

Curriculum Vitae

Goo Jun

October 2024

Current Position

Associate Professor
Human Genetics Center
Department of Epidemiology, Human Genetics & Environmental Sciences
School of Public Health
University of Texas Health Science Center at Houston

Education

- 2010-2014 Postdoctoral Research Fellow, University of Michigan, Ann Arbor, MI.
Center for Statistical Genetics & Department of Biostatistics
- 2005-2010 Ph.D., The University of Texas, Austin, TX.
Electrical and Computer Engineering
Dissertation: Transfer learning for classification of spatially varying data
- 1997-1999 M.S., University of Michigan, Ann Arbor, MI.
Electrical Engineering and Computer Science
- 1993-1997 B.S., KAIST, Daejeon, South Korea.
Electrical Engineering (major) / Computer Science (minor)

Professional Experience

- Sep. 2021 – Present Associate Professor
University of Texas Health Science Center, Houston, TX
Human Genetics Center
Department of Epidemiology, Human Genetics & Environmental Sciences
School of Public Health
- Oct. 2014 – Aug. 2021 Assistant Professor
University of Texas Health Science Center, Houston, TX
Human Genetics Center
Department of Epidemiology, Human Genetics & Environmental Sciences
School of Public Health
- Aug. 2010 – Oct. 2014 Postdoctoral Research Fellow
University of Michigan, Ann Arbor, MI
Center for Statistical Genetics & Department of Biostatistics
PIs: Gonçalo Abecasis and Michael Boehnke
- Jun. 2009 – Aug. 2009 Adv. Short-Term Research Opportunity (ASTRO)
Oak Ridge Nat. Lab., Oak Ridge, TN
Computational Sciences and Engineering Division (CSED)
- Jan. 2007 – Aug. 2010 Research Assistant
University of Texas, Austin, TX

Department of Electrical and Computer Engineering
Advanced Data Exploration and Analysis Laboratory (IDEAL)

Dec. 1999 – Jun. 2010 Engineer/Senior Engineer
Samsung Electronics, Suwon, South Korea.
Digital Media R&D Center

Honors and Awards

- PRIME Award, UTHealth School of Public Health, 2017
- Young Investigator Travel Award, American Diabetes Association, 2017
- Young Investigator Travel Award, American Diabetes Association, 2012

Research Support (Current)

R01DK11863, NIH/NIDDK, Role: PI
2/28/2019 – 01/31/2025, Total direct cost: \$2,653,837
Diabetes Progression with Metabolomic Profiling in Starr County Mexican Americans

R01HL146860, NIH/NHLBI, Role: Co-I
3/17/2020-2/28/2025, Total direct cost: \$3,370,421
Whole-Genome Sequencing Analysis of Coronary Atherosclerosis and Related Traits

NIH/NHLBI, Role: Co-I
05/01/2018-03/31/2029, Total direct cost: \$ 3,276,632
Trans-Omics for Precision Medicine (TOPMed) Informatics Research Center (IRC)

University of Michigan/Chan Zuckerberg Initiative (CHZI), Role: Subcontract PI
08/01/2024-01/31/2026, Total direct cost: \$69,565
Scalable analysis of submicron resolution spatial transcriptomics

Research Support (Completed)

OT2OD002748, BCM/NIH, Role: Co-I
09/25/2018 – 08/31/2023, Total direct cost: \$875,576
The Baylor-Hopkins Clinical Genomics Center for All of Us

U01 AG058589, NIH/NIA Boston University, Role: Co-I
09/30/2018-08/31/2023, Total Direct \$295,046
Therapeutic Target Discovery in ADSP Data via Comprehensive Whole-Genome Analysis
Incorporating Ethnic Diversity and Systems Approaches

R01DK116378-01, NIH/NIDDK, Role: Co-I
09/15/2017 – 06/30/2023, Total direct cost: \$2,297,462
Microbiome and Worsening Glycemia Among Mexican Americans in Starr County, Texas

R03HD098552, NIH/NICHD, Role: PI
6/10/2019 – 5/31/2022, Total direct cost: \$200,000
The Maternal Genome and Risk of Structural Birth Defects

R03DE028340-01, NIH/NIDCR, Role: PI, subcontract from UTHealth School of Dentistry (PI: Iwata)
02/01/2019 – 01/31/2021, Total direct cost: \$22,990
Molecular Regulatory Network in Frontonasal Development

UM1HG008898, BCM/NIH, Role: Co-I
01/14/2016 – 5/31/2022, Total direct cost: \$1,764,436
Genomic Architecture of Common Disease in Diverse Populations

R03DE026509, NIH/NIDCR, Role: PI, subcontract from UTHealth School of Dentistry (PI: Iwata)
09/01/2017 – 08/31/2020, Total subcontract direct cost: \$49,708
Identification of microRNAs associated with cleft lip

1R03DE026208, NIH/NIDCR. Role: PI, subcontract from UTHealth School of Dentistry (PI: Iwata)
09/01/2016 – 08/31/2018, Total subcontract direct cost: \$59,850
Transcripts and Functions Targeted by Non-coding RNAs in Palate Development

CURE Pilot, BCM/CURE, Role: Co-I
02/01/2018 – 07/31/2018, Total direct cost: \$8,557
Pilot in Silico Mortality Risk Attribution in Sudep and Sudden Death in the Young (SUDY) to
Information Precision Molecular Diagnostics of Sudden Death

U54HG003273, BCM/NIH, Role: Co-I
11/01/2011-10/31/2016, Total direct cost: \$1,272,981
The Human Genome Sequencing Center

