

# Johannes Köster

---

Address Überruhrstr. 496, 45289 Essen, Germany  
Phone +49 1520 9045922  
Email johannes.koester@uni-due.de  
Homepage <http://johanneskoester.bitbucket.io>  
Birth May 5, 1985 in Gelsenkirchen, Germany  
Nationality German  
Family status Married, one child

---

## Education

- 2011-2015 **Ph.D. in Computer Science (grade 1.0/A with honors)** at TU Dortmund, Germany. Thesis title: “Parallelization, Scalability, and Reproducibility in Next-Generation Sequencing Analysis”. Supervisor: Sven Rahmann, Chair of Genome Informatics, University of Duisburg-Essen. Second assessor: Axel Mosig, Bioinformatics Group, Ruhr University Bochum.
- 2005 – 2010 **Diploma (equiv. to M.Sc.) in Computer Science (grade 1.0/A with honors)** at TU Dortmund, Germany. Thesis title: “Propagating Interaction Logic toward Predictive Protein Hypernetworks”. Supervisors: Sven Rahmann, Bioinformatics for High-Throughput-Technologies, Computer Science XI, TU Dortmund and Eli Zamir, Systems Biology of Cell-Matrix Adhesion, Department of Systemic Cell Biology, Max Planck Institute of Molecular Physiology Dortmund.

---

## Positions

- from August 2017 **Group leader** (permanent position), Algorithms for reproducible bioinformatics (<https://koesterlab.github.io>), Genome Informatics, Institute of Human Genetics, Faculty of Medicine, University of Duisburg-Essen.
- August 2016 – August 2017 **Researcher** (NWO Veni funded) at Life Sciences, Centrum Wiskunde & Informatica, Amsterdam, Netherlands.
- since August 2016 **Consultant** (data analysis and workflow management) for Juno therapeutics, 307 Westlake Avenue, Seattle, USA.
- since May 2016 **Consulting scientist** in Myles Brown’s group (Division of Molecular and Cellular Oncology, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School), USA.
- May 2016 – July 2016 **Postdoctoral research fellow** under the supervision of Alexander Schönhuth, Life Sciences, Centrum Wiskunde & Informatica, Amsterdam, Netherlands.

- 2015 – 2016 **Postdoctoral research fellow** in Xiaole Shirley Liu’s group (Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard School of Public Health) and Myles Brown’s group (Division of Molecular and Cellular Oncology, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School), USA.
- 2013 – 2014 **Volunteership** at NIH regarding the adaptation of Snakemake for the biowulf computer cluster in cooperation with Sean Davis, Center for Cancer Research, NCI/NIH, USA.
- 2011 – 2015 **Research fellow** at both the Chair of Genome Informatics of Sven Rahmann, University of Duisburg-Essen and in the “European Network for Cancer in Children and Adolescents” project under Alexander Schramm, University Hospital Essen, Germany.
- 2011 **Research fellow** in Sven Rahmann’s group Bioinformatics for High-Throughput-Technologies, Computer Science XI, TU Dortmund, Germany.
- 2009 – 2015 **Guest member** in the group of Eli Zamir, Systems Biology of Cell-Matrix Adhesion, Department of Systemic Cell Biology, Max Planck Institute of Molecular Physiology Dortmund, Germany.

---

## Grants and awards

- 2017 **Lorentz workshop** “Making sense of millions of single cells: approaching the era of Single Cell Data Science, experimental advances and computational challenges”. Amount: 12,000 €
- 2016 **NWO Veni Grant** “Fully reproducible workflows scaling from workstations to the cloud” (016.Veni.173.076). Amount: 250,000 €
- Travel fellowship** for HiTSeq 2016, Orlando, USA. Amount: 1000 \$
- 2015 **Free organizational plan from Continuum Analytics** for the Bioconda project. Amount: 1176 \$
- 2014 **NVIDIA Hardware Grant** received for publication “Massively parallel read mapping on GPUs with the q-group index and PEANUT”. Amount: 2000 € (Tesla GPU)
- 2011 **Hans-Uhde-Preis 2011** by the ThyssenKrupp Uhde GmbH for extraordinary diploma thesis with practical application in Computer Science. Amount: 500 €
- Travel award** of the 9th International Conference on Pathways, Networks and Systems Medicine, Crete, Greece. Amount: 1000 \$
- Honorable mention** at the ”Doktorandenkolleg Informatik Ruhr“, Germany.
- Best poster award** at the science day of the University Hospital Essen, Germany. Amount: 100 €

