

HoJoon Lee

hojoon@stanford.edu

286 Mosher Way, Palo Alto, CA 94304 • (650) 862-1286

Current Position: Senior Research Engineer, Med/Oncology, Stanford University

EDUCATION:

Post-doctoral Training, Med/Oncology (2017) **Stanford University, Stanford, CA**

Ph.D., Molecular & Cellular Biology with Bioinformatics specialty (2012)

Arizona State University, Tempe, AZ

- Thesis: Identification of Neo-antigens for a Cancer Vaccine by Transcriptome Analysis

Master of Science, Computational Biosciences (2005)

Arizona State University, Tempe, AZ

- Thesis: Estimate of Genomewide Mutation Rate Difference between Human and Mouse

Bachelor of Science, Biology (2002)

Yonsei University, Seoul, Korea

RESEARCH GRANTS:

2017 The National Cancer Institute (NCI) Seven Bridges Cancer Genomics Cloud pilot (HHSN261201400008C)

PROFESSIONAL EXPERIENCE:

Senior Research Engineer

School of Medicine, Stanford University

March 2019 - Current

- Manage bioinformatics analysis pipeline for all genomics data
- Lead a team to develop bioinformatics analysis pipeline for single cell immunogenomics data
- Lead a team to develop dynamic representing the human reference genomes, which enable population level sequencing analysis
- Lead a team to develop new cancer immune therapy targets (under IP process)

Project Leader

School of Medicine, Stanford University

March 2017 – March 2019

- Developed bioinformatics pipeline to analyze single cell sequencing of T cell receptors (TCRs)
- Developed bioinformatics pipeline to identify personalized neo-antigens for clinical phase 1 trial of immune therapy by pLADD. (**licensing to Aduro**)
- Developed new algorithms to analyze sequencing reads for substitution, indels, gene fusions, copy number variation, and Cas9 mutagenesis
- Developed analysis pipeline for whole genome sequencing from clinical samples with Intermountain Healthcare

Post-doctoral fellow, Hanlee Ji

School of Medicine, Stanford University

March 2012 – March 2017

- Set up bioinformatics tools on Amazon Web Service through Seven Bridge for immune-genomics analysis of the Cancer Genome Atlas (TCGA)
- Identified clinically relevant genomic/proteomic changes from >10,000 samples of > 32 cancers in TCGA by integrative analysis.
- Developed a web portal for the exploration of the clinical associations of the TCGA data; <http://genomeportal.stanford.edu/pan-tcga>
- Designed the optimal probes for targeted sequencing such as STR-OS seq and digital droplet PCR

Research Associate, Stephen A. Johnston**The Biodesign Institute, ASU**

August 2005 – February 2012

Worked in the cancer vaccine project that funded by Department of Defense (DoD) and Keck grant.

- Developed algorithm to identify tumor-specific frame-shifted mutations derived from gene fusions, alternative splicing and insertion/deletion as neo antigens that could be used in vaccine.
- Validated these putative candidates by molecular biology such as RT-PCR and cloning. Validated candidates were tested in mouse model.
- Constructed database to organize all data using mysql with all available information such as epitopes, MHC binders, gene expression, exon structure, GO annotation and etc to evaluate them as vaccine antigens.

Research Associate, Sudhir Kumar**The Biodesign Institute, ASU**

August 2003 – August 2005

- Worked in neutral molecular evolution project by estimating the neutral substitution rate from repeat elements of human and mouse genome.
- Identified orthologs between human and mouse by using reciprocal blast and estimated divergence of orthologous repeats by several difference approaches.
- Developed sequence assembly algorithm using k-mer, which was the first in the world and adopted by Velvet.

Internship**Arizona Technology Enterprises**

September 2006 – December 2006

- Hand-on experiences by participating real cases at AzTE with university inventors and industry
- Learned about technology assessment, intellectual property analysis, market assessment, and commercialization strategy.

