

Lijing Bu

Center for Evolutionary and Theoretical Immunology, Department of Biology
167 Castetter Hall MSC03 2020, 1 University of New Mexico
Albuquerque, NM 87131-0001, USA
email: lijing@unm.edu office: 505-507-0594

EDUCATION

2015	Ph.D. Biology	University of New Mexico, Albuquerque, NM, USA
2009	M.Sc. Medical Genetics	Wenzhou Medical University, Zhejiang Province, China
2004	B.Sc. Clinical Medicine	Taishan Medical University, Shandong Province, China

RESEARCH INTERESTS

General Interests

Functional and Comparative Genome Informatics, Evolutionary Genomics, Bioinformatics

Specific Interests

Host and Parasite Interactome, Gene Duplication History, Biodata Visualization

RESEARCH EXPERIENCE

2015 ~ Present Senior Research Scientist / Bioinformatics Specialist

Center for Evolutionary and Theoretical Immunology (CETI),
Department of Biology at the University of New Mexico, USA

- ◆ Genome assembly and comparative genomic analysis of two snail strains, which are susceptible and resistant to *Schistosoma mansoni*
- ◆ Mitochondrial genome assembly and comparative analysis for multiple snail strains
- ◆ Transcriptome analysis of vector snail's response to parasites (Dual RNA-Seq)
- ◆ Develop workflow to identify gene duplication from reference genome (worm, human etc.) and analyze their structural and evolutionary features
- ◆ Ultra conservative elements search within blood fluke family use genomes and SRA data
- ◆ Comparative analysis of microbial communities with 16S rRNA gene sequencing
- ◆ Opossum immune cells classification based on single cell RNA-Seq data
- ◆ Bioinformatics support: Local Galaxy server setup and maintenance.
- ◆ Maintain computational hardware associated with CETI molecular biology core

2009 ~ 2015 Ph.D. Candidate, Biology, Advisor: Dr. Vaishali Katju
University of New Mexico, Albuquerque, USA

- ◆ Evolutionary rate asymmetry and its correlation with codon usage and expression of gene duplicates from whole genome duplication and small-scale duplication in *S. cerevisiae*
- ◆ Identify, analyze and visualize the evolutionarily young gene duplicates of reference genome (yeast, worm, human and chimpanzee)
- ◆ Gene conversion rate estimation based on Illumina pair-end sequences of 85 *C. elegans* mutation accumulation strains
- ◆ SNPs and Structural variation (SV) analysis for 85 *C. elegans* mutation accumulation strains, and 21 Fog2 knockout *C. elegans* mutation accumulation strains

2006 ~ 2009 Master Candidate, Medical Genetics, Advisor: Dr. Hongzhi Li
Wenzhou Medical University, Zhejiang Province, China

- ◆ Stable expression of fusion protein anti-erbB2-scFv-Fc-CD28-CD3(ζ) in Jurkat cells
 - fusion protein 2D and 3D structure prediction
 - PCR conformation of fusion gene within reconstructed vector
 - Breast cancer cell culture/transformation with reconstructed vector
 - Selection for breast cancer cells that stably express fusion protein
- ◆ The Repertoire and Evolution of ATP-Binding Cassette Systems in *Synechococcus* and *Prochlorococcus* (collaborated with Dr. Jinyu Wu)
 - Data collection from public database
 - Organize data and build workflow with Perl scripts
 - Phylogenetic analysis with bioinformatics tools and statistical packages: BLAST, ORTHOLOG, MEGA, PAML, SPSS

TRAINING/WORKSHOPS ATTENDED

Jul. 2018 G-OnRamp Workshop (Create Genome Browsers for Genome Annotation)
Washington University in St. Louis, MO, USA

- ◆ Handle RNA-Seq data using Galaxy, Use G-OnRamp to Create Customized Genome Browsers, Explore Genome Browsers, Set up Collaborative Genome Annotation Environment on Apollo, Use Apollo to Edit Genome Annotation, Virtual Machine and Cloud Launch Deployment, Differential Expression Analysis with RNA-Seq

Oct. 2017 Data Visualization in R- Bioinformatics Workshop

Wyoming INBRE Bioinformatics Core

- ◆ R, RStudio, Data Importing, Visualization: Geographical Maps, Heat Maps, Bubble Scatter Plots

Jul. 2016 4-days Short Course on Next-Generation Sequencing: Technology and Statistical Methods, funded by NHGRI, NIH

University of Alabama at Birmingham, AL, USA

HudsonAlpha Institute for Biotechnology, AL, USA

- ◆ NGS Platforms/Chemistry/Library Preparation, Methylation and Bisulfite Sequencing, Variant Calling and Assembly NGA Data, Virtual Machine, Regulatory Genome, Genome Variation Analysis, GWAS, Rare Variants Analysis, ChiP-Seq Data Analysis, RNA-Seq, Single Cell Sequencing, Microbiome Data Analysis

Jun. 2012 ~ 2 Months Paid Bioinformatics Intensive Internship

Aug. 2012 National Center for Genome Resources, Santa Fe, NM, USA

- ◆ Two weeks systematic training on UNIX commands, DNA sequencing technology and data analysis: Unix, Genome Assembly, Annotation, Genome Variants, Differential Gene Expression, and Pathways Analysis
- ◆ Hands on Project 1: Differential gene expression analysis of MCF7 (breast cancer) cells under 4 different treatments (ethanol, Nature hormone 17 alpha-estradiol, GPER-specific agonist and fluorescent estradiol-based estrogen probe)
- ◆ Hands on Project 2: Genome assembly and comparative genomic analysis of Illumina HiSeq 2000 sequencing reads from five *Staphylococcus aureus* strains.