---

## Invited Talks

- November 2017 **Robert-Koch-Institute, Berlin:** Benchmarking RNA-seq data analysis
- November 2017 **ETH Zürich:** Snakemake tutorial
- February 2017 **Gutenberg University Mainz:** Snakemake and Bioconda
- February 2017 **Forschungszentrum Jülich:** Snakemake and Bioconda
- December 2016 **Berlin Institute of Health:** Snakemake tutorial

- November 2016 **Helmholtz Open Science Workshop, Dresden, Germany:** Keynote about sustainable software development in science.
- September 2016 **Netherlands Cancer Institute (NKI), Amsterdam:** Snakemake and Bioconda.
- July 2016 **Rust bay area meeting, Mozilla Foundation Los Angeles, USA:** Bioinformatics with Rust.
- June 2016 **Aquatic Ecosystem Research, University of Duisburg-Essen, Germany:** Snakemake and Bioconda.
- June 2016 **Leiden University Medical Center, Netherlands:** Algebraic variant calling, Snakemake and Bioconda.
- 2015 **Broad Institute, Boston, USA:** Workflow Management with Snakemake
- 2014 **Workshop on Fast Data Processing on GPUs, CUDA Center of Excellence, Dresden:** Massively parallel read mapping on graphics cards.
- DTL Focus Meeting, Utrecht, Netherlands:** Snakemake

---

### Organized events

- June 2018 **Co-Organizer:** Lorentz workshop, Leiden, The Netherlands: Making sense of millions of single cells: approaching the era of Single Cell Data Science, experimental advances and computational challenges
- September 2017 **PC member:** CIBB 2017, Cagliari, Italy
- March 2017 **Organizer:** Snakemake tutorial at CWI Amsterdam
- February 2017 **Co-Organizer and session chair:** workshop “data structures in bioinformatics” (DSB) 2017 (<https://dsb2017.github.io>)

---

### Teaching and Supervision Experience

- 2017 **Co-supervisor of master thesis** about building a pan-genome data structure for viruses, Bastiaan van der Roest
- 2015 **Supervision of intern** for the documentation of a workflow for the analysis of CRISPR/Cas9 knockout screens.  
**Co-supervisor of master thesis** “Entwurf von q-Gramm-basierten Readmappingverfahren für lange Reads”, Bianca Patro, TU Dortmund, Germany.
- 2013 **Guest lecture** about variant calling with next-generation sequencing in the course “Statistic in Genetics”, Faculty of Statistics, TU Dortmund, Germany.
- 2013-2015 **Supervision of student assistants** for the development and maintenance of the web-based Exomate platform for interactive, database-supported exploration and filtration of variant calls, University of Duisburg-Essen, Germany.
- 2013-2015 **Supervision of student assistant** for the development and implementation of logic-based protein network models incorporating interaction dependencies, University of Duisburg-Essen, Germany.
- 2012 **Guest lectures** about protein networks, complex prediction and interaction dependencies in the course “Computational Omics”, Faculty of Computer Science, TU Dortmund, Germany.

