

INVITED LETTER

Envisioning population-scale immune multi-omics atlas projects

1 | THE EVOLUTION OF POPULATION-SCALE BIOLOGY: BEYOND GENOMIC CATALOGUES

The completion of the Human Genome Project (HGP) in 2003 provided the first reference framework for human genetics and initiated a transition toward population-scale genomic research.¹ Over the past decade, this transition has driven the emergence of numerous large-scale population cohort genomic projects. Among these, the UK Biobank (UKB) stands out not only for its scale but as a benchmark model for such initiatives, integrating large-scale genotyping with deep phenotyping, longitudinal follow-up, and open data access.² Notably, UKB also exemplifies a broader technological trajectory observed across the field: an iterative progression from array-based genotyping,² to whole-exome sequencing (WES),³ and ultimately to whole-genome sequencing (WGS),⁴ with each successive phase capturing increasing genomic resolution across a growing number of participants (Figure 1 Bottom panel). This framework has profoundly shaped the design of other large-scale population cohort genomic projects worldwide,⁵ each adapting this paradigm to distinct populations and healthcare systems (Figure 1 Bottom panel).

Together, these initiatives have substantially advanced our ability to associate genetic variation with disease risk across populations, establishing a robust foundation for precision medicine. However, the emphasis on genomic associations has revealed a fundamental limitation: most genetic variants, particularly in non-coding regions, exert their effects through context-dependent regulatory mechanisms that are highly cell type-specific and shaped by chromatin and cellular states.^{6,7} This challenge is amplified in the immune system due to its extreme cellular heterogeneity and dynamic functional responses.⁸ Together, these limitations highlight the need for next-generation reference frameworks that integrate immune-

specific, multi-omic, and single-cell-resolved data beyond population genomics.

2 | POPULATION-SCALE IMMUNE MULTI-OMICS: CONCEPT AND FEASIBILITY

Building on these limitations, the field is witnessing a growing trend toward population-scale, immune-focused reference frameworks, giving rise to a new class of initiatives collectively termed Population-scale Immune Multi-omics Atlas Projects (PIMAs) (Figure 1 Top panel). Analogous to large-scale population cohort genomic projects that transformed population genetics, PIMAs seek to systematically characterise immune cell diversity and regulatory states by integrating molecular information across genomic, epigenomic, transcriptomic, proteomic, and metabolic layers at cellular resolution.^{7,9,10} By focusing on coordinated, population-wide immune profiling rather than isolated datasets, PIMAs provide a foundation for mechanistic interpretation of genetic risk in immune-mediated diseases.

Recent advances in scalable single-cell technologies, high-throughput proteomics, and computational frameworks for multi-omics integration have enabled the generation of high-resolution cellular atlases at unprecedented scale, making such population-scale immune profiling increasingly feasible.^{11,12} Focusing on peripheral blood mononuclear cells (PBMCs) provides a practical and minimally invasive window into systemic immune function, as the diverse repertoire of circulating immune cell subsets they encompass enables near non-invasive, real-time characterisation of the organism-level immune states. These cellular profiles not only capture immune perturbations associated with disease states but also reflect variations driven by genetic background, ancestry, environmental exposures, lifestyle and ageing, thereby serving as scalable

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics.