Peer-reviewed journal articles

1. Rocheleau G, Clarke SL, Auguste G, Hasbani NR, Morrison AC, Heath AS, Bielak LF, Iyer KR, Young EP, Stitzel NO, **Jun G**, Laurie C, Broome JG, Khan AT, Arnett DK, Becker LC, Bis JC, Boerwinkle E, Bowden DW, Carson AP, Ellinor PT, Fornage M, Franceschini N, Freedman BI, Heard-Costa NL, Hou L, Chen YI, Kenny EE, Kooperberg C, Kral BG, Loos RJJ, Lutz SM, Manson JE, Martin LW, Mitchell BD, Nassir R, Palmer ND, Post WS, Preuss MH, Psaty BM, Raffield LM, Regan EA, Rich SS, Smith JA, Taylor KD, Yanek LR, Young KA; NHLBI TOPMed Consortium; Hilliard AT, Tcheandjieu C, Peyser PA, Vasan RS, Rotter JI, Miller CL, Assimes TL, de Vries PS, Do R. Rare variant contribution to the heritability of coronary artery disease. *Nat Commun.* 2024 Oct 9;15(1):8741. doi: 10.1038/s41467-024-52939-6. PMID: 39384761
2. Si Y, Lee C, Hwang Y, Yun JH, Cheng W, Cho CS, Quiros M, Nusrat A, Zhang W, **Jun G**, Zöllner S, Lee JH, Kang HM. FICTURE: scalable segmentation-free analysis of submicron-resolution spatial transcriptomics. *Nat Methods.* 2024 Oct;21(10):1843-1854. doi: 10.1038/s41592-024-02415-2. Epub 2024 Sep 12. PMID: 39266749
3. Weiss MC, Sun J, Jackson BP, Turyk ME, Wang L, Brown EL, Aguilar D, Hanis CL, Argos M, Sargis RM. Predictors of Toxic Metal/Metalloid Exposures Among Mexican Americans in Starr County, Texas. *J Racial Ethn Health Disparities.* 2024 Jul 5. doi: 10.1007/s40615-024-02064-3. Epub ahead of print. PMID: 38969926.
4. de Vries PS, Conomos MP, Singh K, Nicholson CJ, Jain D, Hasbani NR, Jiang W, Lee S, Lino Cardenas CL, Lutz SM, Wong D, Guo X, Yao J, Young EP, Tcheandjieu C, Hilliard AT, Bis JC, Bielak LF, Brown MR, Musharoff S, Clarke SL, Terry JG, Palmer ND, Yanek LR, Xu H, Heard-Costa N, Wessel J, Selvaraj MS, Li RH, Sun X, Turner AW, Stilp AM, Khan A, Newman AB, Rasheed A, Freedman BI, Kral BG, McHugh CP, Hodonsky C, Saleheen D, Herrington DM, Jacobs DR Jr, Nickerson DA, Boerwinkle E, Wang FF, Heiss G, **Jun G**, Kinney GL, Sigurslid HH, Whole-genome sequencing uncovers two loci

- for coronary artery calcification and identifies ARSE as a regulator of vascular calcification. *Nat Cardiovasc Res*. 2023 Dec;2(12):1159-1172. doi: 10.1038/s44161-023-00375-y. Epub 2023 Dec 4. PMID: 38817323
5. Musfee FI, **Jun G**, Mitchell LE, Chen H, Guo D, Prakash SK, Adkar SS, Grove ML, Choi RB, Klarin D; Million Veteran Program; Boerwinkle E, Milewicz DM. X-linked genetic associations in sporadic thoracic aortic dissection. *Am J Med Genet A*. 2024 Sep;194(9):e63644. doi: 10.1002/ajmg.a.63644. Epub 2024 Apr 30. PMID: 38688863
 6. Kim Y, Cheng W, Cho CS, Hwang Y, Si Y, Park A, Schrank M, Hsu JE, Xi J, Kim M, Pedersen E, Koues OI, Wilson T, **Jun G**, Kang HM, Lee JH. Seq-Scope Protocol: Repurposing Illumina Sequencing Flow Cells for High-Resolution Spatial Transcriptomics. bioRxiv [Preprint]. 2024 Apr 1:2024.03.29.587285. doi: 10.1101/2024.03.29.587285. PMID: 38617262, Accepted to *Nature Protocols*
 7. Brown EL, Essigmann HT, Hoffman KL, Petrosino J, **Jun G**, Brown SA, Aguilar D, Hanis CL. C-Reactive Protein Levels Correlate with Measures of Dysglycemia and Gut Microbiome Profiles. *Curr Microbiol*. 2023 Dec 21;81(1):45. doi: 10.1007/s00284-023-03560-1. PMID: 38127093
 8. Weiss MC, Shih YH, Bryan MS, Jackson BP, Aguilar D, Brown EL, **Jun G**, Hanis CL, Argos M, Sargis RM. Arsenic metabolism, diabetes prevalence, and insulin resistance among Mexican Americans: A mendelian randomization approach. *Environ Adv*. 2023 Jul;12:100361. doi: 10.1016/j.envadv.2023.100361. Epub 2023 Mar 24. PMID: 37426694; PMCID: PMC10328543.
 9. Gunasekara CJ, MacKay H, Scott CA, Li S, Laritsky E, Baker MS, Grimm SL, **Jun G**, Li Y, Chen R, Wiemels JL, Coarfa C, Waterland RA. Systemic interindividual epigenetic variation in humans is associated with transposable elements and under strong genetic control. *Genome Biol*. 2023 Jan 12;24(1):2. doi:10.1186/s13059-022-02827-3. PMID: 36631879; PMCID: PMC9835319.
 10. Wheeler MM, Stilp AM, Rao S, Halldórsson BV, Beyter D, Wen J, Mihkaylova AV, McHugh CP, Lane J, Jiang MZ, Raffield LM, **Jun G**, Sedlazeck FJ, Metcalf G, Yao Y, Bis JB, Chami N, de Vries PS, Desai P, Floyd JS, Gao Y, Kammers K, Kim W, Moon JY, Ratan A, Yanek LR, Almasy L, Becker LC, Blangero J, Cho MH, Curran JE, Fornage M, Kaplan RC, Lewis JP, Loos RJJ, Mitchell BD, Morrison AC, Preuss M, Psaty BM, Rich SS, Rotter JI, Tang H, Tracy RP, Boerwinkle E, Abecasis GR, Blackwell TW, Smith AV, Johnson AD, Mathias RA, Nickerson DA, Conomos MP, Li Y, Þorsteinsdóttir U, Magnússon MK, Stefansson K, Pankratz ND, Bauer DE, Auer PL, Reiner AP. Whole genome sequencing identifies structural variants contributing to hematologic traits in the NHLBI TOPMed program. *Nat Commun*. 2022 Dec 8;13(1):7592. doi: 10.1038/s41467-022-35354-7. PMID: 36481753; PMCID: PMC9732337.
 11. Xi J, Lee JH, Kang HM*, **Jun G***. STtools: a comprehensive software pipeline for ultra-high-resolution spatial transcriptomics data, *Bioinformatics Advances*, Volume 2, Issue 1, 2022, vbac061, <https://doi.org/10.1093/bioadv/vbac061>, *Corresponding
 12. Kim HS, Xiao X, Byun J, **Jun G**, DeSantis SM, Chen H, Thrift AP, El-Serag HB, Kanwal F, Amos CI. Synergistic Associations of PNPLA3 I148M Variant, Alcohol Intake, and Obesity With Risk of Cirrhosis, Hepatocellular Carcinoma, and Mortality. *JAMA Network Open*. 2022 Oct 3;5(10):e2234221. doi: 10.1001/jamanetworkopen.2022.34221. PMID: 36190732.
 13. Essigmann HT, Hanis CL, DeSantis SM, Perkison WB, Aguilar DA, **Jun G**, Robinson DA, Brown EL. Worsening Glycemia Increases the Odds of Intermittent but Not Persistent Staphylococcus aureus Nasal Carriage in Two Cohorts of Mexican American Adults. *Microbiol Spectr*. 2022 Jun 29;10(3):e0000922. doi: 10.1128/spectrum.00009-22. Epub 2022 May 18. PMID: 35583495; PMCID: PMC9241628.
 14. Huang L, Rosen JD, Sun Q, Chen J, Wheeler MM, Zhou Y, Min YI, Kooperberg C, Conomos MP, Stilp AM, Rich SS, Rotter JI, Manichaikul A, Loos RJJ, Kenny EE, Blackwell TW, Smith AV, **Jun G**, Sedlazeck FJ, Metcalf G, Boerwinkle E; NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, Raffield LM, Reiner AP, Auer PL, Li Y. TOP-LD: A tool to explore linkage disequilibrium with TOPMed whole-genome sequence data. *Am J Hum Genet*. 2022 Jun 2;109(6):1175-1181. doi: 10.1016/j.ajhg.2022.04.006. Epub 2022 May 2. PMID: 35504290; PMCID: PMC9247832.

