Hannah Yu Ph. D, Clinical Genomics Scientist

e-mail: yuhannah82@gmail.com
github: https://github.com/hannahyu-source/portfolio
linkedin: https://www.linkedin.com/in/hannah-yu-9020ab268/?skipRedirect=true
google scholar: https://scholar.google.com/citations?user=pQJStZ0AAAAJ&hl=ko

Qualifications profile

Clinical genomics scientist with over 10 years of experience in translational research, NGS technologies, and molecular biology. Experienced in clinical genomic interpretation (WES/WGS), cancer genomics, and NGS workflow from experimental design to data interpretation. Conducted cancer genomics research at Samsung Medical Center and clinical variant interpretation at 3billion, supporting precision medicine through genomic analysis and clinical reporting. Previously developed a high-fidelity Nanopore sequencing strategy at The Ohio State University, leading to a first-author publication in *Advanced Science* (IF 14.5). Strong background in multidisciplinary collaboration, international research, and translating genomic data into clinically meaningful insights.

Education

2010-2016: Ph. D study in Dept. of Bio-Analytical Science of University of Science and Technology (UST), South Korea
Thesis Title: Development of reference methodologies for gene measurements: Gene methylation, gene expression, and bacterial community profiling
Supervisor: Prof. Inchul Yang
2009-2010: Internship program in the Center for Bioanalysis of Korea Research Institute of Standards and Science (KRISS), South Korea
Supervisor: Prof. Inchul Yang
2006-2008: MS study in Dept. of Science Education (Biology) of Chungnam National University (CNU), South Korea
Thesis Title: Isolation of Farnesoid X Receptor (*fxr)* gene in zebrafish
Supervisor: Prof. Cheol-Hee Kim
2001-2006: BS study in Dept. of Life Science of Handong Global University, South Korea

Research Experience and Skills

3billion - Clinical Genomics Scientist (Nov 2025 ~ Feb 2026)

- Interpreted WES/WGS variants and generated clinical reports.
- Conducted AI-assisted literature review and variant assessment for rare diseases.
- Integrated genomic and phenotypic data in collaboration with clinicians and genetic counselors.
- Contributed to the development of AI-based genetic diagnostic models.

Samsung Medical Center — Doctoral Researcher (Mar 2025 – Sep 2025)

- Designed and executed experiments for cancer patient sample sequencing.
- Extracted nucleic acids, performed quality assessment, and prepared sequencing libraries from clinical samples.
- Conducted whole-exome sequencing (WES) and total-RNA sequencing using NovaSeq X Plus and AVITI24.
- Collaborated with bioinformatics team to process, analyze, and interpret sequencing data.

Pusan National University — Doctoral Researcher (Sep 2024 – Jan 2025)

- Prepared DNA resources and conducted quality control on blood and tissue specimens obtained from the National Biobank of Korea, including assessment of DNA concentration/purity, integrity, sex determination by qPCR, and bacterial/mycoplasma contamination testing.
- Performed whole-genome sequencing of six Mycobacterium tuberculosis strains provided by the National Institute of Health and carried out comparative genomic analyses to assess the effects of growth-promoting supplements on sequencing outcomes.

The Ohio State University — Postdoctoral Researcher (Apr 2016 – Jul 2024)

- Developed high-fidelity, long-range single-molecule sequencing methods using Oxford Nanopore platforms.
- Engineered circular long-read DNA constructs and twin random barcodes for advanced sequencing accuracy.
- Applied nanopore-based high-fidelity, long-range single-molecule sequencing technologies to:
    - Detect and monitor SARS-CoV-2 genomic variants in wastewater.
    - Discover novel marine RNA viruses.
    - Profile HIV-1 variants under ART drug pressure.
- Conducted in vitro selection of HIV-1 variants,
- CRISPR/Cas9-based genome editing
- Investigated RNA virus epigenetics, including HIV-1 RNA modification landscapes at the single-virus level.
- Engineered and characterized infectious HIV-1 clones and retroviral vectors for functional studies.

