

# Karel Břinda

INRIA French national research institute

Updated on March 11, 2022

INRIA/IRISA, GenScale  
Campus de Beaulieu  
35042 Rennes, France  
✉ karel.brinda@inria.fr  
📄 <http://karel-brinda.github.io>

## Research Interests

Bioinformatics / Computational Biology: Algorithms, Genomics, Rapid Diagnostics, Pathogens, Antibiotic Resistance

## Academic Appointments

- 2022–present **INRIA – National Research Institute for Computer Science and Automation** (Rennes, France).  
*Permanent researcher*, INRIA Starting Faculty Position, INRIA/IRISA Rennes, GenScale team
- 2017–2021 **Harvard Medical School & Harvard TH Chan School of Public Health** (Boston, USA)  
2019–2021 *Research associate*, Department of Biomedical Informatics, HMS  
2018–2019 *Research associate*, Center for Communicable Disease Dynamics, HSPH  
2017–2018 *Postdoctoral research fellow*, Center for Communicable Disease Dynamics, HSPH.  
Advisors: Michael Baym (HMS) & William P. Hanage (HSPH)

## Education

- 2013–2016 **PhD**, Computer Science, LIGM Université Paris-Est (Paris, France).  
Advisors: Gregory Kucherov (CNRS) and Valentina Boeva (Curie Institute & Cochin Institute)  
Dissertation: *Novel computational techniques for mapping and classification of Next-Generation Sequencing data*
- 2011–2013 **MSc (Ing)**, Mathematical Computer Science (with honors), Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering (Prague, Czech Republic).
- 2008–2011 **BSc**, Mathematical Computer Science, Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering (Prague, Czech Republic).

## Publications

Google Scholar: [https://scholar.google.com/citations?user=TTaAd\\_MAAAAJ](https://scholar.google.com/citations?user=TTaAd_MAAAAJ)

### Selected Journal Articles

- [1] **K. Břinda**, M. Baym, and G. Kucherov, “Simplitigs as an efficient and scalable representation of de Bruijn graphs,” *Genome Biology* **22**(96), 2021. doi:10.1186/s13059-021-02297-z.
- [2] **K. Břinda**, A. Callendrello, K. C. Ma, D. R. MacFadden, T. Charalampous, R. S. Lee, L. Cowley, C. B. Wadsworth, Y. H. Grad, G. Kucherov, J. O’Grady, M. Baym, and W. P. Hanage, “Rapid inference of antibiotic resistance and susceptibility by genomic neighbour typing,” *Nature Microbiology* **5**, pp. 455–464, 2020. doi:10.1038/s41564-019-0656-6.
- [3] **K. Břinda**, V. Boeva, and G. Kucherov, “RNF: a general framework to evaluate NGS read mappers,” *Bioinformatics* **32**(1), pp. 136–139, 2016. doi:10.1093/bioinformatics/btv524.
- [4] **K. Břinda**, M. Sykulski, and G. Kucherov, “Spaced seeds improve  $k$ -mer-based metagenomic classification,” *Bioinformatics* **31**(22), pp. 3584–3592, 2015. doi:10.1093/bioinformatics/btv419.

### Other Journal Articles

- [5] N. B. Wikle, T. N.-A. Tran, B. Gentileco, S. Leighow, J. Albert, E. R. Strong, **K. Břinda**, H. Inam, F. Yang, S. Hossain, P. Chan, W. P. Hanage, M. Messick, J. Pritchard, E. Hanks, and M. F. Boni, “SARS-CoV-2 epidemic after social and economic reopening in three U.S. states reveals shifts in age structure and clinical characteristics,” *Science Advances* **8**(4), 2022. doi:10.1126/sciadv.abf9868.
- [6] Y. Che, X. Xu, Y. Yang, **K. Břinda**, W. P. Hanage, C. Yang, and T. Zhang, “High-resolution genomic surveillance elucidates a multilayered hierarchical transfer of resistance between WWTP-and human/animal-associated bacteria,” *Microbiome* **10**(1), pp. 1–16, 2022. doi:10.1186/s40168-021-01192-w.
- [7] Y. Che, Y. Yang, X. Xu, **K. Břinda**, M. F. Polz, W. P. Hanage, and T. Zhang, “Conjugative plasmids

interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes,” *Proceedings of the National Academy of Sciences* **118**(6), 2021. doi:10.1073/pnas.2008731118.

