Zhang et al. *Zool. Res.* 2026, 47(1): 1−3
https://doi.org/10.24272/j.issn.2095-8137.2025.649
https://cstr.cn/32248.14.zr.2025.649

ZR | Zoological Research

Commentary Open Access

Macaque genotype-phenotype resources: A prismatic portal into human biology and disease

Non-human primates serve as indispensable models in biomedical research due to their high degree of genetic, physiological, and behavioral similarity to humans (Yao, 2022). Macaques are among the most widely used species in studies of drug efficacy, disease mechanisms, and neural function (National Academies of Sciences, Engineering, and Medicine et al., 2023). However, recent increases in global demand, compounded by public health emergencies such as the COVID-19 pandemic, have resulted in widespread shortages and escalating costs (Grimm, 2023). A more persistent constraint is the limited availability of macaques with well-documented genotypic and phenotypic profiles, which restricts experimental reproducibility and reduces the reliability of research results (Bimber et al., 2019).

To address this dual quantitative and qualitative crisis, a macaque biobank integrating large-scale phenotypic and multi-omics datasets is imperative to support standardized, reproducible research (Figure 1). Such a resource enables precise stratification and selection of individuals based on defined traits or genotypes relevant to specific drug responses, disease susceptibility, or other experimental variables, thereby minimizing inter-individual variability and enhancing experimental reproducibility and translational relevance. It also facilitates the discovery of naturally occurring macaque models harboring clinically relevant variants or spontaneous pathological phenotypes, reducing reliance on labor-intensive, ethically constrained artificial engineering of disease models. Following the establishment of the Big Monkey Facility (Yao, 2022), the Macaque Biobank project was initiated to generate a comprehensive dataset from a large cohort of originally wild-caught and bred individuals. The first phase has yielded important findings regarding the genetic basis of phenotypic variation and enabled the identification of spontaneous models using a reverse-genomics approach (Zhang et al., 2025).

However, phenotypic data collection in macaques presents distinct methodological challenges that differ fundamentally from those encountered in human biobanking. A key constraint arises from the inability of macaques to cooperate voluntarily with examination procedures, necessitating physical restraint or chemical sedation even for basic measurements. These interventions inevitably induce stress responses that significantly disrupt baseline physiological states, leading to alterations in heart rate, blood pressure, and circulating glucocorticoids, thereby confounding data quality and interpretability (Pfefferle et al., 2018). Compounding this issue is the absence of standardized protocols, validated reference standards, and specialized instrumentation tailored to macaque-specific anatomy and physiology. These deficiencies present major barriers to data harmonization in multi-center or longitudinal studies, where inconsistent measurement conditions across timepoints and facilities severely compromise data comparability and reproducibility. Moreover, ethical restrictions further constrain the frequency and invasiveness of sampling procedures, limiting access to critical biospecimens, such as whole blood or cerebrospinal fluid, to avoid unnecessary harm. All procedures must adhere to stringent animal welfare regulations and require prior approval from institutional ethics committees, especially in studies involving repeated or high-frequency measurements (Carvalho et al., 2018). Collectively, these limitations significantly reduce the scope, resolution, and consistency of phenotypic datasets derived through conventional means.

With recent technological advancements, an increasing number of high-resolution and multidimensional quantitative phenotypes have been explored. The declining costs of high-throughput sequencing have facilitated comprehensive integration of multi-omics data, including genomic profiling of sequence variants, transcriptomic analysis of expression patterns, metabolomic characterization of biochemical states, and proteomic quantification of protein abundance, providing a holistic understanding of the molecular mechanisms underlying phenotypic traits. In parallel, non-invasive imaging modalities, such as magnetic resonance imaging (MRI) and functional MRI (fMRI), allow assessment of brain structure, functional connectivity, and region-specific activation (Matthews & Jezzard, 2004), while computed tomography (CT) supports quantitative evaluation of bone structure, soft tissue composition, and organ morphology (Mazonakis & Damilakis, 2016). Wearable sensor systems enable continuous monitoring of physiological and behavioral parameters in real time (Fletcher et al., 2012), and infrared thermal imaging allows non-invasive monitoring of body temperature dynamics. Recently, a frequency-modulated continuous wave radar system has been established to non-invasively measure heart rate and respiratory rate in conscious, unrestrained macaques (Zhang et al., 2024). These emerging tools provide ethically compliant, high-

Received: 26 December 2025; Accepted: 31 December 2025; Online: 05 January 2026

Foundation items: This work was supported by the Ministry of Science and Technology of China (STI2030-Major Project 2021ZD0200900), Yunnan Revitalization Talent Support Program Young Talent Project to B.L.Z., Yunnan Revitalization Talent Support Program Yunling Scholar Project to Y.G.Y., and Yunnan Province (202305AH340006)

