郑小琪个人简历

出生日期：1981 年 6 月 21 日
地　　址：徐汇区桂林路 100 号，
　　　　　上海师范大学 数学系，三号楼 232 室
电话号码：15618228810 (手机)
　　　　　021-64323734 (办公室)
E-mail: xqzheng@shnu.edu.cn or zheng.shnu@gmail.com

研究概况

　　从事**生物信息学**和**生物统计**方面的研究，擅长高通量数据的**统计分析**和**数学建模**。围绕基于 DNA 甲基化数据的肿瘤纯度估计，考虑纯度的差异甲基化分析、肿瘤样本聚类等问题上取得了创新性的研究成果。主要学术贡献有：开发了基于亚硫酸盐测序数据（BS-seq）的肿瘤细胞纯度估计方法 MethylPurify；开发了基于 450k 芯片数据的肿瘤纯度和差异甲基化方法 InfiniumPurify; 提出了抗癌药物敏感性预测的双层网络模型。

教育经历

✧　**2006 - 2009**　　大连理工大学应用数学系　　博士
✧　**2004 - 2006**　　大连理工大学应用数学系　　硕士
✧　**2000 - 2004**　　山东师范大学数学系　　学士

研究工作经历

✧　**2012 - 至今**　　上海师范大学　数学系　　副教授
✧　**2009 - 2012**　　上海师范大学　数学系　　讲师
✧　**2013 - 2014**　　美国哈佛大学　（**Harvard University**）
　　　　　　　　　　　生物统计与计算生物学系　　访问学者

论文发表情况 (#: 第一作者，*: 通讯作者)

[1] **Zheng X#***, Zhang N#, Wu HJ, Wu H*: Estimating and accounting for tumor purity in the analysis of DNA methylation data from cancer studies. ***Genome Biol*** 2017, 18:17. (BioMed Central 出版集团旗舰杂志，IF = 11.908)

[2] **Zheng X#**, Zhao Q#, Wu H-J#, Li W, Wang H, Meyer CA, Qin QA, Xu H, Zang C, Jiang P, Li F, Hou Y, He J, Wang J, Wang J, Zhang P*, Zhang Y*, Liu XS*: MethylPurify: tumor purity deconvolution and differential methylation detection from single tumor DNA methylomes. ***Genome Biol*** 2014, 15:419. (BioMed Central 出版集团旗舰杂志，IF = 11.908)

[3] **Zheng X#***, Zhong S: From structure to function, how bioinformatics help to reveal functions of our genomes. ***Genome Biol***, 2017, accepted. (BioMed Central 出版集团旗舰杂志，IF = 11.908)

[4] Zhang N#, Wu HJ#, Zhang W, Wang J, Wu H*, **Zheng X***: Predicting tumor purity from methylation microarray data. ***Bioinformatics*** 2015, 31:3401-3405. (生物信息学领域最具影响力的杂志之一，IF = 7.307)

[5] Zhang W#, Feng H#, Wu H*, **Zheng X***, Accounting for Tumor purity improves cancer subtype lassification from DNA methylation data, ***Bioinformatics***, 33(17), 2017, 2651–2657. (生物信息学领域最具影响力的杂志之一，IF = 7.307)

[6] Zhang N#, Wang H#, Fang Y, Wang J*, **Zheng X***, Liu XS*: Predicting Anticancer Drug Responses Using a Dual-Layer Integrated Cell Line-Drug Network Model. ***PLoS Comput Biol*** 2015, 11:e1004498. （计算生物学领域最具影响力的杂志之一，IF = 4.542)

[7] Niu B#, Paulson JN, **Zheng X**, Kolter R*: Simplified and Representative Bacterial Community of Maize Roots. ***P Natl Acad Sci USA*** 2017, 114 (12): E2450-E2459. (著名综合类杂志，IF = 9.661)

[8] Wang F#, Zhang N, Wang J, Wu H#, **Zheng X***: Tumor purity and differential methylation in cancer epigenomics. ***Brief Funct Genomics*** 2016, 15:408-419. (IF = 4.098)

[9] Ni T#, Li XY#, Lu N#, An T, Liu ZP, Fu R, Lv WC, Zhang YW, Xu XJ, Grant Rowe R, Lin YS, Scherer A, Feinberg T, **Zheng XQ**, Chen BA, Liu XS, Guo QL, Wu ZQ*, Weiss SJ*: Snail1-dependent p53 repression regulates expansion and activity of tumour-initiating cells in breast cancer. ***Nat Cell Biol*** 2016, 18:1221-1232. (IF = 20.060)

