SUN, Kun (孙坤)

Room A1921, Gaoke Innovation Center, Guangming District, Shenzhen 518132, China

Email: sunkun@szbl.ac.cn   hellosunking@foxmail.com

Personal website: https://hellosunking.github.io

ORCID: 0000-0002-9883-1616

Google scholar: https://scholar.google.com/citations?user=F_B_a6cAAAAJ

EDUCATION

Ph.D. in Department of Chemical Pathology   2014
The Chinese University of Hong Kong

B.Sci. in Department of Electronic Engineering   2010
University of Science and Technology of China

RESEARCH EXPERIENCE

Assistant Professor/Principal Investigator (faculty)   2019-current
Institute of Cancer Research, Shenzhen Bay Laboratory

Research Assistant Professor (faculty),   2017-2019
Postdoctoral fellow   2014-2017
Department of Chemical Pathology, The Chinese University of Hong Kong
Supervisors: Prof. **Y.M. Dennis Lo (卢煜明)** and Prof. Rossa W.K. Chiu (赵慧君).

Doctoral thesis, Department of Chemical Pathology,   2010-2014
The Chinese University of Hong Kong
Supervisors: Prof. Hao Sun (孙昊) and Prof. Huating Wang (王华婷).

Bachelor thesis, Department of Electric Engineering,   2006-2010
University of Science and Technology of China
Supervisor: Prof. Xiang Chen (陈香) on National Students' Innovative Project 2009 (国家大学生创新计划); Prof. Fanrang Kong (孔凡让) and Prof. Hui Li (李辉) on ABU Robocon 2010, China Division (央视机器人大赛).

RESEARCH INTERESTS

➢ **Cancer liquid biopsy and Noninvasive prenatal testing**
➢ **Computational biology/Bioinformatics**
➢ Mechanisms of cancer metastasis

➤ LncRNA identification and functional investigation
➤ Biomarkers in cancers and diseases
➤ Data mining and visualization
➤ Functional genomics

FEATURED PUBLICATIONS

(#: co-first authorship; *: (co-)corresponding authorship)

1. **Sun K**[#], Jiang P[#], Chan KCA[#], Wong J, Cheng YK, Liang RH, Chan WK, Ma ES, Chan SL, Cheng SH, Chan RW, Tong YK, Ng SS, Mong RSM, Hui DS, Leung TN, Leung TY, Lai PBS, Chiu RWK, Lo YMD. Plasma DNA tissue mapping by genomewide methylation sequencing for noninvasive prenatal, cancer and transplantation assessments. *Proc Natl Acad Sci U S A* 2015 Oct 6; 112(40):E5503-12. (**ESI Highly Cited Paper**; Citation: 710+)

2. **Sun K***, Jiang P, Cheng SH, Cheng THT, Wong J, Wong VWS, Ng SSM, Ma BBY, Leung TY, Chan SL, Mok TSK, Lai PBS, Chan HLY, Sun H, Chan KCA, Chiu RWK, Lo YMD*. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. *Genome Res* 2019 Mar; 29(3):418-427. (Citation: 190+)

3. Jiang P[#], **Sun K**[#], Peng W[#], Cheng SH, Ni M, Yeung PC, Heung MM, Xie T, Shang H, Zhou Z, Chan RW, Wong J, Wong VW, Poon LC, Leung TY, Lam WKJ, Chan JYK, Chan HLY, Chan KCA, Chiu RWK, Lo YMD. Plasma DNA end motif profiling as a fragmentomic marker in cancer, pregnancy and transplantation. *Cancer Discov* 2020 May; 10(5):664-673. (Citation: 190+)

4. An Y, Zhao X, Zhang Z, Xia Z, Yang M, Ma L, Zhao Y, Xu G, Du S, Wu X, Zhang S, Hong X, Jin X*, **Sun K***. DNA methylation analysis explores the molecular basis of plasma cell-free DNA fragmentation. *Nat Commun* 2023 Jan 18; 14(1):287. (Citation: 25+)

5. **Sun K**, Jiang P, Wong AIC, Cheng YKY, Cheng SH, Zhang H, Chan KCA, Leung TY, Chiu RWK, Lo YMD. Size-tagged preferred ends in maternal plasma DNA shed light on production mechanism and show utility in noninvasive prenatal testing. *Proc Natl Acad Sci U S A* 2018 May 29; 115(22):E5106-E5114. (Citation: 130+)

COMPLETE PUBLICATION LIST
Total citation=5074, h-index=31, i100-index=20 at Jul 31, 2024 (Google Scholar)

Refereed journal papers:

1. Liu X#, Yang M#, Hu D, An Y, Wang W, Lin H, Pan Y, Ju J, **Sun K***. Systematic biases in reference-based plasma cell-free DNA fragmentomic profiling. *Cell Rep Methods.* 2024 Jun 17; 4(6):100793. (IF: 4.3)

2. Zhao Y, Yang M, Gong F, Pan Y, Hu M, Peng Q, Lu L, Lyu X, **Sun K***. Accelerating 3D genomics data analysis with Microcket. *Commun Biol.* 2024 Jun 1; 7:675. (IF: 5.9)

