Ryan C. Thompson, Ph. D.

8656 Via Mallorca
Unit G
La Jolla, CA 92037

Phone: (908) 922-7470
Email: rct@thompsonclan.org
Website: https://github.com/DarwinAwardWinner

Summary

- Accomplished, experienced computational biologist with deep knowledge of biology, mathematics, statistics, and programming

- 11 years of bioinformatics experience answering a wide range of complex biological questions through analysis of large NGS- and microarray-based whole transcriptome and epigenetic data sets

- Strong focus on sound statistical analysis, effective data visualization and communication, and making expert knowledge and successful analysis methods available as re-usable software tools

- Supportive team member with a strong ability to foster cross-disciplinary literacy by explaining concepts at the right level of abstraction

Education

2012—2019	**Ph. D. in Bioinformatics**, Skaggs Graduate School of Chemical and Biological Sciences **The Scripps Research Institute**, La Jolla, California Dissertation: *Bioinformatic analysis of complex, high-throughput genomic and epigenomic data in the context of CD4$^+$ T-cell differentiation and diagnosis and treatment of transplant rejection* Advisor: Andrew Su
2005—2009	**B. S. in Biology with High Distinction; B. A. in Mathematics** **University of Virginia**, Charlottesville, Virginia Undergraduate thesis: *Contig Farmer: A tool for extracting maximal-length contiguous Sequences from a Database of Short Sequence Reads* Advisor: Paul J. Rushton

Research Experience

Salomon & Su Labs, The Scripps Research Institute   La Jolla, CA   2012 - 2019

Contact: Dr. Andrew Su, asu@scripps.edu

- Created an open source, reproducible workflow to analyze a large multi-omics next-gen sequencing dataset of 220 RNA-seq and ChIP-seq samples to reveal interactions between differential histone methylation and differential gene expression during T-cell activation, as well as key differences in activation between naïve and memory cells [1, 3, 8].
  Links: Reproducible workflow, Slides, Example results and visualizations

- Significantly improved cross-validated performance of machine learning classifier for identifying biomarkers of transplant rejection by developing appropriate single-sample microarray normalization procedures[5], including training a custom set of frozen RMA normalization vectors. Classifier is currently being developed into a clinical test for transplant dysfunction.
  Links: fRMA example code & plots, Classifier Method Write-up

- Performed a comprehensive comparative evaluation of over 20 subtly different statistical models for differential methylation in Illumina 450k arrays, selecting a model that best explained the observed sources and trends of variation in the data, including cross-domain application of a method originally designed for RNA-seq data.
  Link: Example diagnostic plots

- Evaluated and optimized cost and performance of custom protocol for RNA-seq of human and primate blood samples while minimizing nuisance globin reads. Increased yield of useful reads nearly 2-fold. [7]
  Link: Example results

- Investigated effects of life-span-extending drug on worm gene expression over time revealing that the drug retards age-related "transcriptional drift", preserving a youthful phenotype at the molecular level. [4]
  Link: PCoA Plot

- Implemented a systems biology tool to analyze and efficiently present and summarize differential expression for multiple gene set & pathway methods run on multiple pathway databases, as well as differential expression of individual genes within each pathway.
  Links: Presentation, Example Results, Summary

- Performed comparative analysis of multiple differential expression statistical models to define best practice for optimal sensitivity while maintaining false positive control. Presented on theoretical and practical similarities and differences between methods.
  Links: RNA-seq Presentation, Advanced RNA-seq Presentation, Example plot

- Taught basic RNA-seq theory and practical analysis for the graduate-level introductory bioinformatics course.
  Links: Lecture Slides, Hands-on lab section

- Adapted common normalization methods from RNA-seq to improve performance in analysis of RASL-seq experiments. [10]

- Active member of the Bioconductor community since 2012 and contributing developer for several Bioconductor packages.
  Links: BiocParallel, ChIPPeakAnno, MyGene, Support Profile

- Mentored 3 coworkers in learning bioinformatics, one from a programming background and the other from a biology background.

Summer Internship, Informatics IT, Merck & Co.   Boston, MA   2011

Contact: Adnan Derti, adnan.derti@gmail.com

- Built a transcriptome assembly and quantification pipeline using Cufflinks, including fully-automated cluster job control for high-throughput reproducible analysis, and presented a conceptual overview of Cufflinks' assembly and quantification algorithms to help the team understand Cufflinks.
  Link: Presentation Slides

- Assisted in a molecular genetics study to evaluate peformance of two variant calling algorithms in detection of causal mutations in antibiotic-resistant bacterial genomes.

