

# KAI LI

(+1) 5712509112 | likaicnsh@gmail.com  
7901 Cambridge ST 146, Houston, Texas 77054

## EDUCATION

---

### Qingdao University (QDU)

Sep 2013 - Jun 2017

Biotechnology (Genomics and Bioinformatics) Bachelor

- GPA: 3.36/4
- Courses: Molecular Biology, Genomics and Bioinformatics, Python Programming, Data Mining, Data Structure and Algorithm Basics, Database System, Operating System, Digital Image Processing, Bioinformatics Algorithm, etc.

### BGI College

Mar 2016 - May 2017

- Courses: Statistic Applied in Bioinformatics, New Technology of Genomics, Bioinformatics Technology.

### Missouri State University (MSU) Springfield MO

Sep 2015 - Dec 2015

- Courses: Microbiology, Cellular Biology, Biology Experiments training.

## WORK EXPERIENCE

---

### Baylor College of Medicine

Dec 2018 - Present

Bioinformatics Programmer, Breast Center

Houston

- NeoFlow project: took responsibility to develop and maintain NeoFlow, a proteogenomics data analysis pipeline based on NextFlow.
- Clinical Proteomic Tumor Analysis Consortium (CPTAC) projects: applied NeoFlow to predicate neoantigens and found RNA or protein evidences. Until now, NeoFlow was applied on five cancer projects in CPTAC.
- Evaluation of different quality control methods for variant peptide identification in proteogenomics: took in charge of data analysis and data interpretation. The recommended strategy was implemented in NeoFlow. This study was published on Nature Communication.
- DeepRescore project: took responsibility of data analysis, data interpretation and workflow development. To improve peptide identification in immunopeptidomics, two deep learning-derived features as indicators of PSM quality were combined with other previously used features for PSM rescoring in DeepRescore. DeepRescore increased both the sensitivity and quality of MHC-binding peptide and neoantigen identifications compared to existing methods. This study was published on Proteomics.

### BGI-Shenzhen

Jun 2017 - Oct 2018

Bioinformatician, Bioinformatics platform, MS department

Shenzhen

- PDV project: took in charge of coding of PDV using Java. The goal of this project was to develop an integrated proteomics data viewer. The functions including database searching results visualization, denovo sequencing results visualization, proteogenomics files (proBAM and proBED) visualization, public database data visualization and MS/MS data visualization. In addition, it also provides a GUI to proteoQC for data quality analysis. PDV is platform-independent and can be used through both GUI and command line modes.
- PDV applications: generated spectra annotation figures using PDV for several projects, including published paper (PMID: 29222161) and unpublished projects.

### BGI-Shenzhen

Mar 2016 - May 2017

Bioinformatician (Internship), Bioinformatics platform, MS department

Shenzhen

- SWATH-expert project: took in charge of developing a GUI tool for XIC extraction and visualization for large scale SWATH dataset using Java. It was a part of SWATH-expert which is a new tool for SWATH data analysis. This was a collaboration with Dr. Tiannan Guo from the laboratory of Dr. Ruedi Aebersold in ETH Zurich.
- Explored applications of the cutting-edge deep learning in proteomics data analysis. Predicated the peptide phosphorylation sites by convolutional neural networks (CNNs) based on MXNet.
- Learned to use NextFlow to build pipeline for RNA-Seq and Exon-Seq data analysis.

## HONORS & AWARDS

---

|   |             |
|---|-------------|
| Outstanding graduate student of Shandong Province (5%)                          | June.2017   |
| Third class scholarship in BGI College (7/71).                                  | 2016 - 2017 |
| Model Student of Academic Records in QDU (1/19)                                 | 2015 - 2016 |
| 1st prize of China High School Biology Olympiad (Shandong Province division 3‰) | 2011        |

## LEADERSHIP EXPERIENCE

---

|   |                      |
|---|----------------------|
| <b>Art Work</b>   | Oct 2015             |
| Leader of China exhibition area   | Springfield, MO U.S. |
| <ul style="list-style-type: none"><li>• Art Work was an official exhibition in the downtown in Springfield MO. Led a team of 15 members and took one week to collect materials, design and decorate the exhibition area and attracted more than 500 local visitors.</li></ul> |                      |

## MISCELLANEOUS

---

- **Skills:** Java, Shell, R, Python, Amazon Web Service
- **Languages:** Chinese (Native), English (TOEFL: 99)
- **Conferences:** Speaker on ASMS(American Society for Mass Spectrometry) 2020 Reboot, poster presentation on Clinical Proteomics Tumor Analysis Consortium (CPTAC) 2019 Annual Meeting
- **Activities:** AWSome Day for AWS training (Student), BGI Life Periodic Plan (Principal of dasypodidae )

## PUBLICATIONS

---

- **Li, K.,** Jain, A., Malovannaya, A., Wen, B., & Zhang, B. (2020). DeepRescore: Leveraging Deep Learning to Improve Peptide Identification in Immunopeptidomics. *Proteomics*, 1900334.
- Wen, B., **Li, K.,** Zhang, Y., & Zhang, B. (2020). Cancer neoantigen prioritization through sensitive and reliable proteogenomics analysis. *Nature communications*, 11(1), 1-14.
- **Li, K.,** Vaudel, M., Zhang, B., Ren, Y., & Wen, B. (2019). PDV: an integrative proteomics data viewer. *Bioinformatics*, 35(7), 1249-1251.
- Wang, X., Codreanu, S. G., Wen, B., **Li, K.,** Chambers, M. C., Liebler, D. C., & Zhang, B. (2018). Detection of proteome diversity resulted from alternative splicing is limited by trypsin cleavage specificity. *Molecular & Cellular Proteomics*, 17(3), 422-430.
- Li, B., Li, N., Zhang, L., **Li, K.,** Xie, Y., Xue, M., & Zheng, Z. (2018). Hsa\_circ\_0001859 regulates ATF2 expression by functioning as an MiR-204/211 sponge in human rheumatoid arthritis. *Journal of immunology research*, 2018.
- Ren, Z., Qi, D., Pugh, N., **Li, K.,** Wen, B., Zhou, R., ... & Jones, A. R. (2019). Improvements to the rice genome annotation through large-scale analysis of RNA-Seq and proteomics data sets. *Molecular & Cellular Proteomics*, 18(1), 86-98.
- Clark, D. J., Dhanasekaran, S. M., Petralia, F., Pan, J., Song, X., Hu, Y., ... & Ma, W. (2019). Integrated proteogenomic characterization of clear cell renal cell carcinoma. *Cell*, 179(4), 964-983.
- Dou, Y., Kawaler, E. A., Zhou, D. C., Gritsenko, M. A., Huang, C., Blumenberg, L., ... & Liu, W. (2020). Proteogenomic characterization of endometrial carcinoma. *Cell*, 180(4), 729-748.
- Gillette, M. A., Satpathy, S., Cao, S., Dhanasekaran, S. M., Vasaikar, S. V., Krug, K., ... & Krek, A. (2020). Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell*, 182(1), 200-225.
- Huang, C., Chen, L., Li, Y., Savage, S., Schnaubelt, M., Leprevost, F. V., ... & Zhang, B. (2020). Proteogenomics characterization of HPV-negative head and neck squamous cell carcinomas. (Under review)

## REFERENCES

---

References available upon request.