

# Sangtae Kim

Department of Computer Science and Engineering  
University of California, San Diego  
9500 Gilman Drive, La Jolla, CA

sak008@ucsd.edu  
Tel: (858) 525-5284  
<http://cseweb.ucsd.edu/~sak008>

## Education

---

**University of California, San Diego** Sep. 2006 - Present  
Ph.D. in Computer Science

Dissertation: Generating Functions of Tandem Mass Spectra and Their Applications to Peptide Identifications  
Advisor: Pavel Pevzner

**Seoul National University, Seoul, Korea** Mar. 2000 - Feb. 2002  
Master in Computer Science

Dissertation: An Efficient Algorithm to Find All Normalized Local Alignments Above a Threshold  
Advisor: Kunsoo Park

**Seoul National University, Seoul, Korea** Mar. 1996 - Feb. 2000  
Bachelor in Computer Science

## Professional Experience

---

**Graduate Student Researcher** Sep. 2006 - Present  
Department of Computer Science and Engineering  
University of California, San Diego

**Full-time Researcher** Jul. 2005 - Aug. 2006  
Research Center for Computer Technology  
Hanyang University, Seoul, Korea

**Full-time Instructor** Apr. 2002 - Jun. 2005  
Department of Computer Science  
Korea Military Academy, Seoul, Korea

- Served as an officer in the Republic of Korea army for mandatory military service.

## Publications

---

### Refereed Journal Papers

- Kyowon Jeong, **Sangtae Kim**, Nuno Bandeira and Pavel Pevzner, Gapped Spectral Dictionaries and Their Applications for Database Searches of Tandem Mass Spectra, **Molecular & Cellular Proteomics**, 10, M110.002220, 2011.
- Sangtae Kim**, Nikolai Mischerikow, Nuno Bandeira, J. Daniel Navarro, Louis Wich, Shabaz Mohammed, Albert J. R. Heck and Pavel Pevzner, The Generating Function of CID, ETD and CID/ETD Pairs of Tandem Mass Spectra: Applications to Database Search, **Molecular & Cellular Proteomics**, 9, 2840-2852, 2010.
- Sangtae Kim**, Nuno Bandeira and Pavel Pevzner, Spectral Profiles: A Novel Representation of Tandem Mass Spectra and its Applications for de Novo Peptide Sequencing and Identification, **Molecular & Cellular Proteomics**, 8, 1391-1400, 2009. (*Highlighted on the cover of the June 2009 issue of Molecular & Cellular Proteomics.*)

- **Sangtae Kim**, Nitin Gupta, Nuno Bandeira and Pavel Pevzner, Spectral Dictionaries: Integrating De Novo Peptide Sequencing with Database Search of Tandem Mass Spectra, **Molecular & Cellular Proteomics**, 8, 53-69, 2009.
- Pavel Pevzner, **Sangtae Kim** and Julio Ng, Comment on "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry", **Science**, 321 (5892), 1040, 2008.
- **Sangtae Kim**, Nitin Gupta and Pavel Pevzner, Spectral Probabilities and Generating Functions of Tandem Mass Spectra: A Strike against Decoy Databases. **Journal of Proteome Research**, 7, 3354-3363, 2008. (Noted as one of the 20 most cited articles of the journal in recent 3 years.)
- **Sangtae Kim**, Seungjin Na, Ji Woong Sim, Heejin Park, Jaeho Jeong, Hokeun Kim, Younghwan Seo, Jawon Seo, Kong-Joo Lee and Eunok Paek, MODi : A Powerful and Convenient Web Server for Identifying Multiple Post-translational Peptide Modifications from Tandem Mass Spectra, **Nucleic Acids Research**, 34, W258-W263, 2006.
- **Sangtae Kim**, Jeong Seop Sim, Heejin Park, Kunsoo Park, Hyunseok Park and Jeong-Sun Seo, A Heuristic Algorithm to Find All Normalized Local Alignments Above Threshold, **Genomics & Informatics**, 1, 26-32, 2003.

### Refereed Conference Papers

- Kyowon Jeong, **Sangtae Kim**, Nuno Bandeira and Pavel Pevzner, Gapped Spectral Dictionaries and Their Applications for Database Searches of Tandem Mass Spectra, In Proceedings of the Fourteenth International Conference on Research in Computational Molecular Biology (RECOMB-2010), 208-232, Aug. 2010.

### Working Papers

- **Sangtae Kim** and Pavel Pevzner, MS-GF+: A Sensitive and Adaptive Database Search Engine for Mass Spectrometry, in preparation.
- To-ju Huang, Claudiu Farcas, Jeremy Carver, Natalie Castellana, Ari Frank, **Sangtae Kim**, Jian Wang, Xiaowen Liu, Pavel A. Pevzner, Vineet Bafna, Ingolf Krüger and Nuno Bandeira, ProteoSAFE: A Scalable, Accessible, and Flexible Software Environment for Proteomics Analysis, in preparation.
- **Sangtae Kim** and Nuno Bandeira, False Discovery Rates in Spectral Identification: a Target- decoy approach, submitted.

