

Rick Farouni | Curriculum Vitae

ETH Zurich - Department of Biosystems Science and Engineering (D-BSSE)

✉ rfarouni@protonmail.com • 🌐 <http://rfarouni.github.io/>

I am an applied statistician and bioinformatics scientist with expertise in the statistical modeling and bioinformatics analysis of genomics, single-cell transcriptomics (scRNA-seq), and CRISPR genome editing data.

Experience

Bioinformatics Scientist.....

- **The Swiss Federal Institute of Technology in Zurich (ETH Zurich)** **Basel, Switzerland**
Research Lab of Dr. Platt, Department of Biosystems Science and Engineering 10/2020-Present

Postdoctoral Researcher.....

- **McGill University-Génomique Québec Innovation Centre** **Quebec, Canada**
Research Lab of Dr. Najafabadi, Department of Human Genetics 06/2018-09/2020
- **Massachusetts General Hospital-Harvard Medical School** **Massachusetts, USA**
Research Lab of Dr. Pinello, Molecular Pathology Unit 06/2017-05/2018
- The Broad Institute of MIT and Harvard (Affiliated Member)

Research Intern.....

- **Department of Biomedical Informatics Summer Internship Program** **Ohio, USA**
Research Lab of Dr. Ewy Mathè, The Ohio State University 05/2016-08/2016

University Teaching Assistant.....

- **The Ohio State University** **Ohio, USA**
Graduate Teaching Associate (Statistics) 2013-2017
Served as a Teaching Assistant for three courses:
 - Repeated Measures Models
 - Covariance Structure Models
 - Data Analysis in Psychology

Education

Academic Qualifications.....

- **PhD in Quantitative Psychology (i.e. *Applied Statistics*)** **Ohio, USA**
The Ohio State University 2015–2017
Dissertation Topic: '*Application of Deep Latent Generative Models to the Unsupervised Learning of Chromatin States*'
- **Master of Science in Mathematical Statistics** **Ohio, USA**
The Ohio State University 2012–2015
- **Master's Degree in Psychometrics** **Ohio, USA**
The Ohio State University 2012–2014
Thesis Project: '*Latent Variable Modeling of Categorical Item Responses in a Hierarchical Bayesian Framework*'
- **Bachelor's Degree in Psychology** **Pennsylvania, USA**
The Pennsylvania State University 2011–2012

Publications

Journal Papers (first author/co-author).....

- **Farouni, R.**, Djambazian, H., Ferri, L. E., Ragoussis, J., and Najafabadi, H. S. (2020). Model-based analysis of sample index hopping reveals its widespread artifacts in multiplexed single-cell RNA-sequencing. ***Nature Communications***. Paper available at <https://www.nature.com/articles/s41467-020-16522-z>. Website: <https://csglab.github.io/PhantomPurgeR/>.
- Clement, K.[†], **Farouni, R.**[†], Bauer, D. E., and Pinello, L. (2018). Design and analysis of unique molecular identifiers for deep amplicon sequencing. ***Bioinformatics***. <https://doi.org/10.1093/bioinformatics/bty264>. Preprint available at <https://www.biorxiv.org/content/biorxiv/early/2018/03/23/288118.full.pdf>
- Pinello, L.[†], **Farouni, R.**[†], and Yuan, G-C. (2018). Haystack: systematic analysis of the variation of epigenetic states and cell-type specific regulatory elements. ***Bioinformatics***. <https://doi.org/10.1093/bioinformatics/bty031>. Preprint available at <https://doi.org/10.1101/199067>
- Baskin, E.[†], **Farouni, R.**[†], and Mathè, E. (2017). ALTRE: workflow for defining ALTERed Regulatory Elements using chromatin accessibility data. ***Bioinformatics***. <https://doi.org/10.1093/bioinformatics/btx386>. Preprint available at <http://www.biorxiv.org/content/early/2016/10/14/080564.full.pdf+html>

Journal Papers (middle author).....