- [8] T. N.-A. Tran, N. Wikle, E. Albert, H. Inam, E. R. Strong, **K. Břinda**, S. M. Leighow, F. Yang, S. Hossain, J. R. Pritchard, P. Chan, W. P. Hanage, E. M. Hanks, and M. F. Boni, “Optimal SARS-CoV-2 vaccine allocation using real-time attack-rate estimates in Rhode Island and Massachusetts,” *BMC Medicine* **19**(162), 2021. doi:10.1186/s12916-021-02038-w.
- [9] D. R. MacFadden, B. Coburn, **K. Břinda**, A. Corbeil, N. Daneman, D. Fisman, R. Lee, M. Lipsitch, A. McGeer, R. Melano, S. Mubareka, and W. P. Hanage, “Using genetic distance from archived samples for the prediction of antibiotic resistance in *Escherichia coli*,” *Antimicrobial Agents and Chemotherapy* **64**(5), pp. 455–464, 2020. doi:10.1128/aac.02417-19.
- [10] B. Grüning, R. Dale, A. Sjödin, B. A. Chapman, J. Rowe, C. H. Tomkins-Tinch, R. Valieris, J. Köster, and The Bioconda Team (incl. **K. Břinda**), “Bioconda: sustainable and comprehensive software distribution for the life sciences,” *Nature Methods* **15**(7), pp. 475–476, 2018. doi:10.1038/s41592-018-0046-7.
- [11] **K. Břinda**, E. Pelantová, and O. Turek, “Balances of  $m$ -bonacci words,” *Fundamenta Informaticae* **132**(1), pp. 33–61, 2014. doi:10.3233/FI-2014-1031.
- [12] L. Balková, **K. Břinda**, and O. Turek, “Abelian complexity of infinite words associated with quadratic parry numbers,” *Theoretical Computer Science* **412**(45), pp. 6252–6260, 2011. doi:10.1016/j.tcs.2011.08.016.

### Refereed Full Conference Papers

- [13] P. Červenka, **K. Břinda**, M. Hanousková, P. Hofman, and R. Seifert, “Blind friendly maps: tactile maps for the blind as a part of the public map portal (mapy.cz),” in *Computers Helping People with Special Needs, ICCHP 2016, 2016, Proceedings, Part II*, pp. 131–138, 2016.
- [14] **K. Břinda**, “Languages of lossless seeds,” in *Proceedings 14th International Conference on Automata and Formal Languages (AFL 2014), Electronic Proceedings in Theoretical Computer Science*, vol. 151, pp. 139–150, 2014.

### Technical Reports

- [15] **K. Břinda**, V. Boeva, and G. Kucherov, “Ococo: an online variant and consensus caller,” *arXiv*, vol. 1712.01146, no. q-bio.GN, 2018.
- [16] **K. Břinda**, V. Boeva, and G. Kucherov, “Dynamic read mapping and online consensus calling for better variant detection,” *arXiv*, vol. 1605.09070, no. q-bio.GN, 2016.

### Patents

- [17] W. P. Hanage, **K. Břinda**, and M. Baym, “Rapid identification of strains from sequence data,” *United States Patent Application*, no. 17/251343, 2021.

### Theses

- [18] **K. Břinda**, “Novel computational techniques for mapping and classification of Next-Generation Sequencing data.” PhD thesis, Université Paris-Est, 2016.
- [19] **K. Břinda**, “Lossless seeds for approximate string matching.” MSc thesis, Czech Technical University, 2013.
- [20] **K. Břinda**, “Abelian complexity of infinite words.” BSc thesis, Czech Technical University, 2011.