15. Yang Y, Bartz TM, Brown MR, Guo X, Zilhao NR, Trompet S, Weiss S, Yao J, Brody JA, Defilippi RD, Hoogeveen RC, Lin HJ, Gudnason V, Ballantyne CM, Dorr M, Jukema JW, Petersmann A, Psaty B, Rotter JI, Boerwinkle E, Fornage M, **Jun G***, Yu B*. Identification of Functional Genetic Determinants of Cardiac Troponin T and I in a Multi-Ethnic Population and Causal Associations with Atrial Fibrillation. *Circulation: Genomic and Precision Medicine*. 2021 Nov 4;doi: 10.1161/CIRCGEN121003460. PMID: 34732054. *Corresponding
16. Yoshioka H, **Jun G**, Suzuki A, Iwata J. Dexamethasone Suppresses Palatal Cell Proliferation through miR-130a-3p. *Int J Mol Sci*. 2021 Nov 18;22(22):12453. doi: 10.3390/ijms22212453. PMID: 34830336.
17. Essigmann HT, Hoffman KL, Petrosino JF, **Jun G**, Aguilar D, Hanis CL, DuPont HL, Brown EL. The impact of the Th17:Treg axis on the IgA-Biome across the glycemic spectrum. *PLOS ONE*. 2021 Oct 20;16(10):e0258812. PMID: 34669745.
18. Zhu C, Mohan R, Lin SH, **Jun G**, Yaseen A, Jiang X, Wang Q, Cao W, Hobbs BP. Identifying individualized risk profiles for radiotherapy induced lymphopenia among esophageal cancer patients using machine learning. *JCO Clinical Cancer Informatics*. 2021 Sep;5:1044-1053. PMID: 34665662
19. Cho CS, Xi J, Si Y, Park SR, Hsu JE, Kim M, **Jun G**, Kang HM, Lee JH. Microscopic examination of spatial transcriptome using Seq-Scope. *Cell*. 2021 Jun 7:S0092-8674(21)00627-9. doi: 10.1016/j.cell.2021.05.010. PMID: 34115981.
20. **Jun G***, Sedlazeck F, Zhu Q, English A, Metcalf G, Kang HM; Human Genome Structural Variation Consortium (HGSVC), Lee C, Gibbs R, Boerwinkle E. muCNV: Genotyping Structural Variants for Population-level Sequencing. *Bioinformatics*. 2021 Mar 24:btab199. doi: 10.1093/bioinformatics/btab199. Epub ahead of print. PMID: 33760063. *Corresponding
21. Zarate S, Carroll A, Mahmoud M, Krasheninina O, **Jun G**, Salerno WJ, Schatz MC, Boerwinkle E, Gibbs RA, Sedlazeck FJ. Parliament2: Accurate Structural Variant Calling at Scale. *Gigascience*. Dec. 2020; 9(12), giaa145
22. Yang Y, Suzuki A, Iwata J, **Jun G***. Novel Analytical Strategies Identified Additional Genetic Associations on Cleft Lip with or without Cleft Palate, *Genes*. Oct. 2020; 11(11), 1280. doi: 10.3390/genes11111280.
23. Brown EL, Essigmann HT, Hoffman KL, Palm NW, Gunter SM, Sederstrom JM, Petrosino JF, **Jun G**, Aguilar D, Perkison WB, Hanis CL, DuPont HL, The impact of diabetes on the gut and salivary IgA microbiomes, *Infection and Immunity*, 2020 Nov 16;88(12):e00301-20. doi: 10.1128/IAI.00301-20. PMID: 32900816; PMCID: PMC7671898.
24. **Jun G**, Aguilar D, Evans C, Burant CF, Hanis CL, Metabolomic profiles associated with subtypes of prediabetes among Mexican Americans in Starr County, Texas. *Diabetologia*. 2020 Feb;63(2)287-295. doi: 10.1007/s00125-019-05031-4. PMID: 31802145.
25. Zhu C, Lin SH, Jiang X, Xiang Y, Belal Z, **Jun G**, Mohan R. A novel deep learning model using dosimetric and clinical information for grade 4 radiotherapy-induced lymphopenia prediction. *Phys Med Biol*. 2020 Feb 4;65(3):035014. doi: 10.1088/1361-6560/ab63b6. PMID: 31851954.
26. Suzuki A, Yoshioka H, Summakia D, Desai NG, **Jun G**, Jia P, Loose DS, Ogata K, Gajera MV, Zhao Z, Iwata J. MicroRNA-124-3p suppresses mouse lip mesenchymal cell proliferation through the regulation of genes associated with cleft lip in the mouse. *BMC Genomics*. 2019 Nov 14;20(1):852. doi: 10.1186/s12864-019-6238-4. PMID: 31727022.
27. Gajera M, Desai N, Suzuki A, Li A, Zhang M, **Jun G**, Jia P, Zhao Z, Iwata J. MicroRNA-655-3p and microRNA-497-5p inhibit cell proliferation in cultured human lip cells through the regulation of genes related to human cleft lip. *BMC Med Genomics*. 2019 May 23;12(1):70. doi: 10.1186/s12920-019-0535-2. PMID: 31122291.

