

NING, Kang

➤ Contact Information:

Professor
Computational Biology Team
Functional genomics group
Qingdao Institute of Bioenergy and Bioprocess Technology (QIBEBT)

Room 120, Biology Building
Songling Road 189
Qingdao, Shandong, China 266101

Phone: (+86) 532-80662624, (+86) 15264290838
Email: ningkang@qibebt.ac.cn, albertnk@gmail.com
Website: <http://www.computationalbioenergy.org/>

➤ General Research Interests:

- Genomics, next generation sequencing data analysis, gene regulation network, metagenomics, single-cell omics, etc. Especially in the area of microbiology and bioenergy.
- Proteomics, label-free protein identification and quantification. Especially integrative study combining different types of omics data.
- High-performance computation, algorithm design, applied combinatorial mathematics.

➤ Professional Experience:

◆ 08/2010 ~ present

Associate professor / Professor, Team Leader for Computational Biology
Computational Biology Team
Functional genomics group
Qingdao Institute of Bioenergy and Bioprocess Technology (QIBEBT)

◆ 11/2007 ~ 08/2010

Research lab Specialist and Post Doctoral Research Fellow
Pathology informatics Core, Department of Pathology, University of Michigan
Medical School, Ann Arbor, Michigan, USA
Advisors: Dr. Alexey I. Nesvizhskii

◆ 09/2002 ~ 07/2003

Research assistant
China National Science Foundation project in National High Performance
Computer Center, Hefei, China
Advisor: Dr. Liusheng Huang

➤ Education:

◆ Ph.D. in Computer Science

08/2003 ~ 08/2008

National University of Singapore, Singapore

Thesis Title: Algorithms for peptide and PTM identification using tandem mass spectrometry

Advisors: Dr. Hon Wai Leong and Dr. Louxin Zhang

Area of study: Combinatorial algorithms, Computational Biology and Bioinformatics

◆ **B.Sc. in Computer Science (5-year basis)**

09/1998 ~ 07/2003

University of Science and Technology of China, Hefei, P. R. China

Advisor: Dr. Liusheng Huang

➤ **Research Experiences:**

◆ **05/2013 ~ present: Single-cell platform development for microbial research**

CDZ (China-Germany Research Center) project (694,553 CNY on China side) aimed to establish single-cell platform for microbial research.

◆ **01/2013 ~ present: Method development for metagenome data-mining**

NSFC (Natural Science Foundation of China) project (800,000 CNY) aimed to create large database and data-mining system for metagenomic research.

◆ **11/2012 ~ present: Method development for synthetic biology**

MoST (Ministry of Science and Technology of China) project (863) (8.8M CNY) for foundations of synthetic biology research on microalgae.

◆ **01/2012 ~ present: Method development for multi-omics data analysis**

NSFC project (210,000 CNY) aimed to develop new methods for combining different types of omics data -- genomics, transcriptomic, proteomics, metabolomics and network-omics -- from high-throughput sequencing facilities, and interpret gene regulation process, gene expression dynamics, etc. I lead this project.

◆ **08/2011 ~ present: Algorithms for genome evolution and community evolution**

China-US collaborative research on algorithm development for genome evolution and community evolution. The aim is progressive develop methods for analysis of simple problems such as SNP and indel detection, to more biologically complex problems such as SNP and indel detections from metagenomic data, and to more computationally complex problems such as the fitness of the gene modules in the network. I lead the collaborative project on China side.

◆ **08/2010 ~ present: Method development for metagenomics research on microbial communities**

CAS (Chinese Academy of Sciences) and MoST project (973) for the broad area of metagenomics research (500,000 CNY). I lead the initiative on method

development, which aims to provide a set of methods, from genome assembly, gene annotation to network analysis, which are urgently needed by researchers worldwide.

◆ **08/2010 ~ present: Method development for pan-genome research on microalgae**

A comprehensive analysis on pan-genome's gene prediction, gene annotation, gene expression and comparative genomic study in micro-algae. Supported by CAS, MoST and NSFC of China. I lead the initiative on high-performance and highly-accurate method development for these analysis.

◆ **08/2010 ~ present: High-performance methods and collaborative methods development for genomic data analysis**

CAS supported projects (2.0M CNY) on accelerating the genomic data analysis efficiency by high-performance computing. There are generally two sub-areas. The first is the development and utilization of high-performance computing hardware (such as GPGPU) and software to improve efficiency. The second is the development of collaborative systems for efficient process of multiple sets of genomic data, and establish the online website for open service. I lead the initiative on both areas.