Military Service

Served as Korean Augmentation to US Army (KATUSA)

January 1999 – March 2001

Duty: Military Police

COMPUTATIONAL SKILLS AND TECHNIQS:

Programming languages: Python, R, Perl, SQL, Matlab and Docker

Comfortable working in Linux, Window, iOS, GitHub, Amazon Web Service (AWS) and Google Cloud

Next Generation Sequencing (NGS) analysis

- Variant calling: BWA, GATK, Mutect2, Samtools, Picard
- Structural Variation: CNVkit, TopHat-Fusion
- Transcriptome: STAR, HTseq, Kalisto
- Intra-tumor heterogeneity: Expands
- Immunogenomics: NetMHC, OptiType, Cell Ranger VDJ
- Single Cell Analysis: Cell Ranger (10X Genomics), Seurat
- Molecular evolution: MUMMER, MEGA, BLAST, BLAT

GitHub Page: <https://github.com/compbio>**AWARDS AND FELLOWSHIPS:**2019 2nd place in Stanford HealthAI hackathon

2015 Travel grants for AACR 2015 annual meeting

2014 Best poster award in Oncology retreat (Stanford University)

2009 Travel fellowships of School of Life Sciences (Arizona State University)

2008 GPSA Conference Travel Grant (Arizona State University)

2007 Travel Fellowship, 2007 RECOM Computational Cancer Biology

2007 Travel fellowships of School of Life Sciences (Arizona State University)

2006 Dr. John and Rose Maher Alumni Scholarship for students involved in cancer research

2005 Travel Fellowship, 2005 International Society for Computational Biology (ISMB) meeting

TEACHING EXPERIENCE:**Teaching Undergraduates**

- Biocore Explorations: F14-BIO-41-01 in Stanford University (2016, 2015, 2014)

Mentoring

- Post-doctoral fellow in Ji Lab: Dmitri Pavlichin (2018~)
- Medical Students in Ji Lab: Anjali Jotwani (2018~)
- Stanford **undergraduate internship** in Ji lab: Ahmed (2017~), Alex Fu (2016), Jacob Parker (2014~2016), Ambika Acharya (2013)
- ADVANCE journal club for incoming **bioscience graduates**, Stanford University (2014, 2013)
- ASU Biodesign summer internship for **high school students**: Carrie Lin (2010), Samuel Hooke (2007) and Anita Lin (2007)
- ASU Biodesign lab rotation of 1st year **Ph.D. student**: Muskan Kukreja (2008)

PATENTS:

High Resolution STR analysis using Next Generation Sequencing (co-inventor) – US PTO 62/200,904
 CRISPRpic (co-inventor) – under review
 Super epitopes for cancer immune therapy (co-inventor) – under review

AFFILIATIONS:

American Association for Cancer Research (AACR)	Associate member	2007 – Present
American Association for the Advancement of Science (AAAS)	Member	2006 – Present
Society for Molecular Biology and Evolution (SMBE)	Member	2006 – Present
International Society for Computational Biology (ISCB)	Member	2005 – Present
Korean Life Scientists in the bay (KOLIS)	Stanford Representative	2014
Korean Student Association at ASU	President	2004 July –2005 June

PUBLICATIONS:

1. Journal Publications

- Seung Woo Cho*, **HoJoon Lee***, Howard Y Chang, Hanlee P Ji “**CRISPRpic: Fast and precise analysis for CRISPR-induced mutations via prefixed index counting**” bioRxiv 416768; <http://dx.doi.org/10.1101/416768> (2018). * co-first author
- GiWon Shin, Stephanie U Greer, Li C Xia, **HoJoon Lee**, Jun Zhou, T Christian Boles, Hanlee P Ji “**Assembly of Mb-size genome segments from linked read sequencing of CRISPR DNA targets**” bioRxiv 373142; doi: <https://doi.org/10.1101/373142> (2018)
- Li Charlie Xia, Dongmei Ai, **HoJoon Lee**, Noemi Andor, Chao Li, Nancy R. Zhang, Hanlee P. Ji “**SVEngine: an efficient and versatile simulator of genome structural variations with features of cancer clonal evolution**” Gigascience, 7(7), 2018.
- Christina Wood-Bouwens, Billy T. Lau, Christine M. Handy, **HoJoon Lee**, Hanlee P. Ji “**Single-Color Digital PCR Provides High-Performance Detection of Cancer Mutations from Circulating DNA**” J Mol Diagn 2017, 19: 697

