SHEN, Lishuang Ph.D.（沈利爽 博士)

目前公司：	香港某纳斯达克 NASDAQ 十亿级精准医学公司
目前职位：	VP of Bioinformatics 生物信息学

手　机	(+1) 617-615-1050 (USA), (+852) 5790-2945 (Hong Kong)	邮箱网站	Shen_lishuang@yahoo.com , shen@OmicsExpert.com https://OmicsExpert.com , https://mseqdr.org

优势亮点

- 高产生物信息学与遗传学科学家，在自然杂志及子刊等作为主要作者发表医学多组学与生物信息学论文 80 余篇，被引用 8000 次，4 自然系列文章引用超 4000 次。论文涉及二型糖尿病，阿尔兹默病，乳腺癌动物模型，线粒体病等遗传病机理及群体遗传学，遗传病与肿瘤分子遗传。
- 领导美股精准医学公司生信团队十万级全外显组产品软硬件支持.新基因检测产品调研设计上市
- 20 年丰富经验多组学与生物信息学家技术领导。Illumina 高通量测序技术最早用户。精通各种测序及微芯片数据分析技术原理和应用场景。全基因组突变扫描全转录组单细胞基因表达分析
- 8 年经验遗传病与肿瘤分子遗传诊断全外显子组，基因 Panel 及微芯片的 IVD 产品与液体活检技术研发,成功 CAP 认证申请及其生信平台研发。深入参与二代测序技术遗传诊断病人的数据分析与解释。参与洛杉矶儿童医院新个性化精准医学中心创立及运行。与临床医生密切互动共同诊断。建立维护及运行病人与临床样品与表型的注册体系，系统达到美国病人隐私保护的 HIPAA 标准
- 设立高性能生物信息学分析体系平台整合本地高性能低运营费用的 GPU 计算平台与亚马逊 AWS 云平台，支持十万级全外显子组从测序到解读报告各分析。三年计算成本较云平台降 60%
- 精通生物信息学分析及数据库使用与建设，多组学及医学大数据体系整合，统计与机器学习用于医学遗传学。大规模文献发掘与本地基因组数据表型数据综合，确立从突变到表型和疾病相关性
- ClinGen 官方线粒体病突变致病性解读专家委员会成员,参与线粒体 DNA 突变致病性解读指南制定
- 连续 10 年任线粒体罕见病的国际合作平台 MSeqDR Consortium 首席生物信息科学家,独力创立和维护 MSeqDR.org 网站全栈建设,数据搜集管理,整合疾病,表型,突变及序列数据分析体系和平台。
- 连续五年受邀为 UMDF 主办线粒体医学国际年会培训本人主建 mseqdr.org 数据库及网站工具。
- 担任《Medicine》杂志编辑。暨受邀为各个基因组学医学及生物信息学学术期刊审稿五十余次。
- 多种语言编程：R, Linux bash, Perl, PHP, Python, web/JavaScript/CSS/Ajax/NodeJS, SQL,AI 机器学习 ML。领导开发和产品化 AI 祖源分析与品系分析的算法与产品化包装。基于深度学习的线粒体遗传病突变的致病性预测。
- 新冠病毒(洛杉矶儿童医院)与结核菌等传染性病原微生物基因组进化与传染病学研究及实时分析，15 篇论文含 5 篇一作，快速建立及维护相关微生物基因组数据库及网站平台.

EXECUTIVE SUMMARY

- A productive bioinformatics and human genomics scientist with 80+ papers and 8,000+ citations, H-index 31, 4 high-impact Nature series papers on late-onset Alzheimer's disease and diabetes
- A visionary bioinformatics leader with 20 years of experience optimizing genomic analysis infrastructures and pipelines, managing high-performance team, and delivering cutting edge genetic medicine results in large-scale genomics and multi-omics projects.
- 6 years' experience developing infrastructures for a new Personalized Medicine Center with local HPC, and AWS cloud solutions to integrate phenome & genomic data for genetic disease tumor diagnosis.
- Lead Bioinformatician for MSeqDR.org since 2013 - the international Mitochondrial Disease Sequence Data Resource Consortium. Full-stack development of the public MSeqDR web resources .
- Member of 2 ClinGen Mitochondrial Disease Curation Expert Panels: (1) Genes, (2) Variants.
- Experienced in integrating Clinical and Multi-omics data for genetic disease and tumor diagnosis, preparing CAP/CLIA accreditation applications.

- Pivotal role in CHLA hospital's COVID-19 / SARS-CoV-2 task team: genomic surveillance, epidemiology and evolution. Experience in Tuberculosis and *P. aeruginosa* infectious microbial genomics. 15 papers.
- Programming since 2002 in R, Linux bash, Perl, PHP, Python, web/JavaScript/css/Ajax/NodeJS, SQL, AI/ML.

学位信息 教育经历

1.中国科学院遗传与发育学研究所

专业名称　分子遗传学基因组学　　学历　博士, 博士后

博士论文：水稻分子遗传图谱的构建和基因组分析
(http://wenjin.nlc.cn/allSearch/searchDetail?searchType=&showType=1&indexName=data_408&fid=012002605424)

2.山东大学

专业名称　遗传学 , 生物化学　　学历　硕士，本科

3.美国 德克萨斯理工大学 Texas Tech University

专业名称　交叉学科生物信息学 Bioinformatics / Computer Science /Statistics　　学历　硕士

其它学术总结 Other Scholar Profiles:

- **Google Scholar Profile (8000+ citations):** https://scholar.google.com/citations?user=9HFQBgEAAAAJ
- **NCBI bibliography (88 items):** https://www.ncbi.nlm.nih.gov/myncbi/1FMLoDt0vMnAp/bibliography/public/
- **ResearchGate:** https://www.researchgate.net/profile/Lishuang_Shen/stats
- **Lead Bioinformatician for MSeqDR Consortium since 2013** - Mitochondrial Disease Sequence Data Resource Consortium (https://mseqdr.org)

