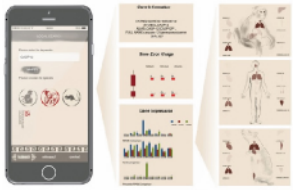
RhesusBase: Understanding human biology in the genomic framework of rhesus macaque



Rhesus Base APP for iPhone (www.rhesusbase.org)

With the support by the National Natural Science Foundation of China, the research group led by Li-Chuanyun (**李川昀**) at the Institute of Molecular Medicine, Peking University published in *Molecular Biology and Evolution* (2016,33(5): 1370—1375), as the new component of their knowledge base for monkey research community, the RhesusBase (http://www.rhesusbase.org).

While rhesus macaque is a unique model for evolutionary and translational studies of human biology, several unresolved issues have limited the current use of this model—inadequate functional genomics annotations, error-prone gene structures, and lack of a platform for visualizing and assessing high-throughput data. To address these key issues, Dr. Li's group has established a tissue bank with 56 types of monkey tissues, annotated monkey genome and transcriptome through comprehensive functional genomics studies, and refined the fine-scale structures of over 20,000 monkey genes. These efforts have culminated in the development of the "RhesusBase", with 7.6 billion functional annotation records to provide an information-rich framework for monkey genomics study (*Nucleic Acids Research*, 2013, 41: D892—905; *Molecular Biology and Evolution*, 2014, 31(5): 1309—1324).

In their recent work, the group further closed a critical gap in the population genetics study of rhesus macaque. Currently the population genetic studies have significantly accelerated the evolutionary and functional interrogations of genes and regulatory regions in Nematoda, Drosophila, mice, and humans. However, rather limited polymorphism data have been reported for non-human primates, despite their uniqueness in evolutionary and biomedical studies. To develop the new gold standard for genetics and comparative evolutionary studies in non-human primates, Li's group profiled the first genome-wide mutational map in rhesus macaque, on the basis of their genome-wide sequencing efforts in 31 rhesus macaque animals. They also incorporated these valuable data resources into RhesusBase in a user-friendly way.

The group has demonstrated important applications of this resource in illuminating the mechanisms and regulations underlying human evolution and disease. In the framework of RhesusBase, they identified 43 human-specific protein-coding genes and proposed a novel model for gene origination, which states that protein-coding genes may generally emerge out of lncRNAs (*PLOS Genetics*, 2012, 8(9): e1002942; *PLOS Genetics*, 2015, 11(7): e1005391). Moreover, they clarified the mechanisms and evolution of primate regulations across multiple regulatory levels such as RNA editing, and identified a catalog of 9,295 human-biased regulatory events (*PLOS Genetics*, 2014, 10(4): e1004274; *Molecular Biology and Evolution*, 2015,32(12):3143—3157). Finally,they also identified a novel drug target through family-based genetics study in a monkey family with extreme disease phenotype. These findings provide novel clues to the understandingof fundamental questions such as "what makes us uniquely human", as well as to translational studies by elucidating novel biomarkers and drug targets.

As China is on the way to be the leading country for non-human primate research, the RhesusBase is sure to make its impact by providing a promising "one-stop"genomic framework for functional interrogation of human biology.