- Co-supervisor of bachelor thesis** “Rekonstruktion von Protein-Interaktionsabhängigkeiten mit dem Quine-McCluskey-Algorithmus”, Bianca Patro, TU Dortmund, Germany.
- 2011 **Teaching assistant** for “Datastructures, Algorithms and Programming 1”, at the Chair of Computer Science X, TU Dortmund, Germany.
- Co-supervisor of diploma thesis** "Entwurf einer Datenstruktur für Pangenome", Christiane Küch, TU Dortmund, Germany.
- Co-supervisor of bachelor thesis** “Konstruktion von Protein-Hypernetzwerken durch Text-Mining in der PubMed Datenbank”, Michael Nimbs, TU Dortmund, Germany.
- 2007-2009 **Student teaching assistant** for “Datastructures, Algorithms and Programming 1”, at the Chair of Computer Science X, TU Dortmund, Germany.

---

### Peer reviewed Publications

- 2018 Cornwell, M., Vangala, M., Taing, L., Herbert, Z., **Köster, J.**, Li, B., Sun, H., Li, T., Zhang, J., Qiu, X., Pun, M., Jeselsohn, R., Brown, M., Liu, X. S., and Long, H. (2018). VIPER: Visualization Pipeline for RNA-seq, a Snakemake workflow for efficient and complete RNA-seq analysis. *BMC Bioinformatics* 19:135.
- 2016 Stöcker, B. K., **Köster, J.**, Rahmann, S. (2016). SimLoRD: Simulation of Long Read Data. *Bioinformatics*, 32(17), 2704–2706.
- Ma, J., **Köster, J.**, Qin, Q., Hu, S., Li, W., Chen, C., ... Liu, X. S. (2016). CRISPR-DO for genome-wide CRISPR design and optimization. *Bioinformatics*.
- Townsend, E. C., Murakami, M. A., Christodoulou, A., Christie, A. L., **Köster, J. (first collaborative author)**, DeSouza, T. A., ... Weinstock, D. M. (2016). The Public Repository of Xenografts Enables Discovery and Randomized Phase II-like Trials in Mice. *Cancer Cell*, 29(4), 574–586.
- 2015 Li, W., **Köster, J. (co-first author)**, Xu, H., Chen, C.-H., Xiao, T., Liu, J. S., ... Liu, X. S. (2015). Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biology*, 16(1), 281.
- Köster, J.** (2015). Rust-Bio: a fast and safe bioinformatics library. *Bioinformatics*.
- Schramm, A., **Köster, J.**, Assenov, Y., Althoff, K., Peifer, M., Mahlow, E., Odersky, A., Beisser, D., Ernst, C., Henssen, A. G., Stephan, H., Schröder, C., Heukamp, L., Engesser, A., Kahlert, Y., Theissen, J., Hero, B., Roels, F., Altmüller, J., Nürnberg, P., Astrahantseff, K., Gloeckner, C., De Preter, K., Plass, C., Lee, S., Lode, H. N., Henrich, K., Gartlgruber, M., Speleman, F., Schmezer, P., Westermann, F., Rahmann, S., Fischer, M., Eggert, A., Schulte, J. H. (2015). Mutational dynamics between primary and relapse neuroblastomas. *Nature Genetics*, 47(8), 872–877.
- Schwermer, M., Lee, S., **Köster, J.**, Maerken, T. van, Stephan, H., Eggert, A., Morik, K., Schulte, J. H., Schramm, A., van Maerken, T. (2015). Sensitivity to cdk1-inhibition is modulated by p53 status in preclinical models of embryonal tumors. *Oncotarget*, 6(17), 15425–15435.