28. Chaisson MJP, Sanders AD, Zhao X, Malhotra A, Porubsky D, Rausch T, Gardner EJ, Rodriguez OL, Guo L, Collins RL, Fan X, Wen J, Handsaker RE, Fairley S, Kronenberg ZN, Kong X, Hormozdiari F, Lee D, Wenger AM, Hastie AR, Antaki D, Anantharaman T, Audano PA, Brand H, Cantsilieris S, Cao H, Cerveira E, Chen C, Chen X, Chin CS, Chong Z, Chuang NT, Lambert CC, Church DM, Clarke L, Farrell A, Flores J, Galeev T, Gorkin DU, Gujral M, Guryev V, Heaton WH, Korlach J, Kumar S, Kwon JY, Lam ET, Lee JE, Lee J, Lee WP, Lee SP, Li S, Marks P, Viaud-Martinez K, Meiers S, Munson KM, Navarro FCP, Nelson BJ, Nodzak C, Noor A, Kyriazopoulou-Panagiotopoulou S, Pang AWC, Qiu Y, Rosanio G, Ryan M, Stütz A, Spierings DCJ, Ward A, Welch AE, Xiao M, Xu W, Zhang C, Zhu Q, Zheng-Bradley X, Lowy E, Yakneen S, McCarroll S, **Jun G**, Ding L, Koh CL, Ren B, Flicek P, Chen K, Gerstein MB, Kwok PY, Lansdorp PM, Marth GT, Sebat J, Shi X, Bashir A, Ye K, Devine SE, Talkowski ME, Mills RE, Marschall T, Korbel JO, Eichler EE, Lee C. Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat Commun.* 2019 Apr 16;10(1):1784. doi: 10.1038/s41467-018-08148-z. PMID: 30992455.
29. Thompson J, Mendoza F, Tan E, Bertol JW, Gaggari AS, **Jun G**, Biguetti C, Fakhouri WD. A cleft lip and palate gene, *Irf6*, is involved in osteoblast differentiation of craniofacial bone. *Dev Dyn.* 2019 Mar;248(3):221-232. doi: 10.1002/dvdy.13. Feb 7. PMID: 30684382.
30. **Jun G***, Manning A, Almeida M, Zawistowski M, Wood AR, Teslovich TM, ..., Altshuler D, McCarthy MI, Duggirala R, Boehnke M, Frayling TM, Abecasis GR, Blangero J. Evaluating the contribution of rare variants to type 2 diabetes and related traits using pedigrees. *Proc Natl Acad Sci U S A.* 2018;115(2):379-84. doi: 10.1073/pnas.1705859115. PMID: 29279374. *Corresponding.
31. Herrington DM, Mao C, Parker SJ, Fu Z, Yu G, Chen L, Venkatraman V, Fu Y, Wang Y, Howard TD, **Jun G**, Zhao CF, Liu Y, Saylor G, Spivia WR, Athas GB, Troxclair D, Hixson JE, Vander Heide RS, Wang Y, Van Eyk JE. Proteomic Architecture of Human Coronary and Aortic Atherosclerosis. *Circulation.* 2018;137(25):2741-56. Epub 2018/06/20. doi: 10.1161/CIRCULATIONAHA.118.034365. PMID: 29915101.
32. Suzuki A, Abdallah N, Gajera M, **Jun G**, Jia P, Zhao Z, Iwata J. Genes and microRNAs associated with mouse cleft palate: A systematic review and bioinformatics analysis. *Mech Dev.* 2018;150:21-7. Epub 2018/02/24. doi: 10.1016/j.mod.2018.02.003. PMID: 29475039.
33. Rashkin S, **Jun G**, Chen S, Genetics, Epidemiology of Colorectal Cancer C, Abecasis GR. Optimal sequencing strategies for identifying disease-associated singletons. *PLoS Genet.* 2017;13(6):e1006811. doi: 10.1371/journal.pgen.1006811. PMID: 28640830.
34. Metwalli KA, Do MA, Nguyen K, Mallick S, Kin K, Farokhnia N, **Jun G**, Fakhouri WD. Interferon Regulatory Factor 6 Is Necessary for Salivary Glands and Pancreas Development. *J Dent Res.* 2017;22034517729803. doi: 10.1177/0022034517729803. PMID: 28898113.
35. Manning A, Highland HM, Gasser J, Sim X, ..., **Jun G**, ..., Abecasis G, Bell GI, Blangero J, Cox NJ, Duggirala R, Seielstad M, Wilson JG, Dupuis J, Ripatti S, Hanis CL, Florez JC, Mohlke KL, Meigs JB, Laakso M, Morris AP, Boehnke M, Altshuler D, McCarthy MI, Gloyn AL, Lindgren CM. A Low-Frequency Inactivating *AKT2* Variant Enriched in the Finnish Population Is Associated With Fasting Insulin Levels and Type 2 Diabetes Risk. *Diabetes.* 2017;66(7):2019-32. doi: 10.2337/db16-1329. PMID: 28341696
36. Jiang ZD, Ajami NJ, Petrosino JF, **Jun G**, Hanis CL, Shah M, Hochman L, Ankoma-Sey V, DuPont AW, Wong MC, Alexander A, Ke S, DuPont HL. Randomised clinical trial: faecal microbiota transplantation for recurrent *Clostridium difficile* infection - fresh, or frozen, or lyophilised microbiota from a small pool of healthy donors delivered by colonoscopy. *Aliment Pharmacol Ther.* 2017;45(7):899-908. doi: 10.1111/apt.13969. PMID: 28220514.
37. Hixson JE, **Jun G**, Shimmin LC, Wang Y, Yu G, Mao C, Warren AS, Howard TD, Heide RSV, Van Eyk J, Wang Y, Herrington DM. Whole Exome Sequencing to Identify Genetic Variants Associated with Raised Atherosclerotic Lesions in Young Persons. *Sci Rep.* 2017;7(1):4091. doi: 10.1038/s41598-017-04433-x. PMID: 28642624.