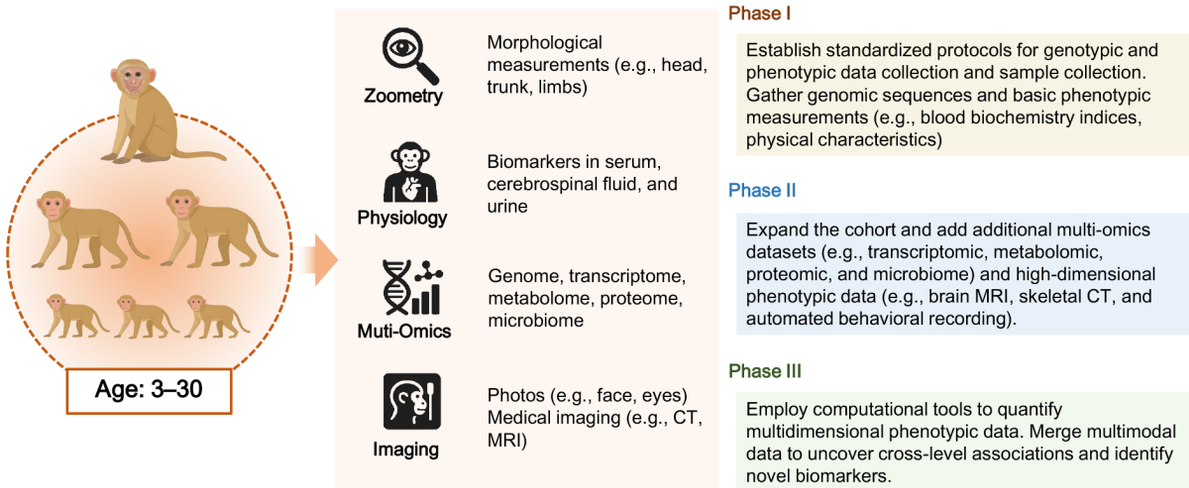

Figure 1 Design and implementation roadmap of the Macaque Biobank

Macaques aged 3–30 years from the breeding facility at the Kunming Institute of Zoology will be enrolled for biobanking. Morphological parameters of the body and head will be measured, and blood, cerebrospinal fluid, and urine will be collected for physiological analysis and multi-omics sequencing. Additional data collection will include facial and ocular imaging, brain MRI, and skeletal CT. Given technical constraints and ethical considerations, the biobank will be established in three phases. The initial phase will establish standardized protocols and collect pilot data, followed by large-scale cohort expansion and multimodal data acquisition. The final phase will focus on the application of artificial intelligence for multimodal data fusion and analysis. Macaque drawings provided by Min-Qiang Deng.

throughput alternatives that enhance data fidelity while minimizing stress-related artifacts in macaque phenotyping.

Artificial intelligence (AI) is revolutionizing macaque phenotyping by transforming subjective observations into scalable, quantitative, and multidimensional data. By leveraging computer vision and deep learning, AI can extract fine-scale facial features from both two-dimensional and three-dimensional images with high precision (Morita et al., 2022), thereby enabling detection of subtle phenotypic variations that remain imperceptible to human observers. In neuroimaging, AI-based analysis can automatically delineate brain regions, quantify structural abnormalities, and track temporal changes in morphology, offering a level of accuracy and throughput that has transformed human diagnostic workflows but remains underutilized in non-human primates. In behavioral phenotyping, deep learning models trained on video recordings can automatically classify and quantify complex social and motor behaviors, such as grooming, aggression, or affiliative interactions, thereby replacing labor-intensive manual scoring with objective and high-throughput pipelines (Zhai et al., 2025). These models are particularly useful in detecting subtle or infrequent behavioral phenotypes relevant to disease modeling, including stereotypic movements in neuropsychiatric disorders or gait anomalies in neurodegenerative studies, while eliminating observer bias. Beyond single-modality analysis, AI also enables integrative analysis across heterogeneous data. Graph neural networks and related architectures can synthesize behavioral, imaging, genomic, and physiological data to resolve multi-scale associations between molecular features and organism-level traits (Nam et al., 2024; Zhang et al., 2023). This integrative framework supports early identification of disease-associated phenotypes, refinement of experimental models through data-driven stratification, and development of novel digital biomarkers with potential translational relevance.

Given the exceptional costs, technical complexities, and ethical constraints associated with macaque research, the development of a macaque biobank requires a staged implementation strategy (Figure 1). The initial phase should prioritize the establishment of standardized procedures for the acquisition, storage, and annotation of phenotypic and genetic data. This foundational step includes construction of a pilot cohort incorporating whole-genome sequencing and baseline phenotypic assessments, such as hematological and biochemical profiles and key morphological characteristics. Subsequent phases should expand the cohort to thousands of individuals, with the addition of diverse data modalities, including transcriptomic, metabolomic, proteomic, and microbiome datasets, alongside high-dimensional phenotypic data derived from brain MRI, skeletal CT, and automated behavioral recording. The final phase should employ advanced computational tools to enable automated and accurate high-throughput quantification of complex phenotypes. Integration of these multidimensional datasets enables dissection of cross-scale biological relationships, supports discovery of novel biomarkers, allows genetic decoding of key traits, and establishes digital indices of physiological state, facilitating the development of robust disease models with translational relevance.