- 2014 **Köster, J.**, Rahmann, S. (2014). Massively parallel read mapping on GPUs with the q-group index and PEANUT. *PeerJ* 2, e606.
- 2013 Schramm, A., **Köster, J.**, Marschall, T., Martin, M., Schwermer, M., Fielitz, K., Büchel, G., Barann, M., Esser, D., Rosenstiel, P., Rahmann, S., Eggert, A., Schulte, J. H. (2013). Next-generation RNA sequencing reveals differential expression of MYCN target genes and suggests the mTOR pathway as a promising therapy target in MYCN-amplified neuroblastoma. *International Journal of Cancer* 132, E106–E115.
- Rahmann, S., Martin, M., Schulte, J. H., **Köster, J.**, Marschall, T., Schramm, A. (2013). Identifying transcriptional miRNA biomarkers by integrating high-throughput sequencing and real-time PCR data. *Methods* 59, 154–163.
- Althoff, K., Beckers, A., Odersky, A., Mestdagh, P., **Köster, J.**, Bray, I.M., Bryan, K., Vandesompele, J., Speleman, F., Stallings, R. L., Schramm, A., Eggert, A., Sprüssel, A., Schulte, J. H. (2013). MiR-137 functions as a tumor suppressor in neuroblastoma by downregulating KDM1A. *International Journal of Cancer* 133, 1064–1073.
- 2012 **Köster, J.**, Rahmann, S. (2012). Snakemake – a scalable bioinformatics workflow engine. *Bioinformatics* 28, 2520–2522.
- Köster, J.**, Zamir, E., Rahmann, S. (2012). Efficiently mining protein interaction dependencies from large text corpora. *Integr. Biol.* 4, 805–812.
- Köster, J.**, Rahmann, S. (2012). Building and Documenting Workflows with Python-Based Snakemake, in: Böcker, S., Hufsky, F., Scheubert, K., Schleicher, J., Schuster, S. (Eds.), *German Conference on Bioinformatics 2012, OpenAccess Series in Informatics (OASISs)*. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, pp. 49–56.
- Schramm, A., Schowe, B., Fielitz, K., Heilmann, M., Martin, M., Marschall, T., **Köster, J.**, Vandesompele, J., Vermeulen, J., de Preter, K., Koster, J., Versteeg, R., Noguera, R., Speleman, F., Rahmann, S., Eggert, A., Morik, K., Schulte, J. H. (2012). Exon-level expression analyses identify MYCN and NTRK1 as major determinants of alternative exon usage and robustly predict primary neuroblastoma outcome. *Br J Cancer* 107, 1409–1417.

---

## Other publications

- 2017 Dijkstra, L. J., **Köster, J. (co-first author)**, Marschall, T., Schönhuth, A. (2017). Enhancing Sensitivity And Controlling False Discovery Rate In Somatic Indel Discovery Using A Latent Variable Model. *bioRxiv* 121954
- 2016 **Köster, J.**, Brown, M., Liu, X. S. (2016). A Bayesian model for single cell transcript expression analysis on MERFISH data. Under review.

---

## Software (excerpt)

- Snakemake Popular workflow management system (<http://snakemake.bitbucket.org>)
- Bioconda Popular distribution of bioinformatics software realized as a channel for the versatile package manager Conda (<https://bioconda.github.io>).
- libprosic Generic library for calling various types of genomic variants from NGS data using a Bayesian latent variable model (<https://github.com/prosic/libprosic>).

MERFISHtools	Bayesian framework for single-cell transcript expression analysis on multiplexed error-robust fluorescence in-situ hybridization data ( <a href="https://merfishtools.github.io">https://merfishtools.github.io</a> ).
Rust-Bio	Bioinformatics algorithm and data structure library for the Rust language ( <a href="https://rust-bio.github.io">https://rust-bio.github.io</a> ).
VISPR	Web-based, interactive visualization framework for CRISPR/Cas9 knockout screen experiments ( <a href="https://bitbucket.org/liulab/vispr">https://bitbucket.org/liulab/vispr</a> ).
ALPACA	Framework for algebraic variant calling with Bayesian control of the false discovery rate and statistical dependencies between samples ( <a href="https://github.com/johanneskoester/alpaca">https://github.com/johanneskoester/alpaca</a> )
PEANUT	Fast GPU-based read mapper exploiting the q-group index data structure ( <a href="https://peanut.readthedocs.org">https://peanut.readthedocs.org</a> ).
Libmodallogic	Implementation of the modal logic K and the corresponding tableau algorithm as a JAVA library ( <a href="https://github.com/johanneskoester/libmodallogic">https://github.com/johanneskoester/libmodallogic</a> ).