Mikhail Kolmogorov

Postdoctoral Fellow
Department of Computer Science and Engineering,
University of California, San Diego

Email: *fenderglass@gmail.com*
Web: *kolmogorov.org*

Research Interests

The focus of my research is *computational biology* - algorithms, mathematical models and tools aimed to answer fundamental questions about living systems through the analysis of large-scale biological data. Specifically, I have extensive experience in computational genomics and proteomics, with a focus on the design and development of efficient algorithmic approaches that bridge theoretical models and real experimental data.

Research Positions

- **Posdoctoral Fellow** *Oct 2019 - present*
  University of California, San Diego, USA
  Advisors: Rob Knight and Pavel Pevner

Education

- **PhD, Computer Science** *Sep 2014 - Sep 2019*
  Department of Computer Science and Engineering,
  University of California, San Diego, USA
  Thesis: Algorithms for long-read assembly
  Advisor: Pavel Pevzner

- **MSc, Bioinformatics** *Sep 2012 – Jul 2014*
  St. Petersburg Academic University, Russia
  Thesis: Genome assembly using multiple references
  Advisor: Son Pham

- **BSc, Applied Mathematics and Informatics** *Sep 2008 - Jul 2012*
  Department of Mathematics, ITMO University,
  St.Petersburg, Russia
  Thesis: Variational method for the minimum energy path calculation
  Advisor: Igor Popov

Journal Publications

- **Mikhail Kolmogorov,** Derek M. Bickhart, Bahar Behsaz, Alexey Gurevich, Mikhail Rayko, Sung Bong Shin, Kristen Kuhn, Jeffrey Yuan, Evgeny Polevikov, Timothy P. L. Smith and Pavel A. Pevzner. "metaFlye: scalable and accurate long-read metagenome assembler". *Nature Methods* (2020) *doi: 10.1038/s41592-020-00971-x*

- **Mikhail Kolmogorov,** Jeffrey Yuan, Yu Lin and Pavel Pevzner. "Assembly of long error-prone reads using repeat graphs". *Nature Biotechnology* 37 540-546 (2019)

- Alla Mikheenko and **Mikhail Kolmogorov. "**Assembly Graph Browser: interactive visualization of assembly graphs". *Bioinformatics* 35 3476-3478 (2019)

- **Mikhail Kolmogorov**, Joel Armstrong, Brian J. Raney, Ian Streeter, Matthew Dunn, Fengtang Yang, Duncan Odom, Paul Flicek, Thomas Keane, David Thybert, Benedict Paten and Son Pham. "Chromosome assembly of large and complex genomes using multiple references". *Genome Research* 28 (11), 1720-1732 (2018)

- Alex Bishara, Eli L Moss, **Mikhail Kolmogorov**, Alma Parada, Ziming Weng, Arend Sidow, Anne E Dekas, Serafim Batzoglou, Ami S Bhatt. "High-quality genome sequences of uncultured microbes by assembly of read clouds". *Nature Biotechnology* 36, 1067–1075 (2018)

- Jingtao Lilue, Anthony G Doran, Ian T Fiddes, Monica Abrudan, Joel Armstrong, (22 authors), **Mikhail Kolmogorov,** (29 authors), Benedict Paten, Thomas M Keane. "Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci". *Nature Genetics 50,* 1574–158 (2018)

- David Thybert, Maša Roller, Fábio CP Navarro, Ian Fiddes, Ian Streeter, Christine Feig, David Martin-Galvez, **Mikhail Kolmogorov**, (33 authors), Benedict Paten, Son Pham, Thomas M Keane, Duncan T Odom, Paul Flicek. "Repeat associated mechanisms of genome evolution and function revealed by the Mus caroli and Mus pahari genomes". *Genome Research 28 (4),* 448-459 (2018)

- **Mikhail Kolmogorov**, Eamonn Kennedy, Zhuxin Dong, Gregory Timp and Pavel Pevzner. "Single-Molecule Protein Identification by Sub-Nanopore Sensors". *PloS Computational Biology* 13 (5), e1005356 (2017)

- Yu Lin*, Jeffrey Yuan*, **Mikhail Kolmogorov***, Max Shen and Pavel Pevzner. "Assembly of long error-prone reads using de Bruijn graphs". *PNAS 113 (52),* E8396-E8405 (2016)*, *equal contribution

- **Mikhail Kolmogorov**, Xiaowen Liu and Pavel Pevzner. "SpectroGene: a tool for proteogenomic annotations using top-down spectra". *Journal of Proteome Research* 15 (1), 144-151 (2015)

Refereed Conference Proceedings

- Evgeny Polevikov and **Mikhail Kolmogorov. "**Synteny paths for assembly graphs comparison". *Workshop on Algorithms in Bioinformatics 2019, Niagara Falls, NY*

- **Mikhail Kolmogorov,** Jeffrey Yuan, Yu Lin and Pavel Pevzner. "Assembly of long error-prone reads using repeat graphs". *RECOMB* 2018, *Paris, France* (journal version in Nature Biotech. 2019)