职业经历与主要业绩

2021.10 至今： VP 生物信息学，VP of Bioinformatics

公司描述：香港, 某纳斯达克 NASDAQ 某十亿元级精准医学公司

职责业绩：

- 领导跨国生物信息学队伍建设, 招聘, 培养, 管理。负责十万级全外显组基因检测产品生信软硬件设计及运运营支持。
- 人类及动物新遗传病与肿瘤基因检测产品（外显子组，panel，测序加芯片）调研开发及运营支持。基于不同技术平台与不同类型组学数据的液体活检产品研发。
- 成功完成年度主打宠物健康与祖系检测 IVD 产品概念化内容设计到实验室验证与产品上市(https://circlepaw.co/)。与计算机工程团队合作自动化基于 AI 机器学习研发的祖源分析算法模型。
- Illumina, Nanopore, PacBio 二代三代测序技术数据自动化分析体系平台建设. 与实验室及计算机工程团队完成二代测序本地化的实验室与数据分析体系建设。
- 本地 GPU-HPC+200TB NAS + Cloud 云硬件及软件供应商遴选和谈判。搭建硬件与软件平台支持人类及动物大规模基因检测数据快速分析，相比纯云平台降低 70%费用。
- 肿瘤易感基因及外显子组产品成功 CAP 认证申请及其生信及数据支持平台设计研发。
- 指导建成公司超十万客户全外显子组与临床与健康数据整合的高速检索数据库，数据量超 70 亿行。
- ·Led the bioinformatics team at a global molecular diagnostic company to analyze and interpret clinical genomics data from over 100K customer exomes in a business environment.
- ·Led the bioinformatics team to create in-house automated analysis and reporting pipelines for short- and long-read sequencing (Illumina, PacBio, Nanopore) data, and collaborated with IT/Engineering/ NGS Lab teams to transform our team's prototypes into automated production pipelines.
- Migration of NGS sequencing from an external provider to an in-house NGS + Bioinformatics platform.

- ·Piloted and optimized the cost-effective NVIDIA GPU-based HPC / storage + Clara Parabricks bioinformatics infrastructures for automated rapid processing of large-scale NGS testing data.
- Provide in-depth technical guidance on bioinformatics and AI / machine learning in supporting molecular diagnostic products. R&D for NGS/ microarray exome/Panels genetic disease & tumor liquid biopsy assays.
- Played a key role in the full cycle of conceptualization, design, and marketing of a non-invasive DNA test for dogs (https://circlepaw.co/), designed to provide 200+ key DNA insights into dogs' genetic health risks and AI /machine learning-based breed predictions.
- Upgraded and developed the AI /machine learning and population genetics-based global ancestry prediction and component analysis workflow.
- Served as a member in 2 ClinGen's Mitochondrial Disease Gene & Variant curation expert panels.
- Team in successful CAP certification application and SOP for clinical sequencing and genetic diagnosis.
- Led the team to establish a comprehensive variant database of 7 billion rows that cross-linked all 150K customers' whole exome variants with health/clinical data at individual variant to person levels, optimized database server for rapid joint query of variants with 2 million rows ACMG/pathogenicity annotations.

2015.08-2021.10：洛杉矶儿童医院 南加州大学医学院 个性化医疗精准医学中心 生物信息高级科学家

公司描述：美国-加利福尼亚州-洛杉矶 **医院 Center for Personalized Medicine, Department of Pathology & Laboratory Medicine, Children's Hospital Los Angeles， USC Keck Medical School**

所在部门：个性化医疗精准医学中心 **Center for Personalized Medicine, Children's Hospital Los Angeles**

职责业绩：

- 参与本精准医学中心生物信息平台从零开始的创建工作,搭建了一个本地高性能计算机与亚马逊云平台结合前瞻性地满足本中心高通量测序和遗传诊断工作需求的硬件与软件平台.
- 遗传病与肿瘤分子遗传诊断产品与液体活检技术研发。从 NGS 数据分析存储阐释到临床医学报告各个环节的各种标准确立, CAP/CLIA SOP 及培训文档的撰写。对实验室科研人员及医生进行病人及临床样品注册系统使用培训.
- 建立个性化医疗精准医学的生物信息系统，开发分子遗传诊断服务的工具与数据库。建立基于标准化字典的保密性病人及临床样品注册系统，临床表型驱动的测序数据分析与阐释本地系统与在线平台.
- 首席生物信息科学家六年独力创立和维护罕见线粒体病的国际合作平台 MSeqDR.org Consortium,整合疾病,表型,突变的序列数据在线分析体系。
- 临床全基因组，全外显子组，线粒体组大规模测序数据的分析与阐释，开发其在线系统和本地系统.
- 阿尔兹默病人单细胞基因表达和突变模式整合分析眼瘤液体活检技术研发。
- 《线粒体病基因合集》一书的在线编辑系统建立及运行管理. 并列作者.
- 参与建立线粒体染色体突变的致病性 ACMG 解读指南.
- ClinGen 线粒体病基因与突变解释专家组成员.
- 基于 RNA-seq 基因融和与基因表达的儿童肿瘤检测系统研发.
- 线粒体病大规模文献挖掘基因，突变，表型与疾病的关联性.
- 建立整合机器学习与临床表型指导的全基因组外显子组数据解读在线系统。
- 新冠病毒的二代测序与进化基因组学，流行病学的数据分析系统.