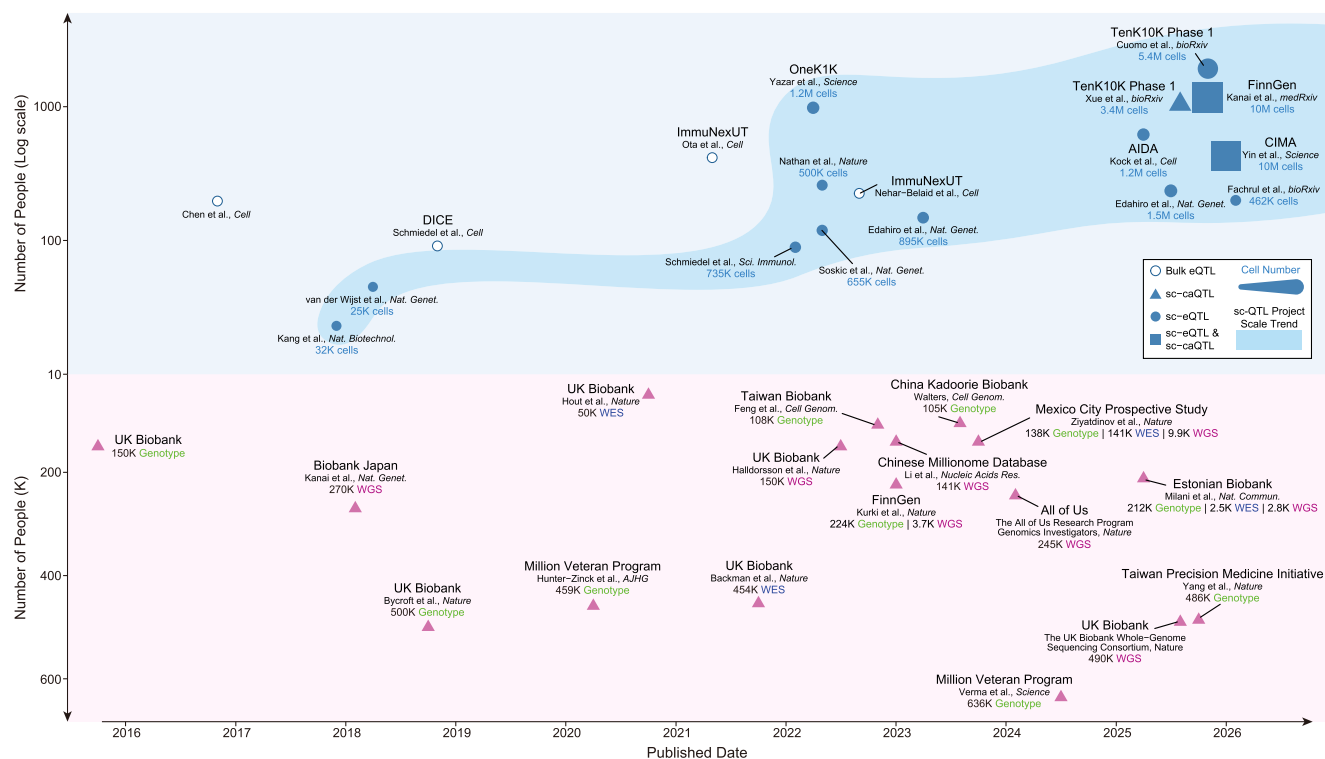


FIGURE 1 Representative large-scale population genomics projects and immune cell profiling studies with QTL analyses over the past decade. (Top panel) Immune cell omics studies with QTL analyses. Marker shapes indicate omics modalities, and marker size reflects the number of single cells (annotated). Scaling trends over time are shown. (Bottom panel) Large-scale population genomics projects. The number of sequenced individuals is indicated for each project, stratified by sequencing technology.

and informative readouts for monitoring immune system dynamics across individuals and populations.¹³ Importantly, the cell type-specific resolution afforded by single-cell approaches addresses a key limitation of bulk analyses, as immune dysfunction often arises from perturbations in defined cellular subsets that are masked when signals are averaged across heterogeneous cell populations.

Crucially, the comprehensive data generated by such initiatives have the potential to advance the development of virtual cells—data-driven computational models that integrate multi-layered molecular information to represent and simulate immune cell states. Emerging models are beginning to enable the simulation of genetic perturbations, supporting the prediction of functional effects during disease progression and the prioritisation of potential therapeutic targets.¹⁴ By linking genetic variation to chromatin accessibility, transcription factor activity, gene expression and downstream protein and metabolic states, population-scale immune multi-omics data may further facilitate the transition of virtual cells toward more quantitative and cell-type-resolved representations. In this context, PIMAs could provide population-level context and regulatory depth, although fully bridging genetic associations and immune phenotypes remains an ongoing challenge.

