JOELLE MBATCHOU PH.D.

- Proficient and innovative researcher with 10+ 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
|
2024

Senior Manager, Statistical Genetics
Regeneron Genetics Center
Tarrytown, New York

- Leading and mentoring a team member in the development of an analytical pipeline for time-to-event data.
- Develop machine learning methods to integrate functional annotations into rare variant association analyses using exome sequencing data.

2023
|
2022

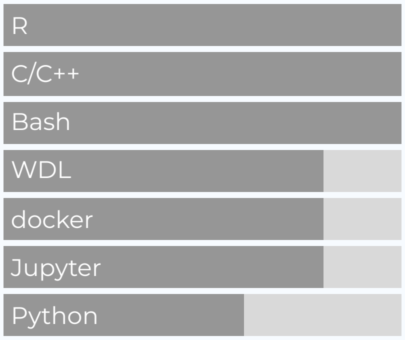
Manager, Statistical Genetics
Regeneron Genetics Center
Tarrytown, New York

- 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

CONTACT

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

TECHNICAL SKILLS



Last updated on 2025-06-05.

2021
|
2019

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).

RESEARCH EXPERIENCE

2019
|
2013

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.

TEACHING EXPERIENCE

Current
|
2022

Instructor

Summer Institute in Statistical Genetics

University of Washington/Georgia Institute of Technology

- 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, REGENIE and R packages GWASTools and bigsnpr.
- Built a website to host the course materials using workflowr R package.

2019
|
2012

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

2018

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.

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 mateirals and supervised weekly hands-on workshops for students.

AWARDS

2022

Selected as one of 35 innovators under 35[1]

MIT Technology Review

2022

Selected as one of 17 Rising Stars in Health Tech[2]

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

2024

JASPER: fast, powerful, multitrait association testing in structured samples gives insight on pleiotropy in gene expression[3]

American Journal of Human Genetics

- Authors: Mbatchou, J. & McPeek, M. S.

2024

Joint testing of rare variant burden scores using non-negative least squares[4]

American Journal of Human Genetics

- Authors: Ziyatdinov, A., Mbatchou, J., Marcketta, A., Backman, J., Gaynor, S., et al.

2024

Genetic risk factors for COVID-19 and influenza are largely distinct[5]

Nature Genetics

- Authors: Kosmicki, J. A., Marcketta, A., Sharma, D., Di Gioia, S. A., Batista, S., Yang, X.-M., Tzoneva, G., Martinez, H., Sidore, C., Kessler, M. D., Horowitz, J. E., Roberts, G. H. L., Justice, A. E., Banerjee, N., Coignet, M. V., Leader, J. B., Park, D. S., Lanche, R., Maxwell, E., Knight, S. C., Bai, X., Guturu, H., Baltzell, A., Girshick, A. R., McCurdy, S. R., Partha, R., Mansfield, A. J., Turissini, D. A., Zhang, M., Mbatchou, J., et al.

2023

Genotyping, sequencing and analysis of 140,000 adults from Mexico City[6]

Nature

- Authors: Ziyatdinov, A., Torres, J., Alegre-Díaz, J., Backman, J., Mbatchou, J., Turner, M., et al.

2023

BRASS: Permutation methods for binary traits in genetic association studies with structured samples[7]

PLoS Genetics

- Authors: Mbatchou, J., Abney, M., & McPeek, M. S.

2023

Rare coding variants in CHRNB2 reduce the likelihood of smoking[8]

Nature Genetics

- Authors: Rajagopal, V. M., Watanabe, K., Mbatchou, J., Ayer, A., Quon, P., et al.

2022

Common and rare variant associations with clonal haematopoiesis phenotypes.[9]

Nature

- Authors: Kessler, M. D., Damask, A., O'Keeffe, S., Banerjee, N., Li, D., Watanabe, K., Marketta, A., Van Meter, M., Semrau, S., Horowitz, J., Tang, J., Kosmicki, J. A., Rajagopal, V. M., Zou, Y., Houvras, Y., Ghosh, A., Gillies, C., Mbatchou, J., et al.

2022

Genome-wide analysis provides genetic evidence that ACE2 influences COVID-19 risk and yields risk scores associated with severe disease.[10]

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.[11]

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.[12]

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.[13]

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.[14]

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. [15]

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.[16]

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
Nature Genetics, Genetic Epidemiology, Bioinformatics.

ORAL PRESENTATIONS AND POSTERS

2023 **Using protein language model annotations to improve the power of exome-wide association studies**
The American Society of Human Genetics Annual Meeting
- Poster presentation

2023 **Powerful gene-based testing for exome-wide association studies**
University of Washington Statistical Genetics Symposium
- Invited Talk

2022 **REGENIE v3: more efficient analysis of rare genetic variation with an extended set of gene-based tests**
The American Society of Human Genetics Annual Meeting
- Poster presentation

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/10.1016/j.ajhg.2024.06.010
4: https://doi.org/10.1016/j.ajhg.2024.08.021
5: https://doi.org/10.1038/s41588-024-01844-1
6: https://doi.org/10.1038/s41586-023-06595-3
7: https://doi.org/10.1371/journal.pgen.1011020
8: https://doi.org/10.1038/s41588-023-01417-8
9: https://doi.org/10.1038/s41586-022-05448-9
10: https://doi.org/10.1038/s41588-021-01006-7
11: https://doi.org/10.1056/NEJMoa2117872
12: https://doi.org/10.1126/science.abf8683
13: https://doi.org/10.1038/s41586-021-04103-z
14: https://doi.org/10.1016/j.ajhg.2021.05.017
15: https://doi.org/10.1038/s41588-021-00870-7
16: https://doi.org/10.1159/000446957