## Software

- RASE** <http://github.com/c2-d2/rase>.  
Tool, pipeline, library and databases (*S. pneumoniae* and *N. gonorrhoeae*) for rapid inference of antibiotic resistance and susceptibility by genomic neighbor typing using nanopore sequencing.
- ProPhyle** <http://prophyle.github.io>.  
Accurate and resource-frugal  $k$ -mer-based metagenomic classification using the Burrows-Wheeler Transform and simplitigs.
- ProphAsm** <http://github.com/prophyle/prophasm>.  
Tool for rapid and memory-efficient computation of simplitigs.

- ProPhex <http://github.com/prophyle/prophex>.  
Tool for rapid  $k$ -mer matching using the Burrows-Wheeler Transform. Co-developed with Kamil Salikhov.
- Ococo <http://github.com/karel-brinda/ococo>.  
Online variant and consensus caller. Call genomic consensus directly from an unsorted SAM/BAM stream.
- RNFtools <http://rnftools.github.io>.  
Generic format for naming simulated DNA reads and the associated toolkit.
- SAMSift <http://github.com/karel-brinda/samsift>.  
Advanced filtering and tagging of SAM/BAM alignments using Python expressions.
- DyMaS <http://github.com/karel-brinda/dymas>.  
Simulator of dynamic mapping (correcting the reference using previously computed alignments).
- Disty <http://github.com/c2-d2/disty>.  
Tool for computing a distance matrix from a core genome alignment.
- NanoSim-H <http://github.com/karel-brinda/nanosim-h>.  
Simulator of nanopore reads (a fork of the NanoSim package).
- SMBL <http://github.com/karel-brinda/smb1>.  
Snakemake bioinformatics library (retired).

## Prizes, Awards, and Honors

- 2019 **Best poster**, Marvin Zelen Data Science Symposium, Dana-Farber Cancer Institute, Boston, USA.
- 2018 **Most innovative poster**, Lake Arrowhead Microbial Genomics Conference (LAMG), Lake Arrowhead, USA.
- 2015 **Travel fellowship of the MSTIC doctoral school**, Université Paris-Est, France.
- 2013 **Offered Scholarship of the French government for PhD “en cotutelle”**, French government’s scholarship program.
- 2013 **MSc (Ing) diploma with honors**, Czech Technical University in Prague.
- 2012 **Fellowship of the RuFiDim organization committee**.
- 2012 **Travel fellowship of the Hlávka Foundation**.

## Teaching

- 2019–2020 **Harvard Medical School, Department of Biomedical Informatics**.  
  - *BMIF01 Concepts in Genome Analysis*, PhD level, TA (1 semester).  
An introductory course of Computational Biology. Recitations, grading, mentoring the students, supervising of the final projects. Close collaboration with four core faculties: Shamil Sunyaev, Michael Baym, Heng Li, and Cheng-Zhong Zhang.
- 2009–2012 **Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering**.  
  - *01ASTE Assistive Technology*, master level, instructor (3 semesters) & co-developed the course.
  - *01SWP Software Project*, master level, supervisor/TA (4 semesters).

## Advising and Supervision

### Master

- 2018 Simone Pignotti, *Species abundance estimation using metagenomic data*, co-supervised with Michael Baym.
- 2016 Mikhail Dubov, *Spaced-seed de Bruijn Graphs*, co-supervised with Gregory Kucherov.

### Dissertation Committees

- 2022–present Khodor Hannoush, comité de suivi individuel

### Others

- 2017–2021 Extensive informal mentoring of PhD students and postdocs in the Baym and Hanage labs at Harvard.

## Media

### Interviews

- 2020/02 **BBC World Service – Science in Action.**  
An interview about our paper on rapid diagnostics of antibiotic resistance using Genomic Neighbor Typing [Link]
- 2020/02 **The Bioinformatics Chat.**  
An interview about our paper on simplifigs for an efficient and scalable representation of de Bruijn graph [Link]

### Other Coverage

- 2020/06 **Massive Science**, *Rapid DNA sequencing of unknown bacteria helps doctors choose which antibiotics to treat it with* [Link].
- 2020/01 **Emory Bacterial Genomics and Metagenomics**, *Seven significant bacterial genomics papers from the 2010s* [Link].