## Selected Presentations

---

### Conference Oral Presentations

- Kyowon Jeong, **Sangtae Kim**, Nuno Bandeira and Pavel Pevzner, Gapped Spectral Dictionaries and Their Applications for Database Searches of Tandem Mass Spectra, 14th International Conference on Research in Computational Molecular Biology (RECOMB 2010), Lisbon, Portugal, Aug. 2010.
- **Sangtae Kim**, Nuno Bandeira, Pavel Pevzner, Spectral Profiles: A Novel Representation of Tandem Mass Spectra and its Application for Gapped Peptide Generation, 57th American Society for Mass Spectrometry Conference (ASMS 2009), Philadelphia, USA, June 2009.
- **Sangtae Kim**, Nuno Bandeira and Pavel Pevzner, Spectral Profiles: A Novel Representation of Tandem Mass Spectra and its Applications for De Novo Peptide Sequencing and Identifications, 5th Annual US HUPO Conference (US-HUPO 2009), San Diego, USA, Feb. 2009.
- **Sangtae Kim**, Nitin Gupta and Pavel Pevzner, Evaluating Statistical Significance of Peptide Identifications Using the Partition Function of Tandem Mass Spectra, HUPO 6th Annual World Congress, Seoul (HUPO 2007), Korea, Oct. 2007.

## Tutorials

- **Sangtae Kim**, False Discovery Rate of Peptide Identifications: Target-Decoy Approach and Generating Function Approach, RECOMB Satellite Conference on Computational Mass Spectrometry (RECOMB-CP 2011), San Diego, USA, Mar. 2011.
- **Sangtae Kim**, Assigning Statistical Significance to Peptide/Protein Identifications, RECOMB Satellite Conference on Computational Mass Spectrometry (RECOMB-CP 2010), San Diego, USA, Mar. 2010.

## Invited Talks

- MS-GF for High-Precision Mass Spectra, Pacific Northwest National Laboratory (PNNL), Richland, USA, Oct. 2010.
- Generating Functions of Tandem Mass Spectra: Estimating Error Rates of Peptide Identifications and their Applications to Database Search, Korea Institute of Science & Technology (KIST), Seoul, Korea, June 2010.
- Bioinformatics: How to Apply Computer Algorithms to Solve Biological Problems?, Korea University, Seoul, Korea, June 2010.

## Poster Presentations

ASMS 2011, ASMS 2010, ASMS 2008, La Jolla Proteomics Conference 2008, UCSD Research Expo 2008, HUPO 2007, Algorithmic Biology 2006, HUPO 2005, ISMB 2005, HUPO 2004, ISMB 2004, RECOMB 2002.

## Teaching Experience

---

### Instructor

Aug. 2002 – Jun. 2005

Department of Computer Science  
Korea Military Academy, Seoul, Korea

- Introduction to Computers
- Programming Language (C)
- Computer Networks

### Teaching Assistant

Mar. 2009 – Jun. 2009

Department of Computer Science  
University of California, San Diego

- Molecular Sequence Analysis

### Teaching Assistant

Sep. 2000 – Dec. 2000

Department of Computer Science and Engineering  
Seoul National University, Seoul, Korea

- Automata Theory

## Professional Service

---

### Program Committee Member

- RECOMB Satellite Conference on Computational Proteomics (RECOMB-CP 2010), Mar. 2010.
- International Workshop on Bio-inspired computing for Hybrid Information Technology (BHIT 2010), Dec. 2010.

## Organizing Committee Chair

- RECOMB Satellite Conference on Computational Proteomics (RECOMB-CP 2012), Apr. 2012.
- RECOMB Satellite Conference on Computational Proteomics (RECOMB-CP 2010), Mar. 2010.
- RECOMB Satellite Conference on Bioinformatics Education (RECOMB-BE 2010), March 2010.

## Organizing Committee Member

- RECOMB Satellite Conference on Bioinformatics Education (RECOMB-BE 2009), March 2009.

## Reviewer

- Journal of Proteome Research
- Journal of Proteomics
- PLOS Computational Biology
- Advances in Bioinformatics
- 15<sup>th</sup> International Conference on Research in Computational Molecular Biology (RECOMB 2011)
- 10<sup>th</sup> Workshop on Algorithms in Bioinformatics (WABI 2010)

## Honors & Awards

---

- Distinguished Graduate Student Scholarship, Chang-Hyun Foundation for Education, 2000-2002
- Merit-based Scholarship, Seoul National University, 1998, 1999.
- Admission Scholarship, Seoul National University, 1996.

## Reference

---

Pavel Pevzner

Professor of Computer Science and Engineering University of California, San Diego

(858) 822-4365

ppevzner@ucsd.edu

Vineet Bafna

Professor of Computer Science and Engineering University of California, San Diego

(858) 822-4978

vbaafna@cs.ucsd.edu

Nuno Bandeira

Assistant Professor of Computer Science and Engineering, Skaggs School of Pharmacy and Pharmaceutical Sciences

University of California, San Diego

(858) 534-8666

bandeira@ucsd.edu

Eunok Paek

Professor of Mechanical and Information Engineering

University of Seoul

+82 (2) 2210-5310

paek@uos.ac.kr