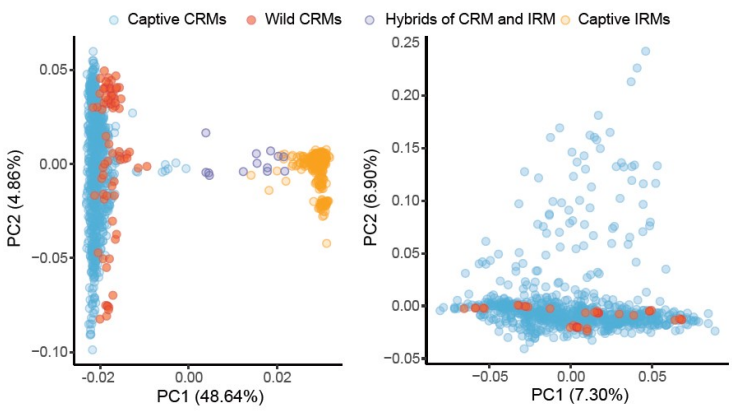


2023 年非人灵长类动物基因组数据集

（国家非人灵长类实验动物资源库编制，云南省昆明市盘龙区茨坝街道龙欣路 17 号， 650201，
0871-68424851， nhp@mail.kiz.ac.cn， 2025-07-14）

资源名称	中文	2023 年非人灵长类动物基因组数据集
	英文	Non-human primate genome dataset for 2023
资源标识		CSTR:13153.11.20250623.OD.RMGseq.01.KIZ
数据集内容		a. 描述摘要 猕猴是目前使用最广泛的非人灵长类实验动物模型之一，其基因组与人的相似度达 93%，且在大脑解剖结构、生殖生理和免疫系统等方面与人类有着极高的相似性，因此被认为是研究人类重大疾病最理想的模型动物。然而受新冠疫情以及生物大分子药物研发等对实验动物需求旺盛的影响，猕猴资源出现严重短缺，甚至出现了“一猴难求”的现象。猕猴生物资源的有效利用依赖于对遗传信息的详尽解析以及对表型的精准描述。同时，清楚的遗传背景也将有助于制定合理的遗传管理策略，避免出现近交衰退和远交衰退的风险。因此，开展系统的中国猕猴表型和遗传研究不仅有利于理解猕猴表型的多样性进化，也是推动和有效利用猕猴作为疾病模型研究最基本的一项任务。 b. 关键词 2023 年，猕猴，遗传队列 c. 数据的时间范围 2023 年 6 月 30 日 d. 数据的空间范围 中国科学院昆明动物研究所 e. 学科范围 18051（动物学） f. 行业范围 M73（研究和试验发展） g. 数据格式 .xlsx，.fastq.gz h. 数据量 332G i. 名词解释与量纲 j. 数据精度 k. 数据更新频度 无

缩略图	
数据质量描述	采用华大基因 DNBseq 高通量测序平台对昆明动物研究所 919 只笼养中国猕猴样本进行全基因组重测序,每个样本的测序深度设定为 30 倍覆盖度(约 90G 数据量),所有测序 reads 经过质量控制后比对至印度猕猴参考基因组(Mmul_10),通过标准的生物信息学分析流程进行变异检测和过滤,最终获得约 8400 万个高质量的遗传变异位点数据集,为后续的群体遗传学分析和进化研究提供了可靠的基础数据。
数据产生方式	实验猴年度健康体检采集约 1ml 血液,采用华大 DNBseq 平台进行基因组重测序,每个样本测序深度约 30X (90G),使用 BWA-MEM 比对到印度猕猴参考基因组(Mmul10),使用 GATK 软件获得 SNP 和 INDEL 遗传变异的数据。
数据采集、加工处理方法	对获得的遗传变异数据进行质量过滤,如: Phred-scaled 评估的碱基质量后验概率值(QUAL)<50.0;基于经过深度校正的质量值(QD)<2.0; RMS 比对质量(MQ)<40.0;使用 Fisher's 精确检验判断是否存在链特异倾向(FS)>60.0;对称比值比检验(SOR)估计的链偏倚<3.0;Alt vs Ref 比对质量差异的 Wilcoxon 秩和检验(MQRankSum)<-12.5;Alt 与 Ref 读取位置偏差(ReadPosRankSum)的 Wilcoxon 秩和检验的 Z-score<-8 等。
数据使用条件、方法	建议使用 GATK 或 bcftools 等变异检测工具对比对后的 BAM 文件进行 SNP 和 INDEL calling,结合 VCFtools 进行变异位点质量过滤和统计分析,最终可利用 PLINK、ADMIXTURE 等群体遗传学软件对获得的高质量变异位点数据开展群体结构分析、主成分分析、连锁不平衡分析等 downstream 分析。
知识产权	a. 标注知识产权说明 使用本数据集时,请在文章中引用以下文献: 1.Bao-Lin Zhang,Yongxuan Chen,Yali Zhang,Yicheng Qiao,Yang Wu,Yi Zhang,Yizheng Lu,Xinran You,Yanling Li,Hong-Di Huang,Qiong Wang,Yijiang Li,Yun Wang,Wenxian Xiao,Hexian Duan,Ming-Hao Qiu,Nan-Hui Chen,Xiaomei Yu,Min-Min Yang,Longbao Lv,David N. Cooper,Ping Zheng,Yong-Gang Yao,Ning Liu,Jian-Hong Wang,Dong-Dong Wu. Forward and reverse genomic screens enhance the understanding of phenotypic variation in a large Chinese rhesus macaque cohort. Nature Communications,(2025) 16:8703. https://doi.org/10.1038/s41467-025-63747-x. b. 数据标注参考以下规范: 数据来源引用参考以下规范: 中文表达方式:数据来源于国家科技基础条件平台—国家非人灵长类实验动物资源库(https://nhp.kiz.ac.cn)。 英文表达方式: National Resource Center for Non-Human Primates,

	National Science & Technology Infrastructure of China (https://nhp.kiz.ac.cn)。 致谢方式参考以下规范： 中文致谢方式：感谢国家科技基础条件平台-国家非人灵长类实验动物资源库 (https://nhp.kiz.ac.cn) 提供数据支撑。 英文致谢方式：Acknowledgement for the data support from National Resource Center for Non-Human Primates, National Science & Technology Infrastructure of China. (https://nhp.kiz.ac.cn)。 c. 数据贡献者信息 姓名：张宝林（吴东东课题组） 单位：中国科学院昆明动物研究所 电话：18787456994 邮箱：zhangbaolin@mail.kiz.ac.cn
其它说明内容	若使用方希望利用该资源的任何材料开展宣传等活动，须事先得到资源管理方的书面授权。