Curriculum Vitae

Sangsoo Lim (임상수)

Postdoctoral Research Associate

Bio and Health Informatics Lab, Bioinformatics Institute
Seoul National University, Seoul, Republic of Korea
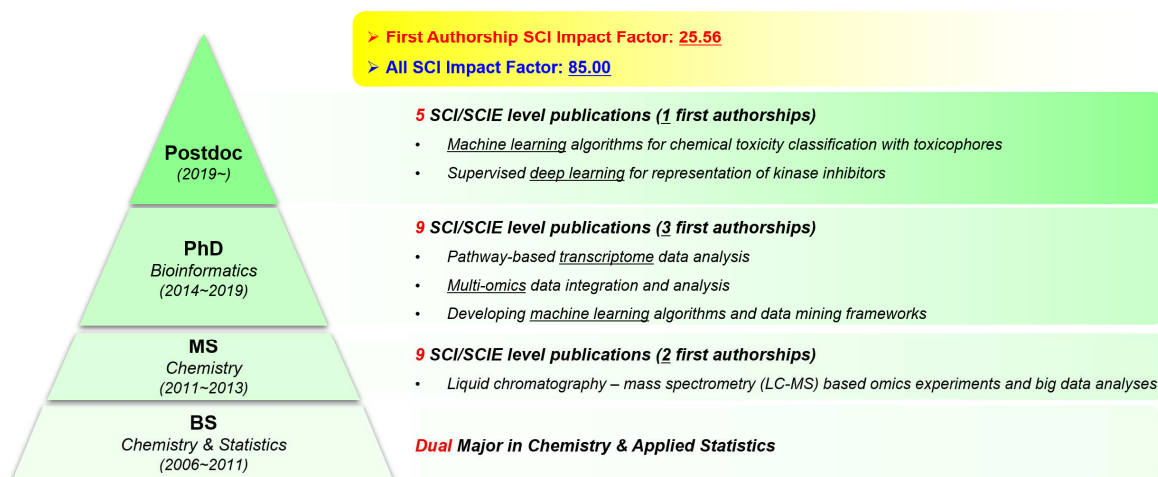**Phone**   010-3371-3078
E-mail   sslim@snu.ac.kr

Profile summary

I am currently a <u>postdoctoral research associate</u> affiliated with Bio & Health Informatics Lab and Bioinformatics Institute at Seoul National University. I received a <u>PhD degree in Bioinformatics</u> at Seoul National University in 2019 and <u>MS degree in Chemistry</u> at Yonsei University in 2013. Having **experienced in both experimental and computational fields of study**, I am currently working on developing machine learning algorithms to model drug-induced liver toxicity of chemical compounds by using graph learning techniques such as supervised Markovian Walk with graph pruning. I am also deeply involved in projects such as representation learning of FDA approved drugs by regularized cascading autoencoders and developing a supervised graph-based Apriori algorithm for frequent subgraph mining.

Education

2014 – 2019	**Ph.D.**, Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Korea

- Research areas: transcriptome data analysis, data mining, computational modeling of biological data
- Advisor: *Sun Kim*
- Thesis title: Quantification of pathway activity using RNA-seq data

2011 – 2013	**M.S.** in Chemistry, Yonsei University, Seoul, Republic of Korea

- Research areas: Lipidomics, LC-MS
- Advisor: *Myeong Hee Moon*
- Thesis title: Development of computational algorithm for structural identification of phospholipids and profiling of phospholipid biomarkers for prostate cancer from human urine

2006 – 2011	**B.S.** in Chemistry & Applied Statistics, Yonsei University, Seoul, Republic of Korea
2003 – 2005	Daeil Foreign Language High School

Professional experiences

2019 – Present	**Postdoc. Research Associate**, Bioinformatics Institute, Seoul National University, Seoul, Republic of Korea

- Research areas:
  cheminformatics – drug toxicity prediction, chemical space representation
  pharmacogenomics – multi-omics integration for drug response prediction