- GiWon Shin, Susan M. Grimes, **HoJoon Lee**, Billy T. Lau, Charlie Xia and Hanlee P. Ji “**CRISPR/Cas9 targeted fragmentation and selective sequencing enable massively parallel microsatellite analysis**” Nat Commun. 2017 8:14291
- **HoJoon Lee**, Jennifer Palm, Susan M. Grimes, Hanlee P. Ji “**The Cancer Genome Atlas Clinical Explorer: A Web and Mobile Interface for Identifying Clinical – Genomic Driver Associations**” Genome Medicine 2015, 7:112
- **HoJoon Lee**, Patrick Flaherty, Hanlee P. Ji “**Systematic genomic identification of colorectal cancer genes delineating advanced from early clinical stage and metastasis**” BMC Med Genomics 2013, 6:54
- Phillip Stafford and **HoJoon Lee** “**Biological Data Centres Rev. 2.0**” Encyclopaedia of Life Sciences (ELS), 2009
- Maulik K. Shah, **HoJoon Lee**, Stephanie A. Rogers, Jeffrey W. Touchman. "An Exhaustive Genome Assembly Algorithm Using K-Mers to Indirectly Perform N-Squared Comparisons in O(N)" *csb*, pp. 740-741, 2004 IEEE Computational Systems Bioinformatics Conference (CSB'04), 2004.

2. Book Chapters

- **HoJoon Lee***, Billy T. Lau*, and Hanlee p. Ji (2013) “**Targeted Sequencing Strategies in Cancer Analysis**” in: Next Generation Sequencing in Cancer Research (Springer) pp 137-163 * co-first author

3. Invited Talks

- **Analysis of cancer genomes for personalized immuno-therapy in clinical phase 1 trial** Annual Translational Bioinformatics Conference, Seoul, Korea, November, 2018
- **Genomic and clonal characterization of immunogenic epitopes across the landscape of the Cancer Genome Atlas (TCGA)** Stanford Oncology and Hematology Annual Research Retreat, Asilomar, CA, August, 2017
- **Epitope Landscape in Stomach Cancer: from Whole Exome Sequencing in TCGA** Avison Biomedical Symposium Cancer Immunotherapy, Yonsei University College of Medicine, Seoul, Korea, August, 2015
- **An Integrative Genomic Analysis for Clinically Relevant Cancer Genetic Aberration and Targeted Therapeutic Prediction** Bimod Seminar, Stanford University, Stanford, April, 2015
- **Systematic Genomic Discovery of Clinical Relevant Gene Aberrations across Twelve Different Cancers** Seoul National University College of Medicine, Seoul, Korea, January, 2014
- **Integrated Genomic Meta-analysis of Colorectal Cancer by Elastic-net** Ewha Research Center for Systems Biology, Seoul, Korea, January, 2013
- **Identification of Tumor-Specific Mutation in Coding Microsatellite DNAs by Using EST Data** RECOM Computational Cancer Biology, University of California, San Diego, September, 2007

4. Poster Presentations

- HoJoon Lee*, Stephanie Greer, Hanlee P. Ji (2018) “Entire landscape of epitopes from all possible missense mutations in coding sequence” AACR Tumor Immunology and Immunotherapy, Miami, USA Nov 27-30
- HoJoon Lee*, Stephanie Greer*, Raunaq Malhotra, Hanlee P. Ji (2018) “Mapping the comprehensive landscape of missense-mutation neoantigens across the human genome” American Association for Cancer Research (AACR) Annual Meeting Chicago, USA April 14-18 * co-first author
- HoJoon Lee*, Li Charlie Xia*, Stephanie Greer, John Bell, Sue Grimes, Noemi Andor, Lincoln Nadauld, Hanlee Ji, Paul Van Hummelen (2018) “High-quality CNV segments from low-throughput whole genome sequencing

from FFPE cancer biopsies base on an evaluation of multiple CNV tools on TCGA data” American Association for Cancer Research (AACR) Annual Meeting Chicago, USA April 14-18 * co-first author

