

# JOELLE MBATCHOU PH.D.

- Proficient and innovative researcher with 9 years of research experience in developing cutting-edge methods to help improve understanding of genetic variation and its relationship to human health and disease
- Avid and charismatic learner who continuously strives for excellence by always meeting deadlines, providing timely solutions, and staying attentive to new developments in the field
- Eager to assist biotech companies in discovering novel targets and indications by building scalable and efficient association analysis tools and pipelines
- Developed a C++ tool which decreased computation costs at the Regeneron Genetics Center by over 50X
- Top Skills: Analytical, Highly Proficient in C++ and R, Pipeline development with WDL & docker, Project Management



## EDUCATION

- 2019
- **PhD., Statistics**  
University of Chicago 📍 Chicago, IL
- 2011
- **BSc Biology & Mathematical Sciences**  
DePaul University 📍 Chicago, IL

## INDUSTRY EXPERIENCE

- Current | 2022
- **Manager, Statistical Genetics**  
Regeneron Genetics Center 📍 Tarrytown, New York
    - Routinely carry out statistical analyses using cloud-based computing platforms on large-scale and high-dimensional human genetics datasets containing millions of genetic variants and 100,000s of individuals
    - Develop statistical methods and computational tools geared for large-scale genetic and genomics studies
    - Build WDL pipelines for data sets with 100,000s of individuals from whole-exome and whole-genome sequencing data
  - **Senior Statistical Geneticist**  
Regeneron Genetics Center 📍 Tarrytown, New York
    - Developed a computational efficient whole genome regression method REGENIE for large-scale genetic association analyses which can be more than 100x faster than current state-of-the-art methods and can handle population structure and imbalanced binary traits
    - Implemented REGENIE into a C++ software which was publicly released on Github
    - Published the REGENIE method as first author in Nature Genetics where it was applied to UK Biobank data (>100 phenotypes, >400K individuals and >10M genetic variants).

## CONTACT

- 📍 Chicago, IL
- 👤 US Citizen
- ✉️ joelle.mbatchou@gmail.com
- 📞 (773)599-2825
- 🐦 joellembatchou
- 🔗 linkedin.com/in/jmbatchou
- 🔗 github.com/joellesophya
- 🔗 joellesophya.github.io

## TECHNICAL SKILLS



Last updated on 2022-08-09.

2019  
|  
2013

## 💻 RESEARCH EXPERIENCE

- **Graduate Student Researcher**  
Department of Statistics 📍 University of Chicago
  - Developed a computationally fast method JASPER to assess significance for a general class of association tests, including tests for high dimensional phenotypes and gene-based tests, adjusting for population structure and family relatedness
  - Designed a permutation-based testing procedure BRASS for assessing significance with binary traits in structured samples for association tests with unknown exact/asymptotic distributions.
  - Built C/C++ software to evaluate JASPER and BRASS through simulation studies & real data applications.

2022

## 👤 TEACHING EXPERIENCE

- **Instructor**  
Summer Institute in Statistical Genetics 📍 University of Washington
  - Taught the association mapping module on genome-wide association studies and sequencing (120 students)
  - Designed coursework as well as hands-on practical exercises using software such as PLINK, RGENIE and R packages GWASTools and bigsnpr.
  - Built a website to host the course materials using workflowr R package.
- **Teaching Assistant**  
Department of Statistics 📍 University of Chicago
  - Assisted in undergraduate courses: Statistical Methods and Applications, Statistical Models/Methods, Applied Regression Analysis and Analysis of Categorical Data
  - Created introductory material for R and STATA through weekly computer sessions
  - Organized weekly office hours where students expanded their understanding of course materials
- **Course Instructor**  
Department of Statistics 📍 University of Chicago
  - Taught an introductory course on statistical methods (STAT 234: Statistical Models/Methods).
  - Designed course materials and delivered bi-weekly lectures to a class of 36 students.
  - Assisted students in office hours with learning statistical concepts as well as R programming skills.

2019  
|  
2012

2018

2017  
|  
2016

- **Statistics Collaborative Learning Team Leader**  
Department of Statistics 📍 University of Chicago
  - Facilitated the design of a new program based on collaborative learning.
  - Designed practical exercises based on class materials and supervised weekly hands-on workshops for students.

## 🏆 AWARDS

- 2022 ● Selected as one of 35 innovators under 35<sup>1</sup>  
MIT Technology Review
- 2022 ● Selected as one of 17 Rising Stars in Health Tech<sup>2</sup>  
Fierce Health Care
- 2021 ● Selected for Reviewers' Choice (top 10% scoring abstracts)  
American Society of Human Genetics
- 2013 ● Department of Education GAANN Fellowship Recipient  
University of Chicago
- 2010 ● Departmental Award for Outstanding Performance in Organic Chemistry  
DePaul University

## ☰ PUBLICATIONS

- 2022 ● Genome-wide analysis provides genetic evidence that ACE2 influences COVID-19 risk and yields risk scores associated with severe disease.<sup>3</sup>  
Nature Genetics
  - Authors: Horowitz, J. E., Kosmicki, J. A., Damask, A., Sharma, D., Roberts, G. H. L., Justice, A. E., Banerjee, N., Coignet, M. V., Yadav, A., Leader, J. B., Marcketta, A., Park, D. S., Lanche, R., Maxwell, E., Knight, S. C., Bai, X., Guturu, H., Sun, D., Baltzell, A., Kury, F. S. P., Backman, J. D., Girshick, A. R., O'Dushlaine, C., McCurdy, S. R., Partha, R., Mansfield, A. J., Turissini, D. A., Li, A. H., Zhang, M., **Mbatchou, J.**, et al.
- 2022 ● Germline Mutations in CIDEB and Protection against Liver Disease.<sup>4</sup>  
New England Journal of Medicine
  - Authors: Verweij, N., Haas, M. E., Nielsen, J. B., Sosina, O. A., Kim, M., Akbari, P., De, T., Hindy, G., Bovijn, J., Persaud, T., Miloscio, L., Germino, M., Panagis, L., Watanabe, K., **Mbatchou, J.**, et al.