- R&D of NGS-based genetic disease and tumor diagnosis assays through Exome/Panel/Liquid biopsy assays using genomic and transcriptomic data.
- Piloted RNA-seq based pediatric cancer diagnosis assay and retinoblastoma liquid biopsy assay (CNV+SNV).
- Whole genome sequencing and single-cell RNA-seq on a Colombian late-onset Alzheimer disease cohort and co-authored 2 Nature Medicine papers.
- Established standards and operation procedures SOP and training documents for clinical sequencing genetic diagnosis; Participated in CAP certification application.
- Clinical sequencing WES, WGS, mitochondrial genome data interpretation in the clinical setting. Setup phenotype-driven clinical sequence data analysis pipelines for CES data interpretation.
- Built a new local HPC+ AWS Cloud hardware and bioinformatics platform. Built center's patient & sample

registries with patient symptom ontology encoding and HIPAA privacy protection.
- MSeqDR.org Consortium Lead Bioinformatician, built mitochondrial disease resources and website of diseases, pathogenic variants, genomic & clinical data with literature mining. Quick-Mitome online WES/WGS interpretation platform. mvTool universal mtDNA annotation tool.
- ACMG Guideline development for mitochondrial DNA variant pathogenicity.
- Served in 2 ClinGen's Mitochondrial Disease gene & variant curation expert panels.
- <Mitochondrial Disease Gene Compendium> book editing system full-stack development/admin.
- Key role in hospital's task team for COVID-19 SARS-CoV-2 virus NGS & epidemiology analysis for genomic surveillance. Promptly build an open COVID-19 virus genomic and phylogenetic data web platform (CARD) in 2 months at the pandemic onset. The discoveries and platforms led to 12 papers

July 2013 – 2023：国际线粒体病数据资源联盟首席生物信息学家 MSeqDR Consortium，https://MSeqDR.org
Lead Bioinformatician for MSeqDR Consortium, the international Mitochondrial Disease Sequence Data Resource Consortium (https://MSeqDR.org)

- Led the full-stack development of MSeqDR's database-driven web portal (https://MSeqDR.org), enabling access to community resources for over 1000 clinicians and researchers globally.
- Served as bioinformatician on 2 successful NIH U24 grant applications totaling ~$2 million in funding for MSeqDR Consortium's projects. Conducted the projects' execution to completion.
- Designed and built web resources that integrate rare mitochondrial diseases and phenotypes with pathogenic variants, genomic/clinical data, built case Virtual Registry through literature mining, Quick-Mitome platform with deep learning variant interpretation, and mvTool universal mtDNA annotation. Invited demo of MSeqDR at 5 annual UMDF Mitochondrial Medicine Symposiums.
- Provided DevOps and administration for the "Mitochondrial Disease Gene Compendium" book authoring/editing system and co-authored book chapters.
- Co-developed ACMG guidelines for scoring mitochondrial DNA variant pathogenicity classification.
- Expert in 2 ClinGen Mitochondrial Disease Curation Expert Panels:(1) Genes, (2) Variants.

2014.01-2015.07：哈佛大学医学院附属麻省眼耳医院，生物信息学家 Bioinformaticist
公司描述：美国-波士顿 医院 Harvard University，Medical School，MASS Eye and Ear Infirmary
职责业绩：

- 线粒体病的国际合作平台 MSeqDR.org Consortium 首席生物信息科学家独力创立和维护罕见病平台，网站建设与数据搜集数据库管理，整合疾病，表型，突变及驱动的序列数据分析体系。
- 眼科基因组研究所生物信息学支持和基于二代测序眼科疾病分子遗传诊断。
- 二代测序全基因组，靶向基因组，RNA-seq，眼科疾病微芯片全基因组突变扫描数据合作分析与阐释
- Lead Bioinformatician for MSeqdr.org Consortium, building integrated rare mitochondrial disease resources of diseases, pathogenic variant database and website, pipelines for variant annotations.
- Bioinformatics support to Ocular Genomics Institute and its NGS eye disease diagnostic services. Data analysis for next generation genomic re-sequencing, whole and targeted exome sequencing, RNA-seq. Genome-wise association study with SNP arrays for eye disease projects.

2013.07-2013.12：**Stritch School of Medicine, Loyola University**，生物信息学家 Bioinformaticist
公司描述：美国-芝加哥 医院 **Bioinformatician, Stritch School of Medicine, Loyola University**
职责业绩：

- 创立线粒体病的国际合作平台 MSeqDR.org Consortium 首席生物信息科学家创立和维护罕见病平台，网站建设与数据搜集数据库管理，整合疾病，表型，突变及驱动的序列数据分析体系。
  Lead Bioinformatician for MSeqdr.org Consortium, building integrated rare mitochondrial disease resources of diseases, pathogenic variant database and website, pipelines for variant annotations.

2010.01-2013 ：康奈尔大学医疗科学系;干细胞再编程中心；康奈尔大学计算服务中心 CBSU，生物信息学家 Bioinformatics Research Associate

公司描述：美国-纽约州 **Cornell University, Dept Biomedical Sciences, Stem Cell Reprogramming Center；CBSU**

职责业绩：

- 生物信息科学家 为康奈尔大学哺乳动物干细胞再编程中心提供生信学服务;
- 生物信息科学家，为康奈尔大学干细胞协会会员提供生信合作咨询与分析;
- 乳腺癌的小鼠染色体构象动物模型的基因组学与表观遗传学研究数据分析;
- 二代测序数据分析：全基因组，定向基因组，RNA-seq，微芯片全基因组突变扫描，CHIP-seq，RRBS 基因甲基化组数据，基因组变异检测与分析，染色体数目变异 CNV 分析,基因组拼装.项目覆盖乳腺癌模型及基础生物学研究。
- Bioinformatics collaboration support to Cornell Mammalian Reprogramming Core.
- Provided bioinformatics consultation and technical support to NYS Stem Cell Consortium.
- Epigenomic and genomic studies in mouse breast cancer model on chromatin modeling.
- Data analysis for next-generation sequencing data (RNA-seq, CHIP-seq, RRBS Methylation, targeted resequencing, SNP and variation, genome assembly), microarray, CNV and genomics data in various breast cancer and fundamental biology projects.