[10] Liu X#, Yang J, Zhang Y, Fang Y, Wang F, Wang J, **Zheng X***, Yang J*: A systematic study on drug-response associated genes using baseline gene expressions of the Cancer Cell Line Encyclopedia. ***Sci Rep*** 2016, 6:22811. (IF = 4.259)

[11] Dong Z#, Zhang N#, Li C, Wang H, Fang Y, Wang J, **Zheng X***: Anticancer drug sensitivity prediction in cell lines from baseline gene expression through recursive feature selection. ***BMC Cancer*** 2015, 15:489. (IF = 3.288)

[12] Li L#, Yu S, Xiao W, Li Y, Huang L, **Zheng X***, Zhou S*, Yang H*: Sequence-based identification of recombination spots using pseudo nucleic acid representation and recursive feature extraction by linear kernel SVM. ***BMC bioinformatics*** 2014, 15:340. (IF =2.448)

[13] Li L#, Yu S, Xiao W, Li Y, Li M, Huang L, **Zheng X***, Zhou S*, Yang H*: Prediction of bacterial protein subcellular localization by incorporating various features into Chou's PseAAC and a backward feature selection approach. ***Biochimie*** 2014, 104:100-107. (IF = 3.112)

[14] Li L#, Yu S, Xiao W, Li Y, Hu W, Huang L, **Zheng X***, Zhou S*, Yang H*: Protein submitochondria localization from integrated sequence repesentation and SVM-based backward feature extraction. ***Mol BioSyst*** 2014. (IF = 2.781)

[15] Li Y#, Liu Q, **Zheng X***, He Pa: UC‑Curve: A highly compact 2D graphical representation of protein sequences. ***Int J Quantum Chem*** 2014, 114:409-415. (IF =2.920)

[16] Zhu J#, Qin Y, Liu T, Wang J, **Zheng X***: Prioritization of candidate disease genes by topological similarity between disease and protein diffusion profiles. ***BMC bioinformatics*** 2013, 14:S5. (IF = 2.448)

[17] Yu X#, **Zheng X***, Liu T, Dou Y, Wang J: Predicting subcellular location of apoptosis proteins with pseudo amino acid composition: approach from amino acid substitution matrix and auto covariance transformation. ***Amino acids*** 2012, 42:1619-1625. (IF = 3.173)

[18] Liu T#, Geng X, **Zheng X***, Li R, Wang J: Accurate prediction of protein structural class using auto covariance transformation of PSI-BLAST profiles. ***Amino acids*** 2012, 42:2243-2249. (IF = 3.173)

[19] Li Y#, Qin Y, **Zheng X***, Zhang Y: Three‑unit semicircles curve: A compact 3D graphical representation of DNA sequences based on classifications of nucleotides. ***Int J Quantum Chem*** 2012, 112:2330-2335. (IF = 2.920)

[20] **Zheng X**#, Liu T, Yang Z, Wang J*: Large cliques in Arabidopsis gene coexpression network and motif discovery. ***J Plant Physiol*** 2011, 168:611-618. (IF = 3.121)

[21] Li C#*, Ma H, Zhou Y, Wang X, **Zheng X**: Similarity analysis of DNA sequences based on the weighted pseudo‑entropy. ***J Comput Chem*** 2011, 32:675-680. (IF = 3.229)

[22] **Zheng X**#, Li C, Wang J*: An information‑theoretic approach to the prediction of protein structural class. ***J Comput Chem*** 2010, 31:1201-1206. (IF = 3.229)

[23] Liu T#, **Zheng X***, Wang J: Prediction of protein structural class for low-similarity sequences using support vector machine and PSI-BLAST profile. ***Biochimie*** 2010, 92:1330-1334. (IF = 3.112)

[24] Liu T#, **Zheng X***, Wang J: Prediction of protein structural class using a complexity-based distance measure. ***Amino acids*** 2010, 38:721-728. (IF = 3.196)

[25] Dou Y#, **Zheng X**, Yang J, Wang J*: Prediction of catalytic residues based on an overlapping amino acid classification. ***Amino acids*** 2010, 39:1353-1361. (IF = 3.196)

[26] Dou Y#, **Zheng X**, Wang J*: Several appropriate background distributions for entropy-based protein sequence conservation measures. ***J Theor Biol*** 2010, 262:317-322. (IF = 2.049)

[27] **Zheng X**#, Qin Y, Wang J*: A Poisson model of sequence comparison and its application to coronavirus phylogeny. ***Math Biosci*** 2009, 217:159-166. (IF = 1.256)

