7865): 107–113.
- Domínguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, et al. 2022. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science*, **376**(6594): eabl5197.
- Eraslan G, Drokhlyansky E, Anand S, Fiskin E, Subramanian A, Slyper M, et al. 2022. Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science*, **376**(6594): eabl4290.
- Ha CWY, Martin A, Sepich-Poore GD, Shi BC, Wang YZ, Gouin K, et al. 2020. Translocation of viable gut microbiota to mesenteric adipose drives formation of creeping fat in humans. *Cell*, **183**(3): 666–683.e17.
- Han L, Wei XY, Liu CY, Volpe G, Zhuang ZK, Zou XX, et al. 2022. Cell transcriptomic atlas of the non-human primate *Macaca fascicularis*. *Nature*, **604**(7907): 723–731.
- Han XP, Wang RY, Zhou YC, Fei LJ, Sun HY, Lai SJ, et al. 2018. Mapping the mouse cell atlas by microwell-seq. *Cell*, **172**(5): 1091–1107.e17.
- Han XP, Zhou ZM, Fei LJ, Sun HY, Wang RY, Chen Y, et al. 2020. Construction of a human cell landscape at single-cell level. *Nature*, **581**(7808): 303–309.
- Kalucka J, de Rooij LPMH, Goveia J, Rohlenova K, Dumas SJ, Meta E, et al. 2020. Single-cell transcriptome atlas of murine endothelial cells. *Cell*, **180**(4): 764–779.e20.
- Kamer CM, Chirumamilla R, Aoki S, Igarashi P, Wallingford JB, Carroll TJ. 2009. Wnt9b signaling regulates planar cell polarity and kidney tubule morphogenesis. *Nature Genetics*, **41**(7): 793–799.
- Krausgruber T, Fortelny N, Fife-Gernedl V, Senekowitsch M, Schuster LC, Lercher A, et al. 2020. Structural cells are key regulators of organ-specific immune responses. *Nature*, **583**(7815): 296–302.
- Liu ZJ, Qian XK, Hong MH, Zhang JL, Li DY, Wang TH, et al. 2021. Global view on virus infection in non-human primates and implications for public health and wildlife conservation. *Zoological Research*, **42**(5): 626–632.
- Salama C, Han J, Yau L, Reiss WG, Kramer B, Neidhart JD, et al. 2021. Tocilizumab in patients hospitalized with Covid-19 pneumonia. *The New England Journal of Medicine*, **384**(1): 20–30.
- Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu WH, et al. 2013. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(9): 3507–3512.
- The Tabula Muris Consortium. 2018. Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature*, **562**(7727): 367–372.
- The Tabula Sapiens Consortium. 2022. The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. *Science*, **376**(6594): eabl4896.
- Wong G, Bi YH, Wang QH, Chen XW, Zhang ZG, Yao YG. 2020. Zoonotic origins of human coronavirus 2019 (HCoV-19 / SARS-CoV-2): why is this work important?. *Zoological Research*, **41**(3): 213–219.