2008-2009：**加拿大多伦多病童医院 Hospital for Sick Children, Toronto,Canada 生物信息学家 Bioinformatics Analyst**

所在部门： Cancer Stem Cell Project

职责业绩：

- 为白血病项目提供微芯片，染色体数目变异 CNV，与基因组数据分析服务;
- 建立基因组和蛋白质组数据库。为癌病干细胞项目提供生物信息学服务;
- 与安大略省癌症研究所合作，建立联邦制的多研究所癌病研究数据库方案(Biomart)数据库
- Established genomics and proteomics database, data analysis for Cancer Stem Cell Project.
- Built federated, cross-institute cancer research data warehousing (Biomart) solution in collaboration with Ontario Institute for Cancer Research.
- Data analysis for microarray, CNV and genomics data in leukemia projects.

2007-2008：加拿大 多伦多大学 CAGEF, University of Toronto 生物信息学家 Bioinformaticist

所在部门： **CAGEF 基因组**中心 **University of Toronto**

职责业绩：

- 作为 Illumina 新的二代测序技术首批用户，开创性进行相关生物信息体系开发，并且自己开发数据分析,存储，质量控制，拼装，突变筛选的算法和工具。全转录组分析，微生物比较基因组.
  As first group Illumina Next Generation Sequencer user, piloted bioinformatics pipeline development, data analysis, warehousing for NGS data. Sequence quality control, alignment, assembling and SNP/INDEL discovery. Transcriptomic profiling. Microbial comparative genomics.

2005-2007：加拿大蒙特利尔 麦吉尔大学医学系，Department of Medicine, McGill University 生物信息学家 Bioinformatician Research Associate

所在部门：**麦吉尔大学医学系，以及** Genome Quebec

职责业绩：

- 采用最新 Illumina 高密度微芯片进行二型糖尿病全基因组扫描研究数据分析。作为第二和第三并列作者发表两篇自然杂志系列文章,引用超过 3700 次;
- 全基因组扫描数据与基因表达和序列数据整合分析分析;
- 设计个性化的基因主导型的全基因组标注 SNP 单核酸 Illumina 芯片 (2 万 SNP 标记)，与商业化的高密度芯片设计互补用于全基因组扫描。
- Bioinformatics and population genetics analysis for a genome-wide association GWAS project of type 2 diabetes with Illumina high-density SNP arrays. Published in Nature.
- Genome-wide SNP data meta-analysis integrating function, expression, and sequence data.

- Gene-centric whole genome SNP tagging, designed 20K custom Illumina SNP chip for whole genome genotyping coverage, complementing commercial SNP array designs.

2003-2005：美国衣阿华州立大学 虚拟现实技术应用中心 Iowa State University

生物信息学博士后 Bioinformatician Research Associate

所在部门：**虚拟现实技术应用中心 (Virtual Reality Application Canter))**

职责业绩：

- 领导一支研究生与半职员工团队，设计，开发，管理并注解公共网络植物微芯片转录组数据库：BarleyBase/ PLEXdb.org
- 设计开发基于网络的微芯片转录组数据分析及图形化工具
- 与生物学家们合作进行微芯片转录组数据及代谢途径分析
- Led a team of students and part-time staff to design, develop, administrate, and curate BarleyBase/PLEXdb, a public web-based plant microarray database.
- Developed web-based tools for microarray data numerical analysis and visualization.
- Collaborated with biologists on the integration of transcriptomic profiling data and pathway data.

发布论文 PUBLICATIONS（English 英文部分 72 篇， 中文部分 19 篇）

73. **Shen, L**., Bard, J. D., Triche, T. J., Judkins, A. R., Biegel, J. A., & Gai, X. (2021). Rapidly emerging SARS-CoV-2 B.1.1.7 sub-lineage in the United States of America with spike protein D178H and membrane protein V70L mutations. Emerging microbes & infections. 10:1, 1293-1299.

72. Arboleda-Velasquez JF, Lopera F, et al, **SHEN L** (co-author), et al. (2019). Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. ***Nature Medicine***. 25(11):1680-1683

71. Lopera, F, Marino, C, Chandrahas, AS, et al, **SHEN L** (co-author) et al. (2023). Resilience to autosomal dominant Alzheimer's disease in a Reelin-COLBOS heterozygous man. ***Nature Medicine,*** *29*(5):1243–1252.

70. **Shen L**, McCormick EM, Muraresku CC, Falk MJ, Gai X. (2020). Clinical Bioinformatics in Precise Diagnosis of Mitochondrial Disease. Clinics in Laboratory Medicine. 40(2):149-161. PMID: 32439066.

69. **Shen, L**., Bard, J. D., Triche, T. J., Judkins, A. R., Biegel, J. A., & Gai, X. (2021). Emerging variants of concern in SARS-CoV-2 membrane protein: a highly conserved target with potential pathological and therapeutic implications. Emerging microbes & infections, 10:1 885–893.

68. **Shen L**, Dien Bard J, Biegel JA, Judkins AR, Gai X. (2020). Comprehensive genome analysis of 6,000 USA SARS-CoV-2 isolates reveals haplotype signatures and localized transmission patterns by state and by country. Frontiers in Microbiology, 2020;11:573430 https://doi.org/10.3389/fmicb.2020.573430

67. Kim ME., Xu LY, Prabakar RK, **Shen L**, Peng CC, Peter Kuhn, Biegel JA, Gai X, Hicks J, Berry JL. (2021). Aqueous Humor as a Liquid Biopsy for Retinoblastoma: Clear Corneal Paracentesis and Genomic Analysis. Journal of visualized Experiments: JoVE, (175), 10.3791/62939. https://doi.org/10.3791/62939.

66. McCormick EM, Kierstin Keller K, Taylor JP, Coffey AJ, **Shen L**, et al, NICHD ClinGen U24 Mitochondrial Disease Gene Curation Expert Panel (2023). Expert panel curation of 113 primary mitochondrial disease genes for the Leigh syndrome spectrum. **Annals of Neurology**. PubMed:37255483.