3. Hu D#, Zhang Z#, Liu X#, Wu Y#, An Y, Wang W, Yang M, Pan Y, Qiao K, Du C, Zhao Y, Li Y, Bao J, Qin T, Pan Y, Xia Z*, Zhao X*, **Sun K***. Generalizable transcriptome-based tumor malignant level evaluation and molecular subtyping towards precision oncology. *J Transl Med.* 2024 May 28; 22:512. (IF: 7.4)

4. Jin X*,Wang Y, Xu J, Li Y, Cheng F, Luo Y, Zhou H, Lin S, Xiao F, Zhang L, Lin Y, Zhang Z, Jin Y, Zheng F, Chen W, Zhu A, Tao Y, Zhao J, Kuo T, Li Y, Li L, Wen L, Ou R, Li F, Lin L, Zhang Y, Sun J, Yuan H, Zhuang Z, Sun H, Chen Z, Li J, Zhuo J, Chen D, Zhang S, Sun Y, Wei P, Yuan J, Xu T, Yang H, Wang J, Xu X, Zhong N, Xu Y*, **Sun K***, Zhao J*. Plasma cell-free DNA promise monitoring and tissue injury assessment of COVID-19. *Mol Genet Genomics.* 2023 Jul; 298(4):823-836. (IF: 2.980)

5. An Y, Zhao X, Zhang Z, Xia Z, Yang M, Ma L, Zhao Y, Xu G, Du S, Wu X, Zhang S, Hong X, Jin X*, **Sun K***. DNA methylation analysis explores the molecular basis of plasma cell-free DNA fragmentation. *Nat Commun* 2023 Jan 18; 14(1):287. (IF: 17.694)

6. Li L, An Y, Ma L, Yang M, Yuan P, Liu X, Jin X, Zhao Y, Zhang S*, Hong X*, **Sun K***. Msuite2: all-in-one DNA methylation data analysis toolkit with enhanced usability and performance. *Comput Struct Biotechnol J* 2022 Mar 10; 20:1271-1276. (IF: 6.155)

7. **Sun K***, Wang H, Sun H*. NAMS webserver: coding potential assessment and functional annotation of plant transcripts. *Brief Bioinform* 2021 May 20; 22(3):bbaa200. (IF: 13.994)

8. **Sun K***, Gu L, Ma L, Duan Y*. Atlas of ACE2 gene expression reveals novel insights into transmission of SARS-CoV-2. *Heliyon* 2021 Jan 1; 7(1):e05850. (IF: 3.776) (**Heliyon Monthly Paper**)

9. **Sun K***, Li L, Ma L, Zhao Y, Deng L, Wang H, Sun H. Msuite: a high-performance and versatile DNA methylation data analysis toolkit. *Patterns (N Y)* 2020 Nov 13; 1(8):100127. (IF: 6.5) (**Editor's Choice Article**)

10. Fan F, Chen D, Zhao Y, Wang H, Sun H, **Sun K***. Rapid preliminary purity evaluation of tumor biopsies using deep learning approach. *Comput Struct Biotechnol J* 2020 Jun 16; 18:1746-1753. (IF: 6.155)

11. **Sun K***. Ktrim: an extra-fast and accurate adapter- and quality-trimmer for sequencing data. *Bioinformatics* 2020 Jun 1; 36(11):3561-3562. (IF: 6.931)

12. Jiang P[#], **Sun K**[#], Peng W[#], Cheng SH, Ni M, Yeung PC, Heung MM, Xie T, Shang H, Zhou Z, Chan RW, Wong J, Wong VW, Poon LC, Leung TY, Lam WKJ, Chan JYK, Chan HLY, Chan KCA, Chiu RWK, Lo YMD. Plasma DNA end motif profiling as a fragmentomic marker in cancer, pregnancy and transplantation. *Cancer Discov* 2020 May; 10(5):664-673. (IF: 38.273) (Highlighted in issue)

13. **Sun K**∗, Jiang P, Cheng SH, Cheng THT, Wong J, Wong VWS, Ng SSM, Ma BBY, Leung TY, Chan SL, Mok TSK, Lai PBS, Chan HLY, Sun H, Chan KCA, Chiu RWK, Lo YMD∗. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. *Genome Res* 2019 Mar; 29(3):418-427. (IF: 9.438)

14. **Sun K**∗, Wang J, Wang H, Sun H∗. GeneCT: a generalizable cancerous status and tissue origin classifier for pan-cancer biopsies. *Bioinformatics* 2018 Dec 1; 34(23):4129-4130. (IF: 6.931)

15. Jiang P[#], **Sun K**[#], Tong YK, Cheng SH, Cheng THT, Heung MMS, Wong J, Wong VWS, Chan HLY, Chan KCA, Lo YMD, Chiu RWK. Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma. *Proc Natl Acad Sci U S A* 2018 Nov 13; 115(46):E10925-E10933. (IF: 12.779)

16. **Sun K**, Jiang P, Wong AIC, Cheng YKY, Cheng SH, Zhang H, Chan KCA, Leung TY, Chiu RWK, Lo YMD. Size-tagged preferred ends in maternal plasma DNA shed light on production mechanism and show utility in noninvasive prenatal testing. *Proc Natl Acad Sci U S A* 2018 May 29; 115(22):E5106-E5114. (IF: 12.779)

