

高通量测序技术在环境微生物学中的应用及发展

吕欣宇*, 孟庆陆, 马玮辛, 郑兆棋

滨州学院化工与安全学院, 山东 滨州

收稿日期: 2023年3月7日; 录用日期: 2023年5月30日; 发布日期: 2023年6月9日

摘要

高通量测序技术的别称为大规模平行测序技术, 采用高效的高通量采样方法和测序方法, 可以快速获得理想的实用菌株。针对高通量测序方法在细菌菌株检测中的应用, 本文总结了行业研究、行业应用发展、高通量测序研究及趋势, 探讨了高通量测序技术在环境微生物学中的应用方法, 以提高检测的准确性和效率。为优化高通量测序技术的准确度、效能和实用范围提供理论指导。还探讨了高通量测序技术行业的发展现状, 预测了高通量测序技术行业的发展方向。

关键词

高通量测序, 菌种检测, 行业现状, 发展趋势

Application and Development of High-Throughput Sequencing Technology in Environmental Microbiology

Xinyu Lv*, Qinglu Meng, Weixin Ma, Zhaoqi Zheng

College of Chemical Engineering and Safety, Binzhou University, Binzhou Shandong

Received: Mar. 7th, 2023; accepted: May 30th, 2023; published: Jun. 9th, 2023

Abstract

High-throughput sequencing technology is called large-scale parallel sequencing technology, and the ideal practical strains can be obtained quickly by using an efficient high-throughput sampling method and sequencing method. For the application of the high-throughput sequencing method in bac-

*第一作者。

文章引用: 吕欣宇, 孟庆陆, 马玮辛, 郑兆棋. 高通量测序技术在环境微生物学中的应用及发展[J]. 生物过程, 2023, 13(2): 85-90. DOI: 10.12677/bp.2023.132012

terial strain detection, this paper summarizes the industry research, industry application development, high-throughput sequencing research and trend, and discusses the application method of high-throughput sequencing technology in strain detection, so as to improve the accuracy and efficiency of detection. In order to optimize the accuracy, efficiency and practical range of high-throughput sequencing technology, and then provide theoretical guidance for the accelerated detection of strains. The development status of the high-throughput sequencing technology industry is also discussed and the development direction of the high-throughput sequencing technology industry is inferred.

Keywords

High-Throughput Sequencing, Strain Detection, Industry Status, Development Trend

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1953年,詹姆斯·沃森和弗朗西斯·克里克共同提出DNA双链结构,标志着生命科学发展向分子生物学的转变,引发了生命科学和生物技术的一场革命[1]。1990年,随着以Sanger测序技术为基础的第一代测序技术不断发展,人类基因组计划正式启动,到2004年,正式进入后基因组时代[2]。Sanger测序技术原理是,核酸模板在核酸聚合酶、引物、四种单脱氧核苷酸存在条件下复制或转录时,如果在四管反应系统中分别按比例引入四种双脱氧核苷酸,只要双脱氧核苷酸掺入链端,该链就停止延长,链端掺入单脱氧核苷酸的片段可继续延长。以分离长短不一的核酸片段(长度相邻者仅差一个核苷酸),根据片段链端的双脱氧核苷酸,便可依次阅读合成片段的核苷酸排列顺序。

深层次和效率等大规模测序技术的要求,促使了以罗氏公司的454、Illumina公司的HiSeq以及Life Technology公司的SOLiD、IonTorrent为代表的新一代测序技术(Next-Generation Sequencing, NGS)的诞生。其显著特点是高通量,在精确测序时,远优于基于Sanger测序法的第一代测序方法,且价格显著降低。454 life sciences是最早的基于焦磷酸序列的商业基因组学系统之一,于2004年左右推出。它具有高速、读取长度较长和高精度等优点,但由于其通量低、成本高等劣势,现在已慢慢退出市场。随着测序技术的发展和优化,Illumina和Life Technoogy等公司现在主导着下一代高通量测量系统的安装。Illumina公司的Solexa和Hiseq测序仪器采用边合成边测序的方法,在使用中具有高通量密度、短时间和低成本的合成和测序技术,目前是主要市场中占主导地位的第二代产品。Life Technology公司推出的IonTorrent的核心技术,通过半导体技术与化学和数字信息相连接,具有高精度、高速、低成本解读等优点,见图1。

