




# Fan Zhang, PhD

 <https://fanzhangharvard.com/>  
 <https://github.com/fzhangcode>

 [fanzhang@broadinstitute.org](mailto:fanzhang@broadinstitute.org)  [google scholar](#)  
 77 Avenue Louis Pasteur, Boston, MA 02115

## Employment

<b>Harvard Medical School (HMS)</b> Research Scientist (Immunogenomics Lab, Dr. Soumya Raychaudhuri)	US 2017 – Present
<b>Broad Institute of MIT and Harvard</b> Associate Scientist	2017 – Present
<b>Brigham and Women's Hospital (BWH) Single Cell Genomics Core</b> Computational Biologist	2019 – Present
<b>Chinese Academy of Sciences</b> Research Associate	China 2012 – 2013

## Education

<b>Ph.D. Biomedical Engineering and Bioinformatics</b> , Worcester Polytechnic Institute (WPI), MA, US Dissertation: Statistical methods for characterizing genomic heterogeneity in mixed samples Advisor: Dr. Patrick Flaherty (now in Statistics Dep. UMass Amherst)	2013 – 2017
<b>M.S. Bioinformatics</b> (Full Fellowship), Jilin University, China	2009 – 2012
<b>B.S. Computer Science</b> (Ranking Top 1%), Jilin University, China	2005 – 2009

## Selected Publications

- Zhang, F.**, Mears, J.R., Shakib, L., Beynor, J.I., Shanaj, S., Korsunsky, I., Nathan, A., the Accelerating Medicines Partnership RA/SLE, Donlin, L.T.\*, Raychaudhuri, S\*. [IFN- \$\gamma\$  and TNF- \$\alpha\$  drive a CXCL10+ CCL2+ macrophage phenotype expanded in severe COVID-19 and other diseases with tissue inflammation](#). *bioRxiv*, 2020
- Fan, J., Slowikowski, K., **Zhang, F.** Single-cell transcriptomics in cancer: computational challenges and opportunities. *Nature Experimental & Molecular Medicine*, In press, 2020
- Orange, D., Yao, V., Sawicka, K., Jak, J., Frank, M.O., Parveen, S., Blachere, N.E., Hale, C., **Zhang, F.**, Raychaudhuri, S., Troyanskaya, O.G., Darnell, R.B. [RNA Identification of PRIME Cells Predicting Rheumatoid Arthritis Flares](#). *New England Journal of Medicine (NEJM)*, 383(3), pp.218-228, 2020
- Zhang, F.\***, Wei, K.\*, Slowikowski, K.\*, Fonseka, C.Y.\*, Rao, D.A.\*, Kelly, S., Goodman, S.M., Tabechian, D., Hughes, L.B., Salomon-Escoto, K., Watts, G.F.M., Jonsson, A.H., Rangel-Moreno, J., Pellett, N.M., Rozo, C., Aprezzese, W., Eisenhaure, T.M., Lieb, D., Boyle, D.L., Mandelin A.M., the Accelerating Medicines Partnership Rheumatoid Arthritis and Systemic Lupus Erythematosus (AMP RA/SLE), Boyce, B.F., DiCarlo, E., Gravallese, E.M., Gregersen, P.K., Moreland, L., Firestein, G.S., Hachohen, N., Nusbaum, C., Lederer, J.A., Perlman, H., Pitzalis, C., Filer, A., Holers, M.V., Bykerk, V.P., Donlin, L.T.\*, Anolik, J.H.\*, Brenner, M.B.\*, Raychaudhuri, S\*. [Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry](#). *Nature Immunology*, 20(7), pp.928-942, 2019
- Zhang, F.**, Wang, C., Trapp, A., and Flaherty, P. [A Global Optimization Algorithm for Sparse Mixed Membership Matrix Factorization](#). In *Contemporary Biostatistics with Biopharmaceutical Applications* (pp. 129-156), ICSA Book Series in Statistics. Springer, Cham, 2019
- Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., **Zhang, F.**, Wei, K., ..., Raychaudhuri, S. [Fast, sensitive and accurate integration of single-cell data with Harmony](#). *Nature Methods*, pp.1-8, 2019
- Kuo, D.\*, Ding J.\*, Cohn, I., **Zhang, F.**, Wei, K., Rao, D., Rozo, C., Sokhi U.K., Shanaj, S., Oliver, D., Echeverria, A.P., DiCarlo E.F., Brenner, M.B., Bykerk, V.P., Goodman, S.M., Raychaudhuri, S., Ratsch, G., Ivashkiv, L.B., Donlin, L.T. [HBEGF+ macrophages in rheumatoid arthritis induce fibroblast invasiveness](#). *Science Translational Medicine*, 11(491), 2019
- Arazi, A.\*, Rao, D.A.\*, Berthier, C.C.\*, Davidson, A., Liu, Y., Hoover, P.J., Chicoine, A., Eisenhaure, T.M., Jonsson, A.H., Li, S., Lieb, D.J., **Zhang, F.**, Slowikowski K., Browne, E.P., Noma, A., Sutherby, D., Steelman, S., Smilek, D.E., Tosta, P., Apruzzese, W., Massarotti, E., Dall, Era, M., Park, M., Kamen, D.L., Furie, R.A., Payan- Schober, F., Pendergraft WF 3rd, McInnis E.A., Buyon, J.P., Petri, M.A., Putterman, C.,