65. Xu L, **Shen L**, Polski A, Prabakar RK, Shah R, Jubran R, Kim JW, Biegel J, Kuhn P, Cobrinik D, Hicks J, Gai X, Berry JL. (2020). Simultaneous identification of clinically relevant RB1 mutations and copy number alterations in aqueous humor of retinoblastoma eyes. Ophthalmic Genetics 41(6):526-532 DOI: 10.1080/13816810.2020.1799417 PMID: 32799607.

64. Truong T, Ryutov A, Pandey U, et al, **Shen L** (co-author). (2021). Increased viral variants in children and young adults with impaired humoral immunity and persistent SARS-CoV-2 infection: a consecutive case series. EBioMedicine https://doi.org/10.1016/j.ebiom.2021.103355 .

63. Moustafa AM, Otto W, et al, **Shen L** (co-author) (2021). Early pandemic molecular diversity of SARS-CoV-2 in children. medRxiv 2021 DOI: https://doi.org/10.1101/2021.02.17.21251960

62. Wong EY, Liya Xu, Lishuang Shen, Mary E. Kim et al (2021). Inter-eye genomic heterogeneity in bilateral retinoblastoma via 1 aqueous humor liquid biopsy. NPJ precision oncology, 5(1), 73. https://doi.org/10.1038/s41698-021-00212-0.

61. Bard, J. D, Bootwalla M, et al, **Shen L** (co-author). (2021). Emergence of SARS-CoV-2 variants of concern in the pediatric population of the United States. https://doi.org/10.1101/2021.05.22.21257660.

60. McCormick EM, Lott MT, Dulik MC**, Shen L**, Attimonelli M, Vitale O, Karaa A, Bai R, Pineda-Alvarez DE, Singh LN,

Stanley CM, Wong S, Bhardwaj A, Procaccio V, Wallace DC, Gai X, Falk MJ (2020). Standards and guidelines for mitochondrial DNA variant interpretation. Human Mut. 41(12):2028-2057.

59. Falk MJ, **Shen L**, Gai X. (2020). Mitochondrial Disease Genes Compendium: connecting with knowledge in the Mitochondrial Disease Sequence Data Resource (MSeqDR). Mitochondrial Disease Genes Compendium, ISBN: 978-0-12-820029-2, 17-23.
58. McAulay K, Schuetz AN, Fauntleroy K, **Shen L**, Merveille YM, Deroncelay A, Cole N, Fitzgerald DW, Ocheretina O, (2021). Multidrug Resistant Pseudomonas aeruginosa in Healthcare Facilities in Port-au-Prince, Haiti, Journal of Global Antimicrobial Resistance, doi:https://doi.org/10.1016/j.jgar.2021.02.016
57. Xu L, Kim ME, Polski A, Prabakar RK, **Shen L**, Peng C-C, Reid MW, Chévez-Barrios P, Kim JW, Shah R, Jubran R, Kuhn P, Cobrinik D, Biegel JA, Gai X, Hicks J, Berry JL. (2021). Establishing the Clinical Utility of ctDNA Analysis for Diagnosis, Prognosis, and Treatment Monitoring of Retinoblastoma: The Aqueous Humor Liquid Biopsy. Cancers. 2021; 13(6):1282. https://doi.org/10.3390/cancers13061282 .
56. **Shen L**, Dien Bard J, Biegel JA, Judkins AR, Gai X. (2020). Comprehensive variant and haplotype landscape of 50,000 global SARS-CoV-2 isolates. bioRxiv https://doi.org/10.1101/2020.07.09.193722
55. **Shen L,** Maglinte DT, Ostrow D, Pandey U, Bootwalla M et al (2020). Children's Hospital Los Angeles COVID-19 Analysis Research Database (CARD) -A Resource for Rapid SARS-CoV-2 Genome Identification Using Interactive Online Phylogenetic Tools. https://doi.org/10.1101/2020.05.11.089763
54. **Shen L** (2023). MSeqDR Quick-Mitome (QM) Web Server – Combining Phenotype-Guided Variant Interpretation and Machine Learning Classifier to Aid in Mitochondrial Disease Diagnosis. Current Protocol (Invited Manuscript, in writing for Nov. 2023).
53. Pandey U, Yee R, **Shen L**, Judkins AR, Bootwalla B, Alex Ryutov A, Maglinte DT, Ostrow D, Precit M, Biegel JA, Jeffrey M Bender, Gai X, Dien Bard J. (2020) High Prevalence of SARS-CoV-2 Genetic Variation and D614G Mutation in Pediatric Patients with COVID-19. Open Forum Infectious Diseases. https://doi.org/10.1093/ofid/ofaa551
52. Pandey U, Yee R, Precit M, Bootwalla M, Ryutov A, **Shen L**, Maglinte DT, Ostrow D, Biegel JA, Judkins AR, Gai X, Dien Bard J, (2020) Pediatric COVID-19 in Southern California: clinical features and viral genetic diversity. medRxiv 2020 doi: https://doi.org/10.1101/2020.05.28.20104539
51. Li C, Debruyne DN, Spencer J, Kapoor V, Liu LY et al, **Shen L**, et al (2020) Highly sensitive and full-genome interrogation of SARS-CoV-2 using multiplexed PCR enrichment followed by next-generation sequencing. bioRxiv, doi: https://doi.org/10.1101/2020.03.12.988246.
50. **Shen L**. (2023). A Rare Disease Virtual Registry of 8,000 Primary Mitochondrial Disease Cases Constructed through Semi-automated Literature Mining Platform at MSeqDR. (Manuscript in writing).
49. Ji J, **Shen L (co-first authors), et al (2019).** Whole exome sequencing in the pediatric setting: A semi-automated and phenotype-driven workflow leads to high diagnostic yield. Cold Spring Harbor Molecular Case Studies.5i:a003756. doi: 10.1101/mcs.a003756
48. Brian Lee, Lily Nasanovsky, **Lishuang Shen, et al**. (2023). Significance Associated with Phenotype (SAP) Score Aids in Variant Prioritization for Exome Sequencing Analysis. Journal of Molecular Diagnosis (in revision).
47. **Shen, L**., Attimonelli, M., Bai, R., Lott, M. T., Wallace, D. C., Falk, M. J., Gai, X. (2018). MSeqDR mvTool: A mitochondrial DNA web and API resource for comprehensive variant annotation, universal nomenclature collation, and reference genome conversion. Human Mutation. 39:806-810.
46. Gustafson MA, McCormick E, Perera L, **SHEN L** (co-author), et al (2019). Mitochondrial single-stranded DNA binding protein novel de novo SSBP1 mutation in a child with single large-scale mtDNA deletion (SLSMD) clinically manifesting as Pearson, Kearns-Sayre, and Leigh syndromes. PloS One. 2019; 14 (9) :e0221829 (https://doi.org/10.1371/journal.pone.0221829).
45. Ung C, Sanchez AV, **Shen L**, Davoudi S, et al (2017). Whole Exome Sequencing Identification of Novel Candidate Genes in Patients with Proliferative Diabetic Retinopathy. Vision Res. 2017 Apr 18. doi:10.1016/j.visres.2017.03.007. PMID: 28431867
44. Ji J, Quindipan C, Parham D, **Shen L**, Ruble D, Bootwalla M, Maglinte DT, Gai X, Saitta SC Biegel JA, Mascarenhas L. (2017). Inherited germline ATRX mutation in two brothers with ATR-X syndrome and osteosarcoma. American Journal of Medical Genetics Part A, DOI: 10.1002/ajmg.a.38184.
43. Clima R, Preste R, Calabrese C, Diroma MA, Santorsola M, Scioscia G, Simone D, **Shen L**, Gasparre G, Attimonelli M. (2017). HmtDB 2016: data update, a better performing query system and human mitochondrial DNA haplogroup predictor. ***Nucleic Acids Research***. 2017 Jan 4;45(D1):D698-D706.
42. Lam J, Oh D, Wong L, Amarnani D, et al, **Shen L**, Gai X, D'Amore P, Kim L, Arboleda-Velasquez J. (2017). Identification of RUNX1 as a mediator of aberrant retinal angiogenesis. Diabetes 66:1950-1956

