

Fangfei Li

✉ fangfeili0525@gmail.com

☎ 631 833-7934

📍 Westampton, NJ

🌐 <https://fangfeili05.github.io/>

🔗 <https://github.com/FangfeiLi05/>

🌐 <https://www.linkedin.com/in/fangfeili/>

SUMMARY & SKILLS

• **PhD in Applied Mathematics & Statistics (Computational Biology Track)** | 2 yrs post-doc + 2 yrs biotech startup + 3 yrs wet-lab (synthetic barcoding-based *in vivo* lineage tracing, molecular biology) • **Skills:** Bayesian optimization, algorithms, machine learning (deep generative model), bioinformatics, data analysis & visualization, Python (NumPy, Matplotlib, pandas, tkinter, Flet), MATLAB, SQL, Git • Strong analytical thinker; detail-oriented, organized, and goal-driven.

INDUSTRIAL & ACADEMIC EXPERIENCE

Online AI Training with Certificates ————— **9/2025-current**
Artificial Intelligence Professional Program, Stanford

- Deep Generative Models (XCS236): Certificate

Computational Biology Staff Scientist ————— **11/2023-8/2025**
BacStitch DNA, South San Francisco, CA

- **Developed an automated end-to-end DNA assembly workflow (Python, web/GUI)** that scores sequences, designs optimal overlapping blocks, checks synthesizability via supplier APIs, and generates automated high-throughput experimental plans.
- **Built a Python tool for optimal DNA block design** that segments sequences into overlapping blocks with minimal structural features in overlaps, under customizable constraints on block count, block/overlap length, and specific regions forced as blocks/overlaps.
- **Created Python visualization tools** to (i) plot BLAST results and structural scores ([Example](#)), (ii) generate *.ab1* files encoding structural scores as ATGC traces at single-nucleotide resolution, and (iii) produce publication-quality figures for publications and internal use.

Postdoctoral Research Fellow ————— **4/2021-10/2023**
Sherlock Lab at Department of Genetics, Stanford University, CA

- Developed a Bayesian optimization algorithm and Python pipeline (*FitMut2*) to identify genome-wide adaptive lineages from fitness phenotypes (without WGS) in experimental evolution and estimate mutation fitness effects and occurrence times.
- Developed an optimization algorithm to segment DNA sequences into overlapping fragments that minimize overlap-region scores (e.g., undesirable DNA structural features) under fixed or ranged constraints on fragment count and fragment/overlap length.

Artificial Intelligence Fellow ————— **9/2019-10/2019**
Insight Data Science, San Francisco, CA

- Developed a deep learning pipeline for portrait editing (a 6-week project), providing continuous adjustment of facial features.

Graduate Researcher ————— **8/2014-12/2020**
Laufer Center for Physical and Quantitative Biology, Stony Brook University, NY

- Developed a fitness estimation tool (*Fit-Seq*) that can estimate unbiased fitness for complex cell pools, using a maximum likelihood algorithm.

- Designed and conducted *in vivo* bar-seq experimental yeast evolution (yeast genetics, cloning, barcoded library construction, lineage-traced evolution, and NGS prep). Performed computational analysis (NGS processing, mathematical modeling for inferring evolutionary parameters from lineage dynamics, and evolutionary simulations for validation) to explore the epistasis of adaptive mutations in yeast evolution.

EDUCATION **PhD, Applied Mathematics and Statistics (Computational Biology)**——8/2013–12/2020
 Stony Brook University, Stony Brook, NY

MS, Systems Theory——9/2010–7/2013
 University of Chinese Academy of Sciences, Beijing, China

BS, Mathematics and Applied Mathematics——9/2006–7/2010
 Beijing Normal University, Beijing, China

PRESENTATIONS **Plenary Talk: Fitness Estimation of Pooled Amplicon Sequencing Studies**——8/2018
 2018 Yeast Genetics Meeting, Stanford CA, 8/22/2018–8/26/2018

PUBLICATIONS

- [1] T Matsui, P-H Hung, H Mei, X Liu, **F Li**, J Collins, W Li, D Miller, N Wilson, E Toro, G J Taghon, G Sherlock, S Levy. High-throughput DNA engineering by mating bacteria. *Preprint* (2024).
- [2] **F Li**, A Mahadevan, G Sherlock. An improved algorithm for inferring mutational parameters from Bar-seq evolution experiments. *BMC Genomics* 24, 246 (2023).
- [3] **F Li**^{*}, J Tarkington^{*}, G Sherlock. Fit-Seq-2.0: an improved software for high throughput fitness measurements using pooled competition assays. *Journal of Molecular Evolution* 91, 334–344 (2023).
- [4] G Avecilla, J N Chuong, **F Li**, G Sherlock, D Gresham, Y Ram. Neural networks enable efficient and accurate simulation-based inference of evolutionary parameters from adaptation dynamics. *PLoS Biology* 20, e3001633 (2022).
- [5] Z Liu, D Miller, **F Li**, X Liu, S Levy. A large accessory protein interactome is rewired across environments. *eLife* 9, e62365 (2020).
- [6] X Liu, Z Liu, A K Dziulko, **F Li**, D Miller, R D Morabito, D Francois, S Levy. iSeq 2.0: a modular and interchangeable toolkit for interaction screening in yeast. *Cell Systems* 8, 338–344 (2019).
- [7] **F Li**, M L Salit, S Levy. Unbiased fitness estimation of pooled barcode or amplicon sequencing studies. *Cell Systems* 7, 521–525 (2018).
- [8] I Frumkin, D Schirman, A Rotman, **F Li**, L Zahavi, E Mordret, O Asraf, S Wu, S Levy, Y Pilpel. Gene architectures that minimize cost of gene expression. *Molecular Cell* 65, 142–153 (2017).
- [9] Y Chen, W Xiong, **F Li**. Convergence of infinite products of stochastic matrices: a graphical decomposition criterion. *IEEE Transactions on Automatic Control* 61, 3599–3605 (2016).
- [10] Y Chen, **F Li**, B Hou, S Tan, H Zhu. Convergence analysis of discrete-time consensus algorithm with both self and transmission delays. *Journal of the Franklin Institute* 353, 2467–2481 (2016).
- [11] **F Li**, Y Chen, J Lü, D Hill. Cluster consensus of Boolean multi-agent systems. *2013 9th Asian Control Conference (ASCC)* (2013).
- [12] Y Chen, W Yu, **F Li**, S Feng. Synchronization of complex networks with impulsive control and disconnected topology. *IEEE Transactions on Circuits and Systems II: Express Briefs* 60, 292–296 (2013).