中国科学家[3]对从陈年葡萄酒中分离出来的奶酪牛奶棒菌株*Lactobacillus casei* BD-II和保加利亚菌株*Lactobacillus delbrueckii* subsp. *bulgaricus* strain NDO2的酸奶棒进行了研究;韩国科学家分析了泡菜发酵的香肠乳杆菌*Lactobacillus farciminis*、棒状乳杆菌*Lactobacillus coryniformis* subsp. *coryniformis*和高丽魏斯氏菌*Weissella koreensis*;德国科学家对从奶酪表面分离的变异棒状杆菌*Corynebacterium variabile*和棒状杆菌*Corynebacterium casei*进行了分析。它们都依次进行了完整的基因组测序,并分析了与它们相关的比较基因组和代谢能力。

2. 高通量测序行业研究

1) 基于颜色或荧光的高通量研究

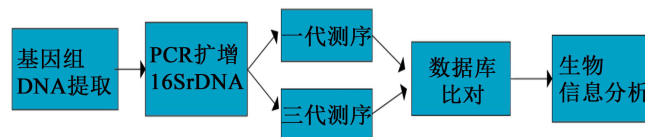


Figure 1. Technical route of strain detection
图 1. 菌种检测技术路线

基于颜色或荧光[2]的细胞筛选是一种非常显而易见的高通量测序技术(图 2),它在细菌进化中广泛有效地用于选择含有有色或荧光代谢物的菌株。对于生产彩色产物(如番茄红素、 β -胡萝卜素和虾青素)菌株,颜色的类型和深度可以初步判断反应代谢产物的类型和产量。在现代机器人技术的帮助下,筛选效率可以达到每次 10^6 个突变体。

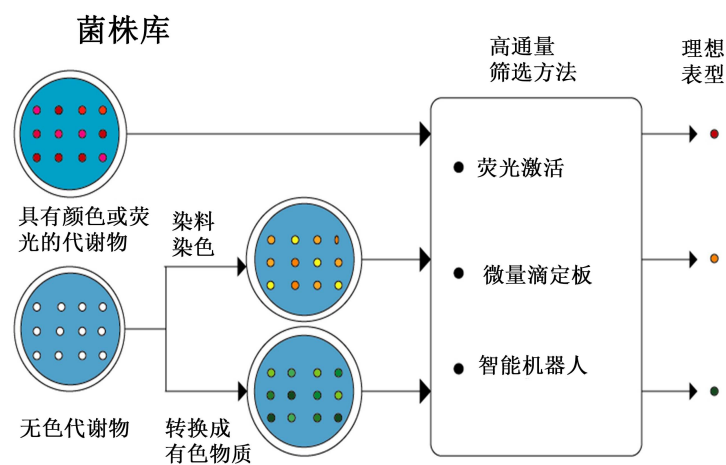


Figure 2. High-throughput sequencing techniques based on color or fluorescence

图 2. 基于颜色或荧光的高通量测序技术

2) 基于转录因子的高通量测序研究

转录因子可以与基因的启动子或增强子区域结合,单独或与其他蛋白质形成复合物,进而刺激或阻断 RNA 聚合酶参与转录过程,从而调节基因表达。在大肠杆菌中发现了 230 多种转录因子,它们可以感知多种代谢物,包括氨基酸、糖类、磷酸糖和脂质。细胞中的一些小分子代谢物可以通过配体结合和磷酸化作用来激活或失活转录因子,进一步统制相关报导基因的表述,从而将代谢物的浓度与荧光强度、细胞生长等检测信号联系起来,最终达到高通量测序的目的。

3) 基于磁珠提取技术的高通量测序研究

固有的 DNA 提取手段有煮沸法、盐析法、苯酚氯仿法、异丙醇沉淀法、离心柱法、磁球法等。本实验应用磁珠法[3]通过高通量测序技术自动地提取血凝块 DNA,将血凝块分成在不同条件下被蛋白酶 K 消化的血凝小块,把用 DNA 裂解液处理后的血凝小块进行比较,分析它们的消化率。通过电泳判断 DNA 的完整性,可以保证 DNA 的品质。研究表明,高通量测序技术的结合可以使提取 DNA 的过程简便高效,同时结合自动化精简人工操作。现阶段,磁珠作为一种精良的固相载体,通常应用于制备核酸、化学发光、细胞分选、纯化蛋白质等领域。硅磁技术普遍应用于研究分子生物学、临床诊断、高通量筛选等领域。

