Jingfa Xiao

xiaojingfa@big.ac.cn

PROFESSIONAL EXPERIENCE

- Professor, **Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS)**, China, 2012−Present
- Associate Professor, **Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS)**, China, 2007−2012
- Postdoctoral Associate, **Department of Chemistry, Utah University, United States of America**, 2005−2007
- Postdoctoral Associate, **Institute of Materia Medica, Chinese Academy of Medical Sciences**, China, 2003−2005
- Assistant Engineering, **Shenyang First Pharmaceutical Co., Ltd**. China, 1995−1998

EDUCATION

- Ph.D. in Computational Chemistry and Biology, **State Key Laboratory of Theoretical and Computational Chemistry, Jilin University**, China, 1998−2003
- B.S. in Biology, **Department of Biology, Sichuan University**, China, 1991−1995

RESEARCH INTERESTS

- Large-scale Genome Variation in Health and Disease
- Big Omics Data Integration and Mining
- Genome Informatics for Precision Medicine

PROJECTS & RESOURCES

- National Key R&D Program of China, Grant No. 2016YFB0201702, 2016-2020, leader
- National Science Foundation of China, Grant No. 31771465, 2018-2021, leader
- Key Research Program of the Chinese Academy of Sciences, Grant No. KJZD-EW-L14, 2015-2018, leader

ACADEMIC ACTIVITIES

- Associate Editor: Frontiers in Plant Science (2016-2018)
- Editorial Board Member: Genomics, Proteomics & Bioinformatics (2012-), China Science & Technology Resources Review (2018-)

ABSTRACT

The advancement of sequencing technology and assembly tools has been facilitating fast acquisition of bacterial genome sequences of various species over the past two decades. However, the second-generation DNA sequencing technologies generate large numbers of short reads, this pose tremendous challenges to the de novo assembly of bacterial genome with repeat sequences. Third-generation, single-molecule sequencing addresses this problem by greatly increasing sequencing read length, which can assemble genomes more continuous. As sequencing technologies development, many assembly tools and annotation softwares are adapting to cope with the increasing sequencing data. The explosive growth of bacterial genome sequence also brought an extremely big challenge to comparative genome analysis. Pan-genome encompasses the entire repertoire of genes accessible to a studied phylogenetic clade or a given species, which is divided into core genome, the dispensable genes, and strain-specific genes. We have been developed a series of pan-genome software, including PAGP, PGAP-X and PGAweb, to perform bacterial pan-genome analysis more efficient. PGAP and PGAP-X supports five main analytic functions including ortholog clustering, pan-genome profiling, sequence variation analysis, species phylogeny, and gene function enrichment. PGAweb a fast and freely available online server, presents features of genomic structural dynamics and sequence diversity with different visualization methods that are helpful for intuitively understanding of dynamics and evolution of bacterial genomes.