Gungor Budak

www.gungorbudak.com

Email: gngrbdk@gmail.com

Mobile: +90 551 987 0525

EXPERIENCE

Velsera, Inc. Ankara, TURKEY
Bioinformatics Engineer *Jan. 2023 – Present*
- Construction and benchmarking of population-specific pangenomes including curation and inspection of genomic variant databases, preprocessing variants for quality control and combining variants into useful collections.
- Dockerizing and organizing NGS analysis tools and workflows using CWL including data conversion, quality control, benchmarking, read mapping, variant calling and variant postprocess tools.

Seven Bridges Genomics, Inc. Ankara, TURKEY
Bioinformatics Analyst *Dec. 2016 – Jan. 2023*
- Leading bioinformatics tasks of a national genome project including raw data QC, preprocessing, genome alignment and variant calling, postprocessing for reporting and visualization.
- Researching and developing bioinformatics algorithms and tools for next-generation sequencing (NGS) data analysis.
- Dockerizing and optimizing various bioinformatics tools for cloud computing on The Seven Bridges Platform and The Cancer Genomics Cloud.

School of Informatics and Computing, IUPUI Indianapolis, IN, USA
Bioinformatics Research Scholar *Oct. 2015 – Jun. 2016*
- Developing a Python package and a React application for systematic identification and comparison of processes, phenotypes and diseases associated with RNA-binding proteins from condition-specific CLIP-seq profiles.
- Transcriptome analysis of developing mouse lens tissue including novel transcript discovery and alternative splicing alterations.
- Developing a database of transcriptome profiles encompassing known and novel transcripts across multiple developmental stages in eye tissues in mouse; the data have been curated from Sequence Read Archive (SRA) and European Nucleotide Archive (ENA) and quality check, alignment, transcript discovery and transcript quantification have been performed.
- Developing a web resource for prognostic analyses of exons across human cancers using clinical transcriptomes obtained from The Cancer Genome Atlas (TCGA).

Evias Pharmaceutical R&D, Co., Ltd. Ankara, TURKEY
Software Specialist *Feb. 2015 – Oct. 2015*
- Developing a cloud-based (AWS EC2) high performance virtual screening for drug discovery web service by integrating different tools such as JSmol, JSME, Google Charts, Autodock 4, Autodock Vina, MolPort Web Services using Django (Python) web framework.

Atlas Biotechnologies, Co., Ltd. Ankara, TURKEY
Intern *May 2014 – Feb. 2015*
- Foreign relations with biotechnology product suppliers, giving orders and tracking the process including customs regulations.
- Detecting possible mutations in sequences obtained from ABI format files by blasting against reference genomes database or aligning to the corresponding reference genome.

Bioinformatics Department, Maastricht University Maastricht, THE NETHERLANDS
Intern *Jun. 2013 – Sep. 2013*
- Analysis of the dataset on four breast cancer cell lines to build network models that represent the active cell signaling pathways in breast cancer.

Bioinformatics Unit, German Cancer Research Center (DKFZ) Heidelberg, GERMANY
Intern *Jun. 2012 – Sep. 2012*
- Contaminant detection and identification in Next-Generation Sequencing data by developing a pipeline which extracts the unmapped reads and compares the possible contaminant genomes in different databases using Perl programming language.

EDUCATION

- **Middle East Technical University** — Ankara, TURKEY
  PhD in Computer Engineering; CGPA: 3.18/4.00 – GPA: 3.00/4.00 — *2016 – Present*
  - Network modeling and analysis of the temporal signaling networks for ovarian cancer patients when there is a delay in freezing tumor tissue after the removal (ongoing study).

- **Middle East Technical University** — Ankara, TURKEY
  MSc in Bioinformatics; CGPA: 3.86/4.00 (95.80/100) – GPA: 3.83/4.00 (94.90/100) — *2016*
  - Network modeling and analysis of the temporal signaling network in *Salmonella*-infected human cells (MSc thesis study).

- **Middle East Technical University** — Ankara, TURKEY
  BSc in Biology; CGPA: 3.32/4.00 (83.20/100) – GPA: 3.75/4.00 (92.50/100) — *2014*
  - Developing a web-based software that constructs functional genetic devices with the input and output parameters which are defined as the activator of promoters and the production of coding sequences, respectively.
  - Design of a biosensor that can detect MRSA (Methicillin-resistant *S. aureus*).

