

# Stanford

---



## Aziz Khan

Sr. Staff Scientist, Stanford Cancer Institute Core

### Bio

---

#### BIO

Aziz Khan is a staff scientist at the Stanford Cancer Institute, where he develops reproducible pipelines and machine learning methods for integrative analysis of multi-omics data at bulk and single-cell resolution to understand tumor evolution and chromatin regulatory dynamics of tumor growth.

Aziz completed his PhD in Bioinformatics at Tsinghua University, China in 2016 followed by a three year postdoctoral training at the University of Oslo, Norway. During PhD and Postdoc his primary research emphasis was on regulatory genomics and epigenomics. He developed computational methods, tools, and resources to understand the (epi)genomic control of gene regulation in development and disease.

Apart from research, he is advocating for open science, open-source, preprints, and reproducibility in research. He is a contributor for Bioconda and also developed several open-source tools and resources such as JASPAR. He is ASAPbio and eLife Community Ambassador and co-founded ECRcentral (ecrcentral.org), a community initiative for early-career researchers.

#### HONORS AND AWARDS

- CSC fully-funded PhD Scholarship, Chinese Scholarship Council (2012 – 2016)
- Erasmus+ mobility grant, European Commission (2019)
- OBF Travel Fellowship, ISMB/ECCB 2019, Basel, Switzerland (May 2019)
- Biocuration Travel Fellowship, Biocuration Conference 2019, Cambridge, UK (Apr 2019)
- Biocuration Travel Fellowship, Biocuration Conference 2018, Shanghai, China (Apr 2018)
- Research Travel Grant, Higher Education Commission (HEC), Pakistan (2016)
- TWAS BIOVISION.Next Fellowship, BioVision conference, Lyon, France (2014)
- TWAS BIOVISION.Next Fellowship, BioVision conference, Lyon, France (2013)
- Research Travel Grant, Higher Education Commission (HEC), Pakistan (2012)

#### EDUCATION AND CERTIFICATIONS

- Teaching Certificate (Associate), CIRTL and Stanford University , Evidence-based STEM Teaching (2024)
- PhD, Tsinghua University, China , Bioinformatics (2016)
- Postdoctoral Fellow, NCMH, University of Oslo, Norway , Computational Biology (2019)

#### LINKS

- Google Scholar: <https://scholar.google.com/citations?user=m23RfrcAAAAJ&hl=en>

- Twitter: <https://twitter.com/khanaziz84>
- ECRcentral: <https://ecrcentral.org>

## Teaching

---

### COURSES

#### 2023-24

- The Art of Reproducible Science: A Hands-on Approach: BIOS 259 (Win)

## Professional

---

### PROFESSIONAL INTERESTS

gene regulation, cancer regulatory genomics and epigenomics, integrative analysis of multi-omics data, machine learning

### PROFESSIONAL AFFILIATIONS AND ACTIVITIES

- Community Ambassador, eLife (2018 - 2020)
- Member, International Society for Computational Biology (ISCB) (2015 - present)
- Co-Chair, Web Committee, ISCB Student Council (2016 - 2019)
- Ambassador, ASAPbio (2019 - present)

## Publications

---

### PUBLICATIONS

- **Germline-mediated immunoediting sculpts breast cancer subtypes and metastatic proclivity.** *Science (New York, N.Y.)*  
Houlahan, K. E., Khan, A., Greenwald, N. F., Vivas, C. S., West, R. B., Angelo, M., Curtis, C.  
2024; 384 (6699): eadh8697
- **JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles.** *Nucleic acids research*  
Rauluseviciute, I., Riudavets-Puig, R., Blanc-Mathieu, R., Castro-Mondragon, J. A., Ferenc, K., Kumar, V., Lemma, R. B., Lucas, J., Cheneby, J., Baranasic, D., Khan, A., Fornes, O., Gundersen, et al  
2023
- **Deterministic evolution and stringent selection during preneoplasia.** *Nature*  
Karlsson, K., Przybilla, M. J., Kotler, E., Khan, A., Xu, H., Karagoyzova, K., Sockell, A., Wong, W. H., Liu, K., Mah, A., Lo, Y. H., Lu, B., Houlahan, et al  
2023
- **Germline-mediated immunoediting sculpts breast cancer subtypes and metastatic proclivity.** *bioRxiv : the preprint server for biology*  
Houlahan, K. E., Khan, A., Greenwald, N. F., West, R. B., Angelo, M., Curtis, C.  
2023
- **Somatic variant detection from multi-sampled genomic sequencing data of tumor specimens using the ith.Variant pipeline.** *STAR protocols*  
Maeser, N., Khan, A., Sun, R.  
2022; 4 (1): 101927
- **Molecular classification and biomarkers of clinical outcome in breast ductal carcinoma in situ: Analysis of TBCRC 038 and RAHBT cohorts.** *Cancer cell*  
Strand, S. H., Rivero-Gutierrez, B., Houlahan, K. E., Seoane, J. A., King, L. M., Risom, T., Simpson, L. A., Vennam, S., Khan, A., Cisneros, L., Hardman, T., Harmon, B., Couch, et al  
2022
- **MITI minimum information guidelines for highly multiplexed tissue images.** *Nature methods*  
Schapiro, D., Yapp, C., Sokolov, A., Reynolds, S. M., Chen, Y., Sudar, D., Xie, Y., Muhlich, J., Arias-Camison, R., Arena, S., Taylor, A. J., Nikolov, M., Tyler, et al  
2022; 19 (3): 262-267

