

Jonas Falck

BIOINFORMATICS - DATA ANALYSIS

Muzenplaats 4, 6525JA Nijmegen, Netherlands

☎ (+31) 655707314 | ✉ jonas@falcken.de | 📷 joe-nas

Skills

IT skills

- **r/bioconductor:** tidyverse, ggplot2, biomaRt, Rcpp, caret, data.table, grid, shiny
- **python:** pandas, numpy, matplotlib, seaborn, scikit-learn, django, SQLAlchemy, flask, jinja2
- **productivity:** shell scripting, regex, ucsc web/command line tools, bedtools, LaTeX, markdown, pandoc, office
- **database:** SQL, elasticsearch
- **operating systems:** linux 10+ years (arch, debian based), windows, macOS
- **version control:** git
- **container and virtualization:** docker, qemu, wsl
- **webserver:** nginx, shiny

Data science/Machine learning

- **supervised:** linear regression, knn, decision trees
- **unsupervised:** HCA, PCA, kmeans
- **statistics:** parametric and non-parametric methods
- **visualization:** publication ready plots, development of custom visualizations using grid
- **datatypes:** 2bit, fasta, fastq, vcf, bigWig, bed
- **databases:** GEO DataSets, dbSNP, RefSeq, OMIM, 1000Genomes

Languages

- **English:** fluent
- **Dutch:** fluent (nt2)
- **German:** native

Experience

Identification and ranking of p63 binding sites putatively involved in the etiology of non-syndromic cleft lip with or without cleft palate

Nijmegen, Netherlands

Radboud University Medical Center Nijmegen RUNMC, Human Genetics

2015-2016

- In-silico prediction and in-vitro validation of clinically relevant transcription factor binding sites using an integrative multiomics approach.
- Took ownership of a wetwork research project, transforming it into bioinformatical research, teaching myself the necessary bioinformatic skills.
- Development of a pipeline to integrate publicly available and in-house multi-omics data (Chip-seq, SNP, GWAS, conservation, linkage disequilibrium), to generate reproducible results. haplotplotR: <http://tinyurl.com/4cqw7s5h>
- Effective communication and presentation of bioinformatical methods and results to researches with no bioinformatical expertise.

CTCF-motif directionality controls CTCF-mediated chromatin interactions and correlates with topological domain structure

Nijmegen, Netherlands

Center for Molecular and Biomolecular Informatics CMBI, Comparative Genomics

2016

- Drove and transformed an explorative research project into a hypothesis driven project resulting in a publication (see writing).
- Hypothesis generation by leveraging multi-omics datasets describing different dimensionalities of the genome, ranging 1D (sequence), 2D (ChIP-Seq), and 3D data (ChIA-PET, Hi-C).
- Hypothesis testing by applying parametric and non-parametric methods, randomization, as well as modeling of chromatin loops.
- Applying unsupervised machine learning techniques e.g. PCA, HCA as well as a multitude of visualizations for data exploration.

Logistics employee

Beuningen, Netherlands

Quantore

2018 - now

- Using warehouse management software to analyse erroneous orders and to troubleshoot and provide technical support to co-workers.
- Solving problems independently or to communicate issues with the superiors.
- Ensuring customer satisfaction by quality controlling orders.

Extracurricular Activity

Machine Learning

Grade: 87%

Stanford University, Online

- 11 week course covering go to methods applied in supervised and unsupervised learning, regularization and how to tackle problems like bias and overfitting.
- Certificate of accomplishment: <http://tinyurl.com/v5ytdzhs>

Intro to Machine Learning

KAGGLE.COM, ONLINE MACHINE-LEARNING COMMUNITY

- This course covered the basics of machine learning using python giving an introduction in topics such as data exploration, building models, model validation, under and overfitting and random forests.
- Certificate of accomplishment: <https://tinyurl.com/3t9xuxdy>

