

# Nan Xiao

---

[me@nanx.me](mailto:me@nanx.me)  
<http://nanx.me>

## Academic degrees

2008 – 2012 Bachelor of Science (Statistics), Central South University. *Changsha, China*

## Studies

2015 – 2016 Ph.D. student (Human Genetics), University of Chicago. *Chicago, IL, U.S.*

2012 – 2015 Ph.D. student (Statistics), Central South University. *Changsha, China*

## Publications

### Journal articles

2015 **Nan Xiao** and Q.-S. Xu (2015). “Multi-step adaptive elastic-net: reducing false positives in high-dimensional variable selection”. *Journal of Statistical Computation and Simulation*. doi: [10.1080/00949655.2015.1016944](https://doi.org/10.1080/00949655.2015.1016944).

2015 **Nan Xiao**, D.-S. Cao, M.-F. Zhu, and Q.-S. Xu (2015). “protr/ProtrWeb: R package and web server for generating various numerical representation schemes of protein sequence”. *Bioinformatics*. doi: [10.1093/bioinformatics/btv042](https://doi.org/10.1093/bioinformatics/btv042).

2014 D.-S. Cao\*, **Nan Xiao**\*, Q.-S. Xu and A. F. Chen (2014). “Rcpi: R/Bioconductor package to generate various descriptors of proteins, compounds, and their interactions”, *Bioinformatics*. \***Joint first authors**. doi: [10.1093/bioinformatics/btu624](https://doi.org/10.1093/bioinformatics/btu624).

2015 D.-S. Cao, **Nan Xiao**, Y.-J. Li, W.-B. Zeng, Y.-Z. Liang, A.-P. Lu, Q.-S. Xu, A. F. Chen (2015). “Integrating multiple evidence sources to predict adverse drug reactions based on systems pharmacology model”. *CPT: Pharmacometrics & Systems Pharmacology*. doi: [10.1002/psp4.12002](https://doi.org/10.1002/psp4.12002).

2015 J.-B. Wang, D.-S. Cao, M.-F. Zhu, Y.-H. Yun, **Nan Xiao**, Y.-Z. Liang (2015). “*In silico* evaluation of logD<sub>7.4</sub> and comparison with other prediction methods”. *Journal of Chemometrics*. doi: [10.1002/cem.2718](https://doi.org/10.1002/cem.2718).

2015 L. Shen, D.-S. Cao, Q.-S. Xu, X. Huang, **Nan Xiao**, Y.-Z. Liang (2015). “A novel local manifold-ranking based *k*-NN for modeling the regression between bioactivity and molecular descriptors”. *Chemometrics and Intelligent Laboratory Systems*. doi: [10.1016/j.chemolab.2015.12.005](https://doi.org/10.1016/j.chemolab.2015.12.005).

### Book translations

2016 Max Kuhn and Kjell Johnson (2015). *Applied Predictive Modeling*. (Hui Lin, Yi-Xuan Qiu, En-Chi Ma, **Nan Xiao**, & Vivian Zhang, Trans.). China Machine Press (Original work published in 2013). ISBN: [978-7-1115-3342-9](https://doi.org/978-7-1115-3342-9).

2014 Winston Chang (2014). *R Graphics Cookbook*. (**Nan Xiao**, Yi-Shuo Deng, Tai-Yun Wei, & Yi-Xuan Qiu, Trans.). Posts and Telecom Press (Original work published in 2013). ISBN: [978-7-115-34227-0](https://doi.org/978-7-115-34227-0).

- 2013 Hadley Wickham (2013). *ggplot2: Elegant Graphics for Data Analysis*. (Tai-Yun Wei, Yi-Xuan Qiu, **Nan Xiao**, Tao Gao, & Wei-Cheng Zhu, Trans.). Xi'an Jiaotong University Press (Original work published in 2010). ISBN: [978-7-5605-4969-9](#).
- 2013 Robert Kabacoff (2013). *R in Action: Data Analysis and Graphics with R*. (Tao Gao, **Nan Xiao**, & Gang Chen, Trans.). Posts and Telecom Press (Original work published in 2011). ISBN: [978-7-115-29990-1](#).