◆ **11/2007 ~ present: Analysis and Statistical Validation of Proteomic ne**

NIH/NCRR project (US\$1.2M) aimed to develop a set of statistical models and algorithms that will enable robust, accurate, and transparent analysis of large-scale quantitative tandem mass-spectrometry (MS/MS) based proteomic datasets from human clinical cancer specimens. I joined the project to computationally analyze and develop novel algorithms for protein identification and quantification problems based on tandem mass spectrometry.

◆ **11/2007 ~ 08/2010: Analysis of protein complexes (Ann Arbor, USA)**

NIH/NCI project in The Proteomics Core for University of Michigan Cancer Center. I have worked on the development of novel method for analysis of protein interaction networks and essential protein prediction.

◆ **08/2003 ~ 11/2007: Algorithms in Computational Biology, Algorithms for Peptide Sequencing and Fragment Assembly (Singapore, Singapore)**

NUS (National University of Singapore) research project aimed to provide novel algorithms for peptide sequencing. I have focused on *de novo* peptide sequencing from multi-charge spectra. I have also worked on the theoretical and experimental analysis of the Longest Common Subsequence and the Shortest Common Supersequence problems.

◆ **09/2002 ~ 07/2003: China national grid project on "Huai River Water Resources Protection and Management" system (Hefei, China)**

China National Science Foundation project in National High Performance Computer Center (Hefei, China), aimed at using grid computing technique for automatic river resources protection. I joined the team as research assistant and developed grid computing software for the system.

➤ **Peer-Reviewed Publications:**

◆ Journal papers

1. Qian Zhou, Xiaoquan Su, Anhui Wang, Jian Xu and **Kang Ning***, QC-Chain: Fast and Holistic Quality Control Method for Next-Generation Sequencing Data, *PLoS ONE*, 2013, doi: 10.1371/journal.pone.0060234. (SCI impact factor 4.09, *Corresponding author)
2. Baoxing Song, Xiaoquan Su and **Kang Ning***. MetaSee: An interactive and extendable visualization toolbox for metagenomic sample analysis and comparison, *PLoS ONE*, 2012. (SCI impact factor 4.09, *Corresponding author)
3. Xiaoquan Su; Jian Xu; **Kang Ning***. Meta-Storms: Efficient Search for Similar Microbial Communities Based on a Novel Indexing Scheme and Similarity Score for Metagenomic Data. *Bioinformatics* 2012, 28(19), 2494-2501. (SCI impact factor 5.47, *Corresponding author)
4. **Kang Ning**, Damian Fermin, and Alexey I. Nesvizhskii. Comparative analysis of different label-free mass spectrometry based protein abundance estimates and their correlation with RNA-Seq gene expression data. *J. Proteome Res* (2012). DOI: 10.1021/pr201052x. (SCI impact factor 5.46)
5. Na You, Gabriel Murillo, Xiaoquan Su, Xiaowei Zheng, Jian Xu, **Kang Ning**, Shoudong Zhang, Jiankang Zhu and Xinping Cui. SNP calling using genotype model selection on next generation sequencing data. *Bioinformatics* (2012) 28 (5): 643-650. (SCI impact factor 4.88)
6. Xiaoquan Su, Jian Xu and **Kang Ning***. Parallel-META: A High-Performance Computational Pipeline for Metagenomic Data Analysis. *BMC Systems Biology*, 2011. (SCI impact factor 3.57, *Corresponding author)
7. Fang Yang, Xiaowei Zeng, **Kang Ning**, Kuan-Liang Liu, Chien-Chi Lo, Wei Wang, Jie Chen, Dongmei Wang, Ranran Huang, Xingzhi Chang, Patrick S Chain, Gary Xie, Junqi Ling and Jian Xu. Saliva microbiomes distinguish caries-active from healthy human populations. *The ISME Journal*, 2011, doi:10.1038/ismej.2011.71 (SCI impact factor 6.40)
8. **Kang Ning**, Damian Fermin and Alexey I. Nesvizhskii. Computational Analysis of Unassigned High Quality MS/MS Spectra in Large-scale Proteomic Datasets. *Proteomics*, 2010, 10 (14): 2712 - 2718. (SCI impact factor 5.48, 8 citations)
9. **Kang Ning** and Damian Fermin. SAW: A Method to Identify Splicing Events from RNA-Seq Data based on Splicing Fingerprints. *PLoS ONE*, 2010, 5 (8): e12047. (SCI impact factor 4.35)
10. **Kang Ning**, Hoong Kee Ng, Sriganesh Srihari, Hon Wai Leong and Alexey I Nesvizhskii. Examination of the Relationship between Essential Genes in PPI Network and Hub Proteins in Reverse Nearest Neighbor Topology. *BMC Bioinformatics*, 2010, 11:505. (SCI impact factor 3.43)
11. Sriganesh Srihari, **Kang Ning*** and Hon Wai Leong*. MCL-CAW: A refinement of MCL for detecting yeast complexes from weighted PPI networks by incorporating core-attachment structure. *BMC Bioinformatics*, 2010, 11:504. (SCI impact factor 3.43, *Corresponding author)