Teaching experiences

2021 Spring	IT Fundamentals for Bioinformatics @SNU (head TA)

- Role: Managing Lecture Materials, Practices, and Assignments

2020 Fall	Bioinformatics 2 @SNU (Lecturer for practice)

- Role: Overview and Practices of ML techniques in Bioinformatics

2017 Fall	Bioinformatics 2 @SNU (Lecturer for practice)

- Role: Finding differentially expressed genes using DEseq package

2016 Fall	Bioinformatics 2 @SNU (Lecturer for practice)

- Role: Finding differentially expressed genes using DEseq package

2016 Spring	IT Fundamentals for Bioinformatics @SNU (Lecturer for practice)

- Role: Data mining in WEKA

2015 Fall	Computer Convergence Applications @SNU (Teaching Assistant)

- Role: Lecture preparation

2012 Fall	General Chemistry Experiments @Yonsei University (Teaching Assistant)

- Role: Managing students' experiment: "Precipitation of Sodium Oxalate using potassium permanganate"
- *Excellence in TA awardee*

2012 Spring	Analytical Chemistry Experiments @Yonsei University (Teaching Assistant)

	- Role: Managing students' experiment: "Separation of alkanes using gas/liquid chromatography"
	- *Excellence in TA awardee*
2011 Spring	<u>Analytical Chemistry</u> @Yonsei University (Teaching Assistant)
	- Role: Lecture preparation

Ongoing Projects

1. <u>Drug-Induced Liver Injury (DILI) Classification using Supervised Random Walk with Graph Pruning</u>

Summary	This work is to classify the liver toxicity information of small molecule drugs with toxicophores using iteratively supervised Markovian walk with subgraph pruning. By iteratively generating and pruning subgraphs, classification of a set of toxic and non-toxic chemical compounds was significantly improved compared to the previous algorithms. This model also provides informative substructures that contribute the most to chemical toxicity.
Role	Overall study design, model implementation, evaluation, interpretation and manuscript writing
Advisor	*Sun Kim*
Status	Manuscript writing

2. <u>An ensemble approach to Compound-Protein Interaction Prediction of Kinase Inhibitors</u>

Summary	This project is to make better classification of kinase inhibitor interactions to protein targets by building an ensemble model of previous tools. The aim of the project is to make encoding and architectural characteristics of the existing tools complement each other to improve data interpretation.
Role	Overall study design, evaluation and interpretation
Advisor	*Sun Kim*
Status	Model implementation

3. <u>Development of a Cascade Autoencoder Model for Differential Representation of FDA Approved Drugs</u>

Summary	While new drugs are released and approved by US FDA, some drugs are discontinued or withdrawn from the market in many reasons. This project is to model the approved /discontinued/withdrawn drugs by devising a supervised representation model in terms of FDA approval and target disease types.
Role	Overall study design, evaluation and interpretation
Advisor	*Sun Kim*
Status	Model implementation

4. <u>Supervised Graph Pattern Mining of Drug-Induced Liver Injury (DILI) Compounds using Apriori Algorithm for Pharmacophore Discovery</u>

Summary	Compound-protein interaction occurs in specific regions of both chemicals and proteins called pharmacophores and pockets, respectively. This project is to learn characteristics and substructures of the DILI compounds that are enriched to liver toxicity. This will

	help understand why liver toxicity occurs and what to avoid when developing new drugs.
Role	Overall study design, evaluation and interpretation
Advisor	*Sun Kim*
Status	Model implementation

5. Query Graph Search on Chemical Graphs using Dynamic Programming on Integrated Directed Acyclic Graph

Summary	Chemical compounds are diverse and theoretically infinite structures are possible. Finding a specific substructure from a given compound is thereby complicated especially when the compound is large. This work is to develop an algorithm for searching a query graph against the set of chemical compounds based on dynamic programming with graph pruning.
Role	Overall study design, evaluation and interpretation
Advisor	*Sun Kim*
Status	Model implementation

6. Classification of Kinase Inhibitor Subgroups

Summary	This project is to model kinase inhibitor properties using deep learning based chemical compound embedding methods and target protein information. The model will be useful in designing kinase inhibitors and predicting pharmacological effectiveness and selectivity.
Role	Overall study design, evaluation and interpretation
Advisor	*Sun Kim*
Status	Model implementation

Awards

2012	Excellent Teaching Assistant (Department of Chemistry, Yonsei University)
2011	Excellent Teaching Assistant (Department of Chemistry, Yonsei University)

