

# Andrey O. Kislyuk

---

CONTACT INFORMATION	415 Pennsylvania Ave., Suite 2 San Francisco, CA 94107 U.S.	Tel: +1 (925) 683-0172 kislyuk@gmail.com <a href="http://kislyuk.com/">http://kislyuk.com/</a>
OBJECTIVE	Director, software engineering, bioinformatics or genomics.	
OVERVIEW	Technical leader in software engineering and genomics. Expert in design and implementation of genome analysis systems and cloud platforms. Bioinformatics specialist with expertise in DNA sequence analysis, machine learning, scalable distributed systems architecture, API design, Linux systems, information security, and development operations.	
EMPLOYMENT	<p><b>DNAnexus, Inc.</b>, 2011 – 2016</p> <p><b>Director, Software Architecture:</b> Technical lead on the full stack of the DNAnexus Platform, a global genomics service. Participated in the design and implementation of every major system component from inception to multi-petabyte, 100M core-hour workload, \$5M+ run rate. Supervisor: Omar Serang, Chief Cloud Officer</p> <p><b>Senior Software Engineer:</b> Architecture and development of AWS-based bioinformatics solutions. Supervisor: Andreas Sundquist, CEO</p> <p><b>Pacific Biosciences of California</b>, 2010 – 2011</p> <p><b>Scientist, Bioinformatics:</b> Research, development, and support of various bioinformatics projects. Supervisor: Stephen Turner, CTO</p>	
SPECIALTIES AND AREAS OF EXPERTISE	<ul style="list-style-type: none"><li>• Team management (technical product management, Agile &amp; schedule driven planning)</li><li>• Software engineering (general with emphasis on software architecture; Python, node.js, Java, C/C++, Linux, networking and distributed systems)</li><li>• Amazon Web Services-based architectures (EC2, S3, IAM, VPC, Lambda, RDS, EFS, CloudWatch/CWL, CodeDeploy)</li><li>• Container-based virtualization (Docker/LXC)</li><li>• Genome sequence analysis (mapping, assembly, annotation, variant calling and interpretation, sequencing technology R&amp;D, metagenomics, phylogenomics/evolutionary genomics, epigenomics/methylation detection, single molecule sequencing, visualization/data mining interface design)</li><li>• Cloud and web information security (Docker/LXC security, EC2 security infrastructure, OAuth2, SAML2, cloud IDS)</li><li>• Machine learning (probabilistic graphical models, Markov models/CRFs)</li><li>• High performance computing (cluster architecture, optimization techniques, cloud HPC)</li><li>• Scalable distributed systems architecture (distributed data stores, HA, load balancing, AWS, GCE)</li><li>• Development operations (Linux, CI/CD and deployment at scale, release engineering, configuration management, Ansible, Chef, Fabric)</li><li>• API design (documentation systems, SDK development)</li></ul>	
EDUCATION	<p><b>Georgia Institute of Technology</b>, Atlanta, Georgia Ph.D. 2010, Bioinformatics</p> <p><b>University of California, Berkeley</b> B.A. 2004, Computer Science, Mathematics, and Statistics</p>	<p>Advisor: Joshua S. Weitz</p> <p>Advisor: Inna Dubchak</p>

## PROJECTS

Technical lead on the full stack of the DNAnexus Platform (2011 –).

Published many popular Python modules (argcomplete, ensure, SignXML, etc.) (2012 –).

Researched and developed a DNA modification detection pipeline. (Pacific Biosciences, 2010 –; 2 publications.)

Researched, developed, and evaluated an unsupervised metagenomic binner application. (Georgia Tech, 2008 – 2009; 1 publication.)

Researched, developed, and evaluated a prokaryotic frameshift finder application for genome sequencing quality control. (Georgia Tech/DOE JGI, 2007 – 2008; 1 publication.)

Developed and integrated a set of algorithms for gene finding, genome assembly, annotation and error correction in the Georgia Tech Neisseria Base. Supported deployment of a pipeline containing these components. (Georgia Tech/CDC, 2008 – 2009; 3 publications.)

Implemented components of the Berkeley Genome Pipeline. (LBL/UC Berkeley, 2004 – 2005; 1 publication.)

