

# 宋捷

您好，我是宋捷。生信及IT相关五年工作经验，期间我在创业团队就职时间较长，技能树相对全面，有管理和部门间协调以及完整的产品开发经验。目前感兴趣的相关领域为生物信息，合成生物学，医疗/生物数据建模与挖掘，以及相关的产品开发，项目管理，统筹等相关工作。

如果您对我的背景和技能感兴趣，请随时与我联系。

## 教育经历

- 2012 | 2016  
● **西北农林科技大学**  
(本科) 生命科学学院 生物技术 陕西
- 2016 | 2019  
● **西北农林科技大学**  
(硕士) 旱区作物逆境生物学国家重点实验室 生物信息学 陕西

## 研究及工作经历

- 2015 | 2019  
● **研究生**  
马闯实验室 旱区作物逆境生物学国家重点实验室 陕西
  - 研究方向主要为机器学习/深度学习在RNA甲基化预测上的应用
  - 生信分析管线构建及可视化
  - 期间发表SCI论文5篇 (1作1篇)
- 2019 | 2020  
● **数据挖掘工程师**  
深圳碳云智能科技有限公司 (iCarbonX) 深圳
  - 使用自编码去对单细胞表达数据的降噪工作
  - 免疫芯片数据的处理及挖掘，进行了基于此技术的疫苗肽构建工作
- 2020 | 2021  
● **实习研究员**  
大数据中心 深圳市中山大学第七附属医院 深圳
  - 生信相关：UKB数据的专项处理
  - IT和运维：大数据中心协助指定管理细则，用户手册，并协助管理。
- 2022 | 2023  
● **生信及IT主管**  
深圳寻竹生物科技有限公司 (Ailurus Ltd.) 深圳
  - 生信相关：技术和管线开发，主要涉及NGS数据，芯片数据处理，和下游的分析工作。
  - 大模型：技术和应用相关，af2, progen及众llm，主要涉及模型部署，finetune，产品设计，效率优化。
  - IT和运维：技术相关(全栈)，主要涉及网站开发，服务部署，产品报告自动化，内部IT工具。
  - 部门管理：进行proposal设计和任务分配，硬件资源调配等。

## 主要项目经验



## CONTACT INFO

✉ [q2516581@126.com](mailto:q2516581@126.com)  
🌐 [lovisongjie.top](http://lovisongjie.top)  
☎ songjie9075  
📅 1994.10.24

## SKILLS

**项目技能：**常规/特殊生信流程构建，报告自动化处理，大模型应用开发，常规网站/网页开发。

**运维及管理：**生信小团队的管理，软硬件以及云服务的运行维护，资源调配等。

**技术栈：**分析处理的主要是R+python，流程控制主要是bash, galaxy, snakeflow等，web和报告报告主要使用 shiny latex flask bootstrap等。

Last updated on 2024-01-19 by R.

2023

### ● 高通量筛选启动子的分析系统

公司设计了一套新的超高通量的启动子筛选的合成生物学实验方案。接手后发现了早期代码的设计错误，重新基于实验设计写了一套完整的分析和报告系统，完善了该产品，并协作写了软著。

📍 深圳 Ailurus

2023

### ● 半自动论文生成系统

重新申请了ukb和tcga数据，结合galaxy系统重新开发并优化了自动化分析管线。通过对结果的结构化处理，并结合gpt api和promot工程，开发了半自动的latex论文初稿生成管线。进行一定的人工矫正可进行低质量投稿。

📍 深圳 Ailurus

2022

### ● 癌症靶向用药自动化报告项目

该项目构建了一个使用latex生成主体报告模版，提取患者突变信息并匹配知识库来展示突变和解析信息。该项目主要难点在：1.使用个体信息自动化生成复杂的latex表格且正常渲染；2.报告基于不同付费要显示不同样式；3.要交付给无计算机背景的销售/生产部门使用；4.数据库维护

📍 深圳某医院

2021

### ● ukbiobank数据的处理及阿尔兹海默症的遗传风险分析

ukbiobank 数据（50w 人+80w 遗传位点+2000 表型），进行数据获取，清洗，以及结构化的预处理。构建了一套 ukb 数据的一键式的表型-基因型关联工具。在院期间对AD和MCI进行了关联分析，以及的风险预测模型构建及评估工作。此阶段编写了大量相关技术及数据说明文档。

📍 深圳某医院

2020

### ● 免疫芯片数据的挖掘和疫苗开发

ImmunoSignature 免疫芯片数据,小组内对该引进技术进行了从零开始的预处理管线和分析管线工作的开发工作。本人主要负责关联医院的临床数据和实验的下机数据，清洗，质控，和矫正工作。在新冠疫情初期，个人负责了探索抗体对病毒蛋白响应区域的预测工作，在小鼠实验阶段，筛选出的肽段能有效产生中和抗体。

📍 深圳 iCarbonX



## 管理经验

2022

|  
2023

### ● 生信及IT部门管理

Ailurus任职期间，管理生信及IT部门。除个人项目外，同时负责任务规划及分配，系统资源调配，技能培训，与其他部门和业务公司沟通等职能。

📍 深圳 Ailurus

2021

### ● 数据中心的组建及管理

21 年 6 月受朋友委托完成区块链chia项目调研并运行(总投资400w)。快速组建创业团队并在 15 天内完成全项目调研、硬件测试、采购、项目选址、电路改造、管理软件编撰等，快速部署了两处数据中心。该项目由于政策风险快速终止。

📍 深圳 洛未信息



## 学术文章

- **Transcriptome-Wide Annotation of m<sup>5</sup>C RNA Modifications Using Machine Learning.**  
Frontiers in Plant Science. Song J(1st)
- **PEA: an integrated R toolkit for plant epitranscriptome analysis**  
Bioinformatics. Song J(2nd).
- **deepEA: a containerized web server for interactive analysis of epitranscriptome sequencing data**  
Plant Physiology. Song J(2nd).
- **Evolution of the RNA N<sup>6</sup>-methyladenosine methylome mediated by genomic duplication**  
Plant Physiology. Song J(4th)
- **A Deep Convolutional Neural Network Approach for Predicting Phenotypes from Genotypes**  
Planta. Song J(3rd).
- **miRLocator: A python package and web server for predicting miRNAs from pre-miRNA sequences**  
Methods in Molecular Biology. Song J(4th).
- **Predicting early-onset COPD risk in adults aged 20–50 using electronic health records and machine learning (In review)**  
PeerJ. Song J(2nd).