

# Anna Zhukova

✉ anna.zhukova@pasteur.fr     🌐 Institut Pasteur, 25-28 Rue du Dr Roux, 75015 Paris France  
🔗 research.pasteur.fr/en/member/anna-zhukova     🔗 linkedin.com/in/annazhukova     🔗 github.com/annazhukova

## Research Interests

Phylogenetics: phylodynamics, phylogeography; dynamics of pathogen spreads; mathematical modelling; metabolic modelling; bio-ontologies; algorithms.

## Professional Experience

### 2015- Research Engineer, Institut Pasteur, Paris, FR.

- (permanent)
  - Modelling drug resistance emergence and transmission in HIV-1
  - phylogenetics
    - Spatio-temporal analysis of virus spreads (phylogeography): ZIKV, HIV-1, DENV, Cholera, Sars-CoV-2, Rabies
    - Epidemiological parameter inference from phylogenies (phylodynamics)
    - Development of algorithms and tools for phylogenetics of large pathogen data sets (*Python: numpy/pandas/scipy/ete3; JS; C; R; bash; Snakemake; Docker/Singularity; Slurm*).

### 2014–2015 Postdoc, IBGC/CNRS UMR 5095/U Bordeaux, FR.

- metabolic
  - Constraint-based modelling (FBA, FVA, EFM) of mitochondrial metabolism
  - modelling
    - (*Python: numpy/pandas/COBRApy; SBML*).

### 2011–2014 Doctoral Researcher, Inria/U Bordeaux/CNRS project-team MAGNOME, Talence, FR.

- metabolic
  - Knowledge-based generalization and zoomable visualization of metabolic models
  - modelling
    - (*Python: libSBML; JS: Leaflet.js; SBML; ChEBI; OBO*).

### 2009, 2011 Intern, EMBL-EBI, Functional Genomics & Comput Systems Neurobiol groups, Cambridge, UK.

- bio-
  - Experimental Factor Ontology extension of search in the Array Express Archive
- ontologies
  - Development of the Kinetic Simulation Algorithm Ontology (KiSAO) and libKiSAO (*OWL/OBO, OWL API, Java, Apache Lucene, XML/XSLT, HTML*).

### 2010 Intern, Inria/U Bordeaux/CNRS joint project-team MAGNOME, Talence, FR.

- misc. bioinfo
  - Development of the genome annotation system Magus (*Perl*).

### 2008–2010 Junior software developer, Swift Teams LLC, St. Petersburg, RU.

- software eng.
  - Amazon Web Services Manager plug-in for IntelliJ IDEA (*Java, IntelliJ IDEA, AWS API*)

## Education

### 2011–2014 PhD in Computer Science, University of Bordeaux, FR (advisor: DJ Sherman).

### 2005–2010 BSc+MSc in Software Engineering and Math, St. Petersburg State University, RU (honours).

## Selected Talks [Total: 38 talks, inc. 2 keynotes and 6 invited seminars]

- 3/5/2023 **Phylodynamics: from mathematical modelling to deep learning**, *phyloseminar.org*, inv. seminar.
- 28/1/2020 **Ancestral Reconstruction in Molecular Epidemiology**, *PhyloSIB*, Lausanne, CH.
- 8/9/2018 **Modelling drug resistance spread in HIV**, *5th ESCS @ ECCB 2018*, Athens GR, *keynote*.
- 28/11/2016 **A modeling workflow in systems biol: overview**, *1<sup>st</sup> RSG Luxembourg congress*, Belval LU, *keynote*.
- 5/8/2015 **Multi-level representation of metabolic networks**, *Cheminfo&Metabol gr., EMBL-EBI*, UK, *inv. seminar*.