[28] **Zheng X**#, Liu T, Wang J*: A complexity-based method for predicting protein subcellular location. ***Amino acids*** 2009, 37:427-433. (IF = 3.173)

[29] **Zheng X**#, Li C, Wang J*: A complexity-based measure and its application to phylogenetic analysis. ***J Math Chem*** 2009, 46:1149-1157. (IF = 1.056)

[30] **Zheng X**#, Dou Y, Wang J*: Phylogenetic inference from binary sequences reduced by primary DNA sequences. ***J Math Chem*** 2009, 46:1137-1148. (IF = 1.308)

[31] Wang J#, **Zheng X***: Comparison of protein secondary structures based on backbone dihedral angles. ***J Theor Biol*** 2008, 250:382-387. (IF = 2.113)

[32] Wang J#, **Zheng X***: WSE, a new sequence distance measure based on word frequencies. ***Math Biosci*** 2008, 215:78-83. (IF = 1.246)

开发软件

✧ **MethylPurify**: a python package for tumor purity deconvolution and differential methylation detection from single tumor DNA methylomes. https://pypi.python.org/pypi/MethylPurify. (Collected in OMICtools: https://omictools.com/methylpurify-tool)

✧ **InfiniumPurify**: a python package for estimating and accounting for tumor purity in the analysis of DNA methylation data, https://bitbucket.org/zhengxiaoqi/infiniumpurify. (Collected in OMICtools: https://omictools.com/infiniumpurify-tool)

✧ **InfiniumPurify**: an R package for estimating and accounting for tumor purity in DM calling and clustering analysis, https://cran.r-project.org/web/packages/InfiniumPurify/index.html.

学术报告

1. "基于 DNA 甲基化的肿瘤纯度估计和相关计算问题"，2017 计算表观遗传学研讨会，内蒙古，8.10-13, 2017.

2. "Tumor purity estimation and differential methylation analysis in cancer research", The 14th China-Japan-Korea Bioinformatics Training Course & Symposium , 厦门, 8.19-20, 2016.

3. "Tumor purity estimation and differential methylation analysis in cancer research", The 4th Institute of Mathematical Statistics Asia Pacific Rim Meeting, 香港中文大学, 6.27-30, 2016.

4. "Exploring genetic and epigenetic data in Cancer research: two case studies", Tsinghua-Sanya Workshop on Big Data: Opportunities, Challenges and Innovations, 清华大学三亚数学中心, 12.27–30, 2014.

5. "MethylPurify: tumor purity deconvolution and differential methylation detection from single tumor DNA methylomes", The 8th International Conference on Systems Biology, 青岛, 10.26-28, 2014

6. "MethylPurify: tumor purity deconvolution and differential methylation detection from tumor DNA methylomes", 哈佛大学公共卫生学院, 1.17, 2014

7. "A statistical approach for detecting tumor subclone through single-base resolution DNA methylation data", CFCE retreat, 哈佛大学公共卫生学院, 7.9, 2013

教学和指导学生情况

✧ 承担本科生课程：高等代数与解析几何、复变函数、线性代数、高等数学、概率论与数理统计等

✧ 承担研究生课程：组合数学与图论、计算生物学计算平台、图论及其应用、统计学习基础、生物信息学等

✧ 指导**硕士生 7 名**、**博士生 6 名**（协助王军教授），所指导研究生以第一作者发表 SCI 论文 14 篇

✧ 一直担任本科及研究生数学建模竞赛的指导工作，累计获得全国二等奖三十余项，三等奖二十余项

承担项目

✧ 基于 DNA 甲基化的肿瘤异质性分析及应用算法研究，**国家自然科学基金面上项目** (61572327)， 2016 年 1 月-2019 年 12 月，在研，主持

✧ 拟南芥基因转录调控关系的识别方法研究，**国家自然科学基金青年项目** (31100953)，2012 年 1 月-2014 年 12 月，已结题，主持

✧ 拟南芥转录调控关系识别的机器学习方法研究，**上海市教委创新项目** (SSD10017)，2012 年 1 月-2014 年 12 月，已结题，主持

社会工作

✧ *Computational and Mathematical Methods in Medicine* 杂志编委

✧ *Genome Biology, Molecular Biology and Evolution, Nucleic Acids Research, Bioinformatics, BMC Bioinformatics, Journal of Theoretical Biology* 等杂志审稿人

✧ 中国计算机协会 (CCF) 生物信息学专业组专业委员会委员

✧ 全国"计算系统生物学会"会员