2011 ~ 2012 Graduate Assistant

Molecular Biology Facility, University of New Mexico, Albuquerque, NM, USA

- ◆ Sanger sequencing sample preparation
- ◆ Wrote Perl scripts for local 454 sequencing results summary report
- ◆ Content Management Platform (Drupal, MediaWiki, DocuWiki) testing

RESEARCH SKILLS

Comparative Genomics and Molecule Evolution

- ◆ Familiar with primary and secondary biology/biomedical database and online bioinformatics tools: Ensemble, NCBI, PFAM, RAST, EBI, UCSC Genomics, Galaxy, WormBase, SUPERFAMILY, STRING, KEGG PATHWAY Database, and Stanford Microarray Database (SMD) etc.
- ◆ High-throughput sequence analysis
 - Pipeline construction and data organization with Unix and Perl scripts
 - Familiar with NGS tools from reads cleaning, quality control, to reference mapping, gene expression analysis, de novo transcriptome or genome assembly
 - Statistical tests and chart generation with R packages
- ◆ NGS tools: SAMtools, CNV detection (Pindel, Delly, CNVnator), RNA-Seq (DESeq, edgeR), IGV, GMap/Gsnap, VCFtools
- ◆ Phylogenetic Analysis: Familiar with sequence phylogeny analysis tools: BLAST, Clustal, muscle, PAML, PHYLIP, MEGA, and MrByes etc.

Computer Skills

- ◆ 9-year experience working with command-lines on UNIX systems
- ◆ Quickly master new computer software, skills for PC and network hardware assembly
- ◆ Parallel computing with GNU-Parallel on computer clusters to elevate work efficiency
- ◆ Statistical tools: Proficient in R (edgeR, DESeq, ggplot2), Excel, SPSS, and Matlab
- ◆ Perl language: proficient in 1) Pipeline construction, 2) data mining, and 3) visualization
- ◆ Genomic rearrangement visualization using GBrowse_syn
- ◆ Other skills include: MySQL, PHP, SmartDraw, Photoshop, CorelDRAW, and Content Management Platform (Drupal, MediaWiki, and DocuWiki)
- ◆ Set up and maintain Galaxy server for users who don't deal with command lines

Molecular Biology

- ◆ *E. coli* culture, preservation, and transformation; human breast cancer cell culture and transformation; extraction and purification of DNA; PCR primer design; PCR and DNA sequencing; agarose gel electrophoresis of DNA; and agarose gel DNA fragment recovery

PEER REVIEWD PUBLICATIONS

Google Scholar Link

https://scholar.google.com/citations?hl=en&user=BFcS-9cAAAAJ&view_op=list_works&sortby=pubdate

- 1) Schultz, J.H., **Bu, L.** and Adema, C.M., 2018. Comparative immunological study of the snail *Physella acuta* (Hydrophila, Pulmonata) reveals shared and unique aspects of gastropod immunobiology. *Molecular Immunology*, 101, pp.108-119.
- 2) Yin, G., Hua, S.S.T., Pennerman, K.K., Yu, J., **Bu, L.**, Sayre, R.T. and Bennett, J.W., 2018. Genome sequence and comparative analyses of atoxigenic *Aspergillus flavus* WRRL 1519.
- 3) Zhang, S.M., **Bu, L.**, Laidemitt, M.R., Lu, L., Mutuku, M.W., Mkoji, G.M. and Loker, E.S., 2018. Complete mitochondrial and rDNA complex sequences of important vector species of *Biomphalaria*, obligatory hosts of the human-infecting blood fluke, *Schistosoma mansoni*. *Scientific reports*, 8(1), p.7341.
- 4) Buddenborg, S.K., **Bu, L.**, Zhang, S.M., Schilkey, F.D., Mkoji, G.M. and Loker, E.S., 2017. Transcriptomic responses of *Biomphalaria pfeifferi* to *Schistosoma mansoni*: Investigation of a neglected African snail that supports more *S. mansoni* transmission than any other snail species. *PLoS neglected tropical diseases*, 11(10), p.e0005984.
- 5) Reid, K. M., Patel, S., Robinson, A. J., **Bu, L.**, Jarungsriapisit, J., Moore, L. J., & Salinas, I. (2017). Salmonid alphavirus infection causes skin dysbiosis in Atlantic salmon (*Salmo salar* L.) post-smolts. *PloS one*, 12(3), e0172856.
- 6) Yin, G., Zhang, Y., Hua, S. S. T., Yu, J., **Bu, L.**, Pennerman, K. K., ... & Bennett, J. W. (2017). Genome Sequencing and Analysis of the Postharvest Fungus *Penicillium expansum* R21. *Genome Announcements*, 5(7), e01516-16.
- 7) **Bu, L.** and Katju, V., 2015. Early evolutionary history and genomic features of gene duplicates in the human genome. *BMC genomics*, 16(1), p.621.
- 8) Katju, V., Packard, LB., **Bu, L.**, Keightley, PD., Bergthorsson, U. 2014. Fitness decline in spontaneous mutation accumulation lines of *Caenorhabditis elegans* with varying effective population sizes. *Evolution: Early View* (DOI: 10.1111/evo.12554).
- 9) Mo, G., Ding, Q., Chen, Z., Li, Y., Yan, M., **Bu, L.**, Song, Y., Yin, G. 2014. A Novel Mutation in the *RPE65* Gene Causing Leber Congenital Amaurosis and Its Transcriptional Expression *In Vitro*. *PLoS One* 9:e112400.
- 10) Bartels, J., Darrow, B. G., Schatzberg, S. J., **Bu, L.**, Carlson, R., and Tipold, A. 2014. MIP-3beta/CCL19 is associated with the intrathecal invasion of mononuclear cells in neuroinflammatory and non-neuroinflammatory CNS diseases in dogs. *BMC Veterinary Research*. Vol. 10, No. 1: 157.
- 11) Zhou, Y., **Bu, L.**, Guo, M., Zhou, C., Wang, Y., Chen, L., and Liu, J. 2013. Comprehensive genomic characterization of *Campylobacter* genus reveals some underlying mechanisms for its genomic diversification. *PloS One*. Vol. 8, No. 8: e70241.