## Selected Teaching [Total: 27 courses, 207h]

### 11 – 12/2022 **Approches comparatives phylogénétiques** module of Master 2 EvoGEM, Paris, FR, 9h.

### 3/2022 PHINDaccess course: **Evolution & Phylogenetics**, I Pasteur Tunis, TN, 8h.

### 10/2019 **Behind Evolutionary Trees: Math&Algorithms in Phylogenetics**, I Pasteur Montevideo, UY, 10h.

10/2019 **Math & Computational methods for Phylogenetics & Evol Epidemiology**, IPK, Havana, CU, 7h.  
3/2015 **Metabolic modelling** workshop @ research school "Advances in Systems&Syntetic Biol", Strasbourg, FR, 2h.

## Selected Tutoring [Total: 3 PhD students, 6 interns, 2 scientists]

- 9/2024 – PhD: **Vincent Garot** (College de France), co-advising w. Drs L Jacob (LCQB) & S Alizon (CdF).  
... o Epidemiological parameter estimation from viral genomes with deep learning
- 11/2020 – PhD: **Andrew Holtz** (I Pasteur/FIRE), co-advised w. Drs H Bourhy (IP) & G Baele (KU Leuven).  
12/2023 o Exploring factors that drive & sustain virus spread over time and geographic space
- 11/2019 – PhD: **Jakub Voznica** (I Pasteur/ENS/CRI), co-advised with Dr H Morlon (ENS).  
11/2021 o Deep learning from phylogenies to uncover the transmission dynamics of epidemics

## Grants and scholarships

- 2023 – 2028 **DEELOGENY: Machine learning for improved phylogenomic inference** (ANR PRC, partner).
- 2022 travel grant for teaching @ 26th VEME workshop (Panama) (PAHO).
- 2021 travel grant for teaching @ Summer school IFUMI (Montevideo) (IFUMI).
- 2019 – 2020 **The dynamics of HIV-1 epidemics in Cuba** (Campus France, collaborator).
- 2019 – 2020 **Outcomes of fetuses/infants exposed to Zika in utero in Vietnam** (INCEPTION, collaborator).
- 2011 – 2014 **Inria CORDI-S fellowship for doctoral studies** (Inria CORDI-S).

## Software & Ontologies [(co-)author]

- PastML** Ancestral scenario reconstruction and visualisation for phylogenies: [pastml.pasteur.fr](http://pastml.pasteur.fr)
- PhyloDeep** Deep-learning parameter estimation and model selection from phylogenies: [pypi.org/project/phylodeep](https://pypi.org/project/phylodeep)
- PyBDEI** Birth-Death Exposed-Infectious model parameter estimation from phylogenies: [pypi.org/project/pybdei](https://pypi.org/project/pybdei)
- Mimoza** Metabolic model visualization and navigation system: [mimoza.bordeaux.inria.fr](https://mimoza.bordeaux.inria.fr)
- MAMO** Mathematical Modelling Ontology: [combine.org/standards/mamo](https://combine.org/standards/mamo)
- KiSAO** Kinetic Simulation Algorithm Ontology: [combine.org/standards/kisao](https://combine.org/standards/kisao)

## Academic activity

- Talk/poster selection: MCEB'18,'19,'20, AICoB'20, aSSB'15
- Conf./seminar organisation: SARS-CoV-2 session @SMPGD'22, Comp Biol Department seminar @ IP (2022-23)
- Journal reviewer: Systematic Biol, MBE, Viruses, GBE, Wellcome Open Res, PLOS One, etc.
- Funding proposal reviewer: University of Oslo (2018), UKRI Medical Research Council (2021)
- Thesis jury: J Voznica (U Paris Cité'21), P de Boissier (U Aix-Marseille'22), A Holtz (U Paris Cité'23)
- Thesis advisory committee: L Nesterenko (Université Lyon 1, 2022)
- Hiring committee: Bioinformatics & Biostatistics Hub research engineer positions @ Institut Pasteur: preselection for interviews (2016,'17,'19,'23), technical jury member (2017)
- Researcher representative in the Computational Biol. Department council of Institut Pasteur (2020-21)

## Languages

Russian (*native*), English (*fluent*), French (*fluent*), Italian (*basics*)

## Other

2019,2021 In top 3% in **Google Code Jam to I/O for Women**'19 (30<sup>th</sup> out of ~ 1000),'21 (95<sup>th</sup> out of ~ 6500).