Pandas

KAGGLE.COM, ONLINE MACHINE-LEARNING COMMUNITY

- This course is dedicated to the pandas python data analysis library and covered useful data manipulation skills.
- Certificate of accomplishment: <https://tinyurl.com/4ec7n95c>

Data Visualization

KAGGLE.COM, ONLINE MACHINE-LEARNING COMMUNITY

- This course is covered the statistical data visualization python library seaborn.
- Certificate of accomplishment: <https://tinyurl.com/bdee5xwk>

Feature Engineering

KAGGLE.COM, ONLINE MACHINE-LEARNING COMMUNITY

- This course on feature engineering covers techniques necessary to create more effective features for ML tasks and delves into topics like mutual information for feature selection, feature creation, clustering, PCA as well as target encoding for categorical data.
- Certificate of accomplishment: <https://tinyurl.com/2u7kbxyh>

Computing for Data Analysis

Grade: 97%

JOHN HOPKINS UNIVERSITY, ONLINE

- This course covered the basics of data analysis and visualization in R in a medical context.
- Certificate of accomplishment: <http://tinyurl.com/2ez3unkj>

Nextflow and nf-core community training

2023

SEQERA.IO, ONLINE

- The training gave an introduction to scientific workflow system Nextflow and nf-core and covered basic deployment as well as development of nextflow nf-core pipelines.

The Complete JavaScript Course 2023: From Zero to Expert!

UDEMY, ONLINE

- This course is a 69 hours complete and in-depth course covering fundamentals and advanced topics of JavaScript development.
- Curriculum: <https://www.udemy.com/course/the-complete-javascript-course/>
- Certificate of accomplishment: <https://tinyurl.com/mrxc5h3x>

Node.js, Express, MongoDB & More

Ongoing

UDEMY, ONLINE

- This course covers backend engineering in JavaScript and focuses on building a secure and versatile RESTful API using Node.js, Express.js, MongoDB, and more.
- Curriculum: <https://www.udemy.com/course/nodejs-express-mongodb-bootcamp/>

Full stack web development

SELF-TAUGHT

- Setup and administration of headless Linux systems.
- Applied SQL-like databases and the NoSQL database elasticsearch in a website environment.
- Used python model-view-controller frameworks django and flask + SQLAlchemy.
- Applied caching technology like memcached and redis to speed up website performance.
- Learned html, css, templating and javascript for front-end use.

Writing

CTCF-mediated chromatin loops enclose inducible gene regulatory domains

BMC Genomics

MARTIN OTI, JONAS FALCK, MARTIJN A. HUYNEN & HUIQING ZHOU

March 2016, BMC Genomics; 17:252

- Available here: <https://doi.org/f8vngs>

CTCF-motif directionality controls CTCT-mediated chromatin interactions and correlates with topological domain structure

Master Internship

COMPARATIVE GENOMICS, CENTER FOR MOLECULAR AND BIOMOLECULAR INFORMATICS NIJMEGEN

2016

- Available here: <https://tinyurl.com/2p9hehex>

Identification, ranking and testing of p63 binding sites putatively involved in the etiology of non-syndromic cleft lip with or without cleft palate

HUMAN GENETICS, UNIVERSITY MEDICAL CENTER NIJMEGEN

Master Internship

2016

- Available here: <https://tinyurl.com/259vt5zk>

Software

haploplotR: Visualizing linkage disequilibrium from 1000 genomes data

[github](#)

JONAS FALCK

- HaploplotR is a project that provides a tool for visualizing linkage disequilibrium patterns in human populations using data from the 1000 Genomes Project. HaploplotR uses haplotype data from the 1000 Genomes Project to generate LD plots, which show patterns of correlation between alleles at different loci across the genome. These plots can help researchers to identify regions of the genome that are associated with particular traits or diseases.
- Available here: <http://tinyurl.com/4cjlw7s5h>

Education

Radboud University

Nijmegen, Netherlands

M.S. IN MEDICAL BIOLOGY

2015 -

Radboud University

Nijmegen, Netherlands

B.S. IN MEDICAL BIOLOGY

2012 - 2015