3 | CHINESE IMMUNE MULTI-OMICS ATLAS AS AN EXEMPLAR OF INTEGRATED POPULATION IMMUNE PROFILING

The Chinese Immune Multi-Omics Atlas (CIMA),¹⁰ recently published in *Science*, exemplifies how PIMAs can extend descriptive immune atlases into integrated, population-specific reference resources. By profiling over 10 million circulating immune cells from 428 Chinese adults, CIMA combines single-cell transcriptomic and chromatin accessibility data with WGS and plasma-level lipidomic, metabolomic, and biochemical measurements.¹⁰ While population-scale immune atlas studies such as OneK1K,⁷ AIDA,⁹ FinnGen¹⁵ and TenK10K^{16,17} have been instrumental in systematically cataloguing transcriptional and regulatory diversity across large cohorts, they reflect diverse study designs that balance cohort size, cellular resolution, and molecular modalities (Figure 1 Top panel). In this context, CIMA places particular emphasis on deep per-individual cellular profiling (with a median of approximately 15 000 cells for scRNA-seq and 9600 cells for scATAC-seq per donor), enabling more comprehensive characterisation of rare immune cell subsets and enhancing statistical power for

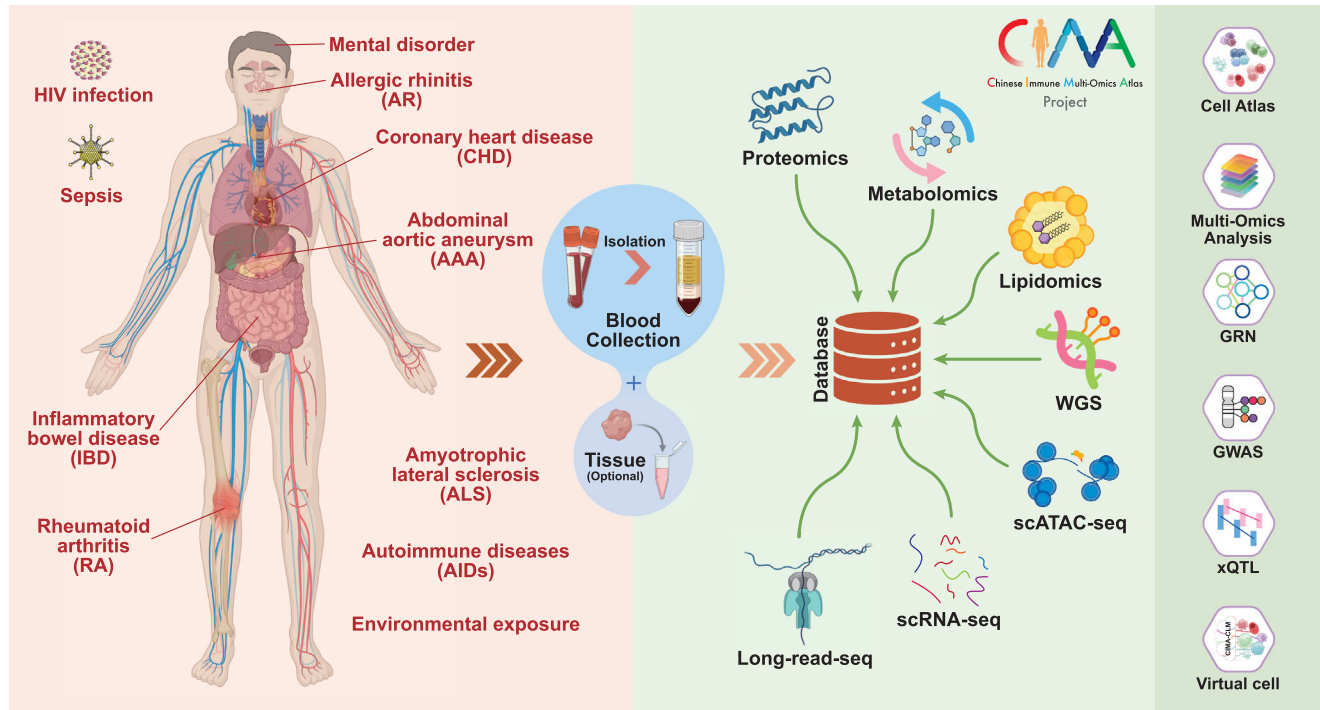


FIGURE 2 Study design framework of the Chinese Immune Multi-Omics Atlas (CIMA) Project. (Left panel) Diseases or conditions involved in the CIMA project. (Right panel) Omics technologies and analytical frameworks involved in the CIMA project.

detecting cell type-specific regulatory effects. This design facilitates the identification of a broader set of genes and cis-regulatory elements influenced by genetic variation, thereby strengthening links between genotype, regulatory activity, and immune cell function at the population scale.