## SOFTWARE ENGINEERING SKILLS

**Programming languages:** Python, Javascript/Coffeescript, C, C++, Java, Perl, Ruby, R, MATLAB, Bash.

**Tools and platforms used:** Extensive knowledge of bioinformatics tools, algorithms, and frameworks for local and whole-genome sequence alignment, assembly, base calling, genome finishing, variant calling, repeat masking, genome browsing, genome data mining, profile alignment, secondary structure prediction, cluster and cloud computing.

**Applications:** Diverse aspects of Linux infrastructure, instrumentation and configuration management toolkits (Gentoo Linux developer), bioinformatics tools (repository maintainer), DBMS (PostgreSQL, MySQL), development tools (Eclipse IDE, Visual Studio), system administration and server management tools, nginx; mathematics, science and engineering suites (Jupyter, R, ggplot, matplotlib, NumPy, BioPython); authoring and content tools (LATEX, BibTeX, MediaWiki).

**Hardware:** Scientific and cluster computing, wide HPC deployments, datacenters, network administration and infrastructure development.

## PUBLICATIONS

D. A. Rasko, D. R. Webster, J. W. Sahl, A. Bashir, N. Boisen, F. Scheutz, E. E. Paxinos, R. Sebra, C.-S. Chin, D. Iliopoulos, A. Klammer, P. Peluso, L. Lee, **A.O. Kislyuk**, J. Bullard, A. Kasarskis, S. Wang, J. Eid, D. Rank, J. C. Redman, S. R. Steyert, J. Frimodt-Møller, C. Struve, A. M. Petersen, K. A. Kroghfelt, J. P. Nataro, E. E. Schadt, and M. K. Waldor, “Origins of the E. coli Strain Causing an Outbreak of Hemolytic–Uremic Syndrome in Germany.” *NEJM*, doi:10.1056/NEJMoa1106920, 2011.

L. S. Katz, J. C. Humphrey, A. B. Conley, V. Nelakuditi, **A.O. Kislyuk**, S. Agrawal, P. Jayaraman, B. H. Harcourt, M. A. Olsen-Rasmussen, M. Frace, N. V. Sharma, L. W. Mayer, and I. K. Jordan, “Neisseria Base: a comparative genomics database for Neisseria meningitidis.” *Database*, doi:10.1093/database/bar035, 2011.

**A.O. Kislyuk**, B. Haegeman, N. H. Bergman, and J. S. Weitz, “Genomic fluidity: an integrative view of gene diversity within microbial populations.” *BMC Genomics*, doi:10.1186/1471-2164-12-32, 2011.

**A.O. Kislyuk**, L. S. Katz, S. Agrawal, M. S. Hagen, A. B. Conley, P. Jayaraman, V. Nelakuditi, J. C. Humphrey, S. A. Sammons, D. Govil, R. D. Mair, K. M. Tatti, M. L. Tondella, B. H. Harcourt, L. W. Mayer, and I. K. Jordan, “A computational genomics pipeline for prokaryotic sequencing

projects,” Bioinformatics, doi:10.1093/bioinformatics/btq284, 2010.

**A.O. Kislyuk**, A. Lomsadze, A.L. Lapidus, and M. Borodovsky. “Frameshift detection in prokaryotic genomic sequences.” International Journal of Bioinformatics Research and Applications, vol. 5, no. 4, pp. 458-477, 2009.

**A.O. Kislyuk**, S. Bhatnagar, J. Dushoff, and J.S. Weitz. “Unsupervised Statistical Clustering of Environmental Shotgun Sequences.” BMC Bioinformatics, doi:10.1186/1471-2105-10-316, 2009.

L.S. Katz, C. Bolen, B. Harcourt, S. Schmink, X. Wang, **A.O. Kislyuk**, R. Taylor, L. Mayer, and I.K. Jordan. “Meningococcus Genome Informatics Platform: a System for Analyzing Multilocus Sequence Typing Data.” Nucleic Acids Research, vol. 37, pp. W606-W611, 2009.

I. Dubchak, A. Poliakov, **A.O. Kislyuk**, and M. Brudno. “Multiple whole genome alignments without a reference organism.” Genome Research, vol. 19, pp. 682-689, 2009.

D. Papatsenko, **A.O. Kislyuk**, M. Levine, and I. Dubchak. “Conservation patterns in different functional sequence categories of divergent drosophila species.” Genomics, vol. 88, no. 4, pp. 431-442, 2006.