41. **Shen L**, Diroma MA, Gonzalez M, Navarro-Gomez D, Leipzig J, Lott MT, van Oven M, Wallace DC, Muraresku CC, Zolkipli-Cunningham Z, Chinnery PF, Attimonelli M, Zuchner S, Falk MJ, Gai X. (2016). MSeqDR: A Centralized Knowledge Repository and Bioinformatics Web Resource to Facilitate Genomic Investigations in Mitochondrial Disease. Hum Mutat. 37(6):540-548. PMID: 26919060.

40. Kartha N, **Shen L**, Maskin C, Wallace M, Schimenti JC. (2016). The Chromatin Remodeling Component Arid1a Is a Suppressor of Spontaneous Mammary Tumors in Mice. Genetics 203(4): 1601–1611.

39. Falk MJ, **Shen L**, Gai X. 2016. From case studies to community knowledge base: MSeqDR provides a platform for the curation and genomic analysis of mitochondrial diseases. Cold Spring Harb Mol Case Stud. 2016 May;2(3):a001065. doi:0.1101/mcs.a001065. PubMed: 27148591.

38. Davoudi S, Navarro-Gomez D, **Shen L**, Ung C, Ren A, Sullivan L, Kwong M, Janessian M, Comander J, Gai X, Lobo AM, Papaliodis GN, Sobrin L. (2016). NOD2 genetic variants and sarcoidosis-associated uveitis. Am J Ophthalmology Case Rep. 3: 39–42.

37. Marni J. Falk, **Lishuang Shen**, Michael Gonzalez, et al. (2015). MSeqDR: Making genomics accessible to the mitochondrial disease community. Mitochondrion 09/2015; 24:S2-S3.

36. Tallmadge LR, **Shen L**, Tseng CT, Miller SC, Barry J,Felippe JB (2015) Bone marrow transcriptome and epigenome profiles of equine common variable immunodeficiency patients unveil block of B lymphocyte differentiation. Clinical Immunology. 160(2):261-276

35. Ocheretina O, **Shen L**, Escuyer VE, Mabou M-M, Royal-Mardi G, Collins SE, et al. (2015) Whole Genome Sequencing Investigation of a Tuberculosis Outbreak in Port-au-Prince, Haiti Caused by a Strain with a "Low-Level" rpoB Mutation L511P - Insights into a Mechanism of Resistance Escalation. PLoS ONE 10(6): e0129207. doi:10.1371/journal.pone.0129207

34. Falk MJ, **Shen L**, Gonzalez M, Leipzig J, Lott MT, et al. (2015) Mitochondrial Disease Sequence Data Resource (MSeqDR): A global grass-roots consortium to facilitate deposition, curation, annotation, and integrated analysis of genomic data for the mitochondrial disease clinical and research communities. Mol Genet Metab. 2015 Mar;114(3):388-396.

33. Navarro-Gomez D, Leipzig J, **Shen L**, Lott M, et al. (2015) Phy-Mer: a novel alignment-free and reference-independent mitochondrial haplogroup classifier. ***Bioinformatics***. 31(8):1310-1312.

32. Sladek R., Rocheleau G., Rung J., Dina, C., **Shen L**., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S., Balkau, B., Heude, B., Charpentier, B., Hudson, T.J., Montpetit, A., Prentki, M., Posner, B.I., Balding, D.J., Meyre, D., Polychronakos, C., Froguel, P., (2007) A Genome-wide Association Study Identifies Novel Susceptibility Loci For Type 2 Diabetes Mellitus, ***Nature*** 445:881-885.

31. Rung J, Cauchi, S, Albrechtsen A, **Shen L**, et al., Polychronakos C, Pedersen O, Froguel P, and Robert Sladek R. (2009) Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia, ***Nature Genetics*** 41:1110-1115.