38. Blangero J, Teslovich TM, Sim X, Almeida MA, **Jun G**, Dyer TD, Johnson M, Peralta JM, Manning A, Wood AR, Fuchsberger C, Kent JW, Jr., Aguilar DA, Below JE, Farook VS, Arya R, Fowler S, Blackwell TW, Puppala S, Kumar S, Glahn DC, Moses EK, Curran JE, Thameem F, Jenkinson CP, DeFronzo RA, Lehman DM, Hanis C, Abecasis G, Boehnke M, Goring H, Duggirala R, Almasy L, T2D-GENES Consortium. Omics-squared: human genomic, transcriptomic and phenotypic data for genetic analysis workshop 19. *BMC Proc.* 2016;10(Suppl 7):71-7. doi: 10.1186/s12919-016-0008-y. PMID: 27980614.
39. Almeida M, Blondell L, Peralta JM, Kent JW, Jr., **Jun G**, Teslovich TM, Fuchsberger C, Wood AR, Manning AK, Frayling TM, Cingolani PE, Sladek R, Dyer TD, Abecasis G, Duggirala R, Blangero J. Independent test assessment using the extreme value distribution theory. *BMC Proc.* 2016;10(Suppl 7):245-9. doi: 10.1186/s12919-016-0038-5. PMID: 27980644.
40. Nicholson AM, Finch NA, Almeida M, Perkerson RB, van Blitterswijk M, Wojtas A, Cenik B, Rotondo S, Inskip V, Almasy L, Dyer T, Peralta J, **Jun G**, Wood AR, Frayling TM, Fuchsberger C, Fowler S, Teslovich TM, Manning AK, Kumar S, Curran J, Lehman D, Abecasis G, Duggirala R, Pottier C, Zahir HA, Crook JE, Karydas A, Mitic L, Sun Y, Dickson DW, Bu G, Herz J, Yu G, Miller BL, Ferguson S, Petersen RC, Graff-Radford N, Blangero J, Rademakers R. Prosaposin is a regulator of progranulin levels and oligomerization. *Nature Communications.* 2016;7:11992. doi: 10.1038/ncomms11992. PMID: 27356620.
41. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, Ma C, Fontanillas P, Moutsianas L, McCarthy DJ, Rivas MA, Perry JRB, Sim X, ..., **Jun G**, *et al.* The genetic architecture of type 2 diabetes. *Nature.* 2016;536(7614):41-7. doi: 10.1038/nature18642. PMID: 27398621.
42. Wood AR, Tuke MA, Nalls M, Hernandez D, Gibbs JR, Lin H, Xu CS, Li Q, Shen J, **Jun G**, Almeida M, Tanaka T, Perry JRB, Gaulton K, Rivas M, Pearson R, Curran JE, Johnson MP, Göring HHH, Duggirala R, Blangero J, McCarthy MI, Bandinelli S, Murray A, Weedon MN, Singleton A, Melzer D, Ferrucci L, Frayling TM. Whole-genome sequencing to understand the genetic architecture of common gene expression and biomarker phenotypes. *Human Molecular Genetics.* 2015;24(5):1504-12. doi: 10.1093/hmg/ddu560. PMID: 25378555.
43. Williams AL, Genovese G, Dyer T, Altemose N, Truax K, **Jun G**, Patterson N, Myers SR, Curran JE, Duggirala R, Blangero J, Reich D, Przeworski M. Non-crossover gene conversions show strong GC bias and unexpected clustering in humans. *eLife.* 2015;4:e04637. doi: 10.7554/eLife.04637. PMID: 25806687.
44. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74. doi: 10.1038/nature15393. PMID: 26432245.
45. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, **Jun G**, *et al.* An integrated map of structural variation in 2,504 human genomes. *Nature.* 2015;526(7571):75-81. doi: 10.1038/nature15394. PMID: 26432246.
46. Mahajan A, Sim X, Ng HJ, Manning A, Rivas MA, Highland HM, Locke AE, Grarup N, Im HK, Cingolani P, Flannick J, Fontanillas P, Fuchsberger C, Gaulton KJ, Teslovich TM, Rayner NW, Robertson NR, Beer NL, Rundle JK, Bork-Jensen J, Ladenvall C, Blancher C, Buck D, Buck G, Burt NP, Gabriel S, Gjesing AP, Groves CJ, Hollensted M, Huyghe JR, Jackson AU, **Jun G**, *et al.* , On Behalf of the T2D-GENES Consortium and GoT2D Consortium. Identification and Functional Characterization of G6PC2 Coding Variants Influencing Glycemic Traits Define an Effector Transcript at the G6PC2-ABCB11 Locus. *PLoS Genet.* 2015;11(1):e1004876. doi: 10.1371/journal.pgen.1004876. PMID: 25625282.
47. **Jun G**, Wing MK, Abecasis GR, Kang HM. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Research.* 2015;25(6):918-25. doi: 10.1101/gr.176552.114. PMID: 25883319.
48. Flickinger M, **Jun G**, Abecasis Gonçalo R, Boehnke M, Kang Hyun M. Correcting for Sample Contamination in Genotype Calling of DNA Sequence Data. *The American Journal of Human Genetics.* 2015;97(2):284-90. doi: http://dx.doi.org/10.1016/j.ajhg.2015.07.002. PMID: 26235984.

