

Ilya Shlyakhter

Address 80 Gardner St Apt 21, Allston, MA 02134 **Mobile Phone** (617) 256 6780
Email ilya_shl@alum.mit.edu

Summary

Highly effective computational biologist with a strong computer science background (Ph.D.) and extensive experience in genomic data analysis. Expert in genome assembly algorithms, NGS data processing methods for DNA-seq and RNA-seq, bioinformatics pipeline development, population genetics simulations and evolutionary genomics. Key contributor to scientific projects reported in leading journals. Go-to person for hard algorithmic questions. Valued by computational and wetlab colleagues for “rigor of thinking on a broad array of computational and analytic problems” as well as “willingness to simply do what most needs doing”.

Technical skills

Languages: Python, R, Java, C/C++, Lisp, bash, SQL. Bioinformatics/genomics tools: Biopython, GATK, sam-tools, Picard, bwa, BMAP, MUMmer, plink and others. Cloud computing: Amazon Web Services (AWS), Google Cloud Platform (GCP), DNAnexus, Terra. Workflow tools: snakemake, Crowell/WDL, Docker, git.

Professional Experience

Feb 2007 - Present Computational R&D Group and Sabeti Lab, Broad Institute of MIT and Harvard, Cambridge, MA
Computational Biologist II (06/16-present), Computational Biologist (02/07-06/16)

- Co-developed a method for finding targets of positive selection in the human genome
- Wrote a flexible and efficient population genetics simulator
- Enabled fitting of a refined population genetics model to 1000 Genomes project data
- Improved the scalability of short-read genome assembly algorithms
- Optimized pipelines for assembling diverse viral genomes from clinical samples
- Automated identification of the most relevant reference genomes
- Created tools to measure optimality of assemblies relative to raw data
- Built a flexible framework for benchmarking assembly methods
- Proposed and built custom visualization tools for exploring genome assemblies
- Served as a resource to many colleagues for computational questions

Oct 2006 - Feb 2007 Department of Molecular Biology, Princeton University, Princeton, NJ
Bioinformatician (volunteer)

- Analyzed the genomic distribution of DNA motifs that induce self-catalyzed DNA damage

Feb 2005 - Feb 2007 Formal Verification Group, NEC Laboratories America, Princeton, NJ
Research Staff Member

- Adapted automatic parallelization and program transformation techniques to Bounded Model Checking of software
- Developed efficient methods for model checking programs with large arrays
- Developed methods for divide-and-conquer analysis of large programs

Education

1993-2005 Massachusetts Institute of Technology, Cambridge, MA
Ph.D. (2005), M.Eng (1999) and B.S. (1997) in Computer Science, B.S. in Mathematics (1997)

2005-2012 Stanford University, Johns Hopkins University, CalTech

11 courses in biology, bioinformatics (Stanford Graduate Certificate), machine learning

Selected publications

- J. Vitti, S. Gosai, I. Shlyakhter, R. Tewhey, S. Reilly, P. C. Sabeti. Identification and characterization of adaptive regulatory variation in diverse human populations. (*in preparation*).
- The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*, 526(7571):68–74, Oct. 2015
- I. Shlyakhter, P. C. Sabeti, and S. F. Schaffner. Cosi2: An efficient simulator of exact and approximate coalescent with selection. *Bioinformatics (Oxford, England)*, 30(23):3427–3429, Dec. 2014
- E. K. Karlsson, J. B. Harris, S. Tabrizi, A. Rahman, I. Shlyakhter, N. Patterson, C. O’Dushlaine, S. F. Schaffner, S. Gupta, F. Chowdhury, A. Sheikh, O. S. Shin, C. Ellis, C. E. Becker, L. M. Stuart, S. B. Calderwood, E. T. Ryan, F. Qadri, P. C. Sabeti, and R. C. Larocque. Natural selection in a bangladeshi population from the cholera-endemic ganges river delta. *Science Translational Medicine*, 5(192):192ra86, July 2013
- S. R. Grossman*, K. G. Andersen*, I. Shlyakhter*, S. Tabrizi*, S. Winnicki, A. Yen, D. J. Park, D. Griesemer, E. K. Karlsson, S. H. Wong, M. Cabili, R. A. Adegbola, R. N. K. Bamezai, A. V. S. Hill, F. O. Vannberg, J. L. Rinn, 1000 Genomes Project, E. S. Lander, S. F. Schaffner, and P. C. Sabeti. Identifying recent adaptations in large-scale genomic data. *Cell*, 152(4):703–713, Feb. 2013
- K. G. Andersen, I. Shlyakhter, S. Tabrizi, S. R. Grossman, C. T. Happi, and P. C. Sabeti. Genome-wide scans provide evidence for positive selection of genes implicated in Lassa fever. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1590):868–877, Mar. 2012
- D. Van Tyne, D. J. Park, S. F. Schaffner, D. E. Neafsey, E. Angelino, J. F. Cortese, K. G. Barnes, D. M. Rosen, A. K. Lukens, R. F. Daniels, D. A. Milner, C. A. Johnson, I. Shlyakhter, and eighteen others. Identification and functional validation of the novel antimalarial resistance locus PF10_0355 in *Plasmodium falciparum*. *PLoS genetics*, 7(4):e1001383, Apr. 2011
- S. R. Grossman*, I. Shlyakhter*, E. K. Karlsson, E. H. Byrne, S. Morales, G. Frieden, E. Hostetter, E. Angelino, M. Garber, O. Zuk, E. S. Lander, S. F. Schaffner, and P. C. Sabeti. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science (New York, N.Y.)*, 327(5967):883–886, Feb. 2010
- International HapMap 3 Consortium. Integrating Common and Rare Genetic Variation in Diverse Human Populations. *Nature*, 467(7311):52–58, Sep. 2010
- I. Maccallum, D. Przybylski, S. Gnerre, J. Burton, I. Shlyakhter, A. Gnirke, J. Malek, K. McKernan, S. Ranade, T. P. Shea, L. Williams, S. Young, C. Nusbaum, and D. B. Jaffe. ALLPATHS 2: Small genomes assembled accurately and with high continuity from short paired reads. *Genome Biology*, 10(10):R103, 2009
- J. Butler, I. MacCallum, M. Kleber, I. A. Shlyakhter, M. K. Belmonte, E. S. Lander, C. Nusbaum, and D. B. Jaffe. ALLPATHS: De novo assembly of whole-genome shotgun microreads. *Genome Research*, 18(5):810–820, May 2008
- A. Zaks, Z. Yang, I. Shlyakhter, F. Ivancic, S. Cadambi, M. K. Ganai, A. Gupta, and P. Ashar. Bitwidth Reduction via Symbolic Interval Analysis for Software Model Checking. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 27(8):1513–1517, Aug. 2008
- I. Shlyakhter. Generating effective symmetry-breaking predicates for search problems. *Discrete Applied Mathematics*, 155(12):1539–1548, June 2007
- F. Ivancic, I. Shlyakhter, A. Gupta, M. Ganai, V. Kahlon, C. Wang, and Z. Yang. Model checking C programs using F-Soft. In *2005 International Conference on Computer Design*, pages 297–308, Oct. 2005
- I. Shlyakhter, R. Seater, D. Jackson, M. Sridharan, and M. Taghdiri. Debugging overconstrained declarative models using unsatisfiable cores. In *18th IEEE International Conference on Automated Software Engineering, 2003. Proceedings.*, pages 94–105, Oct. 2003 (best paper award)