17. **Sun K**, Lun FMF, Leung TY, Chiu RWK, Lo YMD, Sun H. Noninvasive reconstruction of placental methylome from maternal plasma DNA: potential for prenatal testing and monitoring. *Prenat Diagn* 2018 Feb; 38(3):196-203. (IF: 3.242) (**Issue cover article**)

18. **Sun K**, Lun FMF, Jiang P, Sun H. BSviewer: a genotype-preserving, nucleotide-level visualizer for bisulfite sequencing data. *Bioinformatics* 2017 Nov 1; 33(21):3495-3496. (IF: 6.931)

19. Lam WKJ[#], Gai W[#], **Sun K**[#], Wong RSM, Chan RWY, Jiang P, Chan NPH, Hui WWI, Chan AWH, Szeto CC, Ng SC, Law MF, Chan KCA, Chiu RWK, Lo YMD. DNA of erythroid origin is present in human plasma and informs the types of anemia. *Clin Chem* 2017 Oct; 63(10):1614-1623. (IF: 12.167)

20. **Sun K**, Wang H, Sun H. mTFkb: a knowledgebase for fundamental annotation of mouse transcription factors. *Sci Rep* 2017 Jun 8; 7(1):3022. (IF: 4.996)

21. **Sun K**, Chan KCA, Hudecova I, Chiu RWK, Lo YMD, Jiang P. COFFEE: control-free noninvasive fetal chromosomal examination using maternal plasma DNA. *Prenat Diagn* 2017 Apr; 37(4):336-340. (IF: 3.242) (**Issue cover article**)

22. Chan KCA[#], Jiang P[#], **Sun K**[#], Cheng YKY, Tong YK, Cheng SH, Wong AIC, Hudecova I, Leung TY, Chiu RWK, Lo YMD. Second generation noninvasive fetal genome analysis reveals de novo mutations, single-base parental inheritance, and preferred DNA ends. *Proc Natl Acad Sci U S A* 2016 Dec 13; 113(50):E8159-8168. (IF: 12.779)

23. Zhou L[#], **Sun K**[#], Zhao Y[#], Zhang S, Wang X, Li Y, Lu L, Chen X, Chen F, Bao X, Zhu X, Wang

L, Tang LY, Esteban MA, Wang R, Jauch R, Sun H, Wang H. Linc-YY1, a novel functional long non-coding RNA in myogenic differentiation and muscle regeneration. ***Nat Commun*** 2015 Dec 11; 6:10026. (IF: 17.694)

24. **Sun K[#],** Jiang P[#], Chan KCA[#], Wong J, Cheng YK, Liang RH, Chan WK, Ma ES, Chan SL, Cheng SH, Chan RW, Tong YK, Ng SS, Mong RSM, Hui DS, Leung TN, Leung TY, Lai PBS, Chiu RWK, Lo YMD. Plasma DNA tissue mapping by genomewide methylation sequencing for noninvasive prenatal, cancer and transplantation assessments. ***Proc Natl Acad Sci U S A*** 2015 Oct 6; 112(40):E5503-12. (IF: 12.779) (**ESI Highly Cited Paper**)

25. **Sun K,** Yuen YP, Wang H, Sun H. Online Diagnosis System: a webserver for analysis of Sanger sequencing-based genetic testing data. ***Methods*** 2014 Oct 1; 69(3):230-236. (IF: 4.647)

26. **Sun K**, Zhao Y, Wang H, Sun H. Sebnif: an integrated bioinformatics pipeline for the identification of novel large intergenic noncoding RNAs (lincRNAs) - application in human skeletal muscle cells. ***PLoS ONE*** 2014 Jan 6; 9(1):e84500. (IF: 3.752)

27. Lu L[#], **Sun K[#]**, Chen X, Zhao Y, Wang L, Zhou L, Sun H, Wang H. Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. ***EMBO J*** 2013 Oct 2; 32(19):2575-88. (IF: 14.012)

28. **Sun K**, Chen X, Jiang P, Song X, Wang H, Sun H. iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. ***BMC Genomics*** 2013; 14(Suppl 2):S7. (IF: 4.547)

29. Guo T, Li A, Sun P, He Z, Cai Y, Lan G, Liu L, Li J, Yang J, Zhu Y, Zhao R, Chen X, Shi D, Liu Z, Wang Q, Xu L, Zhou L, Ran P, Wang X, **Sun K**, Lu J, Han Y. Astrocyte reactivity is associated with tau tangle load and cortical thinning in Alzheimer's disease. ***Mol Neurodegener.*** 2024 Jul 30; 19(1):58.

30. Liu Z, Shi D, Cai Y, Li A, Lan G, Sun P, Liu L, Zhu Y, Yang J, Zhou Y, Guo L, Zhang L, Deng S, Chen S, Yu X, Chen X, Zhao R, Wang Q, Ran P, Xu L, Zhou L, **Sun K**, Wang Z, Peng Q, Han Y, Guo T. Pathophysiology characterization of Alzheimer's disease in South China's aging population: for the Greater-Bay-Area Healthy Aging Brain Study (GHABS). ***Alzheimers Res Ther.*** 2024 Apr 16; 16(1):84.