12. **Kang Ning** and Alexey I. Nesvizhskii. The Utility of Mass Spectrometry-based Proteomic Data for Validation of Novel Alternative Splice Forms Reconstructed from RNA-Seq Data: A Preliminary Assessment. *BMC Bioinformatics*. (accepted, SCI impact factor 3.43)
13. **Kang Ning**. Deposition and Extension Approach to Find Longest Common Subsequence for Thousands of Long Sequences. *Computational Biology and Chemistry* 34(3):149-157 (2010). (SCI impact factor 1.37)
14. **Kang Ning** and Hon Wai Leong. The multiple sequence sets: problem and heuristic algorithms. *Journal of Combinatorial Optimization*. 2010, doi: 10.1007/s10878-010-9329-3. (SCI impact factor 0.87, premium in Combinatorics)
15. **Kang Ning**, Hoong Kee Ng and Hon Wai Leong. Analysis of the Relationships among Longest Common Subsequences, Shortest Common Supersequences and Patterns and its application on Pattern Discovery in Biological Sequences. *International Journal of Data Mining and Bioinformatics*. (in press, SCI impact factor 0.93)
16. Hoong Kee Ng¹, **Kang Ning**¹ and Hon Wai Leong. Two-phase Filtering Strategy for Efficient Peptide Identification from Mass Spectrometry. *Journal of Proteomics & Bioinformatics* 3: 121-129. doi:10.4172/jpb.1000130. (**¹Co-first authors**)
17. Hon Nian Chua¹, **Kang Ning**¹, Wing-Kin Sung, Hon Wai Leong and Limsoon Wong. Using Indirect Protein-Protein Interactions for Protein Complex Prediction. *Journal of Bioinformatics and Computational Biology*, Vol. 6, No. 3 (2008) page 435-466. (**¹Co-first authors**, Elsevier Scopus impact factor 7.85, 25 citations)
18. **Kang Ning**¹, Nan Ye¹ and Hon Wai Leong. On preprocessing and anti-symmetry in de novo peptide sequencing: Improving efficiency and accuracy. *Journal of Bioinformatics and Computational Biology*, Vol. 6, No. 3 (2008) page 467-492. (**¹Co-first authors**, Elsevier Scopus impact factor 7.85, 2 citations)
19. Ket Fah Chong, **Kang Ning**, Hon Wai Leong and Pavel Pevzner. Modeling and Characterization of Multi-Charge Mass Spectra for Peptide Sequencing. *Journal of Bioinformatics and Computational Biology*, Vol. 4, No. 6 (2006) page 1329-1352. (Elsevier Scopus impact factor 7.85, 4 citations)
20. **Kang Ning** and Hon Wai Leong. Towards a Better Solution to the Shortest Common Supersequence Problem: The Deposition and Reduction Algorithm. *BMC Bioinformatics*, 7 (Suppl 4): S12 (2006). (SCI impact factor 3.43, 5 citations)
21. **Kang Ning**, Kwok Pui Choi, Hon Wai Leong and Louxin Zhang. A Post Processing Method for Optimizing Synthesis Strategy for Oligonucleotide Microarrays. *Nucleic Acids Research*, Vol. 33, No 17 (2005): e144. (SCI impact factor 7.48, 10 citations)

◆ **Journal Papers submitted**

1. **Kang Ning** and Alexey Nesvizhskii. FIT: Tag based method for Fusion proteins Identification. (submitted to *BMC Bioinformatics*)

◆ Book Chapters

1. **Kang Ning** and Jian Xu. Bioinformatics in Microbial Metagenomics: Status and Prospects. *Industrial Biotechnology 2011*. Science Press (China).