Gaasterland Lab, UCSD Bioinformatcs   La Jolla, CA   2010—2012

Contact: Terry Gaasterland, gaasterland@ucsd.edu

- Designed and implemented Deloxer, a critical software step in a new Illumina mate-pair sequencing protocol using Cre recombination. Deloxer is published[6] and now in use in several labs around the world.
  Links: Documentation; Code

- Performed a molecular genetics study to find potential causal mutations for fatal iron overload disease in critically endangered black rhinoceros by *de novo* assembly of transcriptomes for black rhino and closely-related white rhino. Developed a custom pipeline to match up ortholog gene pairs, discover single-nucleotide differences between them, and functionally annotate these differences, and delivered a list of potential causal variants to collaborators for follow-up.
  Link: Example results

- Helped design & implement a large-scale high-throughput exome sequencing pipeline for SNP discovery and functional annotation, including QC and validation of on-target coverage depth and reproducibility of coverage.
  Links: Example 1; Example 2; Example 3

- Created a fully-automated pipeline to produce quality-control metrics and plots for Illumina high-throughput sequencing data for early identification of failed runs or samples.
  Link: Example results

- Investigated the binding motif specificity of ZASC1 transcription factor in mouse T-cells using Affymetrix expression microarrays in ZASC1 siRNA knockdown experiment.

- Analyzed miRNA target predictions using GO & KEGG grouping to identify target pathways of autophagy-related miRNAs for biological validation.
  Link: Example results

Timko Lab, U. of Virginia Biology   Charlottesville, VA   2007—2009

Contact: Paul J. Rushton, Paul.Rushton@sdstate.edu

- Undergraduate thesis: Designed and implemented Contig Farmer, an algorithm for efficient selective contig assembly starting from "seed" sequences of interest, and used Contig Farmer to accelerate transcription factor gene discovery in cowpea and tobacco shotgun genomic sequence data[11].

- Investigated transcription factors mediating plant stress response using expression microarray time-course, and refined the custom microarray design using data from previous runs to identify and eliminate uninformative probes, yielding an improved design for future studies.
  Link: Code

- Acted as interpreter to explain complex biological concepts to programmers and explain complex computational problems to biologists.
  Link: Presentation Slides

Skills

Computing Skills R/BioConductor, Python, Perl, Lisp, Java, C#, C++; Git version control; Remote UNIX system administration & software compilation, computing cluster job management & parallel computation
Public code: https://github.com/DarwinAwardWinner
Stack Overflow profile: http://stackoverflow.com/users/125921

Statistics, Machine Learning, & Data Analysis Multi-omics NGS & microarray analysis, multi-factor linear and generalized linear regression, experimental design and parametrization, empirical Bayesian methods, predictive modeling of clinical outcomes, machine learning classifier training & validation, survival analysis, data visualization & presentation, reproducible research practices

Wet Lab: PCR, molecular cloning, recombinant protein purification, epitope mapping, site-directed mutagenesis, 2D gels, real-time PCR, and associated data analysis

Languages: English: native speaker; German: written and spoken

Other Work Experience

Computing Advisor & Help Desk, U. Va. IT Dept.   Charlottesville, VA   2005—2007

- Provided support via phone and on-site for students, faculty, and staff having problems with computers, phone system, network access, malware, hardware setup, and university web services

- Tasks included support for university-provided software, virus removal, iPod recovery, printer setup and repair, diagnosis of hardware malfunctions, and data recovery from failing hard disks

Summer Sailing Instructor, Raritan Yacht Club   Perth Amboy, NJ   2006—2009

- Instructed children ages 8-18 in sailing skills, safety, seamanship, knots, rigging & de-rigging boats, navigation, and racing strategy and technique, with an emphasis on building character and self-reliance

- Ensured safety of students and staff by maintaining boats and equipment in good repair, by being vigilant to traffic and hazards on a busy waterway, and by communicating and coordinating efficiently with other staff

- Helped organize, run, and referee several regattas per season for students from RYC and neighboring yacht clubs

Peer-reviewed Publications

[1] Sarah A. LaMere, **Ryan C. Thompson**, Xiangzhi Meng, H. Kiyomi Komori, Adam Mark, and Daniel R. Salomon. H3K27 Methylation Dynamics during CD4 T Cell Activation: Regulation of JAK/STAT and IL12RB2 Expression by JMJD3. *The Journal of Immunology*, 199(9):3158–3175, November 2017. `doi:10/gchc9x`.