- Kalunian, L.C., Woodle, E.S., Lederer, J.A., Hildeman, D.A., Nusbaum, C., Raychaudhuri S., Kretzler, M., Anolik, J.H., Brenner, M.B., Wofsy D., Hachohen, N.\*, Diamond, B.\*, the AMP RA/SLE Consortium. [The immune cell landscape in kidneys of patients with lupus nephritis](#). *Nature Immunology*, 20(7), pp.902-914, 2019
9. **Zhang, F.** and Flaherty, P. [Variational inference for rare variant detection in deep, heterogeneous next-generation sequencing data](#). *BMC Bioinformatics*, 18(1), 45, 2017
  10. He, Y., **Zhang, F.**, and Flaherty, P. [RVD2: an ultra-sensitive variant detection model for low-depth heterogeneous next-generation sequencing data](#). *Bioinformatics*, 31(17), 2785-2793, 2015
  11. **Zhang, F.**, Chen, S., Zhang, H., Zhang, X., and Li, G. [Bioelectric signal detrending using smoothness prior approach](#). *Medical Engineering & Physics*, 36(8), 1007-1013, 2014
  12. Geng, Y., **Zhang, F.**, Yang, L., Zhang, Y., and Li, G. [Reduction of the effect of arm position variation on real-time performance of motion classification](#). In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (pp. 2772-2775), 2012
  13. Liang, Y., **Zhang, F.**, Wang, J., Joshi, T., Wang, Y., and Xu, D. [Prediction of drought-resistant genes in Arabidopsis thaliana using SVM-RFE](#). *PLoS one*, 6(7), e21750, 2011
  14. Wang, J., **Zhang, F.**, Wang, Y., Fu, Y., Xu, D., and Liang, Y. Identification of Salt Tolerance Genes in Rice from Microarray Data using SVM-RFE. *The 3rd International Conference on Bioinformatics and Computational Biology (BICoB)*, New Orleans, pp. 30-35, 2011

### In Review Articles

15. Mysore, V., Cullere, X., Mears, J., Rosetti, F., Okubo, K., Liew, P.X., **Zhang, F.**, Madera-Salcedo, I., Rosenbauer, F., Stone, R.M., Aster, J.A., Andrian, U.V., Lichtman, A., Raychaudhuri, S., Mayadas, T.N. FcγR engagement reprograms neutrophils into antigen cross-presenting cells that elicit acquired immunity, In Review, 2020

### Software/Toolkit

---

1. **Toolkit** (R): An integrative pipeline to integrate single-cell RNA-seq, bulk RNA-seq, and mass cytometry data (Zhang, *et al.* 2019)
2. **GOP** (Python): A global optimization model to identify global optimum for a sparse mixed membership matrix factorization for heterogeneous gene expression samples (Zhang, *et al.* 2019)
3. **RVD/RVD2** (Python): A variational Bayesian model to identify rare variants in heterogeneous NGS clinical samples (Zhang, *et al.* 2017)
4. **RA/SLE websites** (R): single-cell data visualization (Zhang, *et al.* 2019; Arazi, *et al.* 2019)