49. Almasy L, Dyer TD, Peralta JM, **Jun G**, Wood AR, Fuchsberger C, Almeida MA, Kent JW Jr, Fowler S, Blackwell TW, Puppala S, Kumar S, Curran JE, Lehman D, Abecasis G, Duggirala R, Blangero J; T2D-GENES Consortium. Data for Genetic Analysis Workshop 18: human whole genome sequence, blood pressure, and simulated phenotypes in extended pedigrees. *BMC Proc.* 2014 Jun 17;8(Suppl1):S2. doi: 10.1186/1753-6561-8-S1-S2. PMID: 25519314.
50. Vrieze SI, Malone SM, Vaidyanathan U, Kwong A, Kang HM, Zhan X, Flickinger M, Irons D, **Jun G**, Locke AE, Pistis G, Porcu E, Levy S, Myers RM, Oetting W, McGue M, Abecasis G, Iacono WG. In search of rare variants: Preliminary results from whole genome sequencing of 1,325 individuals with psychophysiological endophenotypes. *Psychophysiology.* 2014;51(12):1309-20. doi: 10.1111/psyp.12350. PMID: 25387710.
51. TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitzel NO, Lange LA, Lu Y, Tang ZZ, Zhang H, Hindy G, Masca N, Stirrups K, Kanoni S, Do R, **Jun G**, et al. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med.* 2014 Jul 3;371(1):22-31. doi: 10.1056/NEJMoa1307095. Epub 2014 Jun 18. PMID: 24941081.
52. Lange LA, Hu Y, Zhang H, Xue C, Schmidt EM, Tang ZZ, Bizon C, Lange EM, Smith JD, Turner EH, **Jun G**, et al.; NHLBI Grand Opportunity Exome Sequencing Project., Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. *Am J Hum Genet.* 2014 Feb 6;94(2):233-45. doi: 10.1016/j.ajhg.2014.01.010. PMID: 24507775.
53. Zhan X, Larson DE, Wang C, Koboldt DC, Sergeev YV, Fulton RS, Fulton LL, Fronick CC, Branham KE, Bragg-Gresham J, **Jun G**, Hu Y, Kang HM, Liu D, Othman M, Brooks M, Ratnapriya R, Boleda A, Grassmann F, Von Strachwitz C, Olson LM, Buitendijk GHS, Hofman A, Van Duijn CM, Cipriani V, Moore AT, Shahid H, Jiang Y, Conley YP, Morgan DJ, Kim IK, Johnson MP, Cantsilieris S, Richardson AJ, Guymer RH, Luo H, Ouyang H, Licht C, Pluthero FG, Zhang MM, Zhang K, Baird PN, Blangero J, Klein ML, Farrer LA, DeAngelis MM, Weeks DE, Gorin MB, Yates JRW, Klaver CCW, Pericak-Vance MA, Haines JL, Weber BHF, Wilson RK, Heckenlively JR, Chew EY, Stambolian D, Mardis ER, Swaroop A, Abecasis GR. Identification of a rare coding variant in complement 3 associated with age-related macular degeneration. *Nature Genetics.* 2013;45(11):1375-81. doi: 10.1038/ng.2758. PMID: 24036949.
54. **Jun G**, Ghosh J. Semisupervised learning of hyperspectral data with unknown land-cover classes. *IEEE Transactions on Geoscience and Remote Sensing.* 2013;51(1):273-82. doi: 10.1109/TGRS.2012.2198654.
55. Fu W, O'Connor TD, **Jun G**, Kang HM, Abecasis G, Leal SM, Gabriel S, Altshuler D, Shendure J, Nickerson DA, Bamshad MJ, Akey JM. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature.* 2013;493(7431):216-20. doi: 10.1038/nature11690. PMID: 23201682.
56. Duan Q, Liu EY, Auer PL, Zhang G, Lange EM, **Jun G**, Bizon C, Jiao S, Buyske S, Franceschini N, Carlson CS, Hsu L, Reiner AP, Peters U, Haessler J, Curtis K, Wassel CL, Robinson JG, Martin LW, Haiman CA, Le Marchand L, Matise TC, Hindorff LA, Crawford DC, Assimes TL, Kang HM, Heiss G, Jackson RD, Kooperberg C, Wilson JG, Abecasis GR, North KE, Nickerson DA, Lange LA, Li Y. Imputation of coding variants in African Americans: Better performance using data from the exome sequencing project. *Bioinformatics.* 2013;29(21):2744-9. doi: 10.1093/bioinformatics/btt477. PMID: 23956302.
57. Tennessen JA, Bigam AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, **Jun G**, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science.* 2012;336(6090):64-9. doi: 10.1126/science.1219240. PMID: 22604720.