31. Guo T, Zhang M, Shi D, Li Y, Chen X, Zhao R, Lan G, Cai Y, Sun P, Li A, Zhu Y, Liu L, Guo L, Zhou Y, Liu Z, **Sun K**, Wang X, Xu L, Wang Q, Han Y. Plasma biomarker and neuroimaging changes in a Chinese community aging cohort: The Greater-Bay-Area Healthy Aging Brain Study (GHABS). ***Alzheimer's & Dementia*** 2023 Dec; 19(S14):e077578.

32. Li A, Du J, Cai Y, Chen X, **Sun K**, Guo T. Body Mass Index Decrease Has a Distinct Association with Alzheimer's Disease Pathophysiology in APOE ε4 Carriers and Non-Carriers. ***J Alzheimers Dis.*** 2023 Nov 7; 96(2):643-655.

33. Du J, Li A, Shi D, Chen X, Wang Q, Liu Z, **Sun K**, Guo T. Association of APOE4, osteoarthritis, β-amyloid, and tau accumulation in primary motor and somatosensory regions in Alzheimer's

Disease. *Neurology* 2023 May 15; 10.1212/WNL.000000000207369.

34. Liu Y, **Sun K**, Gan Y, Liu H, Yu J, Xu W, Zhang L, Chen D. RNA-sequencing reveals gene expression and pathway signatures in umbilical cord blood affected by birth delivery mode. *Phenomics* 2023 Apr 8; 3(3):228-242.

35. Cai Y, Du J, Li A, Zhu Y, Xu L, **Sun K**, Ma S, Guo T, Alzheimer's Disease Neuroimaging Initiative. Initial levels of β-amyloid and tau deposition have distinct effects on longitudinal tau accumulation in Alzheimer's disease. *Alzheimers Res Ther.* 2023 Feb 7; 15(1):30.

36. Peng Q, Huang Z, **Sun K**, Liu Y, Yoon CW, Harrison RES, Schmitt DL, Zhu L, Wu Y, Tasan I, Zhao H, Zhang J, Zhong S, Chien S, Wang Y. Engineering inducible biomolecular assemblies for genome imaging and manipulation in live cells. *Nat Commun* 2022 Dec 24; 13(1):7933.

37. Wang Y, Li J, Zhang L, Sun HX, Zhang Z, Xu J, Xu Y, Lin Y, Zhu A, Luo Y, Zhou H, Wu Y, Lin S, Sun Y, Xiao F, Chen R, Wen L, Chen W, Li F, Ou R, Zhang Y, Kuo T, Li Y, Li L, Sun J, **Sun K**, Zhuang Z, Lu H, Chen Z, Mai G, Zhuo J, Qian P, Chen J, Yang H, Wang J, Xu X, Zhong N, Zhao J, Li J, Zhao J, Jin X. Plasma cell-free RNA characteristics in COVID-19 patients. *Genome Res* 2022 Feb; 32(2):228-241.

38. Chen X, Wu T, Li L, Lin Y, Ma Z, Xu J, Li H, Cheng F, Chen R, **Sun K**, Luo Y, Zhang C, Chen F, Wang J, Kuo T, Li X, Geng C, Lin F, Huang C, Hu J, Yin J, Liu M, Tao Y, Zhang J, Ou R, Zheng F, Jin Y, Yang H, Wang J, Xu X, Fu S, Jiang H, Jin X, Zhang H. Transcriptional start site coverage analysis in plasma cell-free DNA reveals disease severity and tissue specificity of COVID-19 patients. *Front. Genet.* 2021 May 28; 12:663098.

39. Zhao Y, Zhou J, He L, Li Y, Yuan J, **Sun K**, Chen X, Bao X, Esteban MA, Sun H, Wang H. MyoD induced enhancer RNA interacts with hnRNPL to activate target gene transcription during myogenic differentiation. *Nat Commun* 2019 Dec 19; 10(1):5787.

40. Serpas L, Chan RWY, Jiang P, Ni M, **Sun K**, Rashidfarrokhi A, Soni C, Sisirak V, Lee WS, Cheng SH, Peng W, Chan KCA, Chiu RWK, Reizis B, Lo YMD. *Dnase1l3* deletion causes aberrations in length and end motif frequencies in plasma DNA. *Proc Natl Acad Sci U S A* 2019 Jan 8; 116(2):641-649.

41. Gai W, Ji L, Lam WKJ, **Sun K**, Jiang P, Chan AWH, Wong J, Lai PBS, Ng SSM, Ma BBY, Wong GLH, Wong V, Chan HLY, Chiu RWK, Lo YMD, Chan KCA. Liver- and colon-specific DNA methylation markers in plasma for investigation of colorectal cancers with or without liver metastases. *Clin Chem* 2018 Aug; 64(8):1239-1249.

42. An Y, Wang G, Diao Y, Long Y, Fu X, Weng M, Zhou L, **Sun K**, Cheung TH, Ip NY, Sun H, Wang H, Wu Z. A molecular switch regulating the cell fate choice between muscle progenitor cells and brown adipocytes. *Dev Cell* 2017 May 22; 41(4):382-391.e5.