### Invited Talks

---

1. **Invited Talk. Research Methods Workshop**, Center for Skeletal Research at **Massachusetts General Hospital**, Boston, MA, May, 2020  
Integration of single-cell transcriptomic and proteomic data to study autoimmune diseases
2. **Invited Talk. American College of Rheumatology (ACR)**, Atlanta, GA, Nov. 2019  
System-level multi-modality data integration for single-cell transcriptomics
3. **Oral Talk. American College of Rheumatology (ACR)**, Atlanta, GA, Nov. 2019  
Integration of single cells from inflammatory disease tissues reveals common and unique pathogenic cell states
4. **Oral Talk. The AMP RA/SLE Face to Face**, Washington DC, Sep. 2019  
Integration of single cells from inflammatory disease tissues reveals common and unique pathogenic cell states
5. **Oral Talk. Single Cell Biology Keystone Symposia**, Breckenridge, CA, Jan. 2019  
Integrating single-cell transcriptomics and mass cytometry to define cell states in rheumatoid arthritis
6. **Keynote Speak. IV European Conference “Genomics of Complex Diseases: New Challenges”**. Pfizer-University of Granada-Junta de Andalucía Center for Genomics and Oncological Research (GEYNO), Granada, Spain. Oct. 2018  
Functional genetics and genomics to define rheumatoid arthritis

7. **Invited Talk. *Federation of Clinical Immunology Societies (FOCIS)***, San Francisco, CA. June 2018.  
Defining inflammatory cell states in rheumatoid arthritis joint synovia tissues by single-cell technology
8. **Invited Talk. *American College of Rheumatology (ACR)***, San Diego, CA. Nov. 2017.  
Technology in precision medicine: data analysis and integration of high dimensional data from rheumatoid arthritis joint tissue
9. **Oral Talk. *Immunogenomics***, HudsonAlpha Institute for Biotechnology, Huntsville, AL. Oct. 2017.  
A novel computational method to resolve cellular heterogeneity in disease tissues
10. **Oral Talk. Accelerating Medicines Partnership (AMP) RA/SLE Face-to-Face**, Houston, TX, Feb. 2017.  
Disease relevant signatures identification in rheumatoid arthritis (RA)
11. **Oral Talk. Computational Biomedicine, *Boston University* and *Dana-Farber Cancer Institute***, 2016  
From data to the clinic: statistical methods for characterizing genomic heterogeneity in mixed samples

## **Selected Posters**

---

### *Immunology or Rheumatology Conferences*

1. Jonsson, A.H.\*, Wei, K.\*, **Zhang, F.\***, Nathan, A., Mears, J., Watts, G.F.M., ..., Accelerating Medicines Partnership RA/SLE Network, Anolik, J., Raychaudhuri, S., Brenner, M. A high-throughput single cell multiomics pipeline to obtain multiple types of single-cell data from synovial biopsies: experience from phase 2 of the Accelerating Medicines Partnership (AMP) – RA/SLE network. ***American College of Rheumatology (ACR)***, 2020
2. Jonsson, A.H., **Zhang, F.**, Watts, G., Wei, K., Rao, D., Raychaudhuri, S., Brenner, M. An expanded granzyme K+ CD8 T cell population induces inflammatory responses in rheumatoid arthritis synovium. ***American College of Rheumatology (ACR)***, 2020
3. Meednu, N., **Zhang, F.**, Escalera-Rivera, K., Corsiero, E., Prediletto, E., Bombardieri, M., Dicarlo, E., Orange, E., Goodman, S., Donlin, L., Raychaudhuri, S., Pitzalis, S., McDavid, A., Anolik, J.H. Role of NR4A nuclear receptor family in RA synovial ectopic lymphoid neogenesis revealed by single cell profiling. ***American College of Rheumatology (ACR)***, 2020
4. Wei, K.\*, Jonsson, A.H.\*, **Zhang, F.\***, Nathan, A., Mears, J., Watts, G., Zhu, Z., Korsunsky, I., Donlin, L.T., Rao, D., Filer, A., Boyce, B., Gravalles, E., Holers, V.M., Moreland, L., Gregersen, P., Bykerk, V., Anolik, J., Raychaudhuri, S., Brenner, M. Towards a single cell portrait of rheumatoid arthritis-development of a single cell multiomics pipeline for phase 2 of the Accelerating Medicine Partnership (AMP)-RA network. ***American College of Rheumatology (ACR)***, Washington, DC, 2019
5. **Zhang, F.**, Mears, J., ..., Raychaudhuri, S. Integration of single cells from inflamed tissue in RA and SLE reveals shared immune and stromal cell population. ***Federation of Clinical Immunology Societies (FOCIS)***. Boston, MA, 2019