58. **Jun G**, Flickinger M, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, Boehnke M, Kang HM. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *American Journal of Human Genetics*. 2012;91(5):839-48. doi: 10.1016/j.ajhg.2012.09.004. PMID: 23103226.
59. The 1000 Genomes Consortium, An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56-65. doi: 10.1038/nature11632. PMID: 23128226
60. **Jun G**, Ghosh J. Spatially adaptive classification of land cover with remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*. 2011;49(7):2662-73. doi: 10.1109/TGRS.2011.2105490.
61. **Jun G**, Ghosh J. Spatially adaptive semi-supervised learning with Gaussian processes for hyperspectral data analysis. *Statistical Analysis and Data Mining*. 2011;4(4):358-71. doi: 10.1002/sam.10119.
62. Liu A, **Jun G**, Ghosh J. A self-training approach to cost sensitive uncertainty sampling. *Machine Learning*. 2009;76(2-3):257-70. doi: 10.1007/s10994-009-5131-9.

Peer-reviewed Conference Proceedings

1. Symons CT, Vatsavai RR, **Jun G**, Arel I. Bias selection using task-targeted random subspaces for robust application of graph-based semi-supervised learning. Proceedings - 2012 11th International Conference on Machine Learning and Applications, ICMLA 2012. 2012;1:415-20. doi: 10.1109/ICMLA.2012.75.
2. Vatsavai RR, Symons CT, Chandola V, **Jun G**. GX-means: A model-based divide and merge algorithm for geospatial image clustering. *Procedia Computer Science*. 2011;4:186-95. doi: 10.1016/j.procs.2011.04.020.
3. **Jun G**, Ghosh J, Radosavljevic V, Obradovic Z. Predicting ground-based aerosol optical depth with satellite images via Gaussian processes. KDIR 2010 - Proceedings of the International Conference on Knowledge Discovery and Information Retrieval. 2010:370-5.
4. **Jun G**, Ghosh J. Nearest-manifold classification with Gaussian processes. Proceedings - International Conference on Pattern Recognition. 2010:914-7. doi: 10.1109/ICPR.2010.230.
5. Liu A, **Jun G**, Ghosh J. Active learning of hyperspectral data with spatially dependent label acquisition costs. International Geoscience and Remote Sensing Symposium (IGARSS). 2009;5. doi: 10.1109/IGARSS.2009.5417684.
6. Liu A, **Jun G**, Ghosh J. Spatially cost-sensitive active learning. Society for Industrial and Applied Mathematics - 9th SIAM International Conference on Data Mining 2009, Proceedings in Applied Mathematics. 2009;2:810-21.
7. **Jun G**, Vatsavai RR, Ghosh J. Spatially adaptive classification and active learning of multispectral data with gaussian processes. ICDM Workshops 2009 - IEEE International Conference on Data Mining. 2009:597-603. doi: 10.1109/ICDMW.2009.107.
8. **Jun G**, Ghosh J. Spatially adaptive classification of hyperspectral data with Gaussian processes. International Geoscience and Remote Sensing Symposium (IGARSS). 2009;2. doi: 10.1109/IGARSS.2009.5418067.
9. **Jun G**, Ghosh J. Multi-class boosting with class hierarchies. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2009;5519 LNCS:32-41. doi: 10.1007/978-3-642-02326-2_4.
10. **Jun G**, Ghosh J. Hybrid hierarchical classifiers for hyperspectral data analysis. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2009;5519 LNCS:42-51. doi: 10.1007/978-3-642-02326-2_5.

11. **Jun G**, Ghosh J. An efficient active learning algorithm with knowledge transfer for hyperspectral data analysis. International Geoscience and Remote Sensing Symposium (IGARSS). 2008;1(1). doi: 10.1109/IGARSS.2008.4778790.
12. **Jun G**, Aggarwal JK, Gökmen M. Tracking and segmentation of highway vehicles in cluttered and crowded scenes. 2008 IEEE Workshop on Applications of Computer Vision, WACV. 2008. doi: 10.1109/WACV.2008.4544017.
13. **Jun G**. Home media center and media clients for multi-room audio and video applications. Second IEEE Consumer Communications and Networking Conference, CCNC 2005; pp. 257-260.

Conference Presentations – Oral/Platform

- *Progression and Subtypes of Prediabetes with Metabolomics Profiling in Starr County Mexican Americans*, 77th Scientific Sessions, American Diabetes Association (ADA), Abstract #22-OR, San Diego, CA, 2017
- *Deep whole-genome sequencing in pedigrees to quantify the contribution of private variants to type 2 diabetes and related metabolic traits*, 63rd Annual Meeting of American Society of Human Genetics (ASHG), Abstract #36, Boston, MA, 2013.
- *Detecting Functional Rare Variants Relating to Type 2 Diabetes Using Deep Whole Genome Sequencing*, 72nd Scientific Sessions, American Diabetes Association (ADA), Abstract #140-OR, Philadelphia, PA, 2012.
- *Joint variant calling and analysis across >4,000 exomes of European and African American ancestry*, 12th International Congress of Human Genetics (ICHG) / 61st Annual Meeting of American Society of Human Genetics (ASHG), Abstract #192, Montreal, Quebec, Canada, 2011

Conference Presentations – Poster

- *Identification of novel structural variations affecting common and complex disease risks with >16,000 whole genome sequences from ARIC and HCHS/SOL*, 68th Annual Meeting of American Society of Human Genetics (ASHG), Abstract #3186W, San Diego, CA, 2018. ***Reviewer's Choice**
- *Understanding progression and subtypes of prediabetes with metabolomics and genomic profiling in Starr County Mexican Americans*, 67th Annual Meeting of American Society of Human Genetics (ASHG), Abstract #2764W, Baltimore, MD, 2017.
- *Integrated Genetic Analysis Platform (IGAP) for Web-based Interactive Association Analysis and Visualization of Large Scale Genotype/Phenotype Data*, The Allied Genetics Conference (TAGC), Abstract #P2087C, Orlando, FL, 2016.
- *Scalable multi-sample variant caller (MultiVAC) with fast and efficient local de novo assembly*, Biology of Genomes, Cold Spring Harbor, NY, 2016.
- *Robust Multi-sample Calling of Structural Variations from Complete Genomics Sequencing Data*, Biology of Genomes, Cold Spring Harbor, NY, 2014
- *Analyzing Deep Whole Genome Sequence and Genotype Data of >1,000 Individuals from Large Mexican-American Pedigrees in the T2D-GENES Study*, 62nd Annual Meeting of American Society of Human Genetics (ASHG), Abstract #1513F, San Francisco, CA, 2012.