- Yu Lin*, Jeffrey Yuan*, **Mikhail Kolmogorov***, Max Shen and Pavel Pevzner. "Assembly of long error-prone reads using de Bruijn graphs". *RECOMB 2016, Santa Monica, CA* (journal version in PNAS 2016), *equal contribution

- **Mikhail Kolmogorov**, Eamonn Kennedy, Zhuxin Dong, Gregory Timp and Pavel Pevzner. "Single-Molecule Protein Identification by Sub-Nanopore Sensors". *RECOMB-SEQ 2016, Santa Monica, CA* (journal version in PLoS Comp. Biol 2017)

- **Mikhail Kolmogorov**, Brian Raney, Benedict Paten, Son Pham. "Ragout - a reference-assisted assembly tool for bacterial genomes". *ISMB 2014, Boston, MA*

- Ilya Minkin, Anand Patel, **Mikhail Kolmogorov**, Nikolay Vyahhi, Son Pham**. "**Sibelia: a scalable and comprehensive synteny block generation tool for closely related microbial genomes". *Workshop on Algorithms in Bioinformatics 2013, Antibes, France*

Conference Oral Presentations

- "Long-read assembly: are metagenomes easier than diploid genomes?". *PAG 2020, San Diego, CA* (invited)

- "metaFlye: Scalable long-read metagenome assembly using repeat graphs". *Genome Informatics 2019, Cold Spring Harbor, NY* (selected abstract)

- **"**Synteny paths for assembly graphs comparison". *Workshop on Algorithms in Bioinformatics 2019, Niagara Falls, NY* (proceedings author)

- "Assembly of long error-prone reads using repeat graphs". *AGBT 2019, Marco Island, FL* (selected abstract)

- "Single-Molecule Protein Identification by Sub-Nanopore Sensors". *HUPO 2018, Orlando, FL* (selected abstract)

- "Assembly of long error-prone reads using repeat graphs". *RECOMB* 2018, *Paris, France* (proceedings author)

- "Assembly of long error-prone reads using repeat graphs". *PAG 2018, San Diego, CA* (invited)

- "Single-Molecule Protein Identification by Sub-Nanopore Sensors". *RECOMB-SEQ 2016, Santa Monica, CA* (proceedings author)

- "SpectroGene: a tool for proteogenomic annotations using top-down spectra". *ASMS 2015, St. Louis, MO* (selected abstract)

- "Ragout - a reference-assisted assembly tool for bacterial genomes". *ISMB 2014, Boston, MA* (proceedings author)

Seminar and Symposium Talks

- "Assembly of long error-prone reads using repeat graphs". Center for Microbiome Innovation at UCSD (2019)

- "Assembly of long error-prone reads using repeat graphs". Australian National University, Long-read, long-range workshop (2019)

- "Completing human genome and microbiome: algorithms for long-read assembly", NYU Center for Human Genetics and Genomics symposium (2019)

- "Metagenomic assembly using long reads", ITMO University, bioinformatics seminar series (2020)

- "Metagenomic assembly using long reads", Universidad Miguel Hernandez, metagenomic seminar series (2020)

Mentoring and Advising

- Mentor in the Science Mentors program for Russian-speaking students (*https://sites.google.com/view/sciencementors*), 2020

- Evgeny Polevikov. "Synteny paths for assembly graphs comparison" Master thesis, ITMO University, St. Petersburg, Russia, 2018-2019

- Alla Mikheenko. "Assembly Graph Browser: interactive visualization of assembly graphs" Research project, St. Petersburg Center for Algorithmic Bioinformatics, Russia, 2018-2019

Teaching

- Teaching assistant   *Winter 2016*
  Algorithms in Molecular Biology. UC San Diego

- Teaching assistant   *Spring 2013*
  Bioinformatics Seminar. St. Petersburg Academic University

Fellowships

- CSE Doctoral Fellow Award (UCSD)   Sep *2014 - Aug 2016*

Open Source Projects

- **Flye** – De novo assembler for single molecule sequencing reads (*main developer*)
  https://github.com/fenderglass/Flye

- **Ragout** – Chromosome-level scaffolding using multiple references (*main developer*)
  https://github.com/fenderglass/Ragout

- **Nano-Align** – Protein identification using a nanopore *(main developer)*
  https://github.com/fenderglass/Nano-Align

- **SpectroGene** – Gene prediction using top-down mass-spectrometry *(main developer)*
  https://github.com/fenderglass/SpectroGene

Industry Experience

- Research Intern   *Summer 2017*
  Illumina, San Diego, CA, USA
  Mentors: Dorna Kashef and Serafim Batzoglou

Summer Schools

- Bioinformatics Summer School 2016 (Lecturer)
  Moscow, Russia

- Bioinformatics Summer School 2014 (Lecturer)
  St. Petersburg, Russia