

## Jeremy Gray

---

Address:	55 Front Street East, Unit 807, Toronto, ON, M5E0A7
Phone:	(647) 537-8950
Email:	jeremygray@gmail.com
Website:	www.jeremy.kiwi.nz
GitHub:	www.github.com/jeremycg
StackOverflow:	jeremycg

---

### Summary

Senior Data Scientist, Data Science Instructor, Biologist, Statistician, Bioinformatician, R and Python Programmer.

Significant Expertise in Amazon AWS technologies, Data Wrangling, Big Data, Course Instruction and Design, Statistical Computing, Web App Design, Data Analysis Automation and Reproducible Research.

### Research Interests

Python and R programming, specifically in the creation and optimization of statistical tools and APIs.

Current methods used include algorithm design, hidden Markov models, MCMC, non-linear optimization, supervised clustering, agent based modelling, as well as pedagogical methods in data science.

### Key Technical Skills

**R:** 9 years experience, expertise in ggplot2, dplyr, tidyr, knitr, nlme, caret, PerformanceAnalytics. **Python:** 3 years experience, expertise in scikit-learn, matplotlib, numpy, pandas. **SQL:** 9 years experience, working in SQLite, MySQL, RedShift and Netezza. Query construction and database construction and maintenance. Significant experience (3 years+) in **Bash**, **Git** and **Mathematica**. Expert in AWS migration, EC2 use, linux.

### Current Position

- **Senior Data Scientist.** Precima. May 2016 - Present.
  - Construction of a modelling engine for internal use, written primarily in Python, interacting with SQL databases, AWS and EC2.
  - Migrating existing code into AWS compliant versions.
  - Sped up existing code by 2 orders of magnitude, as well as bugfixing and feature addition
  - Internal consulting on Python and programming to other team members.
  - Introduced and drove adoption of version control and other good programming practices.

### Education and Academic Experience

- **Postdoctoral Fellow in Ecology and Evolutionary Biology** University of Toronto. May 2012 - May 2015.
  - 66% Computational - Developing two major R packages for high throughput statistical analysis of experimental data, and agent based modelling of evolution. Routinely carrying out statistical tasks (ANOVA, linear and non-linear regression, survival analysis, image analysis). Working with gigabyte scale bioinformatic data - creating, automating and maintaining pipelines, carrying out evolutionary analysis of genomic data. Maintaining the lab SQL databases.

- 33% Wet Lab - Carrying out a wide range of contemporary molecular biological techniques - CRISPR/Cas9, microinjection, molecular cloning, microscopy.
- **PhD in Biology.** University of Auckland, March 2007 - January 2011. “Testing the major theories concerning the evolution of sex using experimental evolution”
  - Major computational methods - Mathematica - Creating evolutionary dynamic simulations (stochastic agent based modelling) and likelihood based statistical analysis. R graphics, statistics and population genetics.
  - A wide of range of microbiological techniques - molecular cloning, chemostat culture, microscopy.
- **BSc (Hons I) in Genetics.** University of Otago, 2001-2004 “PRP8 inteins of the fungi of the genus *Aspergillus*”
  - Phylogenetics and bioinformatics. Courses included a large range of statistical and computational training.

## Work History

- **Data Science Instructor - Python.** Precima. February 2016 - April 2016.
  - Developing, designing and presenting a course in Python programming for data science to the 50 people in the Precima data science R&D and Applied Statistics teams.
  - Available online on the course website or github
- **Data Science Consultant.** Corporate Knights. September 2015 - November 2015.
  - Refactoring, bug fixing and completing an R Shiny Web App to compare investment portfolios with and without polluting investments. Available online at decarbonizer.co. R, JavaScript, HTML and CSS.
- **Rena Oil Spill Wildlife Response Team.** Maritime New Zealand. October 2011 - February 2012.
  - Responsible for cleaning and caring for oiled wildlife, public outreach and education, as well as collecting wildlife for treatment.
- **Lab Technician.** Institute of Environmental Science and Research Ltd (ESR). “Horizontal Gene Transfer in the New Zealand Environment”. June 2005-June 2006
  - Responsible for a small government lab’s statistical analysis and MS Access/SQL data storage, as well as microbiological research.

## Teaching Experience

- **Software Carpentry Instructor.** 2015.
  - Teaching data science skills to graduate level scientists. Technologies include Git, Python, R, SQL and Bash. 2 day workshops, Taught at Sick Children’s Hospital Toronto, and University of Toronto.
- **Course Instructor.** University of Toronto. EEB225H1S - Biostatistics 2015
  - Lectured 0.5 of a course, Biostatistics, for second year EEB majors. Course content involved t-tests, ANOVA, linear and multiple linear regression, logistic regression, ANCOVA and experimental design. Developed new R labs and statistical lectures. A total of 12 hours lecturing, and 30 hours leading computational labs.
- **Laboratory Tutor and TA.** 8 courses at the University of Auckland, 3 at the University of Otago, including computational biology and statistics.

## Publications

Cutter AD, **Gray JC**. Ephemeral ecological speciation and the latitudinal biodiversity gradient. **Evolution** 2016;70(10):2171-2185. doi: 10.1111/evo.13030.

Bilotta GS, Burnside NG, **Gray JC**, Orr HG. The Effects of Run-of-River Hydroelectric Power Schemes on Fish Community Composition in Temperate Streams and Rivers. **PLoS One**. 2016; 11(5): e0154271.

Jovelin R, Krizus A, Taghizada B, **Gray JC**, Phillips PC, Claycomb JM, Cutter AD. Comparative genomic analysis of upstream miRNA regulatory motifs in *Caenorhabditis*. **RNA**. 2016;22(7):968-78. doi: 10.1261/rna.055392.115.

Vielle A, Callemeyn-Torre N, Gimond C, Poulet N, **Gray JC**, Cutter AD, Braendle C. Convergent evolution of sperm gigantism and the developmental origins of sperm size variability in *Caenorhabditis* nematodes. **Evolution**. 2016;70(11):2485-2503. doi: 10.1111/evo.13043

**Gray JC**, Cutter AD. Mainstreaming *Caenorhabditis elegans* in experimental evolution. **Proceedings of the Royal Society B: Biological Sciences**. 2014;281(1778) doi:10.1098/rspb.2013.3055

**Gray JC**, Goddard MR. Gene-flow between niches facilitates local adaptation in sexual populations. **Ecology Letters**. 2012;15(9):955-62 doi:10.1111/j.1461-0248.2012.01814.x.

**Gray JC**, Goddard MR. Sex enhances adaptation by unlinking beneficial from detrimental mutations in experimental yeast populations. **BMC Evolutionary Biology**. 2012;12:43. doi: 10.1186/1471-2148-12-43.

Butler MI, **Gray J**, Goodwin TJ, Poulter RT. The distribution and evolutionary history of the PRP8 intein. **BMC Evolutionary Biology**. 2006;6:42. doi:10.1186/1471-2148-6-42

## Awards, Memberships and Positions

- Graduate Student Representative, Research Advisory Board, School of Biological Sciences, University of Auckland. 2008-2010 (two terms).
- Faculty of Science Postgraduate Student/Staff Consultative Committee, Biological Sciences representative, University of Auckland. 2008-2009.
- Epsom delegate, Green Party of Aotearoa New Zealand list ranking conference, 2011.