

## EDUCATION

- PhD** University of Copenhagen, Denmark  
Faculty of Health and Medical Sciences, Globe Institute, Section for GeoGenetics 2022/1 - 2025/3 (expected)  
**Title:** Statistical Methods for Population Genetic Inference Based on Low-Depth Sequencing Data  
**Advisors:** Thorfinn Sand Korneliussen and Rasmus Nielsen
- BSc in Biology** Hacettepe University, Turkey  
Faculty of Science, Department of Biology 2015/8 - 2020/7

## RESEARCH EXPERIENCE

- Visiting Researcher at the Moltke Lab** University of Copenhagen, Denmark  
Faculty of Science, Department of Biology, Section for Computational and RNA Biology 2024/09 - 2024/12
  - Collaborated with Ida Moltke on a research project on IBD segment identification with low-depth sequencing data.
- PhD Fellow** University of Copenhagen, Denmark  
Faculty of Health and Medical Sciences, Globe Institute, Section for GeoGenetics 2022/1 - Present (expected 2025/3)
  - Supervised by Thorfinn Sand Korneliussen and Rasmus Nielsen.
  - Developed statistical methods and frameworks for population genetic inference, with a focus on low-depth NGS data.
  - Implemented computational tools for genomic analysis emphasizing efficiency and accuracy.
  - Authored three first-author papers: one under review and two manuscripts in preparation.
  - Presented research findings at three conferences and one poster session.
  - Contributed to collaborative projects outside of the primary PhD projects, resulting in one paper published in *Nature*, and two papers submitted to *Nature* (under review).
  - Delivered training on workload management using Slurm in the *Introduction to Slurm* workshop for the section.
  - Gained and actively applied expertise in C, C++, Python, and R programming, Git, Conda/Bioconda, Snakemake, Make,  $\LaTeX$ , sed, AWK, and Bash, as well as utilizing Slurm and HPC environments for efficient execution of large-scale research and computational tasks.
- Research Assistant** University of Copenhagen, Denmark  
Faculty of Health and Medical Sciences, Globe Institute, Section for GeoGenetics 2020/10 - 2021/12
  - Contributed to developing and evaluating the Lundbeck Center's ancient DNA bioinformatics pipeline, facilitating the processing of hundreds of ancient genomes, resulting in one preprint.
  - Contributed to the development and maintenance of the [ANGSD](#) toolkit.
  - Contributed to a collaborative research project, resulting in a publication.
  - Managed the Geogenetics servers used by many research groups as a server administrator, supporting large-scale computational research and bioinformatics workflows.
  - Utilized skills in diverse research and computational tasks, including sequence alignment, read trimming and quality assessment, detecting contamination, performing analyses such as PCA and admixture inference, and developing scalable pipelines for high-throughput genomic data processing.
- Research Intern at the Fumagalli Lab** Imperial College London, United Kingdom  
Department of Life Sciences 2019/7 - 2019/9
  - Supervised by Matteo Fumagalli.
  - Contributed to the HMMPloddy project by implementing methods such as genotype likelihood estimation for arbitrary N-ploidies, resulting in a publication.
  - Supported by the Erasmus+ Traineeship Grant.
  - Enhanced proficiency in Python programming, Bash/Shell scripting, and Git.
- Research Intern at the Biogeography Research Lab** Hacettepe University, Turkey  
Faculty of Science, Department of Biology 2019/2 - 2020/7
  - Supervised by Barış Özüdoğru.
  - Developed and optimized a Snakemake pipeline for phylogenomic analyses using RADSeq data, resulting in a manuscript in preparation.

- Delivered presentations on Linux/Bash and phylogenetic methods during internal lab meetings and taught the practical session in the *Introduction to Bioinformatics and Computational Biology* workshop, refining teaching and presentation skills.
- Enhanced proficiency in Python, Snakemake, R, Bash/Shell scripting, Git, and gained experience in executing large analyses on HPC clusters with Slurm.

- Research Intern at the Comparative and Evolutionary Biology Lab** Middle East Technical University, Turkey  
 Department of Biological Sciences: Molecular Biology and Genetics 2018/8 - 2019/6
  - Supervised by Mehmet Somel.
  - Performed analyses to evaluate the effects of the properties of probe sequences in the 1240K SNP capture method.
  - Applied and further developed skills in data analysis using R.
- Research Intern at the Evolutionary and Quantitative Genetics Lab** Hacettepe University, Turkey  
 Faculty of Science, Department of Biology 2015/10 - 2018/12
  - Supervised by Ergi Deniz Özsoy.
  - Conducted statistical analyses for quantitative genetics projects using Drosophila Genetic Reference Panel 2 data, which resulted in a poster presentation and acknowledgments in two MSc and one PhD thesis.
  - Gained and applied skills in data analysis using R.

## PUBLICATIONS

---

- [6] Fulya E. Yediay [et al, including **Isin Altinkaya**]. Ancient genomics support deep divergence between Eastern and Western Mediterranean Indo-European languages. (under review by *Nature*). *bioRxiv* (2024).
- [5] **Isin Altinkaya**, Thorfinn S. Korneliussen, Rasmus Nielsen. vcfgl: A flexible genotype likelihood simulator for VCF/BCF files. (under review by *Bioinformatics*). *bioRxiv* (2024).
- [4] Hugh McColl [et al, including **Isin Altinkaya**]. Steppe Ancestry in western Eurasia and the spread of the Germanic Languages. (under review by *Nature*). *bioRxiv* (2024).
- [3] Morten E. Allentoft [et al, including **Isin Altinkaya**]. Population genomics of post-glacial western Eurasia. *Nature* (2024).
- [2] Samuele Soraggi, Johanna Rhodes, **Isin Altinkaya**, Oliver Tarrant, François Balloux, Matthew C. Fisher, Matteo Fumagalli. HMMploidy: inference of ploidy levels from short-read sequencing data. *Peer Community Journal* (2022).
- [1] Alex Mas-Sandoval, Nathaniel S. Pope, Knud Nor Nielsen, **Isin Altinkaya**, Matteo Fumagalli, and Thorfinn S. Korneliussen. Fast and accurate estimation of multidimensional site frequency spectra from low-coverage high-throughput sequencing data. *GigaScience* (2022).