[2] S. M. Kurian, E. Velazquez, **R. Thompson**, T. Whisenant, S. Rose, N. Riley, F. Harrison, T. Gelbart, J. J. Friedewald, J. Charette, S. Brietigam, J. Peysakhovich, M. R. First, M. M. Abecassis, and D. R. Salomon. Orthogonal Comparison of Molecular Signatures of Kidney Transplants With Subclinical and Clinical Acute Rejection: Equivalent Performance Is Agnostic to Both Technology and Platform. *American Journal of Transplantation*, 17(8):2103–2116, August 2017. `doi:10/gbp6vr`.

[3] S. A. LaMere, **R. C. Thompson**, H. K. Komori, A. Mark, and D. R. Salomon. Promoter H3K4 methylation dynamically reinforces activation-induced pathways in human CD4 T cells. *Genes & Immunity*, 17(5):283–297, July 2016. `doi:10/f97x85`.

[4] Sunitha Rangaraju, Gregory M. Solis, **Ryan C. Thompson**, Rafael L. Gomez-Amaro, Leo Kurian, Sandra E. Encalada, Alexander B. Niculescu, Daniel R. Salomon, and Michael Petrascheck. Suppression of transcriptional drift extends C. elegans lifespan by postponing the onset of mortality. *eLife*, 4(December2015):1–39, December 2015. `doi:10/ggcxmg`.

[5] S M Kurian, a N Williams, T Gelbart, D Campbell, T S Mondala, S R Head, S Horvath, L Gaber, **R Thompson**, T Whisenant, W Lin, P Langfelder, E H Robison, R L Schaffer, J S Fisher, J Friedewald, S M Flechner, L K Chan, A C Wiseman, H Shidban, R Mendez, R Heilman, M M Abecassis, C L Marsh, and D R Salomon. Molecular Classifiers for Acute Kidney Transplant Rejection in Peripheral Blood by Whole Genome Gene Expression Profiling. *American Journal of Transplantation*, 14(5):1164–1172, May 2014. `doi:10/f5xswg`.

[6] Filip Van Nieuwerburgh, **Ryan C Thompson**, Jessica Ledesma, Dieter Deforce, Terry Gaasterland, Phillip Ordoukhanian, and Steven R Head. Illumina mate-paired DNA sequencing-library preparation using Cre-Lox recombination. *Nucleic acids research*, 40(3):e24, February 2012. `doi:10/fmzd3r`.

Other Works

[7] **Ryan C. Thompson**, Terri Gelbart, Steven R Head, Phillip Ordoukhanian, Courtney Mullen, Dongmei Han, Dora M Berman, Amelia Bartholomew, Norma S Kenyon, and Daniel R Salomon. Optimizing yield of deep RNA sequencing for gene expression profiling of peripheral blood samples from cynomolgus monkeys (Macaca fascicularis). (In preparation). Institution: The Scripps Research Institute, 2019.

[8] **Ryan C. Thompson**, Sarah A. Lamere, and Daniel R. Salomon. Reproducible genome-wide epigenetic analysis of H3K4 and H3K27 methylation in naïve and memory CD4+ T-cell activation. (In preparation). Institution: The Scripps Research Institute, 2019.

[9] **Ryan C. Thompson**. *Bioinformatic Analysis of Complex , High-Throughput Genomic and Epigenomic Data in the Context of CD4+ T-Cell Differentiation and Diagnosis and Treatment of Transplant Rejection*. PhD thesis, The Scripps Research Institute, 2019.

[10] Erick R Scott, H Benjamin Larman, Ali Torkamani, Nicholas J Schork, Nathan Wineinger, Max Nanis, **Ryan C. Thompson**, Reza B. Beheshti Zavareh, Luke L Lairson, Peter G Schultz, and Andrew I. Su. RASLseqTools: Open-source methods for designing and analyzing RNA-mediated oligonucleotide Annealing, Selection, and, Ligation sequencing (RASL-seq) experiments. *bioRxiv*, 2016. `doi:10/ggcxmn`.

[11] **Ryan C. Thompson**, Paul J. Rushton, Tom W. Laudeman, and Michael P. Timko. Contig Farmer : A tool for extracting maximal-length contiguous sequences from a database of short sequence reads (Undergraduate Thesis), June 2009. Institution: University of Virginia. URL: `http://darwinawardwinner.github.io/resume/examples/UVa/contigfarmer.pdf`.

[12] **Ryan C. Thompson**. The Sources and Limits of Geometric Rigor from Euclid Through Descartes, May 2008. Institution: University of Virginia. URL: `http://darwinawardwinner.github.io/resume/examples/UVa/math-history-paper.pdf`.