

# Early bioinformatics research in China

生信 2001 张子栋 2020317210101

这篇文章是由陈润生院士撰写的，发表在 *Quantitative Biology* 期刊上，是纪念人类基因组计划 20 周年系列文章的第三篇。文章回顾了我国早期生物信息学的发展历程，包括参与国际人类基因组计划的相关专家和团队，以及陈院士自己课题组如何加入国内人类基因组研究，以及克服各种困难从事非编码序列研究的经历。文章展示了我国生物信息学领域的先驱者们在理论和实践上的创新和贡献，也反映了我国生物信息学的发展与国际接轨的过程。

阅读完这篇文章，我体会到：生物信息学是一门交叉学科，需要不断吸收和融合其他学科的知识和方法，如数学、统计、计算机、密码学等，以应对生命科学中日益复杂的数据分析和挑战。此外，生物信息学不仅要依赖于实验数据，也要有自己的理论基础和创新思维，如陈院士提出的密码学方法用于基因识别，以及他对非编码序列功能的探索。生物信息学要与生命科学紧密结合，关注生命现象的本质和规律，为生命科学提供有效的工具和解决方案，如陈院士参与的泉生热袍菌 B4 基因组、人类 3 号染色体短臂等项目。生物信息学要有开放和合作的精神，与国内外同行交流和分享数据和经验，参与国际重大项目和平台，如陈院士建立的非编码 RNA 数据库 NONCODE 和 NPInter 等。最近，中南大学承办了第二十一届亚太生物信息学会议（APBC 2023），邀请了来自亚太地区的一百余位专家学者和行业人士参加，共同探讨生物信息学领域的前沿科技和研究成果。

在查阅更多资料后，我了解到中国在生物信息学的部分研究已经处于国际前列。例如高通量数据处理能力，我国科学家在基因组测序、单细胞测序、crRNA 测序等领域获得的丰富的数据，也开发了高效、精确和创新性的数据分析和组装方法，例如 CIRIquant、wtdbg2 等。另外我国在构建人类细胞图谱、利亚人工智能度量学习鉴定单细胞类型等方面也取得了突出成果，有望为精准医疗提供重要的数据制成和技术手段，通过解析遗传基因大数据、提前了解人体疾病的危险因素，并加以预防、干涉，以保证人体健康。

此外我国生物信息学还有一些方面需要更长远的发展。比如生物数据库方面，在平时的课程学习，例如生物信息学软件基础课程中，在完成课程项目时，我们使用的数据库、软件均为国外开发。在生物信息学进展课程中，虽然提到了国内的几个生物数据库，但是建立时间、数据量和使用体验上暂时没有达到国外生物数据库的水准。另外生物信息学的部分研究方法和技术还不够成熟和标准化，需要更多的创新和验证。我国生物信息学的人才培养和队伍建设也可以更加完善，需要加强多学科的交叉和融合。同样，生物信息学的科研成果在应用和转化上还有很大的空间和潜力，需要加强与生物医药、农业、环境科学的领域的合作和交流。