Comprehensive biobanking of macaques with integrated genomic and quantitative phenotypic datasets offers a transformative platform for enhancing the value of existing animal resources while addressing persistent concerns regarding data quality and reproducibility. Successful execution of this initiative depends on coordinated strategic planning, sustained financial investment, cross-disciplinary collaboration, and global partnerships. This framework will strengthen the reliability and efficiency of biomedical research, accelerate therapeutic innovation, and contribute to global health advancement.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

B.L.Z. wrote the original draft and drew the figure. Y.G.Y. supervised and edited the manuscript. S.L. and D.D.W. reviewed and edited the paper. All

authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We would like to thank the National Research Facility of Phenotypic & Genetic Analysis of Model Animals (Primate Facility) (http://cstr.cn/31137.02.NPRC) and National Resource Center for Non-Human Primates (http://nhp.kiz.ac.cn) of Kunming Institute of Zoology, Chinese Academy of Sciences, for providing technical support for the Macaque Biobank.

Bao-Lin Zhang[1,2,3], Shu Liu[1,2], Dong-Dong Wu[1,2,5], Yong-Gang Yao[1,2,4,5,*]

[1] *State Key Laboratory of Genetic Evolution & Animal Models*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650201, *China*

[2] *National Research Facility for Phenotypic & Genetic Analysis of Model Animals (Primate Facility) and National Resource Center for Non-Human Primates*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650107, *China*

[3] *Yunnan Key Laboratory of Biodiversity Information*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650201, *China*

[4] *Yunnan Key Laboratory of Animal Models and Human Disease Mechanisms*, *and Yunnan Engineering Center on Brain Disease Models*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650201, *China*

[5] *KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650201, *China*

*Corresponding author, E-mail: yaoyg@mail.kiz.ac.cn

REFERENCES

Bimber BN, Yan MY, Peterson SM, et al. 2019. mGAP: the macaque genotype and phenotype resource, a framework for accessing and interpreting macaque variant data, and identifying new models of human disease. *BMC Genomics*, **20**(1): 176.

Carvalho C, Gaspar A, Knight A, et al. 2018. Ethical and scientific pitfalls concerning laboratory research with non-human primates, and possible solutions. *Animals (Basel)*, **9**(1): 12.

Fletcher RR, Amemori K, Goodwin M, et al. 2012. Wearable wireless sensor platform for studying autonomic activity and social behavior in non-human primates. *In*: Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society. San Diego: IEEE, 4046–4049.

Grimm D. 2023. U. S. , European researchers face monkey shortage crisis. *Science*, **380**(6645): 567–568.

Matthews PM, Jezzard P. 2004. Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery, and Psychiatry*, **75**(1): 6–12.

Mazonakis M, Damilakis J. 2016. Computed tomography: what and how does it measure?. *European Journal of Radiology*, **85**(8): 1499–1504.

Morita T, Ito T, Koda H, et al. 2022. Analyzing and visualizing morphological features using machine learning techniques and non-big data: a case study of macaque mandibles. *American Journal of Biological Anthropology*, **178**(1): 44–53.

Nam Y, Kim J, Jung SH, et al. 2024. Harnessing artificial intelligence in multimodal omics data integration: paving the path for the next frontier in precision medicine. *Annual Review of Biomedical Data Science*, **7**(1): 225–250.

National Academies of Sciences, Engineering, and Medicine, Division on Earth and Life Studies, Health and Medicine Division, et al. 2023. Nonhuman Primate Models in Biomedical Research: State of the Science and Future Needs. Washington (DC): National Academies Press.

Pfefferle D, Plümer S, Burchardt L, et al. 2018. Assessment of stress responses in rhesus macaques (*Macaca mulatta*) to daily routine procedures in system neuroscience based on salivary cortisol concentrations. *PLoS One*, **13**(1): e0190190.

Yao YG. 2022. Towards the peak: the 10-year journey of the National Research Facility for Phenotypic and Genetic Analysis of Model Animals (Primate Facility) and a call for international collaboration in non-human primate research. *Zoological Research*, **43**(2): 237–240.

Zhai H, Yan HY, Zhou JY, et al. 2025. InteBOMB: integrating generic object tracking and segmentation with pose estimation for animal behavior analysis. *Zoological Research*, **46**(2): 355–369.

Zhang BL, Chen YX, Zhang YL, et al. 2025. Forward and reverse genomic screens enhance the understanding of phenotypic variation in a large Chinese rhesus macaque cohort. *Nature Communications*, **16**(1): 8703.

Zhang JJ, Hu RJ, Chen LC, et al. 2024. Contactless vital signs monitoring in macaques using a mm-wave FMCW radar. *Scientific Reports*, **14**(1): 13863.

Zhang YJ, Luo ZY, Sun YW, et al. 2023. From beasts to bytes: revolutionizing zoological research with artificial intelligence. *Zoological Research*, **44**(6): 1115–1131.