**Journal articles** [<sup>f</sup>co-first authors, <sup>c</sup>co-corresponding authors]

1. Morel, M., Zhukova, A., Lemoine, F., & Gascuel, O. (2024). Accurate Detection of Convergent Mutations in Large Protein Alignments with ConDor. *Genome Biology and Evolution*, evae040. doi: 10.1093/gbe/evae040
2. Maestri, R., Perez-Lamarque, B., Zhukova, A., & Morlon, H. (2024). Recent evolutionary origin and localized diversity hotspots of mammalian coronaviruses. *eLife*, 13, RP91745. doi: 10.7554/eLife.91745.1
3. Anjou, C., Lotoux, A., Zhukova, A., Royer, M., Caulat, L. C., Capuzzo, E., Morvan, C., & Martin-Verstraete, I. (2024). The multiplicity of thioredoxin systems meets the specific lifestyles of Clostridia. *PLOS Pathogens*, 20(2), e1012001. doi: 10.1371/journal.ppat.1012001
4. Zhukova<sup>c</sup>, A., Hecht, F., Maday, Y., & Gascuel<sup>c</sup>, O. (2023). Fast and Accurate Maximum-Likelihood Estimation of Multi-Type Birth–Death Epidemiological Models from Phylogenetic Trees. *Systematic Biology*, syad059. doi: 10.1093/sysbio/syad059
5. Holtz<sup>c</sup>, A., Baele, G., Bourhy, H., & Zhukova<sup>c</sup>, A. (2023). Integrating full and partial genome sequences to decipher the global spread of canine rabies virus. *Nature Communications*, 14(1), 4247. doi: 10.1038/s41467-023-39847-x
6. Zhukova<sup>c</sup>, A., Dunn, D., Gascuel<sup>c</sup>, O., & the UK HIV Drug Resistance Database & the Collaborative HIV Anti-HIV Drug Resistance Network. (2023). Modeling Drug Resistance Emergence and Transmission in HIV-1 in the UK. *Viruses*, 15(6), 1244. doi: 10.3390/v15061244
7. Voznica<sup>c</sup>, J., Zhukova<sup>c</sup>, A., Boskova, V., Saulnier, E., Lemoine, F., Moslonka-Lefebvre, M., & Gascuel<sup>c</sup>, O. (2022). Deep learning from phylogenies to uncover the epidemiological dynamics of outbreaks. *Nature Communications*, 13(1), 3896. doi: 10.1038/s41467-022-31511-0
8. Shaikh, B., Smith, L. P., Vasilescu, D., Marupilla, G., Wilson, M., Agmon, E., Agnew, H., Andrews, S. S., Anwar, A., Beber, M. E., Bergmann, F. T., Brooks, D., Brusch, L., Calzone, L., Choi, K., Cooper, J., Detloff, J., Drawert, B., Dumontier, M., Ermentrout, G. B., Faeder, J. R., Freiburger, A. P., Fröhlich, F., Funahashi, A., Garny, A., Gennari, J. H., Gleeson, P., Goelzer, A., Haiman, Z., Hasenauer, J., Hellerstein, J. L., Hermjakob, H., Hoops, S., Ison, J. C., Jahn, D., Jakubowski, H. V., Jordan, R., Kalaš, M., König, M., Liebermeister, W., Sheriff, R. S. M., Mandal, S., McDougal, R., Medley, J. K., Mendes, P., Müller, R., Myers, C. J., Naldi, A., Nguyen, T. V. N., Nickerson, D. P., Olivier, B. G., Patoliya, D., Paulevé, L., Petzold, L. R., Priya, A., Rampadarath, A. K., Rohwer, J. M., Saglam, A. S., Singh, D., Sinha, A., Snoep, J., Sorby, H., Spangler, R., Starruß, J., Thomas, P. J., van Niekerk, D., Weindl, D., Zhang, F., Zhukova, A., Goldberg, A. P., Schaff, J. C., Blinov, M. L., Sauro, H. M., Moraru, I. I., & Karr, J. R. (2022). BioSimulators: a central registry of simulation engines and services for recommending specific tools. *Nucleic Acids Research*, 50(W1), W108-W114. (gkac331) doi: 10.1093/nar/gkac331
9. Seabra<sup>f</sup>, S. G., Libin<sup>f</sup>, P. J. K., Theys<sup>f</sup>, K., Zhukova, A., Potter, B. I., Nebenzahl-Guimaraes, H., Gorbalyena, A. E., Sidorov, I. A., Pimentel, V., Pingarilho, M., de Vasconcelos, A. T. R., Dellicour, S., Khouri, R., Gascuel, O., Vandamme, A.-M., Baele, G., Cuypers, L., & Abecasis, A. B. (2022). Genome-wide diversity of Zika virus: Exploring spatio-temporal dynamics to guide a new nomenclature proposal. *Virus Evolution*, 8(1), veac029. doi: 10.1093/ve/veac029
10. Fillâtre, P., Dufour, M.-J., Behillil, S., Vatan, R., Reusse, P., Gabellec, A., Velmans, N., Montagne, C., Du Coudret, S. G., Droumaguet, E., Merour, V., Enouf, V., Buzele, R., Valence, M., Guillotel, E., Gagniere, B., Baidaliuk, A., Zhukova, A., Tourdjman, M., Thibault, V., Grolier, C., Pronier, C., Lescure, F.-X., Simon-Loriere, E., Costagliola, D., Van Der Werf, S., Tattevin, P., & Massart, N. (2021). A new SARS-CoV-2 variant poorly detected by RT-PCR on nasopharyngeal samples, with high lethality: an observational study. *Clinical Microbiology and Infection*, 28(2), 298.e9-298.e15. doi: 10.1016/j.cmi.2021.09.035