- **Kirami Refia Alemdaroglu High School** — Ankara, TURKEY
  Mathematics and Science; Diploma Grade: 83/100 — *2009*

PUBLICATIONS

- Zou, T., Sethi, R., Wang, J., **Budak, G.**, ..., Watters, R., Weiss, K. (2023). Whole Genome Sequencing for Metastatic Mutational Burden in Extraskeletal Myxoid Chondrosarcoma. *Frontiers in Molecular Medicine.*
  https://www.frontiersin.org/articles/10.3389/fmmed.2023.1152550/full

- Tetikol, H. S., Turgut, D., Narci, K., **Budak, G.**, ..., Jain, A., Davis-Dusenbery, B. N. (2022). Pan-African genome demonstrates how population-specific genome graphs improve high-throughput sequencing data analysis. *Nature Communications.*
  https://www.ncbi.nlm.nih.gov/pubmed/35927245

- Olson, N. D., ..., **Budak, G.**, ... (2022). PrecisionFDA Truth Challenge V2: Calling variants from short and long reads in difficult-to-map regions. *Cell Genomics.*
  https://www.ncbi.nlm.nih.gov/pubmed/35720974

- **Budak, G.**, Dash, S., Srivastava, R., Lachke, S. A., Janga, S. C. (2018). Express: A database of transcriptome profiles encompassing known and novel transcripts across multiple development stages in eye tissues. *Experimental eye research.*
  https://www.ncbi.nlm.nih.gov/pubmed/29337142

- Srivastava, R., **Budak, G.**, Dash, S., Lachke, S. A., Janga, S. C. (2017). Transcriptome analysis of developing lens reveals abundance of novel transcripts and extensive splicing alterations. *Scientific Reports*, 7(1), 11572.
  https://www.ncbi.nlm.nih.gov/pubmed/28912564

- **Budak, G.**, Srivastava, R., Janga, S. C. (2017). Seten: A tool for systematic identification and comparison of processes, phenotypes and diseases associated with RNA-binding proteins from condition-specific CLIP-seq profiles. *RNA*, rna-059089.
  https://www.ncbi.nlm.nih.gov/pubmed/28336542

- Hashemikhabir, S., **Budak, G.**, Janga, S. C. (2016). ExSurv: A Web Resource for Prognostic Analyses of Exons Across Human Cancers Using Clinical Transcriptomes. *Cancer Informatics*, 15(Suppl 2), 17.
  https://www.ncbi.nlm.nih.gov/pubmed/27528797

- Hill, S. M., Heiser, L. M., Cokelaer, T., Unger, M., Nesser, N. K., Carlin, D. E., **The HPN-DREAM Consortium**, ... & Graim, K. (2016). Inferring causal molecular networks: empirical assessment through a community-based effort. *Nature methods*, 13(4), 310-318.
  https://www.ncbi.nlm.nih.gov/pubmed/26901648

- **G. Budak**, O. Eren-Ozsoy, Y. Aydin-Son, T. Can, and N. Tuncbag, Reconstruction of the temporal signaling network in *Salmonella*-infected human cells, *Front. Microbiol.*, 6:00730, 2015. doi:10.3389/fmicb.2015.00730.
  https://www.ncbi.nlm.nih.gov/pubmed/26257716

Oral Presentations

- 2017 – Oral presentation titled "Identification of the Ischemic Pathway Level Changes by Integrating Temporal Phosphoproteome in Ovarian Cancer" at 5th International Congress of the Molecular Biology Association of Turkey / Bogazici University, Istanbul, TURKEY.

- 2012 – Oral presentation titled "METU Biology and Genetics Student Club Introduction" at FBI - DIYbio Outreach Workshop / Marriott Hotel, Walnut Creek, CA, USA.

Poster Presentations

- 2016 – Poster presentation titled "Seten: A tool for systematic identification and comparison of processes, phenotypes and diseases associated with RNA-binding proteins from condition-specific CLIP-seq profiles" at GLBIO/CCBC Great Lakes Bioinformatics and the Canadian Computational Biology Conference 2016 / University of Toronto, Toronto, CANADA.
  Published on F1000Research: https://f1000research.com/posters/5-2586

- 2015 – Poster presentation titled "Reconstruction of the temporal signaling network in Salmonella-infected human cells" at ISMB/ECCB 2015, Intelligent Systems for Molecular Biology / European Conference on Computational Biology 2015 / The Convention Centre Dublin, Dublin, IRELAND.
  Published on F1000Research: http://f1000research.com/posters/4-292

- 2013 – Poster presentation titled "CONTAMINATION: detect contaminating sequences in high-throughput sequencing data" at ISMB/ECCB 2013, Intelligent Systems for Molecular Biology / European Conference on Computational Biology 2013 / Messe Berlin, Berlin, GERMANY (contributed study was presented by co-authors).
  Published on F1000Research: http://f1000research.com/posters/1094020

- 2012 – Paper presentation titled "M4B: A Novel Method for Designing and Ordering Genetic Constructs" at HIBIT 2012, 7th International Symposium on Health Informatics and Bioinformatics / Cappadocia, Nevsehir, TURKEY (contributed study was presented by co-authors).
  Published on IEEE Xplore: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6209054

Skills

- **Programming**: Python, R, JavaScript, C/C++, Java, PHP, Perl, HTML, CSS, MySQL, PostgreSQL, SQLite, NoSQL (MongoDB), LaTeX.