43. Chen X, He L, Zhao Y, Li Y, Zhang S, **Sun K**, So K, Chen F, Zhou L, Lu L, Wang L, Zhu X, Bao X, Esteban MA, Nakagawa S, Prasanth KV, Wu Z, Sun H, Wang H. Malat1 regulates myogenic differentiation and muscle regeneration through modulating MyoD transcriptional activity. *Cell Discov* 2017 Mar 14; 3:17002.

44. Jiang P, Tong YK, **Sun K**, Cheng SH, Leung TY, Chan KCA, Chiu RWK, Lo YMD. Gestational age assessment by methylation and size profiling of maternal plasma DNA: a feasibility study. *Clin Chem* 2017, Feb; 63(2):606-608.

45. Wong FCK, **Sun K**, Jiang P, Cheng YKY, Chan KCA, Leung TY, Chiu RWK, Lo YMD. Cell-free DNA in maternal plasma and serum: A comparison of quantity, quality and tissue origin using genomic and epigenomic approaches. *Clinical Biochemistry* 2016 Dec; 49(18):1379-1386.

46. Jiang P, Peng X, Su X, **Sun K**, Yu SC, Chu WI, Leung TK, Sun H, Chiu RWK, Lo YMD, Chan KCA. FetalQuant-SD: Accurate quantification of fetal DNA fraction by shallow-depth sequencing of maternal plasma DNA. *npj Genomic Medicine* 2016 May 11; 1:16013.

47. Cheng SH, Jiang P, **Sun K,** Cheng KY, Chan KCA, Leung TK, Chiu RWK, Lo YMD. Noninvasive prenatal testing by nanopore sequencing of maternal plasma DNA: feasibility assessment. *Clin Chem* 2015 Oct; 61(10):1305-6.

48. Fu X, Xiao J, Wei Y, Li S, Liu Y, Yin J, **Sun K,** Sun H, Wang H, Zhang Z, Zhang BT, Sheng C, Wang H, Hu P. Combination of inflammation-related cytokines promotes long-term muscle stem cell expansion. *Cell Res* 2015 Jun; 25(6):655-73.

49. Wang L, Zhao Y, Bao X, Zhu X, Kwok YK, **Sun K,** Chen X, Huang Y, Jauch R, Esteban MA, Sun H, Wang H. LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. *Cell Res* 2015 Mar; 25(3):335-50.

50. Li L, Zhang Y, Fan Y, **Sun K,** Su X, Du Z, Tsao SW, Loh TK, Sun H, Chan AT, Zeng YX, Chan WY, Chan FK, Tao Q. Characterization of the nasopharyngeal carcinoma methylome identifies aberrant disruption of key signaling pathways and methylated tumor suppressor genes. *Epigenomics* 2015; 7(2):155-73.

51. Zhao Y, Yang Y, Trovik J, **Sun K,** Zhou L, Jiang P, Lau TS, Hoivik EA, Salvesen HB, Sun H, Wang H. A novel wnt regulatory axis in endometrioid endometrial cancer. *Cancer Res* 2014 Sep 15; 74(18):5103-17.

52. Jiang P, **Sun K,** Lun MF, Guo AM, Wang H, Chan KCA, Chiu RWK, Lo YMD, Sun H. Methy-Pipe: An integrated bioinformatics pipeline for whole genome bisulfite sequencing data analysis. *PLoS ONE* 2014 Jun 19; 9(6):e100360.

53. Chan KCA, Jiang P, Chan CWM, **Sun K**, Wong J, Hui EP, Chan SL, Chan WC, Hui DSC, Ng SSM, Chan HLY, Wong CSC, Ma BBY, Chan ATC, Lai PBS, Sun H, Chiu RWK, Lo YMD. Noninvasive detection of cancer-associated genomewide hypomethylation and copy number aberrations using massively parallel bisulfite sequencing of plasma DNA. *Proc Natl Acad Sci U S A* 2013 Nov 19; 110(47):18761-8.

54. Lun FMF, Chiu RWK, **Sun K**, Leung TY, Jiang P, Chan KC, Sun H, Lo YMD. Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. *Clin Chem* 2013 Nov; 59(11):1583-94.

55. Guo AM, **Sun K**, Su X, Wang H, Sun H. YY1TargetDB: an integral information resource for Yin Yang 1 target loci. *Database (Oxford)* 2013 Feb 14; 2013:bat007.

56. Lu L, Zhou L, Chen EZ, **Sun K**, Jiang P, Wang L, Su X, Sun H, Wang H. A Novel YY1-miR-1 regulatory circuit in skeletal myogenesis revealed by genome-wide prediction of YY1-miRNA network. *PLoS ONE* 2012; 7(2):e27596.

57. Diao Y, Guo X, Li Y, **Sun K**, Lu L, Jiang L, Fu X, Zhu H, Sun H, Wang H, Wu Z. Pax3/7BP is a Pax7- and Pax3-binding protein that regulates the proliferation of muscle precursor cells by an epigenetic mechanism. *Cell Stem Cell* 2012 Aug 3; 11(2):231-241.