◆ Invited talks

1. Xiaoquan Su, Jian Xu and **Kang Ning**. Parallel-META: A High-Performance Computational Pipeline for Metagenomic Data Analysis. nVIDIA GTC China conference, 2011, Beijing, China.
2. **Kang Ning**. Progress in metagenomics. Monsanto Forum China, 2011, Beijing, China.
3. **Kang Ning**. Recent topics in metagenomics. International Bioinformatics workshop (IBW), 2012, Jilin, China.
4. **Kang Ning**. Comparative Analysis of Different Label-Free Mass Spectrometry Based Protein Abundance Estimates and Their Correlation with RNA-Seq Gene Expression Data. Second China Workshop on Computational Proteomics (CNCP2012), 2012, Beijing, China.
5. **Kang Ning**. Meta-Mesh: Metagenome Database and Data Analysis System. 7th China Industrial Biotechnology Summit, 2013, Tianjin, China.

◆ Peer reviewed conference papers

1. Fang, Wei, Xiaoquan Su, Jian Xu and **Kang Ning**¹. A machine learning framework of functional biomarker discovery for different microbial communities based on metagenomic data. *Systems Biology (ISB), 2012 IEEE 6th International Conference on*. IEEE, 2012. (¹Corresponding author)
2. Xiaoquan Su, Jian Xu and **Kang Ning**¹. An Open-source Collaboration Environment for Metagenomics Research. *IEEE e-Science 2011*, 2011. (¹Corresponding author)
3. Xiaoquan Su, Jian Xu and **Kang Ning**¹. Parallel-META: A High-Performance Computational Pipeline for Metagenomic Data Analysis. *IEEE ISB 2011*, 2011. (¹Corresponding author)
4. **Kang Ning** and Alexey I. Nesvizhskii. The Utility of Mass Spectrometry-based Proteomic Data for Validation of Novel Alternative Splice Forms Reconstructed from RNA-Seq Data: A Preliminary Assessment. *GIW 2010*, 2010.
5. Sriganesh Srihari, **Kang Ning**¹ and Hon Wai Leong. Refining Markov Clustering for Protein Complex Prediction by Core-Attachment. *GIW 2009*, 2009. Genome Informatics Series Vol. 23, page 159-168. (¹Corresponding author, 1 citation)

6. Sriganesh Srihari, Hoong Kee Ng, **Kang Ning** and Hon Wai Leong. Detecting Hubs and Quasi Cliques in Scale-free Networks. *ICPR 2008*, 2008. IEEE Catalog Number: CFP08182; ISBN: 978-1-4244-2175-6; ISSN: 1051-4651. (EI indexed)
7. **Kang Ning**, Hoong Kee Ng and Hon Wai Leong. An Accurate and Efficient Algorithm for Peptide and PTM Identification by Tandem Mass Spectrometry. *GIW 2007*, 2007. Genome Informatics Series Vol. 19, page 119-130.
8. **Kang Ning** and Hon Wai Leong. Algorithm for Peptide Sequencing by Tandem Mass Spectrometry Based on Better Preprocess and Anti-symmetric Computational Model. *CSB 2007*, 2007. Computational Systems Bioinformatics: CSB 2007 Conference Proceedings, page 19-30. (2 citations)
9. Hon Nian Chua¹, **Kang Ning**¹, Wing-Kin Sung, Hon Wai Leong and Limsoon Wong. A Novel algorithm for Protein Complex Prediction based on PPI Networks. *CSB 2007*, 2007. Computational Systems Bioinformatics: CSB 2007 Conference Proceedings, page 97-110. (¹**Co-first authors**)
10. Hoong Kee Ng, **Kang Ning** and Hon Wai Leong. A New Approach for Similarity Queries of Biological Sequences in Databases. *PAKDD 2007*, 2007. Lecture Notes in Computer Science, Vol. 4426, page 728-736. (EI indexed, 1 citation)
11. **Kang Ning**, Ket Fah Chong and Hon Wai Leong. De Novo Peptide Sequencing for Mass Spectra Based on Multi-Charge Strong Tags. *APBC 2007*, 2007. Series on Advances in Bioinformatics and Computational Biology, Vol. 5, page 287-296. (3 citations)
12. **Kang Ning** and Hon Wai Leong. The distribution and deposition algorithm for the multiple oligo nucleotide arrays. *GIW 2006*, 2006. Genome Informatics Vol. 17, No.2, page 89-99. (1 citation)
13. **Kang Ning**, Hoong Kee Ng and Hon Wai Leong. PepSOM: An algorithm for peptide identification by tandem mass spectrometry based on SOM. *GIW 2006*, 2006. Genome Informatics Vol. 17, No.2, page 194-205. (3 citations)
14. **Kang Ning**, Hoong Kee Ng and Hon Wai Leong. Finding Patterns in Biological Sequences by Longest Common Subsequences and Shortest Common Supersequences. *BIBE 2006*, 2006. Sixth IEEE International Symposium on Bioinformatics and BioEngineering (BIBE 2006), ISBN 0-7695-2727-2, page 53-60. (EI indexed, 2 citations)
15. **Kang Ning** and Hon Wai Leong. Towards a Better Solution to the Shortest Common Supersequence Problem: A Post Processing Approach. Proceeding of the First International Multi-symposiums on Computer and Computational Sciences (IMSCCS|06), IEEE Computer Society Press, ISBN 0-7695-2581-4, Vol. 1, page 84-90. (EI indexed, 4 citations)
16. **Kang Ning** and Hon Nian Chua. Automated Identification of Protein Classification and Detection of Annotation Errors in Protein Databases using Statistical Approaches. *KDLL 2006*, 2006. Lecture Notes in Computer Science, Vol. 3886, page 123-138.