- Anderson-Trocmé, L., **Farouni, R.**, Bourgey, M., Kamatani, Y., Higasa, K., Seo, J., Kim, C., Matsuda, F., Gravel, S. (2019). Legacy Data Confounds Genomics Studies. ***Molecular Biology***

and Evolution. <https://doi.org/10.1093/molbev/msz201>

- Seruggia, D., Oti, M., Tripathi, P., Canver, M. C., LeBlanc, L., Di Giammartino, D. C., Bullen, M.J., Nefzger C.M., Sun, Y.B.Y., **Farouni, R.**, Polo, J.M., Pinello, L., Apostolou, E., Kim J., Orkin, S.H., and Das, P.P. (2019). TAF5L and TAF6L Maintain Self-Renewal of Embryonic Stem Cells via the MYC Regulatory Network. **Molecular Cell**.
- Clement K., Rees H., Canver, M.C., Gehrke J. M., **Farouni, R.**, Hsu, J.Y., Cole, M., Liu D.R., Joung J.K., Bauer, D.E., Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. **Nature Biotechnology**.
- Hsu, J. Y., Fulco, C.P., Cole, M., Canver, M.C., Pellin D., Sher, F, **Farouni, R.**, Clement K., Biasco L., Engreitz, J.M., Lander, E.S., Joung J.K., Bauer, D.E., Pinello, L. (2018). CRISPR-SURF: Exploratory and interactive software for analyzing CRISPR-based tiling screens. **Nature Methods**.

Consortium Journal Papers (contributor).....

- Grüning, B., Dale, R., Sjödin, A., Rowe, J., Chapman, B. A., Tomkins-Tinch, C. H., The Bioconda Team, and Köster, J.(2018). Bioconda: A sustainable and comprehensive software distribution for the life sciences. **Nature Methods**. (Bioconda team member contributor, ranked 94 in the consortium author list). .

Preprints.....

- **Farouni, R.** (2017). A Contemporary Overview of Probabilistic Latent Variable Models. **arXiv preprint**. Preprint available at <https://arxiv.org/abs/1706.08137>

Dissertation and Thesis.....

- **Farouni, R.** (2017). Application of Deep Latent Generative Models to the Unsupervised Learning of Chromatin States. *PhD Dissertation*. Manuscript available at https://etd.ohiolink.edu/!etd.send_file?accession=osu1492189894812539&disposition=inline
- **Farouni, R.** (2014). Latent Variable Models of Categorical Responses in the Bayesian and Frequentist Frameworks. *Masters Thesis*. Manuscript available at https://etd.ohiolink.edu/!etd.send_file?accession=osu1412374136&disposition=inline

Conference Presentations.....

- **Joint Statistical Meetings** **Seattle**
Poster Presentation **2015**
Poster Title: Across-Subject Predictive Modeling of fMRI BOLD Responses to Faces using a sparse Bayesian Group Factor Analysis Model (available at <https://rfarouni.github.io/assets/posters/jsm2015.pdf>).

Software Development

- PhantomPurgeR: The optimal purging of phantom molecules by the robust estimation of the sample index hopping rate in multiplexed droplet-based single-cell RNA-seq data. GitHub Repo: <https://csglab.github.io/PhantomPurgeR/>.

- Haystack: A Python bioinformatics pipeline for the identification of genomic regions of epigenetic variability across different cell-types, cell-type specific cis-regulatory elements, and their associated transcription factors. GitHub Repo: https://github.com/pinellolab/haystack_bio.
- ALTRE: A Workflow for Identifying ALTerred Regulatory Elements using Chromatin Accessibility Data. GitHub Repo: <https://github.com/Mathelab/ALTRE>.

Awards and Fellowships

- **Graduate Student Conference Presentation Award**
The Ohio State University 2015
- **The Center for Cognitive and Brain Sciences Summer Graduate Fellowship**
The Ohio State University 2015
Project Proposal: 'Decoding the Pixels of the Face Image from the Voxels of fMRI BOLD Activity Patterns'
- **The Social and Behavioral Sciences Summer Fellowship**
The Ohio State University 2014
- **University Fellowship**
The Ohio State University 2012

Technical Skill Set

- **Statistics and Machine Learning**
 - **Scientific Programming Languages:** Proficient in and comfortable transitioning between *R* and *Python* depending on computing goals. Familiar with *Julia*.
 - **Deep Learning Frameworks:** Experienced in using Pytorch, Tensorflow, and Keras.
 - **Probabilistic Programming Languages:** Proficient in Stan, familiar with Edward.
- **Bioinformatics Software:** Analysis of next generation sequencing data, R's Bioconductor core packages, Bowtie2, bedtools.
- **Cluster and High-Performance Computing:** Good knowledge of running bioinformatics analysis pipelines on super-computing clusters.
- **Web and Software Development Tools:** \LaTeX , Linux OS, Git, Docker, and Bash. Basic knowledge in website development tools such as HTML, CSS, and Jekyll.

Journal Review Service

- **Psychometrika**
Ad Hoc Reviewer 2015
- **Psychological Methods**
Ad Hoc Reviewer 2016