11. Bennedbæk, M., Zhukova, A., Tang, M.-H. E., Bennet, J., Munderi, P., Ruxrungtham, K., Gisslen, M., Worobey, M., Lundgren, J. D., & Marvig, R. L. (2021). Phylogenetic Analysis of HIV-1 Shows Frequent Cross-Country Transmission and Local Population Expansions. *Virus Evolution*, 7(2), veab055. doi: 10.1093/ve/veab055
12. Grant, R., Nguyen, T. T. T., Dao, M. H., Pham, H. T. T., Piorkowski, G., Pham, T. D. T., Cao, T. M., Huynh, L. T. K., Nguyen, Q. H., Vien, L. D. K., Lemoine, F., Zhukova, A., Hoang, D. T. N., Nguyen, H. T., Tran, N. T., Le, L. B., Ngo, M. N. Q., Tran, T. C., Le, N. N. T., Nguyen, M. N., Pham, H. T., Hoang, T. T. D., Dang, T. V., Vu, A. T., Nguyen, Q. N. T., de Lamballerie, X., Pham, Q. D., Luong, Q. C., & Fontanet, A. (2021). Maternal and neonatal outcomes related to Zika virus in pregnant women in Southern Vietnam: An epidemiological and virological prospective analysis. *The Lancet Regional Health - Western Pacific*, 11, 100163. doi: 10.1016/j.lanwpc.2021.100163
13. Zhukova<sup>c</sup>, A., Voznica, J., Felipe, M. D., To, T.-H., Pérez, L., Martínez, Y., Pintos, Y., Méndez, M., Gascuel<sup>c</sup>, O., & Kouri<sup>c</sup>, V. (2021). Cuban history of CRF19 recombinant subtype of HIV-1. *PLOS Pathogens*, 17(8), e1009786. doi: 10.1371/journal.ppat.1009786
14. Oprea<sup>f</sup>, M., Njamkepo<sup>f</sup>, E., Cristea<sup>f</sup>, D., Zhukova<sup>f</sup>, A., Clark, C. G., Kravetz, A. N., Monakhova, E., Ciontea, A. S., Cojocaru, R., Rauzier, J., Damian, M., Gascuel, O., Marie-Laure, Q., & Weill, F.-X. (2020). The seventh pandemic of cholera in Europe revisited by microbial genomics. *Nature Communications*, 11(1), 5347. doi: 10.1038/s41467-020-19185-y
15. Keating, S. M., Waltemath, D., König, M., Zhang, F., Dräger, A., Chaouiya, C., Bergmann, F. T., Finney, A., Gillespie, C. S., Helikar, T., Hoops, S., Malik-Sheriff, R. S., Moodie, S. L., Moraru, I. I., Myers, C. J., Naldi, A., Olivier, B. G., Sahle, S., Schaff, J. C., Smith, L. P., Swat, M. J., Thieffry, D., Watanabe, L., Wilkinson, D. J., Blinov, M. L., Begley, K., Faeder, J. R., Gómez, H. F., Hamm, T. M., Inagaki, Y., Liebermeister, W., Lister, A. L., Lucio, D., Mjolsness, E., Proctor, C. J., Raman, K., Rodriguez, N., Shaffer, C. A., Shapiro, B. E., Stelling, J., Swainston, N., Tanimura, N., Wagner, J., Meier-Schellersheim, M., Sauro, H. M., Palsson, B., Bolouri, H., Kitano, H., Funahashi, A., Hermjakob, H., Doyle, J. C., Hucka, M., & SBML Level 3 Community members. (2020). Sbml level 3: an extensible format for the exchange and reuse of biological models. *Molecular Systems Biology*, 16(8), e9110. doi: 10.15252/msb.20199110
16. Ishikawa<sup>f</sup>, S. A., Zhukova<sup>f</sup>, A., Iwasaki, W., & Gascuel, O. (2019). A fast likelihood method to reconstruct and visualize ancestral scenarios. *Molecular biology and evolution*, 36(9), 2069–2085. doi: 10.1093/molbev/msz131
17. Zhukova, A., Fernandes, L. G., Hugon, P., Pappas, C. J., Sismeiro, O., Coppée, J.-Y., Becavin, C., Malabat, C., Eshghi, A., Zhang, J.-J., Yang, F. X., Picardeau, M., Coppée, J.-Y., Becavin, C., Malabat, C., Eshghi, A., Zhang, J.-J., Yang, F. X., Picardeau, M., & Others. (2017). Genome-wide transcriptional start site mapping and sRNA identification in the pathogen *Leptospira* interro-gans. *Frontiers in cellular and infection microbiology*, 7, 10. doi: 10.3389/fcimb.2017.00010
18. Waltemath, D., Karr, J. R., Bergmann, F. T., Chelliah, V., Hucka, M., Krantz, M., Liebermeister, W., Mendes, P., Myers, C. J., Pir, P., Alaybeyoglu, B., Aranganathan, N. K., Baghalian, K., Bittig, A. T., Burke, P. E. P., Cantarelli, M., Chew, Y. H., Costa, R. S., Cursons, J., Czauderna, T., Goldberg, A. P., Gomez, H. F., Hahn, J., Hameri, T., Gardiol, D. F. H., Kazakiewicz, D., Kiselev, I., Knight-Schrijver, V., Knupfer, C., Konig, M., Lee, D., Lloret-Villas, A., Mandrik, N., Medley, J. K., Moreau, B., Naderi-Meshkin, H., Palaniappan, S. K., Priego-Espinosa, D., Scharf, M., Sharma, M., Smallbone, K., Stanford, N. J., Song, J.-H., Theile, T., Tokic, M., Tomar, N., Toure, V., Uhlendorf, J., Varusai, T. M., Watanabe, L. H., Wendland, F., Wolfien, M., Yurkovich, J. T., Zhu, Y., Zardilis, A., Zhukova, A., & Schreiber, F. (2016). Toward Community Standards and Software for Whole-Cell Modeling. *IEEE Transactions on Biomedical Engineering*, 63(10), 2007–2014. doi: 10.1109/TBME.2016.2560762
19. Loira, N., Zhukova, A., & Sherman, D. J. (2015). Pantograph: A template-based method for genome-scale metabolic model reconstruction. *Journal of bioinformatics and computational biology*, 13(02), 1550006. doi: 10.1142/S0219720015500067