17. **Kang Ning**, Ket Fah Chong and Hon Wai Leong. A Database Search Algorithm for Identification of Peptides with Multiple Charges using Tandem Mass Spectrometry. *BioDM 2006*, 2006. Lecture Notes in Computer Science, Vol. 3916, page 2-13. (EI indexed, 3 citations)
18. Ket Fah Chong, **Kang Ning**, Hon Wai Leong and Pavel Pevzner. Characterization of multi-charge mass spectra for peptide sequencing. *APBC 2006*, 2006. Proceedings of 4th Asia-Pacific Bioinformatics Conference, page 109-119. (2 citations)

◆ Posters

1. **Kang Ning** and Alexey Nesvizhskii. Computational Analysis of Unassigned High Quality Spectra from Human T Leukemic Cells. ASMS 2009, 2009.
2. **Kang Ning**, Xia Cao, Hoong Kee Ng, Hon Wai Leong and Alexey Nesvizhskii. Two-phase Filtering Strategy for Identification of Peptide with Post-Translational Modifications. ASMS 2008, 2008.
3. **Kang Ning**, Ket Fah Chong and Hon Wai Leong. De novo Peptide Sequencing for Multi-charge Mass Spectra based on Strong Tags. RECOMB 2006, 2006.
4. **Kang Ning** and Hon Wai Leong. The distribution and deposition method for the multiple oligo nucleotide arrays. RECOMB 2006, 2006.
5. **Kang Ning**, Kwok Pui Choi, Hon Wai Leong and Louxin Zhang. A Comparative Study of the Methods on Effective Synthesis of Oligo Nucleotides. APBC 2005, 2005.

➤ Professional Activities:

- Editorial board member
 - *Journal of Glycobiology*
- Funding reviewer
 - *United Kingdom, Nature Environment Research Council (UK-NERC)*
 - *NSFC (Natural Science Foundation of China)*
- Reviewer for journals:
 - *Nucleic Acid Research* (1 article)
 - *Analytical Chemistry* (1 article)
 - *Proteomics* (3 article)
 - *Bioinformatics* (4 article)
 - *BMC Bioinformatics* (6 articles)
 - *BMC Systems Biology* (2 article)
 - *BMC Genomics* (1 article)
 - *Journal of Bioinformatics and Computational Biology* (3 article)
 - *International Journal of Computational Biology and Drug Design* (1 article)

- *IEEE Spectrum* (1 article)
- *IET Information Security* (2 article)
- *Computers & Operations Research* (1 article)
- *Applied and Environmental Microbiology* (1 article)
- *Expert Review of Proteomics* (1 article)
- *IEEE Transactions on Computational Biology and Bioinformatics* (1 article)
- *International Journal of Computational Biology and Drug Design (IJCBD)* (1 article)
- Reviewer for conferences:
 - *ISMB 2006*
 - *BioDM 2007*
 - *ISMB 2009*
 - *IEEE ISB 2011*
 - *IEEE ISB 2012*

➤ **Membership**

- Member of International Society for Computational Biology (ISCB)
- Member of Institute of Electrical and Electronics Engineers (IEEE)
- Member of American Society for Mass Spectrometry (ASMS)

➤ **Awards and Honors:**

- Chinese Academy of Science, Young Innovative Researcher Award (2011)
- 1st prize project for CAS High-Performance Computation conference (2011)