Publications

Total 1st author SCI papers:	6 (Impact Factor: 25.56)
Total SCI papers:	23 (Impact Factor: 85.00)

[At Seoul National University]

1. Sungjoon Park, Dohoon Lee, Youngkuk Kim, **Sangsoo Lim**, Heejoon Chae and Sun Kim
   "BioVLAB-Cancer-Pharmacogenomics: Tumor Heterogeneity and Pharmacogenomics Analysis of Multi-omics Data from Tumor on the Cloud", *Bioinformatics* Under review, 2021 **[IF: 5.61]**

2. Minsu Kim, Sangseon Lee, **Sangsoo Lim**, Doh Young Lee and Sun Kim
   "Subnetwork representation learning for discovering network biomarkers in predicting lymph node metastasis in early oral cancer", *Bioinformatics* Under review, 2021 **[IF: 5.61]**

3. Dabin Jeong, **Sangsoo Lim**, Sangseon Lee, Minsik Oh, Changyun Cho, Hyeju Seong, Woosuk Jung, Sun Kim

"Construction of Condition-Specific Gene Regulatory Network using Kernel Canonical Correlation Analysis", *Frontiers in Genetics* In Press, 2021 **[IF: 3.79]**

4. Yoon Jeong Park, Sangseon Lee*, **Sangsoo Lim***, Hahn Nahmgoong, Yul Ji, Jin Young Huh, Assim A. Alfadda, Sun Kim, and Jae Bum Kim
"DNMT1 maintains metabolic fitness of adipocytes through acting as an epigenetic safeguard of mitochondrial dynamics", *Proceedings of the National Academy of Sciences* 118(11), 2021 **[IF: 9.41]**

5. **Sangsoo Lim***, Yijinxiu Lu*, Chang Yun Cho, Inyoung Sung, Jungwoo Kim, Youngkuk Kim, Sungjoon Park and Sun Kim
"A review on compound-protein interaction prediction methods: Data, format, representation and model", *Computational and Structural Biotechnology Journal* 19:1541, 2021 **[IF: 6.02]**

6. Minsik Oh, Sungjoon Park, Sangseon Lee, Dohoon Lee, **Sangsoo Lim**, Dabin Jeong, Kyuri Jo, Inuk Jung and Sun Kim
"DRIM: A web-based system for investigating drug response at the molecular level by condition-specific multi-omics data integration", *Frontiers in Genetics* 11:564792, 2020 **[IF: 3.79]**

7. Sangseon Lee, **Sangsoo Lim**, Taeheon Lee, Inyoung Sung and Sun Kim
"Cancer subtype classification and modeling by pathway attention and propagation", *Bioinformatics*, In press, 2020 **[IF: 5.61]**

8. **Sangsoo Lim**, Sangseon Lee, Inuk Jung, Sungmin Rhee and Sun Kim
"Comprehensive and critical evaluation of individualized pathway activity measurement tools on pan-cancer data", *Briefings in Bioinformatics* 21(1), 2020 **[IF: 8.99]**

9. Aeran Lim[*], **Sangsoo Lim**[*] and Sun Kim
"Enhancer prediction with histone modification marks using a hybrid neural network model", *Methods* 166, 2019 **[IF: 3.78]**

10. Minsu Kim, Sangseon Lee, **Sangsoo Lim** and Sun Kim
"SpliceHetero: An information theoretic approach for measuring spliceomic intratumor heterogeneity from bulk tumor RNA-seq", *PloS one* 14(10), 2019 **[IF: 2.74]**

11. Ji Hwan Moon, **Sangsoo Lim**, Kyuri Jo, Sangseon Lee, Seokjun Seo and Sun Kim
"PINTnet: construction of condition-specific pathway interaction network by computing shortest paths on weighted PPI", *BMC systems biology* 11(2), 2017 **[IF: 2.05]**

12. Sungmin Rhee, **Sangsoo Lim** and Sun Kim
"Iterative segmented least square method for functional microRNA-mRNA module discovery in breast cancer", *International Journal of Data Mining and Bioinformatics,* 2017 **[IF: 0.77]**

13. Benjamin Hur, **Sangsoo Lim**, Heejoon Chae, Seokjun Seo, Sangseon Lee, Jaewoo Kang and Sun Kim
"CLIP-GENE: a web service of the condition specific context-laid integrative analysis for gene prioritization in mouse TF knockout experiments", *Biology direct* 11(1), 2016 **[IF: 2.19]**