### *Computational Biology Conferences*

6. **Zhang, F.**, Wei, K., Slowikowski, K., Fonseka, C.Y., Rao, D.A., ..., Raychaudhuri, S. Multi-model single-cell transcriptomics and mass cytometry integration to define cell states in rheumatoid arthritis. ***Single Cell Biology Keystone Symposia***, Breckenridge, CA, Jan. 2019
7. **Zhang, F.**, Wang, C., Trapp, A.C., Flaherty, P. A global optimization algorithm for sparse mixed membership matrix factorization. ***Conference of Reproducibility in Personalized Medicine Research***. Department of Biostatistics at the Harvard TH Chan School of Public Health, Boston, MA, 2016
8. Saddiki, H., **Zhang, F.**, Trapp, A.C., Flaherty, P. A Deterministic Global Optimization Method for Variational Inference. ***Workshop at Neural Information Processing Systems (NIPS)***, Barcelona, Spain, 2016
9. **Zhang, F.**, and Flaherty, P. Variational inference for rare variant detection in deep, heterogeneous next-generation sequencing data. ***New England Statistics Symposium (NESS)***, Yale, CT, 2016
10. **Zhang, F.**, and Flaherty, P. Variant detection model with improved robustness and accuracy for low-depth targeted next-generation sequencing data. The 22th Annual International Conference on ***Intelligent Systems for Molecular Biology (ISMB)***, Boston, MA, 2014
11. **Zhang, F.**, and Flaherty, P. An ultra-sensitive variant detection model for low-depth heterogeneous next-generation sequence. The 18th Annual International Conference on ***Research in Computational Molecular Biology (RECOMB)***. Pittsburgh, PA, 2014

## Grant Applications

---

1. Integrating single-cell genomics and proteomics to define pathogenic cell states shared between RA and RA-ILD Grant K99 (NIH NIAMS, In Review)  
PI: Dr. Fan Zhang 2020
2. Differentiation of immune cells and fibroblasts in inflamed tissue in RA and SLE – Computational systems immunology core (NIH NIAID, In Review)  
Key Personnel (PI: Dr. Soumya Raychaudhuri) 2019
3. A genome-wide survey of point mutations modulating methotrexate  
PhRMA Foundation Informatics Grant 2013080079  
Key Personnel (PI: Dr. Patrick Flaherty) 2015

## Professional Skill

---

**Computation/Statistics:** Bayesian statistics, mixture modeling, convex and global optimization, deep learning

**Genomic data analysis:**

- Multi-omics: single-cell RNA-seq, single-cell ATAC-seq, and single-cell CITE-seq data integration
- Rare variant detection: time-series next-generation sequencing (NGS) DNA-seq data

**Programming:** Proficient in R and Python, and shell scripting in Linux

## Teaching Experience

---

Teaching Assistant, Biomedical Data Analysis and Programming, WPI 2015, 2016  
Teaching Assistant, Biomedical Engineering Design, WPI 2015  
Teaching Assistant, Introduction of Biomedical Engineering, WPI 2014

## Mentorship

---

Kathryn Weinand, PhD student at Bioinformatics and Integrative Genomics Program at Harvard  
“Method development for integrating single-cell ATAC-seq with single-cell RNA-seq” 2019 – Present  
Joseph Mears, Bioinformatics Analyst II at HMS  
“Characterization of macrophage heterogeneity using single-cell transcriptomics” 2018 – Present  
Gaurav Luthria, PhD student at Bioinformatics and Integrative Genomics Program at Harvard  
“Text mining methods for single-cell RNA-seq data annotation” Summer 2017  
Yifan Zhao and Tete Zhang, Master student at WPI  
“NGS data analysis in mixed tumor samples” 2015 – 2016

## Awards & Honors

---

Full Fellowship for Graduate Student, Jilin University 2009 – 2012  
Outstanding Graduate Student, Jilin University (Top 1%) 2011  
Graduate Student Entrance Scholarship, Jilin University 2009  
“Outstanding Star” of Software Engineering for Undergraduate, IBM 2007

## Professional Affiliations

---

**Member**, American College of Rheumatology (ACR) Sep. 2017 – Present  
**Member**, International Society for Computational Biology (ISCB) 2014 – 2015  
**Member**, American Association for the Advancement of Science (AAAS) Sep. 2013 – Present

## Peer Review

---

Ad Hoc Reviewer (*General Journal*) 2017 – 2020  
*Nature, Nature Communication, Cell Reports, iScience, Scientific Reports*  
Ad Hoc Reviewer (*Computational Biology Journal*) 2017 – 2020  
*Genomics, Proteomics & Bioinformatics, Journal of Computational Statistics, NESS (New England Statistics Symposium)*  
Ad Hoc Reviewer (*Rheumatology/Immunology Journal*) 2018 – 2019  
*Arthritis Research & Therapy, Cancer Letters*