

# Scott Clark

Graduate Student, DOE Computational Science Graduate Fellow  
657 Rhodes Hall, Ithaca, NY, 14853

June 12, 2014  
sc932@cornell.edu  
cam.cornell.edu/~sc932

## Education

- **Cornell University** Ithaca, NY  
*Ph.D. Applied Math (current), M.S. Computer Science* 2008 - 2012(projected)
  - Department of Energy Computational Science Graduate Fellow (Full Scholarship, 4 years)
  - Emphasis on machine learning/data mining and algorithm design/software development related to bioinformatics and optimization
  - Committee: Peter Frazier (advisor), Steve Strogatz, Bart Selman, Zhong Wang
- **Oregon State University** Corvallis, OR  
*B.Sc. Mathematics, B.Sc. Computational Physics, B.Sc. Physics* 2004 - 2008
  - Graduated Magna Cum Laude with minors in Actuarial Sciences and Mathematical Sciences
  - Strong emphasis on scientific computing, numerical analysis and software development
  - Advisors: Rubin Landau, Malgorzata Peszynska

## Industry Experience

- **Bloomberg LP** New York, NY  
*Financial Software Development Intern* Summer 2011
  - Developed end-to-end reporting software in C++ and javascript
  - Implemented statistical models to perform forward and backward portfolio analysis

## Research Experience

- **DOE Joint Genome Institute (Lawrence Berkeley National Lab)** Walnut Creek, CA  
*Researcher in Analysis Group under Dr. Zhong Wang* Summer 2010
  - Created **open source** genome validation software tool in python and C
  - Used machine learning to mine TBs of genome data efficiently using novel likelihood function
- **Los Alamos National Laboratory** Los Alamos, NM  
*Researcher in Metagenomics Group under Dr. Nick Hengartner* Summer 2009
  - Wrote **open source** alignment algorithm software tool in python, C and CUDA
  - Used statistical models to discover sequence alignments using parallel algorithms on GPUs
- **Oregon State University** Corvallis, OR  
*Research Assistant under Prof. Malgorzata Peszynska and Prof. Rubin Landau* 2005-2008
  - Finite element analysis with uncertainty and web-based teaching in Java
- **Max Plank Institute for the Physics of Complex Systems** Dresden, Germany  
*NSF REU Research Assistant under Prof. Steven Tomsovic* Summer 2007
  - Research on extreme value statistics in MATLAB and FORTRAN

- **University of California: Davis**  
*NSF REU Research Assistant under Prof. Daniel Cox*  
 – Computational biophysics research as applied to protein folding in Java

Davis, CA  
 Summer 2006

## Publications (available at [cam.cornell.edu/~sc932/Research.html](http://cam.cornell.edu/~sc932/Research.html))

1. SC Clark, R Egan, P Frazier, and Z Wang, *ALE: an Assembly Likelihood Evaluation Framework to Assess the Accuracy of Metagenome Assemblies* in preparation for *Genome Research* (2012).
2. SC Clark, N Hengartner, and J Berendzen, *Velvetropo: a parallel, bitwise algorithm for nding homologous regions within multiple sequences* submitted to *BMC Bioinformatics* (2010).
3. SC Clark, *Solving Genomic Jigsaws*. DEIXIS Magazine **8** 30–32 (2010).
4. KC Kunes, SC Clark, DL Cox, and RR Singh, *Left handed helix models for mammalian prion fibrils*. *Prion* **2**, 2, 81–90 (2008).

## Presentations

- *Learning for Metagenomic Assembly Validation and Optimization*
  - Supercomputing 2011 (SC11) (Poster) – Seattle, WA (November 2011)
  - SIAM CSE conference (Talk) – Reno, NV (February 2011)
  - INFORMS annual meeting (Talk) – Austin, TX (November 2010)
  - Cornell Math Sciences Seminar (Talk) – Ithaca, NY (November 2010)
  - DOE Joint Genome Institute Visiting Speaker (Talk) – Walnut Creek, CA (August 2010)
- *Development and Exploration of Velvetropo: a bitwise, parallel alignment algorithm on biological sequences*
  - Supercomputing 2010 (SC10) (Poster, ACM student competition) – New Orleans, LA (November 2010)
  - Cornell Math Sciences Seminar (Talk) – Ithaca, NY (November 2010)
  - DOE CSGF Annual Conference (Poster, award finalist) – Washington D.C. (June 2010)
- *Redundancy in random  $k$ -SAT*
  - Cornell Engineering Research Conference (Poster) – Ithaca, NY (May 2010)
  - DOE CSGF Annual Conference (Poster, award finalist) – Washington D.C. (June 2009)
- *Finite Element Analysis of Uncertain Interfaces*
  - Oregon State University Computational Mathematics Seminar (Talk) – Corvallis, OR (June 2008)