- HoJoon Lee, Stephanie Greer, Sue M. Grimes, Jae-Ho Cheong, Hanlee P. Ji (2017) “Genomic and clonal characterization of immunogenic epitopes across the landscape of the Cancer Genome Atlas” American Association for Cancer Research (AACR) Annual Meeting Washington, D.C., USA April 1-5
- HoJoon Lee, Jacob J. Parker, John Bell, Hanlee P. Ji (2016) “Dynamic genomic indexing enables accurate somatic variant detection from cancer genome sequencing without sequence alignment limitations” American Association for Cancer Research (AACR) Annual Meeting New Orleans, USA April 16-20
- HoJoon Lee, Jennifer Palm, Hanlee P. Ji (2015) “The Stanford-TCGA Portal: An Interactive Web/Mobile Interface for Exploring the Clinical Phenotypic Relevance of Specific Cancer Drivers” American Association for Cancer Research (AACR) Translation of the Cancer Genome San Francisco, USA February 7-9
- HoJoon Lee, Hanlee P. Ji (2014) “An Integrative Genomic Analysis for Clinically Relevant Cancer Genetic Aberration and Targeted Therapeutic Prediction” American Association for Cancer Research (AACR) Annual Meeting San Diego, USA April 6-9
- HoJoon Lee, Patrick Flaherty, Hanlee P. Ji (2013) “Systematic Identification of Genomic Markers for Clinical Stage of Colorectal Cancer” American Association for Cancer Research (AACR) Annual Meeting Washington D.C., USA April 6-10
- HoJoon Lee, John-Charles Rodenberry, Kristen Day, Jose Cano-Buendia, Felicia Craciunescu, Jean-Charles Chapuis, Phillip Stafford, Kathryn Sykes, Patricia E. Carrigan, Stephen Albert Johnston (2010) “Neo-peptides from Gene Fusions as Cancer Vaccine Antigens” American Association for Cancer Research (AACR) Annual Meeting Washington D.C., USA April 17-21
- HoJoon Lee, Jose Cano-Buendia, Kristen Day, Patricia E. Carrigan, Phillip Stafford, Kathryn Sykes and Stephen A. Johnston (2009) “Systematic characterization of gene fusion transcripts by using information from the express sequence tag database” 13th Internal Conference on Research in Computational Molecular Biology (RECOMB) Tucson AZ, USA May 18-21
- HoJoon Lee, Douglas Lake, Phillip Stafford and Stephen Johnston (2008) “Translocation fusion transcripts as a source of cancer vaccine antigens” American Association for Cancer Research (AACR) Annual Meeting San Diego CA, USA April 12-16
- HoJoon Lee and Stephen Albert Johnston (2007) “Detection of tumor-specific mutations in coding microsatellite DNAs from expressed sequence tag (EST)” RECOM Computational Cancer Biology San Diego CA, USA Sept 18 – 20
- HoJoon Lee, Sankar Subramanian and Sudhir Kumar (2005) “Estimate of Genomewide Mutation Rate Difference between Human and Mouse Ancestral Repeats” Intelligent Systems for Molecular Biology (ISMB) Detroit MI, USA Jun 25 – 29
- HoJoon Lee, Stephanie Rogers, Maulik Shah and Jeffrey Touchman (2004) “A Novel Genome Assembler: Using K-Mers to Indirectly Perform N-squared Comparisons in O(N)” Computational Systems Bioinformatics Conference. Stanford University Aug16-19.
- HoJoon Lee, Sehee Hwang , Insuk Sohn, Sujong Kim, Young Sik Lee and Yong Sung Lee (2004) “Dinucleotide relative abundance (DrA)-based Classification of Viruses; full-length genome sequence analysis of whole Papillomaviridae(PV) family” Intelligent Systems for Molecular Biology (ISMB). Glasgow, UK Jul 31-Aug 4
- HoJoon Lee, Taeyun Oh, Insuk Sohn, Sehee Hwang, Sujong Kim, Seok-Ki Lee, Ikjin Kim (2002) “Analysis of Viral Genome by k-mer Method” 2002 Korean Society for Bioinformatics Annual Meeting. Pusan, Korea Nov 15-17