30. **Shen L**, Gong J, Caldo RA, Nettleton D, Cook D, Wise RP, Dickerson JA. (2005) Barleybase - An Expression Profiling Database for Plant Genomics. ***Nucleic Acid Research***. 33: D614-618.

29. Wallace MD, Pfefferle AD**, Shen L**, McNairn AJ, Cerami EG, Fallon BL, Rinaldi VD, Southard TL, Perou CM, Schimenti JC. (2012) Comparative oncogenomics implicates the neurofibromin 1 gene (NF1) as a breast cancer driver. Genetics. 192(2):385-396.

28. Bolcun-Filas EW, Bannister LA, Barash A, Schimenti KJ, Hartford SA, Eppig J, Handel MA**, Shen L**, and Schimenti JC (2011) A-MYB (MYBL1) transcription factor is a master regulator of male meiosis. Development 138:3319-30.

27. Ching YH, Munroe RJ, Moran JL, et al, **Shen L**, Shindo M, Shou JY, Stenson EK, Stover PJ and Schimenti JC (2010) High resolution mapping and cloning of ENU-induced mutations in the Rw region of mouse Chromosome 5. BMC Genetics 2010, 11:106doi:10.1186/1471-2156-11-106.

26. Wise RP, Caldo RA, Hong L, **Shen L**, Cannon EK, Dickerson JA. (2007) BarleyBase/PLEXdb: A Unified Expression Profiling Database for Plants and Plant Pathogens, Methods in Molecular Biology, 406:347-363, Plant Bioinformatics - Methods and Protocols. Edwards D. ed. Humana Press. 2007.

25. Tang X, **Shen L**, Dickerson JA (2005) BarleyExpress: A Web-Based Submission Tool for Enriched Microarray Database Annotations. ***Bioinformatics*** 21:399-401.

24. Courtois B, **Shen L**, Petalcorin W, Carandang S, Mauleon R, Li Z (2003) Locating QTLs controlling constitutive root traits in the rice population IAC 165 x Co39. Euphytica 134 (3): 335-345 2003

23. **Shen L**, Courtois B, McNally K, Robin S, Li Z(2001) Evaluation of near-isogenic lines of rice introgressed with QTLs for root depth through marker-aided selection. Theor. Appl.Genet. 103:75-83.

22. He, P., J.Z. Li, X.W. Zheng, LS. Shen, C.F. Lu, Y. Chen and L.H. Zhu (2001). Comparison of molecular linkage maps and agronomic trait loci between DH and RIL populations derived from the same rice cross. Crop Science 41(4): 1240-1246.

21. Bach Jensen L, Courtois B, **Shen L**, Li Z, Olofsdotter M, Mauleon R (2001) Locating genes controlling allelopathic effects against *Echinochloa crus-galli* in upland Rice. Agronomy J., 93:21-26.
20. Courtois B, McLaren G, Sinha P K, Prasad K Yadav R and **Shen L** (2000) Mapping QTLs associated with drought avoidance in upland rice. Molecular Breeding, 6:55-60.
19. **Shen LS**, Zheng X and Zhu LH. (2000) Mapplotter - a program for getting graphic output for genetic maps, graphical genotypes and QTL curves. Hereditas (Beijing) 22(3):172-174.
18. **Shen L**, Courtois B, McNally K, McCouch S, Li Z (1999) Developing near-isogenic lines of IR64 introgressed with QTLs for deeper and thicker roots through marker-aided selection. In: Ito O, O'Toole J, Hardy B (eds) Genetic improvement of rice for water-limited environments. IRRI, Philippines, p275-289
17. Gong JM, He P, Qian QA, **Shen LS**, Zhu LH, Chen SY (1999) Identification of salt-tolerance QTL in rice (Oryza sativa L.). Chinese Science Bulletin 44 (1): 68-71.
16. **Shen LS**, He P, Xu YB, Tan ZB, Lu CF and Zhu LH (1998) Genetic molecular linkage map construction and genome analysis of rice doubled haploid population. J Integrative Plant Biology (Acta Botanica Sinica) 40(12): 1115-1122.
15. **Shen LS** and Zhu LH (1998) Direct PCR-based genetic mapping of rice telomeric repeats associated sequences. Genome, 41 (2): 193-198.
14. Lu CF, **Shen LS**, Tan ZB, Xu YB, He P, Chen Y and Zhu LH (1997) Comparative mapping of QTLs for agronomic traits of rice across environments by using a doubled haploid population, Theor. Appl.Genet. 94(1): 145-150.
13. Gong XQ, Hong DJ, Chen SY, **Shen LS**, Li P and Yan CJ (1998) Isolation and characterization of five rice telomere-associated sequences. Science in China (series C), 41 (4): 372-380.
12. Xu YB, **Shen LS**, McCouch SR and Zhu LH. (1997) Extending the genetic linkage map of rice doubled haploid population with microsatellite markers. Chinese Science Bulletin 20:2220-2223.
11. He P, **Shen LS**, Lu CF, Chen Y, Zhu LH. (1998). Genetic analysis and mapping the anther culture response genes in rice (Oryza sativa L.). Acta Genet Sinica, 1998, 25(4):337-344.
10. Tan ZB, **Shen LS**, Yuan ZK, Zhu LH (1997). Identification of QTLs for ratooning ability and grain yield traits of rice and the analysis of their genetic effects. Acta Agronimica Sinica. 23(3): 289-295.
9. He P, **Shen LS**, Lu CF, Chen Y and Zhu LH (1998) Analysis of quantitative trait loci which contribute to anther culturability in rice (*Oryza sativa* L.). Molecular Breeding, 4(2): 165-172.
8. He P, Chen Y, **Shen LS**, Lu CF and Zhu LH (1997) Inheritance of callus induction ability in rice anther culture. Chinese Science Bulletin, 42(14): 1209-1211.
7. Liu YS, Zhu LH, Sun JS, Wang YZ, He P, **Shen LS** (1997) Mapping QTLs for reproductive barriers occurring in hybrid between indica and japonica rice (*Oryza sativa* L). J Integrative Plant Biology (Acta Botanica Sinica) 39(12):1099-1104.
6. Liu YS, Zhu LH, Sun JS, **Shen LS**, Wang YZ, He P, Lu CF and Chen Y (1998) Inheritance and mapping of embryo sac abortion in hybrid between indica and japonica rice (*Oryza sativa* L.). Sci. in China (Series C). 41(3): 272-277.
5. Tan, ZB, **Shen LS**, Kuang HC, Lu CF, Chen Y, Zhou, KD and Zhu LH (1997) Identification of QTLs for lengths of the top internodes and other traits in rice and analysis of their genetic effects. Chinese J. Genet. 24(1): 15-22.
4. Chen Y, Lu CF, He P, **Shen LS**, Xu JC (1997) Gametic selection in a doubled haploid population derived from anther culture of indica/japonica hybrid of rice (*Oryza sativa* L.) Chinese J. Genet. 24(3): 201-209.
3. Xing JP, S Chen, L Zhu**, L Shen**, J Li, S Bi, X Jiang (1995) Molecular biological analysis of large-grained rice mutant induced outer space condition in recoverable satellite. Space Medicine and Medical Engineering 8 (2), 109-112.
2. Zhu LH, Lu CF, Li P, **Shen LS** and Xu YB (1995) Using doubled haploid populations of rice for quantitative trait locus mapping. In: Rice genetics III: proceedings of the Third International Rice Genetics Symposium, ed. by G. S. Khush, p. 631-635.
1. **Shen LS** and Zhu LH (1995) The comparative genome analysis and comprehensive genetic system of plants, Progress in Biotechnology. 15:23-27.