## Participation in Workshops and Conferences

- **Supercomputing 2011** Seattle, WA  
*Association for Computing Machinery/IEEE* *November 12-18, 2011*
- **SIAM Conference on Computational Science and Engineering** Reno, NV  
*Society for Industrial and Applied Mathematics* *Feb 28-March 4, 2011*
- **Supercomputing 2010** New Orleans, LA  
*Association for Computing Machinery/IEEE* *November 13-19, 2010*
- **INFORMS Annual Meeting 2010** Austin, TX  
*Institute for Operations Research and the Management Sciences* *November 7-10, 2010*
- **ICCS Computational Science Summer School (Many-core)** Oakland, CA  
*International Center for Computational Science* *August 2-6, 2010*
- **Computational Science Graduate Fellowship Annual Conference** Washington D.C.  
*Krell Institute* *June 22-24, 2010*
- **High Performance Computing Workshop** Washington D.C.  
*Krell Institute* *June 21, 2010*
- **Cornell Engineering Research Conference** Ithaca, NY  
*Cornell College of Engineering* *March 16, 2010*
- **Supercomputing 2009** Portland, OR  
*Association for Computing Machinery/IEEE* *November 14-20, 2009*
- **q-bio Conference on Cellular Information Processing,** Los Alamos/Sante Fe, NM  
*Los Alamos National Laboratory Center for Non-Linear Studies* *August 5-9, 2009*
- **q-bio Summer School on Cellular Information Processing** Los Alamos/Sante Fe, NM  
*Los Alamos National Laboratory Center for Non-Linear Studies* *July 20 - August 4, 2009*
- **Computational Science Graduate Fellowship Annual Conference** Washington D.C.  
*Krell Institute* *July 14-16, 2009*
- **High Performance Computing Workshop** Washington D.C.  
*Krell Institute* *July 13, 2009*
- **DOE ACTS Workshop (Computational Science Tools)** Berkeley, CA  
*NERSC/LBNL/ASCR/DOE Office of Science* *August 19-22, 2008*
- **Computational Science Graduate Fellowship Annual Conference** Washington D.C.  
*Krell Institute* *June 16-19, 2008*
- **Risk Analysis: Perception, Policy and Practice Workshop** Research Triangle Park, NC  
*Statistical and Applied Mathematical Sciences Institute* *October 3-4, 2008*
- **Kickoff Workshop of the SAMSI program on Risk Analysis** Research Triangle Park, NC  
*Statistical and Applied Mathematical Sciences Institute* *September 16-19, 2007*
- **SAMSI/CRSC Undergraduate Modeling Workshop** Raleigh, NC  
*SAMSI/NCSU/Duke* *May 21-25, 2007*

## Awards, Grants & Honours

Department of Energy Computational Science Graduate Fellowship (CSGF) (\$300 000)	2008-2012
NERSC Production Allocation (PI) (100 000 Cray XT4 hours)	2012
NERSC Startup Allocation Renewal (PI) (15 000 Cray XT4 hours)	2012
Cornell University Conference Travel Grant (\$390)	2011
NERSC Startup Allocation Renewal (PI) (15 000 Cray XT4 hours)	2011
Cornell University Conference Travel Grant (\$390)	2010
NERSC Startup Allocation (PI) (15 000 Cray XT4 hours)	2010
DOE CSGF Essay Contest Honorable Mention (\$500)	2010
Cornell University Sage Fellowship (\$55 000, declined)	2008-2009
Joel Davis Award in Mathematics (\$1 000)	2007-2008
URISC Undergraduate Research Fellowship (\$1 500)	2007-2008
NSF Research Experience for Undergraduates Program (MPI PKS) (\$6 000)	2007
Paul Copson Memorial Scholarship in Physics (\$1 000)	2006-2007
NSF Research Experience for Undergraduates Program (UC Davis) (\$6 000)	2006
Nicodemus Scholarship in Physics (\$1 000)	2005-2006
Diversity Achievement Scholarship (\$4 000)	2004-2008

## Selected Open Source Projects ([github.com/sc932](https://github.com/sc932))

- **ALE: Assembly Likelihood Estimator** C, Python  
*Probabalistic evaluation of genome assemblies* 2010 - Current
- **Velvetrope** Python, C, CUDA  
*A parallel statistical algorithm for finding homologous regions within sequences* 2009 - 2010
- **BetaHelix** Java  
*Computes various statistics about a left or right handed beta helix* 2006 - 2007

## Skills

- **Development:** C/C++, Python, CUDA, JavaScript, Ruby (Rails), Java, FORTRAN, MATLAB
- **Numerical Analysis:** Optimization, Linear Algebra, ODEs, PDEs, Monte Carlo, Computational Physics, Complex Systems, Iterative Methods, Tomology
- **Computer Science:** Machine Learning, Data Mining, Parallel Programming, Data Structures, Artificial Intelligence, Operating Systems
- Discovering and implementing new ideas. Give me an API and a problem and I will figure it out.
- Diverse background in Math, Computer Science, Physics and Biology allows me to communicate to a wide scientific and general audience and begin contributing to any group immediately.
- I have worked in many places in a myriad of fields. I can readily learn and adapt to a new discipline, area or environment and start pushing real results quickly.