- 12) **Bu, L.**, Bergthorsson, U., and Katju, V. 2011. Local synteny and codon usage contribute to asymmetric sequence divergence of *Saccharomyces cerevisiae* gene duplicates. *BMC Evolutionary Biology*. Vol. 11, No. 1: 279-287.
- 13) Bai J., Wang, J., Xue F., Li, J., **Bu, L.**, Hu, J., Xu, G., Bao, Q., Zhao, G., Ding, X., Yan, J., and Wu, J. 2010. proTF: a comprehensive data and phylogenomics resource for prokaryotic transcription factors. *Bioinformatics*. Vol. 26, No. 19: 2493-2495.
- 14) **Bu, L.**, Xiao, J., Lu, L., Xu, G., Li, J., Zhao, F., Li, X., and Wu, J. 2009. The repertoire and evolution of ATP-binding cassette systems in *Synechococcus* and *Prochlorococcus*. *Journal of Molecular Evolution*. Vol. 69, No. 4: 300-310.
- 15) **Bu, L.**, Su, S., Li, H., Lin, B., and Zhang, Y. 2009. Relationship between the methylation of ER α gene promotor in breast cancer and the protein expression. *Journal of Wenzhou Medical College*. Vol. 1, No. 2: 0133-0134.

PEER REVIEWD PUBLICATIONS IN PREPARATION

- 1) **Bu, L.**, Hanna, B., Schilkey, F., Loker, E., Zhang, SM. Genome Sequence of a homozygous line of *Biomphalaria glabrata*, the intermediate host of *Schistosoma mansoni*, with implications for understanding resistance to infection.
- 2) **Bu, L.** The Gene Duplication Database: Identification and Structural Classification of Gene Duplication Events

INVITED PRESENTATIONS

2017 New Mexico Consortium Biolab Seminar, Los Alamos, NM, USA

The Gene Duplication Database: Identification and Structural Classification of Gene Duplication Events

2016 New Mexico Cyberinfrastructure Capstone Event, Albuquerque, NM, USA

Researcher Panel member represent the University of New Mexico, talked about the challenge, current work and future need for the next generation sequencing data analysis.

2016 Data to knowledge Interdisciplinary Research Symposium held by the College of Arts and Sciences (A&S) at The University of New Mexico, Albuquerque, NM, USA

Lightening talk title: Local Cloud for NGS Data with A Public Galaxy Server: Open source, Web-based Bioinformatics Platform

TEACHING EXPERIENCE

2016 Workshop for RNA-Seq analysis using Galaxy, Center for Advanced Research Computing, University of New Mexico, NM, USA

Galaxy platform setup, 3 hours workshop presentation and hands on activity for users from multiple departments of UNM, including Biology, Biomedical school, Chemistry, and computer science.

2016 University of New Mexico Bioinformatics Club

Brief introduction to Galaxy, a web-based tool for bioinformatics without command line

2013 University of New Mexico Bioinformatics Club

Tutorial of Unix command lines

2010~ Teaching Assistant of Molecular Biology Lab (x4) and Genetics Lab (x3)

2014

In the 75 minutes lab, I organize undergraduate students to perform hands on lab activities, practice and review knowledge points from lecture, discuss questions related to lecture, homework and exams. Outside of the lab, I grade homework and exams and hold office hours to help undergraduate students with their questions.