This integrative framework allows CIMA to trace how genetic variation propagates through regulatory networks to influence immune phenotypes (Figure 2). Extensive mapping of 338 036 cis-regulatory elements and construction of enhancer-driven gene regulatory networks delineates transcription factor programs that define immune cell identity across fine-grained cell types, while cell-type-resolved quantitative trait locus analyses connect genetic variants to 9600 eGenes and 52 361 caPeaks. These analyses enrich descriptive immune cell maps with regulatory context and population specificity, as illustrated by the asthma-associated variant rs34415530, which modulates *IKZF4* expression specifically in CD4⁺ FOXP3⁺ regulatory T cells and is linked to IL-12B secretion differences. Moreover, the development of CIMA-CLM, a cell language model that predicts chromatin accessibility from DNA sequences and gene expression, demonstrates how population-scale data can support the construction of predictive virtual cell models capable of evaluating variant effects computationally, transforming static associations into dynamic, mechanistically interpretable frameworks. The CIMA study highlights a major advance in human

immunology by showing that population-scale single-cell multi-omics can resolve cell type-specific regulatory mechanisms underlying genetic effects, thereby enabling more precise links between genetic variation and immune phenotypes.¹⁸

4 | TOWARDS A GLOBALLY INTEGRATED AND DISEASE-EXPANDED PIMA FRAMEWORK

With population-scale immune atlases now emerging across multiple regions, the central challenge has shifted from establishing feasibility to achieving coordination and interoperability. Efforts such as CIMA, alongside parallel initiatives across Europe, Asia and beyond, collectively demonstrate that population-specific immune reference frameworks are both technically achievable and biologically informative.^{7,9,10,15,16,19} The next imperative is therefore to organise these independent efforts into a globally integrated framework — one that enables systematic harmonisation of multi-omics data and immune cell-resolved regulatory models across populations, rather than isolated interpretation. Central to such a framework is the adoption of shared principles for data generation, processing and multi-omics integration, allowing virtual cells constructed from different populations to become directly comparable and mutually informative.

Equally important is the extension of this framework beyond healthy population cohorts into disease-focused settings. The majority of existing PIMAs have been established on predominantly healthy populations, which, while essential for defining regulatory baselines, provides limited direct insight into the mechanisms underlying immune-mediated disease. Expanding PIMAs to encompass disease cohorts (spanning autoimmune, inflammatory, infectious and oncological conditions) would enable the systematic dissection of how genetic and regulatory variation contribute to immune dysfunction in specific pathological contexts. Such an extension would transform population immune atlases from descriptive reference resources into mechanistically actionable frameworks, ultimately bridging the gap between the regulatory architecture defined in healthy populations and the cellular and molecular perturbations that drive immune-mediated diseases.

ACKNOWLEDGEMENTS

We thank all of our team members. This work was supported by the National Key Research and Development Program of China (2025YFC3409300), Guangdong Basic and Applied Basic Research Foundation (2026A1515012050 and 2024B1515230003), and Shanxi Province Higher Education “Billion Project” Science and Technology Guidance Project.

Yuhui Zheng^{1,2} 

Jianhua Yin^{1,3,4}

Chuanyu Liu^{1,3,4} 

¹State Key Laboratory of Genome and Multi-omics Technologies, BGI Research, Shenzhen, China

²College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

³Shenzhen Proof-of-Concept Center of Digital Cytopathology, BGI Research, Shenzhen, China

⁴Shanxi Medical University-BGI Collaborative Center for Future Medicine, Shanxi Medical University, Taiyuan, China

Correspondence

Yuhui Zheng, Jianhua Yin and Chuanyu Liu, State Key Laboratory of Genome and Multi-omics Technologies, BGI Research, Shenzhen, China.