## Manuscripts in Preparation

- [3] **Isin Altinkaya**, Lei Zhao, Rasmus Nielsen, Thorfinn S. Korneliussen. ngsAMOVA: A genotype likelihood framework for analysis of molecular variance (AMOVA) with low-depth sequencing data.
- [2] **Isin Altinkaya**, Ida Moltke, Rasmus Nielsen, Thorfinn S. Korneliussen. IBDGL: A method for the accurate detection of IBD segments in low- and medium-depth genomic data without phasing.
- [1] **Isin Altinkaya**, Emrullah Yılmaz, İsmail K. Sağlam, Barış Özüdoğru. Understanding the phylogenetic relationships within the *Noccaea* species complex from RADSeq data using the multi-species coalescent.

## PROFESSIONAL EXPERIENCE

---

- Freelance Programmer** Remote  
 Self-Employed 2019/1 - 2020/10
  - Applied various **AI** and **ML** techniques such as **Random Forest** and **K-Means Clustering** to solve complex problems.
  - Utilized programming languages and tools such as **Python**, **R**, **JavaScript**, **HTML&CSS**, **AWK**, and **Bash scripting** to perform data analysis and customized program development.
  - Delivered customized solutions for clients, developing strong problem-solving and programming skills.

## SKILLS

---

**Programming languages:** <sup>Expert</sup> C , <sup>Expert</sup> R , <sup>Advanced</sup> Python , <sup>Advanced</sup> C++ , <sup>Intermediate</sup> Rust , <sup>Intermediate</sup> Perl , <sup>Beginner</sup> Java  
**Scripting and typesetting:** L<sup>A</sup>T<sub>E</sub>X, Markdown, HTML&CSS, JavaScript, Angular, Google Apps Script, sed, AWK, Bash  
**Operating systems:** GNU/Linux (including server administration and HPC environments)  
**Other tools and frameworks:** Git, Conda/Bioconda, Snakemake, Nextflow, Docker, RShiny, Dash, Django, Make

## OPEN SOURCE PROJECTS AND CONTRIBUTIONS

---

- [gptchatteR](#): An **R** package for interacting with **OpenAI's language models** in the R console, enabling AI-driven data analysis, visualization, and learning. (2022 - 2023).
- [ngsAMOVA](#): A **C/C++** toolkit for AMOVA,  $d_{xy}$ , Neighbor-Joining, and IBD segment detection from NGS data using probabilistic frameworks (see Publications/Manuscripts in Preparation). (2022 - present).
- [google-slides\\_scripts](#): Custom **Google Apps Script** functions to enhance Google Slides' features. (2022).
- [vcfgl](#): A **C/C++** tool for fast and efficient simulation of genotype likelihoods (see Publications/5). (2022 - present).
- [ANGSD](#): Contributed to the development and maintenance of ANGSD, a **C/C++** toolkit for analyzing next-generation sequencing data. (2021 - 2024).
- [bash-functions](#): A collection of custom **Bash** functions for various tasks, including 'runda' for safe **conda** environments and 'treemd' for generating **Markdown**-formatted Git repository structures. (2021 - present).
- [decluster](#): A **C++** tool for identifying and stratifying duplicate reads in high-throughput sequencing data, including functionality for detecting cluster duplicates, PCR duplicates, and calculating sequence complexity and GC content. (2021 - 2022).
- [PopulasyonBuyumesi](#): An **RShiny** app for teaching population ecology through population growth simulations and interactive plots (in Turkish). Available at [popbsim.isinaltinkaya.com](http://popbsim.isinaltinkaya.com). (2020).
- [cheatsheets](#): Custom **Markdown** tutorials with tips and examples for tools like **Bash**, **Git**, **Jupyter**, **Make**, **Conda**, and **R**. (2020 - present).
- [HMMploidy](#): A **Python** tool to calculate ploidy levels from genotype likelihoods and coverage using Hidden Markov Models. (see Publications/2), (2019).
- [i3gcalendar](#): Google Calendar integration with i3blocks, written in **Python**, for the i3 window manager. (2018).




## LANGUAGES

---

- **Turkish**: Native
- **English**: Native-like fluency (ILR level R-5, S-4+, L-5, W-5)
- **German**: Elementary proficiency (ILR level R-1, S-0+, L-0, W-0+)
- **Danish**: Elementary proficiency (ILR level R-1, S-0+, L-0+, W-1)

## REFERENCES

---

- [Rasmus Nielsen](#)  • [rasmus\\_nielsen@berkeley.edu](mailto:rasmus_nielsen@berkeley.edu)  
University of California, Berkeley, Department of Integrative Biology and Department of Statistics  
University of Copenhagen, Faculty of Health and Medical Sciences, Globe Institute, Section for GeoGenetics
- [Thorfinn Sand Korneliussen](#)  • [tskorneliussen@sund.ku.dk](mailto:tskorneliussen@sund.ku.dk)  
University of Copenhagen, Faculty of Health and Medical Sciences, Globe Institute, Section for GeoGenetics
- [Matteo Fumagalli](#)  • [m.fumagalli@qmul.ac.uk](mailto:m.fumagalli@qmul.ac.uk)  
School of Biological and Behavioural Sciences, Queen Mary University of London, United Kingdom