20. Zhukova<sup>c</sup>, A., & Sherman, D. J. (2015). Mimoza: web-based semantic zooming and navigation in metabolic networks. *BMC Systems Biology*, 9(1), 10. doi: 10.1186/s12918-015-0151-5
21. Zhukova<sup>c</sup>, A., & Sherman, D. J. (2014a). Knowledge-based generalization of metabolic models. *J Comp Biol.*, 21(7), 534–47. doi: 10.1089/cmb.2013.0143
22. Zhukova<sup>c</sup>, A., & Sherman, D. J. (2014b). Knowledge-based generalization of metabolic networks: a practical study. *Journal of Bioinformatics and Computational Biology*, 12(2), 1441001. doi: 10.1142/S0219720014410017
23. Zhukova, A., Adams, R., Laibe, C., & Le Novère, N. (2012). LibKiSAO: a Java Library for Querying KiSAO. *BMC Research Notes*, 5(1), 520+. doi: 10.1186/1756-0500-5-520
24. Courtot<sup>f</sup>, M., Juty<sup>f</sup>, N., Knüpfer<sup>f</sup>, C., Waltemath<sup>f</sup>, D., Zhukova<sup>f</sup>, A., Dräger, A., Dumontier, M., Finney, A., Golebiewski, M., Hastings, J., Hoops, S., Keating, S., Kell, D. B., Kerrien, S., Lawson, J., Lister, A., Lu, J., Machne, R., Mendes, P., Pocock, M., Rodriguez, N., Villeger, A., Wilkinson, D. J., Wimalaratne, S., Laibe, C., Hucka, M., & Le Novère, N. (2011). Controlled vocabularies and semantics in systems biology. *Molecular systems biology*, 7(1), 543. doi: 10.1038/msb.2011.77
25. Malone, J., Holloway, E., Adamusiak, T., Kapushesky, M., Zheng, J., Kolesnikov, N., Zhukova, A., Brazma, A., & Parkinson, H. (2010). Modeling sample variables with an Experimental Factor Ontology. *Bioinformatics*, 26(8), 1112–1118. doi: 10.1093/bioinformatics/btq099