- **UniBind: maps of high-confidence direct TF-DNA interactions across nine species.** *BMC genomics*  
Puig, R. R., Boddie, P., Khan, A., Castro-Mondragon, J. A., Mathelier, A.  
2021; 22 (1): 482
- **Pakistan: anger mounts over threat to higher education.** *Nature*  
Khan, A.  
2021; 592 (7856): 685
- **Changing scientific meetings for the better.** *Nature human behaviour*  
Sarabipour, S. n., Khan, A. n., Seah, Y. F., Mwakilili, A. D., Mumoki, F. N., Sáez, P. J., Schwessinger, B. n., Debat, H. J., Mestrovic, T. n.  
2021
- **JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles.** *Nucleic acids research*  
Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., Fornes, O., Leung, T. Y., Aguirre, et al  
2021
- **A call to eradicate non-inclusive terms from the life sciences.** *eLife*  
Khan, A. n.  
2021; 10
- **The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution.** *Cell*  
Rozenblatt-Rosen, O., Regev, A., Oberdoerffer, P., Nawy, T., Hupalowska, A., Rood, J. E., Ashenberg, O., Cerami, E., Coffey, R. J., Demir, E., Ding, L., Esplin, E. D., Ford, et al  
2020; 181 (2): 236–49
- **COVID-19: students caught in Pakistan's digital divide.** *Nature*  
Khan, A. n.  
2020; 587 (7835): 548
- **BiasAway: command-line and web server to generate nucleotide composition-matched DNA background sequences.** *Bioinformatics (Oxford, England)*  
Khan, A. n., Puig, R. R., Boddie, P. n., Mathelier, A. n.  
2020
- **JASPAR 2020: update of the open-access database of transcription factor binding profiles.** *Nucleic acids research*  
Fornes, O., Castro-Mondragon, J. A., Khan, A., van der Lee, R., Zhang, X., Richmond, P. A., Modi, B. P., Correard, S., Gheorghe, M., Baranasic, D., Santana-Garcia, W., Tan, G., Cheneby, et al  
2019
- **Modeling RNA-Binding Protein Specificity In Vivo by Precisely Registering Protein-RNA Crosslink Sites** *MOLECULAR CELL*  
Feng, H., Bao, S., Rahman, M., Weyn-Vanhentenryck, S. M., Khan, A., Wong, J., Shah, A., Flynn, E. D., Krainer, A. R., Zhang, C.  
2019; 74 (6): 1189+
- **A map of direct TF-DNA interactions in the human genome** *NUCLEIC ACIDS RESEARCH*  
Gheorghe, M., Sandve, G., Khan, A., Cheneby, J., Ballester, B., Mathelier, A.  
2019; 47 (4): e21
- **Integrative modeling reveals key chromatin and sequence signatures predicting super-enhancers** *SCIENTIFIC REPORTS*  
Khan, A., Zhang, X.  
2019; 9: 2877
- **High OGT activity is essential for MYC-driven proliferation of prostate cancer cells** *THERANOSTICS*  
Itkonen, H. M., Urbanucci, A., Martin, S. S., Khan, A., Mathelier, A., Thiede, B., Walker, S., Mills, I. G.  
2019; 9 (8): 2183–97
- **Making genome browsers portable and personal** *GENOME BIOLOGY*  
Khan, A., Zhang, X.  
2018; 19: 93
- **JASPAR RESTful API: accessing JASPAR data from any programming language** *BIOINFORMATICS*

- Khan, A., Mathelier, A.  
2018; 34 (9): 1612–14
- **Put science first and formatting later** *EMBO REPORTS*  
Khan, A., Montenegro-Montero, A., Mathelier, A.  
2018; 19 (5)
  - **JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework** *NUCLEIC ACIDS RESEARCH*  
Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J. A., van der Lee, R., Bessy, A., Cheneby, J., Kulkarni, S. R., Tan, G., Baranasic, D., Arenillas, D. J., Sandelin, et al  
2018; 46 (D1): D260–D266
  - **Super-enhancers are transcriptionally more active and cell type-specific than stretch enhancers** *EPIGENETICS*  
Khan, A., Mathelier, A., Zhang, X.  
2018; 13 (9): 910–22
  - **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nature methods*  
Grüning, B. n., Dale, R. n., Sjödin, A. n., Chapman, B. A., Rowe, J. n., Tomkins-Tinch, C. H., Valieris, R. n., Köster, J. n.  
2018; 15 (7): 475–76
  - **Intervene: a tool for intersection and visualization of multiple gene or genomic region sets** *BMC BIOINFORMATICS*  
Khan, A., Mathelier, A.  
2017; 18: 287
  - **dbSUPER: a database of super-enhancers in mouse and human genome.** *Nucleic acids research*  
Khan, A., Zhang, X.  
2016; 44 (D1): D164-71