书籍章节 **Book Chapters**

61. **Shen L**, McCormick EM, Muraresku CC, Falk MJ, Gai X.(2020) Clinical Bioinformatics in Precise Diagnosis of Mitochondrial Disease. Clin Lab Med. 2020 Jun;40(2):149-161. doi: 10.1016/j.cll.2020.02.002. PMID: 32439066.

51. Falk MJ, **Shen L**, Gai X. (2020) Mitochondrial Disease Genes Compendium: connecting with knowledge in the Mitochondrial Disease Sequence Data Resource (MSeqDR). Mitochondrial Disease Genes Compendium, 17-23.

51.2. **Shen L** (2020) Mitochondrial Disease Genes Compendium. (Author of ten genes' chapters).

24. Roger P. Wise, Rico A. Caldo, Lu Hong, Lishuang Shen, Ethalinda Cannon, Julie A. Dickerson: BarleyBase/PLEXdb. Plant Bioinformatics, 02/2008: pages 347-363.

近期受邀讲座 Conference Presentation:

1. **Lishuang Shen**, An integrated genomic surveillance platform reveals multiple introductions and accelerating localization of SARS-CoV-2 into California, USA, and worldwide countries. Data Con LA, 2020-10, Los Angeles
2. **Lishuang Shen e**, MSeqDR Tutorial Workshop, UMDF Mitochondrial Medicine 2019 - Washington DC.
3. **Lishuang Shen**, MSeqDR Hands-on Workshop, UMDF Mitochondrial Medicine 2018- Nashville, TN.
4. **沈利爽** 线粒体突变的致病性解读，2018 生命科学与人类生活国际论坛，2018.05 浙江省金华市.
5. **Lishuang Shen**, MSeqDR Hands-on Workshop, UMDF Mitochondrial Medicine 2017-Washington DC.
6. **Lishuang Shen**, MSeqDR Tutorial Workshop, UMDF Mitochondrial Medicine 2016 –Seattle, WA.
7. **Lishuang Shen**, MSeqDR Hands-on Workshop, UMDF Mitochondrial Medicine 2015- Pittsburgh, PA.

中文发布论文（**PUBLICATIONS IN CHINESE,** 19 篇知网: http://search.cnki.com.cn/Search/Result?author=沈利爽 ）

[1]沈利爽,郑先武,朱立煌.Mapplotter -一个输出遗传图谱、图示基因型和 QTL 曲线图形的软件.遗传,2000(03):172-174.

[2]沈利爽,何平,徐云碧,谭震波,陆朝福,朱立煌.水稻 DH 群体的分子连锁图谱及基因组分析.植物学报,1998(12):3-5.

[3]沈利爽,朱立煌.应用 PCR 进行水稻染色体末端区域作图.遗传学报,1998(06):508-516.

(** 为篇幅简略计，另外 16 篇中文论文略。The other 16 papers are omitted for brevity)

领域技能 SKILLS

- Leadership: bioinformatics team building, coaching, and management.
- Fields: Bioinformatics, Genomics / Transcriptomics / Epigenomics / Genetics, Cancer, Mitochondrial Disease, Personalized Medicine, Statistics, Molecular Biology, Protein 3D structure modeling.
- Programming Languages: R, Python, Perl, Bash / Awk Shell , , SQL, PHP, HTML/CSS/JavaScript, full stack database-driven reactive website (https://mseqdr.org) with RESTful API.
- Bioinformatics Computational Biology: GPU-based HPC + Parabricks, Cloud computing, OpenCGA genomic Big data platform, Illumina Dragen, GATK, DeepVariant, VarScan, SnpEff, NGS, NCBI, EBI, ClinGen/ClinVar, TCGA, ICGC.
- Cloud Computing: AWS Platform for NGS data processing and MSeqDr.org database-driven websites.
- Databases: MySQL, MongoDB
- AI /Machine Learning: h2o, xgboost, GBM, GLM, deep-learning
- Operating Systems: Linux (Ubuntu, CentOS, AWS, WSL2), Windows
- Version and Workflow Control: GitHub, NexFlow, SnakeMake
- Statistical Analysis: R, SAS