- **Bioinformatics tools**: NGS data QC (FastQC, Trim Galore!, Cutadapt), data manipulation (SAMTools, BCFTools), data analysis (BWA, HISAT, STAR, GATK, MACS2, StringTie, rMATS, IGV, NetworkX, Cytoscape, Autodock4, PyMol, UCSC Chimera, Ensembl API, Rabix).

- **Technologies & Libraries**: Git, svn, AWS, Google Cloud, Wordpress, Django, Laravel, React, D3.js, Twitter Bootstrap, BioPython, Numpy, Pandas, Scipy, Scikit-learn, Matplotlib, pysam, igraph, ggplot2, circlize, ComplexHeatmap.

- **Softwares/OSes**: RStudio, Postman, Android Studio, Adobe Photoshop/Illustrator, MS Office tools, MS Windows, Unix/GNU/Linux, macOS.

Languages

- Turkish (native)

- English (fluent) – METU EPE score: 90.0/100 (2017)

- French (beginner)

Awards and Certificates

- 2017 – Best Poster Award at HIBIT 2017, 10th International Symposium on Health Informatics and Bioinformatics, by ISCB Student Council Regional Study Group Turkey.

- 2016 – Thesis of the Year by METU Prof. Dr. Mustafa N. Parlar Education and Research Foundation.

- 2015 – Travel funding to ISMB/ECCB 2015, Intelligent Systems for Molecular Biology / European Conference on Computational Biology 2015 by BMC Bioinformatics.

- Certificate of high honor / certificate of honor: 2012 – 2013 spring, 2013 – 2014 fall / 2010 – 2011 fall, 2011 – 2012 fall, 2011 – 2012 spring, 2012 – 2013 fall.

- 2011 – iGEM (International Genetically Engineered Machines) Competition World Championship "Software Division: Best Use of Registry" Award / Massachusetts Institute of Technology, Cambridge, MA, USA.
- 2011 – iGEM (International Genetically Engineered Machines) Competition European Jamboree Silver Medal / Vrije University, Amsterdam, THE NETHERLANDS.
- 2011 – Hacettepe University Technopolis Technovation Competition Mention Award / Hacettepe University, Ankara, TURKEY.

Volunteering

- 2015 – Lilly AMECA 2015, Lilly Africa, Middle East and Central Asia (AMECA) 2015 / Indianapolis, IN, USA.
- 2015 – ISMB/ECCB 2015, Intelligent Systems for Molecular Biology / European Conference on Computational Biology 2015 / Dublin, IRELAND.
- 2013 – TEDxMaastricht / Maastricht, THE NETHERLANDS.
- 2010 – 2014: Organizing various scientific conferences on biodiversity, evolutionary biology and many other fields in biology as a member of METU Biology and Genetics Student Club / Ankara, TURKEY.

Memberships

- International Society for Computational Biology, ISCB (as a member since 2015).
- Tree of Evolution Movement (as the co-founder and developer since November, 2010) http://www.evrimagaci.org.
- METU Biology and Genetics Student Club (as a member since September, 2009, as a member of administrative board between September, 2010 – February, 2011, as the head of the administrative board between February, 2011 – June, 2012, as a member of supervising council between June, 2012 – January, 2014).

Interests and Hobbies

- Biology, molecular biology, genetics, computer science, statistics, bioinformatics, coding, programming, website development.
- Blogging on bioinformatics and English language; playing guitar; doing sports; playing basketball and football, volunteering.

References

- Assoc. Prof. Nurcan TUNCBAG
  Informatics Institute, Middle East Technical University (METU), Ankara, TURKEY
  Office: +90 312 210 7882 – ntuncbag@metu.edu.tr
- Prof. Dr. Rengul CETIN-ATALAY
  Informatics Institute, Middle East Technical University (METU), Ankara, TURKEY
  Office: +90 312 210 7887 – rengul@metu.edu.tr
- Assoc. Prof. Dr. Yesim AYDIN SON
  Head of the Medical Informatics Program
  Informatics Institute, Middle East Technical University (METU), Ankara, TURKEY
  Office: +90 312 210 7708 – yesim@metu.edu.tr
- Karl-Heinz Glatting
  Head of the Bioinformatics Unit, Genomics and Proteomics Core Facility
  Bioinformatics Unit (HUSAR), Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), Heidelberg, GERMANY
  Office: +49-6221-42-2349 – genome@dkfz.de