**ARNON BENSHAHAR**

Rehovot Hanahar 8, Ramat Gan, Israel C: 00972-524541543 | [arnon.benshahar@gmail.com](mailto:arnon.benshahar@gmail.com) | id: 300538949

# Education

**MSc in Computer Science-Bioinformatics track.** 2014 **–** 2016

Ben-Gurion University of the Negev, Be'er Sheva, Israel.

* Thesis – **“A Biclique Approach to Reference Anchored Gene Blocks**”, supervised by Prof. Michal Ziv-Ukelson and Dr. Danny Hermelin.

## BSc in Computer Science, Bioinformatics Track, 2011 - 2014

**Honor's Student program.**

Ben-Gurion University of the Negev, Be'er Sheva, Israel.

## High School Diploma, Biology and Computer science. 2005

**Atid High School for Sciences**- Lod, Israel.

# Publications

My thesis is a result of a very wide interdisciplinary research, consisting of the following disciplines: (1) Theory: parameterized complexity and string algorithms, (2) Microbiology and computational genomics, and (3) Distributed computing applications. The research conducted during my M.Sc yielded several publications, mainly: a **journal paper (ranked Q1)**, a **conference proceedings paper (ranked A)**, and **12 pending patents**.

In addition, the theoretical and bioinformatics results from my thesis were presented (by me, [Link to the video lecture](https://www.youtube.com/watch?v=ZSX1DIDGP9M&feature=youtu.be)) at the **WABI conference which is part of Algo2016** in Denmark, and the results of the industry-oriented collaborative work were presented (by the Dell-EMC head of our project) in two meetings in the USA: **President Obama’s Climate-Change Day** meeting at the White House, and the **Strata+Hadoop conference in New York**.

Finally, I implemented the algorithms purposed in my thesis as a software tool, and made both the **code and the results globally available to the wide public of medical, microbiological and bioinformatics researchers** (via GitHub).

In what follows, I elaborate in detail on the inter-disciplinary research done during my M.Sc. My thesis research was inspired by a fruitful collaboration with industry, via a joint project with the DELL-EMC Inc. company. In the scope of the collaboration, we developed a prototype distributed application for the global surveillance of disease outbreaks, based on genomic sequencing.

This collaboration yielded the following 12 pending patents:

1. Metagenomics-based biological surveillance system using big data profiles, U.S.

Application No. 14/983,914

1. Automated metagenomic epidemiological investigation, U.S. Application No.

14/983,920

1. Reasoning on data model for disease monitoring, characterization and

investigation, U.S. Application No. 14/983,932

1. Distributed data processing platform for metagenomic epidemiological

investigation, U.S. Application No. 14/983,943

1. Distributed data processing platform for biological surveillance using big data

profiles, U.S. Application No. 14/983,952

1. Metagenomics-based biological surveillance system with distributed sequencing

centers, U.S. Application No. 14/983,958

1. Automated metagenomic monitoring and characterization, U.S. Application No.

14/983,971

1. Distributed data processing platform for metagenomic monitoring and

characterization, U.S. Application No. 14/983,991

1. Metagenomics-based characterization using genomic and epidemiological

comparisons, U.S. Application No. 14/983,981

1. Distributed data processing platform for biological surveillance using genomic

and epidemiological comparisons, U.S. Application No 14/984,004

1. Methods and apparatus implementing a data model for disease monitoring,

characterization and investigation, U.S. Application No. 14/983,932

**12**. Distributed Data Analytics, Dell EMC Ref. EMC-16-0899.02; File No. 4414-817

The results of this project were presented by the project manager from Dell-EMC in two high-visibility public conferetnces\meetings:

1. The Symposium on Data & Innovation at the Climate-Health Nexus panel, Climate-Change Day meeting at the White House, April 7, 2015.
2. The Strata+Hadoop conference in New York, September 27-29, 2016 https://conferences.oreilly.com/strata/strata-ny-2016/public/schedule/detail/54860

The project with EMC inspired my interest in biclustering algorithms and in molecular microbiology, which evolved into the focus of my thesis. Thus, it was natural for me to continue studying and researching these subjects more thoroughly after the completion of our industry-funded collaboration. This yielded new theoretical results on the subject of multi-genome gene-team discovery (a biclustering problem variant), and the implementation of my new algorithms in a bioinformatics tool, which I made publicly available via Github. The bioinformatics tool was published on OMIC TOOLS ([link to the publication](https://omictools.com/reference-anchored-gene-blocks-tool)), search engine for omic data analysis tool. I also applied my tool to the data-mining of microbial genomes in search of horizontally transferred gene-teams involved in virulence, pathogenicity and antibiotic resistance pathways, and made the discovered gene teams (and their analysis) available to the microbiology community.

This research was first published in a peer-reviewed conference proceedings paper, which

I presented at the WABI conference in Aarhus, Denmark, as part of ALGO 2016 (22-24 August,

2016). **(This conference is ranked A according to the CS department conference ranking**

**Chart.)**

* Benshahar, A., Chalifa-Caspi, V., Hermelin, D., & Ziv-Ukelson, M. A Biclique Approach to Reference Anchored Gene Blocks and Its Applications to Pathogenicity Islands. *In proc. of the* *International Workshop on Algorithms in* Bioinformatics (pp. 14-26). Springer International Publishing.

After the conference, we further extended our research and submitted it as a journal paper,

which was published in the Journal of Computational Biology (JCB). (**This journal is ranked Q1**

**according to the journal paper ranking engine used in BGU).** This journal paper included some

important research extensions, such as an extensive framework, developing an open source,

Publicly available tool of our algorithm, and new significant biological results and their analysis.

* Benshahar, A., Chalifa-Caspi, V., Hermelin, D., & Ziv-Ukelson, M. (2017). A Biclique Approach to Reference-Anchored Gene Blocks and Its Applications to Genomic Islands. Journal of Computational Biology (JCB), 2018.