

Noora J. Al-Muftah

E-mail: noa043@mail.harvard.edu

Academic website: <https://scholar.harvard.edu/njm>

Higher Education

Harvard T.H. Chan School of Public Health

2016 - May 2018 (expected)

Master of Science in Computational Biology and Quantitative Genetics

Carnegie Mellon University

2012 – 2016

Bachelor of Science in Computational Biology

Cumulative GPA: 3.49/4.0

Graduated with *College Honors* and *University Honors*

Undergraduate Senior Thesis: *The Discovery of Hidden Relatedness and Population Structure in the 1000 Genome Project*

Work and Research Experience

Van Allen Lab, Dana Farber Cancer Institute

Boston, MA

Graduate Research Student, PI: Professor Eliezer Van Allen

September 2017 - present

- Currently working on my Master's thesis. Our goal is to attain a better understanding of the link between the mutations in the BAF/PBAF chromatin remodeling complexes and outcome of immunotherapy treatment
- Our first step is to classify the BAF and PBAF mutations in whole exome sequencing, RNA seq, and methylation data of cancer patients from The Cancer Genome Atlas

Hanage Lab, Harvard T.H. Chan School of Public Health

Boston, MA

Summer Research Intern, PI: Professor William Hanage

June 2017 – present

Project: *Analysis of Beta Lactamase Sequence Similarity Networks*

- Assembled a dataset of *Beta Lactamase* (BL) sequences and a representative dataset of bacterial genomes
- Applied BLAST to calculate sequence identities then constructed sequence similarity networks of the BLs with other proteins
- Our next steps are to create phylogenetic trees of the BL clusters and identify the functions of the hypothetical proteins in the networks

Qatar Computing Research Institute

Doha, Qatar

Undergraduate Research Student, PI: Dr. Halima Bensmail

January 2016 – May 2016

Project (also my undergraduate thesis): *The Discovery of Hidden Relatedness and Population Structure in the 1000 Genome Project*

- Utilized the genetic concept of linkage disequilibrium to eliminate millions of redundant SNPs in the 1000 Genomes data
- Estimated kinship matrices for each of the 26 populations by three methods: identity-by-state, identity-by-descent, and the Balding-Nichols model
- Visualized the kinship matrices as multidimensional scaling plots to show differences in genetic heterogeneity, hidden relatedness, and population structure in the populations

Qatar Computing Research Institute

Doha, Qatar

Summer Research Intern, PI: Dr. Halima Bensmail

May 2015 - August 2015

Project: *Applying Statistical Machine Learning Techniques to Cancer Genomics Challenges*

- Identified statistically significant genomic features of cancer cell lines that engender resistance to cancer drugs by applying the Lasso shrinkage and selection techniques on the data matrices
- Simulated the cancer gene co-expression network of the gene expression data as a dense, undirected graph by estimation of the precision matrix

Department of Genetics, Weill Cornell Medical College in Qatar

Doha, Qatar

Undergraduate Research Student, PI: Dr. Khalid Fakhro

January 2015 – May 2015

Project: *SNV-check: A Quality Control Tool For Familial Exome Sequencing Data Based On The Sharing Of Rare Genetic Mutation*

- Implemented a pipeline in Python that utilizes the rare-variant sharing phenomenon of SNPs to calculate the likelihood that two individuals are related and distinguish the type of their relationships (e.g. parent-child, sibling, first/second cousin, or unrelated)

Papers

E. Ullah, M. Aupetit, **N. Al Muftah**, R. Rawi, M. Saad, and H. Bensmail. Discovery of Relatedness in the 1000 Genomes Project Data. *PLOS Genetics*. Submitted: April 2017.

E. Ullah, S. M. Shama, **N. Al Muftah**, I. R. Thompson, R. Rawi, R. Mall, and H. Bensmail. Integrative Statistical Inferences for Drug Sensitivity Biomarkers in Cancer. *Bioinformatics*. Submitted: October 2017.

E. Ullah, M. Aupetit, **N. Al Muftah**, R. Rawi, M. Saad, and H. Bensmail. KinVis: A visualization tool to detect cryptic relatedness in genetic datasets. *Bioinformatics*. Submitted: October 2017.

Posters

E. Ullah, S. Shama, **N. Al Muftah**, I. Thompson, R. Rawi, R. Mall, and H. Bensmail. Identification of Cancer Drug Sensitivity Biomarkers. IEEE BIBM 2017.

N. Al Muftah, E. Ullah, M. Aupetit, R. Rawi, and H. Bensmail. The Discovery of Hidden Relatedness and Population Structure in the 1000 Genome Project. 2nd Annual Next Generation Sequencing USA Congress. October 2016.

N. Al-Muftah, R. Rawi, R. Thompson, and H. Bensmail. Cell Line Screens Identify Biomarkers of Drug Sensitivity in Glioma Cancer. ICBCSB 2016: 18th International Conference in Bioinformatics, Computational and Systems Biology. February 2016.

N. Al-Muftah (advised by Dr. Khalid Fakhro). SNV-check: A quality control tool for familial Exome Sequencing data based on the sharing of rare genetic mutations. CMUQ Meeting of the Minds research symposium. April 29, 2014.

N. Al-Muftah (advised by Professor Saquib Razak). PyExoplanets: A Computed Application for Detecting Exoplanet Transits on Stars Light Curves. CMUQ Meeting of the Minds research symposium. April 30, 2013.

Awards

- Qatar Research Leadership Scholarship**, Qatar Foundation, 2013 - current
- Thanaa Award for Professional Excellence**, Qatar Foundation, 2014, 2015, and 2016
- 3rd Best Undergraduate Research Project**, Meeting of the Minds symposium at Carnegie Mellon University in Qatar, 2014
- 2nd Best Undergraduate Research Project**, Meeting of the Minds symposium at Carnegie Mellon University in Qatar, 2013

Skills

Programming Languages: Strongly experienced in R, Standard ML, Python, and MATLAB. Familiar with C, c0, and Perl.

Other tools: Bioconductor, SPSS statistical regression, LaTeX, and Unix shell scripting

Referees

Please contact me via email for referees' details.