Invited materials (reviews/commentaries/others):

1. Chen Q, Zeng Y, Kang J, Hu M, Li N, **Sun K**, Zhao Y. Enhancer RNAs in transcriptional regulation: recent insights. *Front. Cell Dev. Biol.* 2023 May 17; 11:1205540.

2. Ju J, **Sun K***. Plasma cell-free DNA analysis for COVID-19 and beyond. *Clin Transl Disc* 2022; 2:e122.

3. An Y, Fan F, Jiang X*, **Sun K***. Recent advances in liquid biopsy of brain cancers. *Front Genet* 2021 Sep 17; 12:720270. (IF: 4.772)

4. Li L, Guo X, **Sun K***. Recent advances in blood-based and AI-enhanced approaches for gastrointestinal cancer diagnosis. *World J Gastroenterol* 2021 Sep 14; 27(34):5666-5681. (IF: 5.374)

5. **Sun K***. Clonal hematopoiesis: background player in plasma cell-free DNA variants. *Ann Transl Med* 2019 Dec; 7(Suppl 8):S384. (IF: 3.616)

6. Gai W, **Sun K***. Epigenetic biomarkers in cell-free DNA and applications in liquid biopsy. *Genes (Basel)* 2019 Jan 9; 10(1):32. (IF: 4.141) (**Editor's Choice Article**)

7. **Sun K**, Zhou L, Zhao Y, Wang H, Sun H. Genome-wide RNA-seq and ChIP-seq reveal Linc-YY1 function in regulating YY1/PRC2 activity during skeletal myogenesis. *Genom Data* 2016 Feb 2; 7:247-9.

8. **Sun K,** Jiang P, Chan KCA. The impact of digital DNA counting technologies on noninvasive prenatal testing. *Expert Rev Mol Diagn* 2015; 15(10):1261-1268. (IF: 5.670)

9. **Sun K,** Lu L, Wang H, Sun H. Genome-wide profiling of YY1 binding in skeletal muscle cells. *Genom Data* 2014 May 23; 2:89-91.

Book chapters:

1. Liu X, Yuan P, **Sun K***. DNA methylation data analysis using Msuite. *Methods Mol Biol* 2023; 2624:1-6.

2. Sablok G[#], **Sun K[#]**, Sun H. NAMS: Noncoding Assessment of long RNAs in Magnoliophyta Species. *Methods Mol Biol* 2019; 1933:257-264.

3. Peng X, **Sun K**, Zhou J, Sun H, Wang H. Bioinformatics for novel long intergenic noncoding RNA (lincRNA) identification in skeletal muscle cells. *Methods Mol Biol* 2017 Mar 1; 1556:355-362.

Conference papers:

1. **Sun K**, Yuen YP, Wang H, Sun H. The Online Diagnosis System for Sanger sequencing based genetic testing. In: *Big Data and Smart Computing (BIGCOMP), 2014 IEEE International Conference on*, 2014: 47-51.

2. Lu L, Zhou L, Chen EZ, **Sun K**, Jiang P, Su S, Wang L, Sun H, Wang H. Genome-wide identification of TF-miRNA regulatory networks in myogenesis. In: *Bioinformatics and Biomedicine Workshops (BIBMW), 2011 IEEE International Conference on*, 201: 941.

3. Jiang P, Sun X, Chen EZ, **Sun K**, Chiu RWK, Lo YMD, Sun H. Methy-Pipe: an integrated bioinformatics data analysis pipeline for whole genome methylome analysis. In: *Bioinformatics and Biomedicine Workshops (BIBMW), 2010 IEEE International Conference on*, 2010: 585-590.

PUBLIC PATENTS

1. Chiu RWK, Chan KC, Lo YMD, Jiang P, and **Sun K**. METHYLATION PATTERN ANALYSIS OF TISSUES IN A DNA MIXTURE. U.S. Patent 11062789, filed Jul 20, 2015, and **granted** Jul 13, 2021.

2. Lo YMD, Chan KCA, Chiu RWK, Jiang P, and **Sun K**. METHYLATION PATTERN ANALYSIS OF HAPLOTYPES IN TISSUES IN A DNA MIXTURE. U.S. Patent 10689706, filed Jul 20, 2016, and **granted** Jun 23, 2020.

3. Lo YMD, Chiu RWK, Chan KCA, and **Sun K**. DETECTING HEMATOLOGICAL DISORDERS USING CELL-FREE DNA IN BLOOD. U.S. Patent 10781490, filed May 30, 2017, and **granted** Sep 22, 2020.

4. Namsaraev E, Lo YMD, Chiu RWK, Chan KCA, Jiang P, **Sun K**, and Lam WKJ. METHODS AND SYSTEMS FOR TUMOR DETECTION. U.S. Patent 11459616, filed Oct 25, 2017, and **granted** Oct 4, 2022.

5. Lo YMD, Chiu RWK, Chan KCA, Jiang P, **Sun K**. SIZE-TAGGED PREFERRED ENDS AND ORIENTATION-AWARE ANALYSIS FOR MEASURING PROPERTIES OF CELL-FREE MIXTURES. U.S. Patent application 20190341127, filed May 3, 2019, and issued Nov 7, 2019.