Invited Talks and Lectures

- Lecture, *Analysis of High Throughput Sequencing Data in Population Scale*, 13th Asian Institute in Statistical Genetics and Genomics, Seoul, South Korea, July 2018

- Lecture, *Variant calling and filtering for SNPs*, Sequence Analysis Workshop, University of Michigan, Ann Arbor, Michigan, June 2014
- Invited Talk, *Methods for modeling and correcting sample contamination in next-generation sequencing data*, 4th International Conference for Statistical Methods in the Biomedical Science and Bioinformatics, Choong-Ang University, Seoul, Korea, June 2013
- Lecture, *Sequencing: Mapping, alignment, and variant calling with practical/Visualization and QC of variant calls*, 2013 International Workshop on Statistical Genetic Methods for Human Complex Traits, Boulder, Colorado, March 2013

Software

- muCNV
 - Multi-sample SV genotyping for large-scale WGS data
 - <https://github.com/gjun/mucnv>
- verifyIDintensity
 - Detecting and estimating sample contamination in genotype array data
 - <http://github.com/gjun/verifyIDintensity>

Patents (Inventor)

- US Patent 9,342,141 and 8,504,637, KR Patent 1005773620000, Audio/video device, apparatus and method for controlling audio/video device
- US Patent 8,260,843, KR Patent 1007728610000, Apparatus and method for providing remote user interface
- US Patent 8,196,170, KR Patent 1007556950000, Method of connecting to internet via broadcast receiving device and apparatus for the same
- KR Patent 1004777050000, Method for displaying main screen of home media center
- KR Patent 1005974040000, Method and apparatus for transferring home media center's media contents apart from home media center
- KR Patent 1007886520000, Apparatus and method for dialing auto sound

Teaching

- PH2612 Epidemiology I
 - Fall 2024 (100%)
- PH2998 (Special Topics) / PH2781 (Regular) Practical Python Programming / Algorithms for Data Analysis
 - Summer 2016 (50%), Summer 2017-2024, Fall 2023 (100%)
- PH2998 (Special Topics) / PH2872 Practical Computational Genomics & Bioinformatics
 - Spring 2016-2018 (20%), Spring 2019-2024 (100%)
- PH2998 (Special Topics) Computational Genomics and NGS Data Analysis
 - Spring 2017 (80%), Fall 2017 (100%)
- PH2615 Epidemiology II
 - Fall 2017-2018 (10%)

- PH2815 Genetics and Human Disease
 - Fall 2017-2024 (5%)
- PH1998 (Special Topics) Data Science and Analytic Method
 - Summer 2016 (33%)
- PH1980 Intro to Genomics and Bioinformatics
 - Fall 2016 (16%)

Service to the profession

Editorial Board

- BMC Cardiovascular Disorders, *Associate Editor*
- Frontiers in Genetics, *Associate Editor*
- PEERJ, *Editor*
- IEEE Trans. on Molecular, Biological, and Multi-scale Communications, *Associate Editor*

Journal Reviewer

- Bioinformatics (Oxford)
- Nature Genetics
- Nature Protocols
- Nature Communications
- PLOS Computational Biology
- PLOS Genetics
- BMC Bioinformatics
- Genome Research
- Patter Recognition
- Remote Sensing
- IEEE Transactions on Neural Network and Learning Systems
- IEEE Transactions on Image Processing
- IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing
- IEEE Geoscience and Remote Sensing Letters

Program Committee

- International Workshop on Spatial and Spatiotemporal Data Mining (SSTDM), held in conjunction with IEEE International Conference on Data Mining (ICDM)
- International Workshop on Knowledge Discovery from Sensor Data (SensorKDD), held in conjunction with ACM SIGKDD Conference on Knowledge Discovery and Data Mining
- ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (BigSpatial)

Other Services

- Co-coordinator, NHGRI's Genome Sequencing Program (GSP) Methods Working Group
- Peer Reviewer, Wellcome Trust grants
- Conference Chair, KSEA West Gulf Coast Regional Conference / Korean-American Biomedical Scientist Symposium, November 2017

Service to the institution

- Faculty Council, EHGES Alternate Representative, UTHealth School of Public Health

- Student Scholarship Committee, MD Anderson UTHealth Graduate School of Biomedical Sciences, 2020-2022
- Candidacy Exam Committee, Quantitative Science Program, MD Anderson UTHealth Graduate School of Biomedical Sciences, 2018-2022
- Membership Committee, Quantitative Science Program, MD Anderson UTHealth Graduate School of Biomedical Sciences, 2018-2022
- Reviewer, 2020 John J. Kopchick Research Award, MD Anderson UTHealth GSBS
- Reviewer, 2018 PRIME Award, UTHealth School of Public Health

Service to the community

- Online Independent Study and Mentorship Course Mentor, Katy Independent School District
- Membership Director, Korean-American Scientists and Engineers Association (KSEA), Vienna, VA, 2021-2023
- Chapter President, South Texas Chapter, Korean-American Scientists and Engineers Association (KSEA), 2019-2020
- Council Member, Korean-American Scientists and Engineers Association, 2019-2020
- Organizing Chair, KSEA National Math & Science Competition, April 2019
- President Elect, South Texas Chapter, Korean-American Scientists and Engineers Association, 2018-2019