- 2021
- **Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity.<sup>5</sup>**  
Science  
• Authors: Akbari, P., Gilani, A., Sosina, O., Kosmicki, J. A., Khrimian, L., Fang, Y. Y., Persaud, T., Garcia, V., Sun, D., Li, A., **Mbatchou, J.**, et al.
- 2021
- **Exome sequencing and analysis of 454,787 UK Biobank participants.<sup>6</sup>**  
Nature  
• Authors: Backman, J. D., Li, A. H., Marcketta, A., Sun, D., **Mbatchou, J.**, et al.
- 2021
- **Pan-ancestry exome-wide association analyses of COVID-19 outcomes in 586,157 individuals.<sup>7</sup>**  
American Journal of Human Genetics  
• Authors: Kosmicki, J. A., Horowitz, J. E., Banerjee, N., Lanche, R., Marcketta, A., Maxwell, E., Bai, X., Sun, D., Backman, J. D., Sharma, D., Kury, F. S. P., Kang, H. M., O'Dushlaine, C., Yadav, A., Mansfield, A. J., Li, A. H., Watanabe, K., Gurski, L., McCarthy, S. E., Locke, A. E., Khalid, S., O'Keeffe, S., **Mbatchou, J.**, et al.
- 2021
- **Computationally efficient whole-genome regression for quantitative and binary traits.<sup>8</sup>**  
Nature Genetics  
• Authors: **Mbatchou, J.**, Barnard, L., Backman, J., Marcketta, A., Kosmicki, J. A., Ziyatdinov, et al.
- 2015
- **Retrospective Association Analysis of Binary Traits: Overcoming Some Limitations of the Additive Polygenic Model.<sup>9</sup>**  
Human Heredity  
• Authors: Jiang, D., **Mbatchou, J.** & McPeek, M. S.
- 2011
- **New insights into relationships of lichen-forming Dothideomycetes.**  
Fungal Diversity  
• Authors: Nelsen, M. P., Lücking, R., **Mbatchou, J. S.**, Andrew, C. J., Spielmann, A. A. & Lumbsch, H. T.
- 2010
- **Heiomasia, a new genus in the lichen-forming family Graphidaceae (Ascomycota: Lecanoromycetes: Ostropales) with disjunct distribution in Southeastern North America and Southeast Asia.**  
The Bryologist  
• Authors: Nelsen, M. P., Lücking, R., Plata, E. R. & **Mbatchou, J. S.**

2009

- **A class-wide phylogenetic assessment of Dothideomycetes.**

Studies in Mycology

- Authors: Schoch, C. L., Crous, P. W., Groenewald, J. Z., Boehm, E. W., Burgess, T. I., Gruyter, J. de, Hoog, G. S. de, Dixon, L. J., Grube, M., Gueidan, C., Harada, Y., Hatakeyama, S., Hirayama, K., Hosoya, T., Huhndorf, S. M., Hyde, K. D., Jones, E. B., Kohlmeyer, J., Kruys, A., Li, Y. M., Lucking, R., Lumbsch, H. T., Marvanova, L., **Mbatchou, J. S.**, et al.

## PROFESSIONAL ACTIVITIES

- **Manuscript reviewer**

Genetic Epidemiology, Bioinformatics.

## ORAL PRESENTATIONS AND POSTERS

2022

- **Computationally efficient whole-genome regression framework for association analyses in large-scale biobanks.**

International Common Disease Alliance

- Virtual Scientific Plenary

2021

- **Gene-burden tests, gene-environment interactions and time-to-event data analysis within an efficient whole genome regression framework for large-scale biobanks.**

The American Society of Human Genetics Annual Meeting

- Poster presentation

2021

- **Fast and robust methods to detect gene-environment interactions in large-scale biobanks.**

International Genetic Epidemiology Society Conference

- Lightning talk

2020

- **Computationally efficient whole-genome regression for quantitative and binary traits.**

The American Society of Human Genetics Annual Meeting

- Platform talk

2017

- **A generalized permutation testing method for binary trait association in structured samples.**

The American Society of Human Genetics Annual Meeting

- Poster presentation

2017

- **Genetic association analysis of binary traits in structured samples.**

Eastern North American Region Spring Meeting

- Poster presentation

## LINKS

- 1: <https://www.technologyreview.com/innovator/joelle-mbatchou/>
- 2: <https://www.fiercehealthcare.com/health-tech/rising-stars-health-tech>
- 3: <https://doi.org/10.1038/s41588-021-01006-7>
- 4: <https://doi.org/10.1056/NEJMoa2117872>
- 5: <https://doi.org/10.1126/science.abf8683>
- 6: <https://doi.org/10.1038/s41586-021-04103-z>
- 7: <https://doi.org/10.1016/j.ajhg.2021.05.017>
- 8: <https://doi.org/10.1038/s41588-021-00870-7>
- 9: <https://doi.org/10.1159/000446957>