6. Lo YMD, Chiu RWK, Chan KCA, Jiang P, Chan WY, **Sun K**. CELL-FREE DNA END CHARACTERISTICS. U.S. Patent application 20200199656, filed Dec 19, 2019, and issued Jun 25, 2020.

7. <u>孙坤</u>。一种基于游离 DNA 的癌症诊断模型和应用。中国发明专利 202210496595.9，优先权日期 2022 年 5 月 9 日；公开日期 2022 年 9 月 6 日；PCT 申请号 PCT/CN2022/116248。

8. <u>孙坤</u>、胡丁雪、刘晓懿。一种肿瘤恶性程度评估方法、电子设备和存储介质。中国发明专利 202310370089.X，优先权日期 2023 年 3 月 31 日；公开日期 2023 年 6 月 23 日。

9. <u>孙坤</u>、胡丁雪、刘晓懿、杨梦琦。肝细胞癌预后评估以及分子分型的方法、标志物及其应用。中国发明专利 202310362560.0，优先权日期 2023 年 3 月 31 日；公开日期 2023 年 7

月 11 日。

10. <u>孙坤</u>、鞠佳、王琦。一种无创产前检测胎儿染色体拷贝数异常的优化方法。中国发明专利 202311138167.X，优先权日期 2023 年 9 月 1 日；公开日期 2023 年 12 月 15 日。

11. <u>孙坤</u>。基于谷氨酰胺代谢模式的癌症风险评估方法及其应用。中国发明专利 202310642438.9，优先权日期 2023 年 6 月 1 日；公开日期 2023 年 12 月 26 日。

12. <u>孙坤</u>、鞠佳。一种泛癌种血液诊断标志物及其应用。中国发明专利 202211361261.7，优先权日期 2022 年 11 月 2 日；公开日期 2024 年 5 月 14 日，PCT 申请号 PCT/CN2023/093013。

13. <u>孙坤</u>、郭腾飞。游离 DNA 标志物的筛选方法、DNA 标志物及其应用。中国发明专利 202311712915.0，优先权日期 2022 年 12 月 13 日；公开日期 2024 年 6 月 14 日。

COMPETITIVE GRANTS

1. <u>国家重点研发计划</u>，合成生物学专项（青年科学家项目），2022YFA0912700，**多模态、无创疾病诊断系统的研发与应用**，2022-11 至 2027-10，人民币 400 万元，在研，<u>首席科学家</u>。
(National Key Research and Development Program, serving as PI)

2. 国家自然科学基金委员会，<u>优秀青年基金（海外）</u>，**癌症液体活检**，2022-01 至 2024-12，人民币 300 万元，在研，<u>主持</u>。
(National Talents Program, serving as PI)

3. 国家自然科学基金委员会，青年基金，82101763，**胎盘中的融合基因以及与妊娠期疾病的关联**，2022-01 至 2024-12，人民币 30 万元，在研，<u>主持</u>。
(National Natural Science Foundation, as PI)

4. 广东省基础与应用基础研究基金委员会，省市联合基金，<u>重点项目</u>，2023B1515120073，**基于液体活检的乳腺癌精准诊疗技术开发及应用**，2023-11 至 2026-10，人民币 100 万元，<u>主持</u>。
(Provincial Key Research Program, as PI)

5. 广东省基础与应用基础研究基金委员会，青年基金，2019A1515110173，**一种基于机器学习技术的快速、方便的肿瘤组织纯度评估方法**，2020-01 至 2021-12，人民币 10 万元，已结题，<u>主持</u>。
(Provincial Natural Science Foundation, as PI)

6. 深圳湾实验室，重大项目（培育），S241101004，**阿尔茨海默病无创早期诊断技术开发及应用**，2024-05 至 2026-05，人民币 1000 万元，在研，<u>主持</u>。

7. 深圳湾实验室，概念验证基金，S231802002，**液体活检技术在癌症实时监控与指导精准治疗中的应用**，2024-01 至 2024-12，人民币 100 万元，在研，<u>主持</u>。

8. 深圳华大生命科学研究院，开放基金，BGIRSZ2020007，**血浆游离 DNA 的组织溯源以及在癌症诊断中的应用**，2020-11 至 2022-12，人民币 50 万元，已结题，<u>主持</u>。

9. 深圳市医学科医学院，医学研究专项资金，**染色质上互作蛋白的靶基因鉴定新技术**，2024-01 至 2025-12，人民币 100 万元，在研，参与。

SOFTWARE COPYRIGHTS

1. **高速无偏 cfDNA 片段组学分析软件**，登记号 2024SR0586811，2024 年 4 月 29 日
2. **DNA 甲基化测序大数据分析软件**，登记号 2021SR0279962，2021 年 2 月 23 日
3. **高通量测序大数据预处理软件**，登记号 2021SR0279963，2021 年 2 月 23 日
4. **植物转录本编码潜力评估及功能注释软件**，登记号 2022SR1530495，2022 年 11 月 17 日
5. **三维基因组数据分析软件**，登记号 2022SR1530496，2022 年 11 月 17 日