14. Youngjune Park, **Sangsoo Lim**, Jun-Wu Nam and Sun Kim
"Measuring intratumor heterogeneity by network entropy using RNA-seq data", *Scientific reports* 6, 2016 **[IF: 4.00]**

15. **Sangsoo Lim**, Youngjune Park, Benjamin Hur, Minsu Kim, Wonshik Han and Sun Kim
"Protein interaction network (PIN)-based breast cancer subsystem identification and activation measurement for prognostic modeling", *Methods* 110, 2016 **[SCI; IF: 3.78]**

16. Jinwoo Park, Benjamin Hur, Sungmin Rhee, **Sangsoo Lim**, Minsu Kim, Kwangsoo Kim, Wonshik Han

and Sun Kim

"Information theoretic sub-network mining characterizes breast cancer subtypes in terms of cancer core mechanisms", *Journal of bioinformatics and computational biology* 14(5), 2016 **[SCIE; IF: 0.90]**

[At Yonsei University]

17. **Sangsoo Lim**, Dae Young Bang, Koon Ho Rha and Myeong Hee Moon
"Rapid screening of phospholipid biomarker candidates from prostate cancer urine samples by multiple reaction monitoring of UPLC-ESI-MS/MS and statistical approaches", *Bulletin of the Korean Chemical Society* 35(4), 2014 **[IF: 0.61]**

18. Ju Yong Lee, **Sangsoo Lim**, Sungha Park and Myeong Hee Moon
"Characterization of oxidized phospholipids in oxidatively modified low density lipoproteins by nanoflow liquid chromatography-tandem mass spectrometry", *Journal of Chromatography A* 1288, 2013 **[IF: 4.17]**

19. Ki Hun Kim, Ju Yong Lee, **Sangsoo Lim** and Myeong Hee Moon
"Top-down lipidomic analysis of human lipoproteins by chip-type asymmetrical flow field-flow fractionation-electrospray ionization-tandem mass spectrometry", *Journal of Chromatography A* 1280, 2013 **[IF: 4.17]**

20. Seul Kee Byeon, Ju Yong Lee, **Sangsoo Lim**, Donghoon Choi and Myeong Hee Moon
"Discovery of candidate phospholipid biomarkers in human lipoproteins with coronary artery disease by flow field-flow fractionation and nanoflow liquid chromatography-tandem mass spectrometry", *Journal of Chromatography A* 1270, 2012 **[IF: 4.17]**

21. **Sangsoo Lim**, Seul Kee Byeon, Ju Yong Lee and Myeong Hee Moon
"Computational approach to structural identification of phospholipids using raw mass spectra from nanoflow liquid chromatography-electrospray ionization-tandem mass spectrometry", *Journal of mass spectrometry* 47(8), 2012 **[IF: 2.38]**

22. Dae Yong Bang, **Sangsoo Lim** and Myeong Hee Moon
"Effect of ionization modifiers on the simultaneous analysis of all classes of phospholipids by nanoflow liquid chromatography/tandem mass spectrometry in negative ion mode", *Journal of Chromatography A* 1240, 2012 **[IF: 4.17]**

23. Rae Ung Jeong, **Sangsoo Lim**, Myoung Ok Kim and Myeong Hee Moon
"Effect of D-allose on prostate cancer cell lines: phospholipid profiling by nanoflow liquid chromatography-tandem mass spectrometry", *Analytical and bioanalytical chemistry* 401(2), 2011 **[IF: 3.29]**

24. Ju Yong Lee, **Sangsoo Lim** and Myeong Hee Moon
"Effects of column lengths and particle diameter on phospholipid analysis by nanoflow liquid chromatography-electrospray ionization-mass spectrometry", *Mass Spectrometry Letters* 2(3), 2011 **[IF: 0.94]**

25. Hye Kyeong Min, **Sangsoo Lim**, Bong Chul Chung and Myeong Hee Moon
"Shotgun lipidomics for candidate biomarkers of urinary phospholipids in prostate cancer", *Analytical and bioanalytical chemistry* 399(2), 2011 **[IF: 3.29]**

* Equal Contribution