Email: zhengyuhui@genomics.cn, yinjianhua@genomics.cn and liuchuanu@genomics.cn

ORCID

Yuhui Zheng  <https://orcid.org/0000-0001-5760-0048>

Chuanyu Liu  <https://orcid.org/0000-0003-2258-0897>

REFERENCES

1. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860-921. doi:[10.1038/35057062](https://doi.org/10.1038/35057062)
2. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203. doi:[10.1038/s41586-018-0579-z](https://doi.org/10.1038/s41586-018-0579-z)
3. Backman JD, Li AH, Marcketta A, et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature*. 2021;599:628-634. doi:[10.1038/s41586-021-04103-z](https://doi.org/10.1038/s41586-021-04103-z)
4. The UK Biobank Whole-Genome Sequencing Consortium. Whole-genome sequencing of 490,640 UK Biobank participants. *Nature*. 2025;645:692-701. doi:[10.1038/s41586-025-09272-9](https://doi.org/10.1038/s41586-025-09272-9)
5. Claussnitzer M, Cho JH, Collins R, et al. A brief history of human disease genetics. *Nature*. 2020;577:179-189. doi:[10.1038/s41586-019-1879-7](https://doi.org/10.1038/s41586-019-1879-7)
6. Phanstiel DH, Wang GG. Cell type-specific chromatin topology and gene regulation. *Trends Genet*. 2022;38:413-415. doi:[10.1016/j.tig.2022.02.008](https://doi.org/10.1016/j.tig.2022.02.008)
7. Yazar S, Alquicira-Hernandez J, Wing K, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science*. 2022;376:eabf3041. doi:[10.1126/science.abf3041](https://doi.org/10.1126/science.abf3041)
8. Ginhoux F, Yalin A, Dutertre CA, Amit I. Single-cell immunology: past, present, and future. *Immunity*. 2022;55:393-404. doi:[10.1016/j.immuni.2022.02.006](https://doi.org/10.1016/j.immuni.2022.02.006)
9. Kock KH, Tan LM, Han KY, et al. Asian diversity in human immune cells. *Cell*. 2025;188:2288-2306.e24. doi:[10.1016/j.cell.2025.02.017](https://doi.org/10.1016/j.cell.2025.02.017)
10. Yin J, Zheng Y, Huang Z, et al. Chinese immune multi-omics atlas. *Science*. 2026;391:eadt3130. doi:[10.1126/science.adt3130](https://doi.org/10.1126/science.adt3130)
11. Baysoy A, Bai Z, Satija R, Fan R. The technological landscape and applications of single-cell multi-omics. *Nat Rev Mol Cell Biol*. 2023;24:695-713. doi:[10.1038/s41580-023-00615-w](https://doi.org/10.1038/s41580-023-00615-w)
12. Lim J, Park C, Kim M, et al. Advances in single-cell omics and multiomics for high-resolution molecular profiling. *Exp Mol Med*. 2024;56:515-526. doi:[10.1038/s12276-024-01186-2](https://doi.org/10.1038/s12276-024-01186-2)
13. Mogilenko DA, Shchukina I, Artyomov MN. Immune ageing at single-cell resolution. *Nat Rev Immunol*. 2022;22:484-498. doi:[10.1038/s41577-021-00646-4](https://doi.org/10.1038/s41577-021-00646-4)
14. Ma C, Zhang H, Rao Y, et al. AI-driven virtual cell models in pre-clinical research: technical pathways, validation mechanisms, and clinical translation potential. *NPJ Digit Med*. 2025;9:25. doi:[10.1038/s41746-025-02198-6](https://doi.org/10.1038/s41746-025-02198-6)
15. Kanai M, Delorey TM, Honkanen J, et al. Population-scale multi-ome immune cell atlas reveals complex disease drivers. *medRxiv*. 2025;1-56. doi:[10.1101/2025.11.25.25340489](https://doi.org/10.1101/2025.11.25.25340489)
16. Cuomo ASE, Spenceley E, Tanudisastro HA, et al. Impact of rare and common genetic variation on cell type-specific gene expression in human blood. *medRxiv*. 2025. doi:[10.1101/2025.03.20.25324352](https://doi.org/10.1101/2025.03.20.25324352)
17. Xue A, Fan J, Dong OA, et al. Genetic regulation of cell type-specific chromatin accessibility shapes immune function and disease risk. *medRxiv*. 2025. doi:[10.1101/2025.08.27.25334533](https://doi.org/10.1101/2025.08.27.25334533)
18. Powell JE. Charting immune variation through genetics and single-cell genomics. *Gigascience*. 2026;15:giaf161. doi:[10.1093/gigascience/giaf161](https://doi.org/10.1093/gigascience/giaf161)
19. Kang JB, Raveane A, Nathan A, et al. Methods and insights from single-cell expression quantitative trait loci. *Annu Rev Genomics Hum Genet*. 2023;24:277-303. doi:[10.1146/annurev-genom-101422-100437](https://doi.org/10.1146/annurev-genom-101422-100437)