INVITED TALKS

1. 2023/02/23   P4China 第六届国际肿瘤精准医疗大会
2. 2021/07/28   2021 GCOG Annual Meeting/2021 全球抗癌协作组年会
3. 2021/04/23   深圳微生物与生命健康产业沙龙
4. 2020/11/06   广东省恶性肿瘤表观遗传与基因调控重点实验室
5. 2019/12/31   香港中文大学（深圳）第一附属医院
6. 2019/12/04   中国科学院深圳先进技术研究院

PRESENTATIONS AND POSTER SESSIONS

1. Accelerating 3D genomics data analysis with Microcket. *Poster presented at: The 10th international conference on 3D genomics, Jul 15-18, 2023; Hangzhou, China.*

2. Fragmentation patterns of cell-free DNA and applications in cancer liquid biopsy. *Presented at: The 8th National Conference on Computational Biology and Bioinformatics, Jul 22-25, 2022; Guangzhou, China.*

3. Msuite2: all-in-one DNA methylation data analysis toolkit with enhanced usability and performance. *Poster presented at: The 10th National Annual Conference on Bioinformatics and System Biology, Oct 26-28, 2021; Chengdu, China.*

4. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. *Presented at: 11th International Symposium on Circulating Nucleic Acids in Plasma and Serum (CNAPS), Sep 23-25, 2019; Jerusalem, Israel.*

5. Noninvasive reconstruction of fetal methylome by sequencing of maternal plasma DNA. *Poster presented at: The American Society of Human Genetics Annual Meeting, Oct 17-21, 2017; Orlando, Florida, U. S. A.* (**Reviewers' Choice Abstract**)

6. Sebnif: an integrated bioinformatics pipeline for the identification of novel large intergenic noncoding RNAs (lincRNAs). *Poster presented at: Long Noncoding RNAs: Marching toward Mechanism, Feb 27-Mar 4, 2014; Santa Fe, New Mexico, U. S. A.*

7. iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. *Paper presented at: ISCB-Asia/SCCG 2012, Dec 17-19, 2012; Shenzhen, China.*

ACADEMIC SERVICES

1. Reviewer, Ministry of Science and Technology (talents program), 2024
2. Reviewer, *Guangdong Basic and Applied Basic Research Foundation,* 2020, 2021, 2022
3. Associate editor, *BMC Cancer*, since 2022
4. Guest editor, *Frontiers in Genetics,* 2021
5. *Ad hoc* reviewer, *Nature Biotechnology, Nature Communications, eLife, Cell Reports Medicine, Clinical Chemistry, Clinical and Translational Medicine, etc.*

OTHER SERVICES

1. 2022/11/18 – 2022/11/20 深圳湾实验室肿瘤研究所，第一届 ICR Symposium，大会主席
2. 2022/03/08 – 2022/03/31 深圳市疾病预防控制中心，抗击新冠疫情

HORNORS AND AWARDS

深圳湾实验室宣传之星 2020

Postdoc Research Fellowship (Rank **1st** in 50+ applicants) 2016-2017
The Chinese University of Hong Kong

Postgraduate studentship 2010-2014
The Chinese University of Hong Kong

Outstanding student scholarship 2007-2009
University of Science and Technology of China

STUDENTS/POST-DOCS (CO-)MENTORED

Dr. Lishi Li, Dr. Yunyun An, Dr. Jia Ju, Dr. Dingxue Hu, Dr. Wanqiu Wang, Dr. Yunxia Bai
Miss. Chuhan Li, Mr. Chuan Gao, Miss. Yuqi Pan, Miss. Xiaoxiao Wei, Mr. Zheqi Yan

COMPUTATIONAL SKILLS

➢ **Highly proficient in GNU/Linux operating system**
➢ **Highly skilled in C/C++, Java, Perl, and R programming languages**
➢ Highly skilled in HTML, PHP, and SQL programming languages
➢ Familiar with Photoshop, Illustrator and GUI development

PUBLIC SOFTWARE AND RESOURCES

Long noncoding RNA analysis

1. **iSeeRNA**: accurate and extra-fast lincRNA/mRNA classifier **(citation: 192)**
2. **Sebnif**: comprehensive novel lincRNA filter pipeline for RNA-seq data **(citation: 28)**
3. **NAMS webserver**: coding potential classifier and functional annotation for plants

DNA methylation data analysis

4. **Methy-Pipe**: integrated, multi-functional data analysis pipeline **(citation: 80)**
5. **BSviewer**: lightweight, genotype-aware, nucleotide-level visualizer
6. **Msuite/Msuite2**: high-performance and versatile toolkit for various protocols

Noninvasive prenatal testing (NIPT)

7. **COFFEE**: control-free and cross-platform algorithm for T21 test **(citation: 23)**
8. **FEMER**: high resolution and accuracy fetal methylome reconstruction **(citation: 22)**

Others

9. **Microcket:** rapid and versatile toolkit for 3D genome data analysis
10. **Ktrim:** extra-fast and accurate sequencing data preprocessor **(citation: 31)**
11. **GeneCT**: generalizable pan-cancer cancerous status/tissue origin classifier **(citation: 19)**
12. **ODS**: integrated data analysis platform for Sanger-based genetic testing
13. **mTFkb**: functional annotation of mouse transcription factors **(citation: 14)**