University of Science and Technology / Korea Research Institute of Standards and Science — Ph.D. Candidate (Sep 2010 – Feb 2016)

- Designed and validated reference methodologies for gene measurement.
- Developed certified reference materials (CRMs): dNMP/NMP, human genomic DNA, miRNA, gene methylation, and bacterial 16S rDNA.
- Established NGS-based experimental strategies; prepared and sequenced libraries (Illumina MiSeq).
- Operated advanced analytical platforms:
    - NGS: DNA methylation, metagenomics, transcriptomics.
    - CE: DNA methylation, nucleic acid quantification.
    - MALDI-TOF: protein identification, genotyping.
    - FACS: cell sorting, single-cell analysis.
- Gained expertise in molecular biology methods: cloning, site-directed mutagenesis, qPCR/dPCR, High Resolution Melting(HRM), bisulfite conversion, in vitro methylation, and library preparation for NGS.

Chungnam National University — Research Assistant (2007 – 2008)

- Maintained zebrafish breeding colonies.

- Performed microinjection and whole-mount in situ hybridization on zebrafish embryos.

Publications

Peer-Reviewed Research Papers – Published

1. Yu H, Golconda S, Xue D, Zhu M, Lee GE, Viola C, Dominguez-Huerta G, Bolduc B, Sullivan M and Kim S, CLAE: A High-Fidelity Nanopore Sequencing Strategy for Read-Level Viral Variant Detection and Environmental RNA Virus Discovery, ***Advanced Science***. (2025) Sep 11:305978. doi:10.1002/advs.202505978. Online ahead of print. PMID: 40936099 ***IF:14.5***
2. Baek A, Lee GE, Golconda S, Rayhan A, Manganaris A, Chen S, Tirumuru N, Yu H *et al*., Single-molecule epitranscriptomic analysis of full-length HIV-1 RNAs reveals functional roles of site-specific m6As, ***Nature Microbiology***, (2024) May;9(5):1340-1355. PMCID: PMC11087264. ***IF:28.3***
3. Baek A, Rayhan A, Lee GE, Golconda S, Yu H *et al*., Mapping m6A Sites on HIV-1 RNA Using Oligonucleotide LC-MS/MS, ***Methods Protoc***. (2024) Jan 10;7(1). PubMed PMID: 38251200; PubMed Central PMCID: PMC10801558. doi:10.3390/mps7010007. ***IF:2.4***
4. Kim S, Shukla RK, Yu H *et al*., CD3e-immunotoxin spares CD62Llo Tregs and reshapes organ-specific T-cell composition by preferentially depleting CD3ehi T cells**,** ***Front Immunol***. (2022); 13:1011190. PubMed PMID: 36389741; PubMed Central PMCID: PMC9643874. doi: 10.3389/fimmu.2022.1011190. eCollection, ***IF: 8.8***
5. Kim S, Shukla RK, Kim E, Cressman SG, Yu H *et al*., Comparison of CD3e Antibody and CD3e-sZAP Immunotoxin Treatment in Mice Identifies sZAP as the Main Driver of Vascular Leakage, ***Biomedicines*** (2022), 10, 1221. https://doi.org/10.3390/biomedicines10061221, ***IF: 6.0***
6. Suryawanshi GW, Khamaikawin W, Wen J, Shimizu S, Arokium H, Xie Y, Wang E, Kim S, Choi H, Zhang C, Yu H *et al*., The clonal repopulation of HSPC gene modified with anti-HIV-1 RNAi is not affected by preexisting HIV-1 infection***, Science Advances***, 2020, DOI: 10.1126/sciadv.aay9206, ***IF: 13.1***
7. Yu H, Hahn Y, Park SR *et al.,* Normalization of human RNA-seq experiments using chimpanzee RNA as a spike-in standard, Sci. Rep., (2016) ***IF: 4.3***
8. Yu H, Hahn Y, and Yang I, Reference materials for calibration of analytical biases in quantification of DNA methylation, PLOS One, (2015) ***IF: 3.2***
9. Kang MJ, Yu H, Kim SK, Park SR, and Yang I, Quantification of trace-level DNA by real time whole genome amplification, PLOS One, (2011) ***IF: 3.2***

Peer-Reviewed Research Papers – Submitted

1. Lee GE, Golconda S, Yu H, Kim S, and Kim S, “Universal RNA Linearization Enables Full-Length Nanopore DRS Across Whole Transcriptomes” manuscript submitted.