### **Conference & workshop presentations**

- 2016 **hdnom.io: High-Dimensional Survival Modeling with Shiny**.  
Lightning talk. Shiny Developer Conference, Stanford University. January 2016.
- 2015 **Introduction to Reproducible Research in Bioinformatics**.  
Invited talk. 2015 Bioinformatics Workshop, Center for Research Informatics, University of Chicago. December 3, 2015.
- 2013 **Web Scraping with R**  
The 6<sup>th</sup> China R Conference, Renmin University of China, Beijing, China. May 2013.

### **Software and web applications**

#### **R/Bioconductor packages**

- grex**  
2016 – R package for mapping Ensembl gene IDs in Genotype-Tissue Expression (GTEx) data to Entrez IDs, HGNC symbols, and UniProt IDs. Project website: <http://nanx.me/grex/>. Available from CRAN: <https://cran.r-project.org/package=grex>.
- ggsci**  
2016 – R package offering a collection of ggplot2 color palettes inspired by scientific journals and science fiction TV shows. Project website: <http://ggsci.net>. Available from CRAN: <https://cran.r-project.org/package=ggsci>.
- hdnom**  
2015 – R package building nomograms for high-dimensional survival data, with support for model validation, calibration, and comparison. Project website: <http://hdnom.org>. Available from CRAN: <https://cran.r-project.org/package=hdnom>.
- liftr**  
2015 – R package for dockerizing R Markdown documents with support for Rabix in reproducible research. Project website: <http://liftr.me>. Available from CRAN: <https://cran.r-project.org/package=liftr>.
- sbgr**  
2015 – R Client for Seven Bridges Genomics API. Available from Bioconductor: <https://bioconductor.org/packages/sbgr>.

### **Rcpi**

- 2013 – 2014 R/Bioconductor package as a comprehensive molecular informatics platform for systems pharmacology and drug discovery.  
Available from Bioconductor: <https://bioconductor.org/packages/Rcpi>.

### **protr**

- 2012 – 2014 R package generating various numerical representation schemes of protein sequence for bioinformatics and proteochemometrics research.  
Available from CRAN: <https://cran.r-project.org/package=protr>.

### **enpls**

- 2014 – R package for ensemble partial least squares regression, a unified framework for feature selection, outlier detection, and ensemble learning. Project website: <http://enpls.org>.  
Available from CRAN: <https://cran.r-project.org/package=enpls>.

### **RECA**

- 2012 – 2014 R package for relevant component analysis (RCA) in supervised distance metric learning and supervised dimensionality reduction.  
Available from CRAN: <https://cran.r-project.org/package=RECA>.

## **Web applications**

### **hdnom.io**

- 2015 Web application for building penalized Cox models for high-dimensional data with survival outcomes. Streamlined the process of nomogram building, model validation, calibration, and comparison. Available at: <http://hdnom.io>.

### **ProtrWeb**

- 2013 – 2014 Web application for computing 20 types of protein sequence-derived structural and physicochemical features in bioinformatics. Available at <http://protr.org>.

### **ImgSVD**

- 2014 Web application for image compression via singular value decomposition. Collaborated with *Yihui Xie* (RStudio Inc.), *Yixuan Qiu* (Purdue University), and *Tong He* (Simon Fraser University). Available at <https://yihui.shinyapps.io/imgsvd>.

## **Honors and awards**

- 2013 National scholarship for graduate students (5%).  
Highest award the government established for graduate students in China.
- 2013 Outstanding graduate student award (5%). Central South University, China.
- 2011 1<sup>st</sup> Prize (top 5%) in Hunan Contest District in China Undergraduate Mathematical Contest in Modeling (CUMCM). With *Guan-Chun Liu* and *Tian-Yu Zhao*.
- 2010 Meritorious award (2<sup>nd</sup> in 1,000+) of Central China Undergraduate Mathematical Contest in Modeling (CCUMCM). With *Tao Gao* and *Chen-Xi Guo*.

Last revision: August, 2016