### Review articles [<sup>f</sup> co-first authors]

26. Bassel<sup>f</sup>, L., Zhukova<sup>f</sup>, A., Villabona-Arenas, C. J., Atkins, K. E., Hué, S., & Gascuel, O. (2021). Drug resistance mutations in HIV: new bioinformatics approaches and challenges. *Current Opinion in Virology*, 51, 56–64. doi: 10.1016/j.coviro.2021.09.009
27. Zhukova, A., Bassel, L., Lemoine, F., Morel, M., Voznica, J., & Gascuel, O. (2020). Origin, evolution and global spread of SARS-CoV-2. *Comptes Rendus. Biologies*, 343(0), 1–20. doi: 10.5802/crbiol.29
28. Zhukova, A., Cutino-Moguel, T., Gascuel, O., & Pillay, D. (2017). The Role of Phylogenetics as a Tool to Predict the Spread of Resistance. *The Journal of Infectious Diseases*, 216(suppl\_9), S820–S823. doi: 10.1093/infdis/jix411

### Conference proceeding articles

29. Zhukova, A., & Mazat, J.-P. (2015). The mitochondrion metabolic model iAS253 revisited. *advances in Systems and Synthetic Biology*, 179–188.
30. Zhukova, A., & Mazin, M. (2010). Actor extension of Java language in MPS Environment. *Sci Tech J Inf Technol Mech Opt*, 10(2), 72–77.

### Book chapters

31. Zhukova, A., Gascuel, O., Duchene, S., Ayres, D., Lemey, P., & Baele, G. (2020). Efficiently Analysing Large Viral Data Sets in Computational Phylogenomics. In C. Scornavacca, F. Delsuc, & N. Galtier (Eds.), *Phylogenetics in the genomic era* (pp. 5.3:1–5.3:43). No commercial publisher — Authors open access book.