Peer-Reviewed Research Papers – In preparation

1. Yu H, Golconda S, Baek A, Kim S and Kim S, Novel twin barcoding for ultra-high-accuracy, Nanopore sequencing for HIV-1 whole genome analysis, manuscript in preparation
2. Yu H, Golconda S, Baek A, Kim S and Kim S, Recombination-free, long-range PCR for full-length HIV-1 DNA genotyping, manuscript in preparation
3. Kim S, Yu H, Golconda S, Reneu J, Hanel W and Kim S, “Treatment resistance of Follicular T helper cells in T-cell-targeting antibody and immunotoxin therapies” manuscript in preparation
4. Golconda S, Lee GE, Yu H, Kim S and Kim S, “Context dependent and HIV-1-specific regulation of RNA and protein expression by HIV-1’s site specific m6A RNA modifications” manuscript in preparation.

Peer-reviewed Abstracts

1. Yu H, Baek A, Golconda S and Kim S, "*Novel twin barcoding for ultrahigh-accuracy, Nanopore sequencing for HIV-1 whole genome analysis*."; NHGRI Advanced Genomic Technology Development Meeting, 2019, Boston, MA.
2. Yu H, Baek A, Golconda S and Kim S, "*Ultra-high-accuracy Nanopore target-sequencing for innovations in HIV diagnosis and surveillance*."; NHGRI Advanced Genomic Technology Development Meeting, 2019, Boston, MA.

Patents

1. Yang I, Yu H, Yang HJ, Kim SK, Park SR, Quantitative analysis method using microorganism 16S rDNA gene having single nucleotide polymorphism, US Patent 10,253,378, (2019)
2. Kim S, Yu H, Baek A, Methods of making and using tandem, twin barcode molecules, PCT/US2018/049203, (2018)
3. Yang I, Kang MJ, Yu H, Park SR, Kim SK, Methods and kits for the quantification of nucleic acids, US/9,297,038 (2016)
4. Yang I, Yu H, Yang H, Kim SK, Park SR, Bacterial 16s rDNA sequences with defined single nucleotide polymorphisms, PCT/KR2014/011040, (2014)
5. Yang I, Kang MJ, Yu H, Park SR, Kim SK, Methods and kits for the quantification of nucleic acids, PCT/KR2011/008494, (2011)

Research Fund

1. 2018 – 2024 NHGRI R01 (PI: Kim S) "On-site, high-fidelity target sequencing and absolute quantitation for HIV-1 surveillance"
2. 2019 – 2023 Department of Energy (PI: Sullivan M; Co-I: Kim S) "Virus in soils: key modulators of microbiomes and nutrient cycling?"
3. 2018 – 2022 NHGRI R21 (PI: Kim S) "High-accuracy, long-range sequencing for HIV genotyping"

Conference

1. 2019 Advanced Genomic Technology Development Meeting, Boston, USA, short talk
2. 2018 Nanopore Community Meeting, San Francisco, USA, Poster
3. 2016 Nanopore Community Meeting, New York, USA, Poster
4. 2013 Nucleic Acid Summit, San Francisco, USA, Poster

AI & Data Science Training

1. 2026 AI Convergence Data Analysis Professional Training (WISET × AWS)
   : Python, Google Colab, API-based data acquisition, data visualization, AWS project.
2. 2026 Generative AI Professional Instructor Training (WISET × AWS)
   : Prompt engineering, RAG, Agentic AI, AWS GenAI, AI educational content development.
3. 2026 AI Practical Training (WISET, In Progress)
   : LLM-assisted productivity, Python/Colab, AI workflow automation, research data analysis project.