#### ACADEMIC EXPERIENCE

- |             |  |
|-------------|--|
| 2008 – 2010 | Georgia Institute of Technology (Advisor: Joshua S. Weitz)<br><b>Research assistant:</b> Research, development, and support of various bioinformatics projects.                      |
| 2005 – 2008 | Georgia Institute of Technology (Advisor: Mark Borodovsky)<br><b>Research assistant:</b> Research, development, and support of various bioinformatics projects.                      |
| 2004 – 2005 | Lawrence Berkeley National Laboratory (Advisor: Inna Dubchak)<br><b>Research assistant:</b> Tools for genome sequence alignment and comparative analysis.                            |
| 2003 – 2004 | Northern Arizona University Mathematics REU Program (Advisor: Nikolay Sirakov)<br><b>Research assistant:</b> Computer vision, image contour and support tracing and comparison.      |
| 2003        | UC Berkeley Space Sciences Laboratory (Advisor: Stephen Mende)<br><b>Lab assistant:</b> Software and hardware solutions for a CCD IR telescope performing gravity wave observations. |

#### POSTERS AND PRESENTATIONS

**A.O. Kislyuk**, J.S. Weitz, and I.K. Jordan. “Comparative Sequence Analysis for Predictive Assembly Gap Closure.” Sequencing, Finishing, and Analysis in the Future 2009, Santa Fe, NM.

**A.O. Kislyuk**, S. Bhatnagar, J. Dushoff, and J.S. Weitz. “Unsupervised Statistical Clustering of Environmental Shotgun Sequences.” Metagenomics 2008, La Jolla, CA.

**A.O. Kislyuk** and J.S. Weitz. “Working with metagenomic data.” DARPA Fundamental Laws of Biology 2008, San Francisco, CA.

**A.O. Kislyuk**, A. Lomsadze, A. Lapidus, and M. Borodovsky. “Frameshift detection in prokaryotic genomic sequences.” Finishing in the Future 2008, Santa Fe, NM.

**A.O. Kislyuk**, A. Lomsadze, A. Lapidus, and M. Borodovsky. “Frameshift detection in prokaryotic genomic sequences.” Microbial Genomics (IMMG) 2007, College Park, MD.

**A.O. Kislyuk** and M. Borodovsky. “Frameshift detection in prokaryotic genomic sequences.” 6th Georgia Tech-ORNL International Conference on Bioinformatics, in silico Biology 2007, Atlanta,

GA.

TEACHING  
EXPERIENCE

*Teaching Assistantships at Georgia Institute of Technology*

- Spring 2008      **Theoretical Ecology** (BIOL 4422/6422).
- Fall 2007        **Bioinformatics** (BIOL 7023). Developed a new set of laboratory manuals.
- Fall 2006        **Biological Principles** (BIOL 1502).
- Summer 2006    **Applied Combinatorics** (MATH 3012).
- Spring 2006     **Calculus II and Linear Algebra** (MATH 1510).  
Fall 2005
- Guest Lectures at Georgia Institute of Technology*
- Spring 2009     Special Topics/**Computational Genomics** (BIOL 8803): 3 guest lectures.

PROFESSIONAL  
EXPERIENCE

- Supervised one undergraduate and three graduate research assistants. (Georgia Tech, 2008 – 2009.)
- President** of ABiS (Association of Bioinformatics Students), a professional student society. (Georgia Tech, 2008 – 2009.)
- Developer**, Gentoo Linux. Science project lead, Google Summer of Code mentor. (2008 –)
- Procured, administered and upgraded a multi-user HPC cluster environment. (2008 –)
- Judge**, Siemens Westinghouse Science and Technology Competition regional finals. (2007)
- Intern**, Department of Energy Joint Genome Institute. (2007)
- Programmer**, Web data mining applications for economics research (part-time). (2005 – 2008)
- Software Engineer**, Versant Corporation. Performance instrumentation for object relational database management systems. (2005)

AWARDS

- Bioinformatics Research Excellence Award (Georgia Tech, 2009).
- CDC and ATSDR Honor Award, Partners in Public Health Improvement Category (Centers for Disease Control and Prevention, 2009).
- Pacific Biosciences Rock Star Award (2011).

CITIZENSHIP

US citizen.

PERSONAL

- Languages:** English (fluent), Russian (native).
- Interests:** Running, cycling, martial arts, photography, aviation, robotics.

REFERENCES

*Available upon request.*