## Professional Activities

### Invited Talks

- 2021 **ICCMG 5**, Fifth International Conference on Clinical Metagenomics. *Rapid inference of antibiotic resistance and susceptibility by genomic neighbor typing.* Geneva, Switzerland (remote).
- 2021 **LIRMM Montpellier**, Methods and Algorithms for Bioinformatics Seminar. *Rapid inference of antibiotic resistance and susceptibility by genomic neighbor typing.* Montpellier, France (remote).
- 2020 **BAARN 2020**, Boston Area Antimicrobial Research Network Symposium. *Rapid inference of antibiotic resistance and susceptibility by genomic neighbor typing.* Boston, USA (remote).
- 2019 **Harvard MSI**, Microbiology Chalk Talk Seminar. *Rapid inference of antibiotic resistance and susceptibility by genomic neighbor typing.* Boston, USA.
- 2019 **MIT CSAIL/MIT Department of Mathematics**, Bioinformatics Seminar. *Lineage calling can identify antibiotic resistant clones within minutes.* Boston, USA.
- 2016 **INRIA/IRISA Rennes, France**, Bioinformatics Seminar. *BWT-based indexing structure for metagenomic classification.* Rennes, France.
- 2015 **Institut Curie**, Systems Biology Seminar. *Spaced seeds improve metagenomic classification.* Paris, France.

### Other Talks

- 2021 **Data Structure in Bioinformatics 2021.** *All microbes on a flash drive.* Milano, Italy (remote).
- 2020 **Human and Biotechnology at TU Graz.** *Algorithms to combat antibiotic resistance.* Graz, Austria (remote).
- 2018 **16th Annual Jonathan Freeman Symposium on the Epidemiology of Infectious Disease.** *Lineage calling can identify antibiotic resistant clones within minutes.* Boston, USA.
- 2017 **Permafrost workshop on bacterial evolution.** *Efficient phylogeny-based metagenomic classification.* Bormio, Italy.
- 2016 **SeqBio 2016.** *ProPhyle: a memory efficient BWT-based metagenomic classifier using k-mer propagation.* Nantes, France.
- 2016 **Quantitative Genomics 2016.** *BWT-based indexing structure for metagenomic classification.* University College London, UK.
- 2016 **Data Structures in Bioinformatics 2016.** *Indexing structures for metagenomic classification.* Bielefeld, Germany.
- 2015 **Algorithmics, Bioinformatics and Statistics for NGS data analysis.** *RNF: a general framework to evaluate NGS read mappers.* Paris, France.
- 2015 **SeqBio 2015.** *Dynamic mappers of NGS reads.* Montpellier, France.
- 2014 **Automata and Formal Languages.** *Languages of lossless seeds.* Szeged, Hungary.
- 2012 **Russian-Finnish Symposium in Discrete Mathematics II.** *On the balance of d-bonacci word.* Turku, Finland.
- 2011 **Analytic and Algebraic Methods in Physics VII.** *Abelian complexity of infinite words.* Prague, Czech Republic.

## Organization and Service

- 2021 **ISMB/ECCB Program Committee.**
- 2014–2016 **Czech Students and Young Professionals Organization in France.**  
Organized events for incoming scholars, in collaboration with the *Embassy of the Czech Republic in Paris.*
- 2011 **CSUN Program Committee.**
- 2010–2013 **Academic Senate of the Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University.**  
Elected student representative.  
Elected vice-chair of the senate.
- 2010 **Ministry of the Interior of the Czech Republic.**  
Served as a member of the Working Group for Accessible Documents.

## Referee for Journal Articles

Nature Biotechnology  
Nature Communications  
Genome Biology  
Bioinformatics  
ASM mSystems  
Microbial Genomics  
PeerJ  
Journal of Bioinformatics and Computational Biology

## Referee for Conference Proceedings

RECOMB – Research in Computational Molecular Biology  
ECCB – European Conference on Computational Biology  
HITSEQ – High Throughput Sequencing  
WABI – Workshop on Algorithms in Bioinformatics  
SPIRE – String Processing and Information Retrieval  
ICCHP – International Conference on Computers Helping People with Special Needs  
WALCOM – International Conference and Workshop on Algorithms and Computation