Certification

ADsP (Advanced Data Analytics Semi-Professional), Korea Data Agency, 2026

Universal RNA Linearization Enables Full-Length Nanopore DRS Across Whole Transcriptomes

(revision for *Nature Methods*, 2025)

Abstract
We present universal reverse transcription methods (UmRT and UmRT+) that enhance Nanopore direct RNA sequencing of long and structured RNAs. These protocols enable high-efficiency, full-length recovery of challenging transcripts—including the 20 Kb XIST RNA—and resolve isoform-specific expression, m6A modifications, and poly(A) tail variation. This advancement enables high-resolution profiling of native RNAs, with broad utility in transcriptomics, virology, and RNA-based diagnostics.

Novel twin barcoding for ultra-high-accuracy, Nanopore sequencing for HIV-1 whole genome analysis
(manuscript in preparation)

Abstract
Human Immunodeficiency Virus Type-1 (HIV-1) genotyping has been hindered by short read-length and errors in the current sequencing platforms. We have developed a novel twin-barcoding approach to improve long-range sequencing for HIV-1 genotyping using Nanopore sequencing. Nanopore sequencing can read extremely long nucleotide sequences (company record: 2 Mb), but its application is significantly limited by high error rates (approximately 5-10%). Conventional "single" barcoding is impractical for the error-prone Nanopore sequencing because of the high error rates already associated with reading the short barcodes themselves; for example, when using 10-20bp barcodes, nearly all the barcodes in Nanopore data will have read errors. Our twin-barcoding approach uses a library of Tandem Twin Barcode (TTB) DNA, in which every TTB DNA molecule has two identical barcodes placed next to each other in the same direction. After Nanopore $1D^2$ (or 2D) sequencing of TTB-labeled DNA, barcode read errors can be recognized and removed by comparing 4 identical barcodes in the same double-stranded DNA. Given that chemical synthesis of a high-quality TTB is practically impossible, we have developed a novel twin-barcode preparation method that mimics the DNA duplication steps of retroviral integration. With 10^{14} copies of starting materials (single barcode oligos), we can generate $>10^{11}$ twin-barcode molecules with 10^9 variants. To minimize PCR bias, no PCR is included until the last step. The quality of the TTB library was tested with both Sanger and Illumina sequencing methods, showing > 90% twins in the current TTB library. Using a TTB library specific for HIV-1 *nef*, HIV-1 plasmid and cellular proviral DNA (NL4-3 strain) were sequenced and up to 99.99% accuracy was achieved in our preliminary study. Our twin barcoding approach enables high-accuracy, full-genome HIV-1 sequencing. When using a portable Nanopore sequencer (MinION), our approach has the potential to significantly improve point-of-care HIV diagnosis and real-time surveillance even in resource-limited settings.

Recombination-free, long-range PCR for full-length HIV-1 DNA genotyping
(manuscript in preparation)

Abstract
Given the rapid development of 3rd generation Nanopore sequencing, long-range single molecule target sequencing is expected to deliver great advancements in genomic research. Single molecule target sequencing often requires a PCR step for high-sensitivity and high-accuracy sequencing. However, a long-range PCR (e.g. 9 Kb HIV DNA amplification) is challenging. Previous studies identified that PCR recombination occurs as frequently as 5-40% depending on PCR conditions and template types. These events are difficult to identify and correct during sequence analysis. We have extensively tested different lengths (1-3-6-9Kb) of PCR for HIV DNA and found that PCR recombination occurs more frequently with longer-range PCR. Modifying PCR conditions can partially reduce the recombination events, but such modifications were not effective in full-length PCR for the HIV genome. In order to effectively reduce PCR recombination in long-range PCR, we employed a droplet digital PCR (ddPCR) system that partitions the PCR reaction to 20,000 separate droplets. This in turn enabled single target molecule amplification in nano-liter volume droplets with adjusted input template copy numbers. We compared the DNA sequencing results from regular PCR with those from ddPCR and demonstrated that ddPCR can effectively resolve the PCR recombination issue. With this new PCR approach, we were able to sequence full-length, individual HIV DNA molecules in a high-throughput fashion without PCR-induced recombination. The new approach will help accurately assess the HIV proviral genome in reservoir cells and identify critical information that has been inaccessible with conventional short-read genotyping, including intact vs. defective viruses; recombination hotspots at the whole genome level; and non-drug target mutations.