3. 高通量测序行业应用

3.1. 高通量测序在微生物生态学中的应用

目前, 高通量测序技术越来越多地用于研究海洋微生物生态学[4]。高通量测序技术可以揭示海洋中极少数量的微生物, 并发现更多新的微生物群落, 但在每一级分类中都有不确定的分类信息。Inoue 等人[5]使用高通量测序技术研究 Kath 山谷浅层地下水的生物污染。结果表明, 11 个地下水样品中有 10 个被大肠杆菌严重污染。此外, DNA 微阵列分析结果还揭示了 37 种病原菌的存在。Lu 等人[6]利用 454 焦磷酸测序技术研究农村生活废水和制革厂、服装厂、纽扣厂排放的废水对河流微生物生态系统的影响, 研究表明, 微生物群落的结构受温度、pH 值等环境条件的影响。Maugeri 等人使用 Illumina 测序技术研究意大利帕纳雷阿岛浅水热液系统中微生物群落结构和优势菌的组成。研究表明, 沉积物和流体中存在不同的群落结构, 优势菌分别为小红卵菌属和绿菌属, 序列衍生的原核生物在浅水热液系统的 C、Fe 和 S 循环中起关键作用[7]。

王鹏等人[8]采用高通量测序技术分析鄱阳湖典型湿地土壤细菌群落特征。研究表明, 沿湖面至坡地, 空间位置相近的土壤细菌群落结构具有更大的相似性, 苔草-藨草带、苔草带和芦苇带的细菌群落结构相近, 泥滩带和藜蒿带的细菌群落结构差异较大。湿地土壤平均相对丰度最高的门是变形菌门, 其次为酸杆菌门和绿弯菌门; 多数门分类细菌相对丰度沿湖面至坡地存在一定变化趋势。硝化螺菌属是第一大门类水平细菌群落。

3.2. 高通量测序在菌种筛选中的应用

徐瑛等[9]将脱硫菌的筛选与新一代高通量测序技术相结合, 探索脱硫菌筛选的新方法, 为生物甲烷的高值化利用提供技术支持。实验将富含硫细菌的环境土样及经过培养基富集培养后的液样分别进行 16S rDNA V6 区的高通量测序, 并分析了物种组成和丰度、Alpha 多样性和菌群结构。结果表明, 高通量测序的平均数据利用率达 99.177%, 测序量能够充分反映样品在该区域细菌的群落组成和结构。通过高通量测序技术, 能够在筛选脱硫菌之前就了解实验样品的菌群结构和组成, 为有针对性地设计筛选脱硫菌实验提供了一种新的方法。

何志刚等[10]采集不同低温菌源的样品, 进行低温连续富集继代培养, 利用第二代高通量测序方法分析不同处理的细菌群落结构和数量, 并测定不同世代的纤维素酶、秸秆失重率等数据, 筛选获得一组高效稳定分解玉米秸秆的复合菌群。

3.3. 高通量测序在病毒检测中的应用

2009 年首次报道了应用高通量测序技术检测植物病毒[9]。Al Rwahnih 等人使用以植物总 RNA 为模板的 NGS 平台在西拉子葡萄上找到了三种已知病毒、三种已知类病毒和一种新病毒——西拉葡萄病毒 1 号(Grapevine Syrah Virus-1, GSyV-1) [11]。Kreutze 等人对单独感染或复合感染甘薯羽斑病毒(SPFMV)和甘薯褪绿病毒(SPCSV)的甘薯小 RNA 进行测序。他们不仅鉴定了 SPFMV 和 SPCSV, 还发现了一些其他新的双链 DNA 病毒和单链 DNA 病毒[12]。并且 Adams 等人使用文库抑制和减法的手段通过 NGS 发现了一种新的 RNA 病毒——蛇鞭菊属轻斑驳病毒(Gayfeather Mild Mottle Virus, GMMV) [13]。

4. 高通量研究现状及发展动态

4.1. 高通量测序研究现状

随着病原体宏基因组测序(mNGS)技术的不断发展和优化, mNGS 检测技术不仅在感染的临床诊断中

得到广泛实验和赞同,而且在越来越多的医院中实现本地化,能够快速地满足临床早期传染性病原体诊断的需求。但现阶段 mNGS 的国产化仍存在技术复杂度较高、操作流程较为复杂、报告解读较为复杂、临床应用标准有待提高、质量评价结果不理想、行业中目前缺乏统一规范的行业培训机制等诸多困难。

随着高通量测序技术的发展,大量的基因组测序已经不再是问题。数据规模庞大,类型多样,包括转录组、基因组和蛋白质组等等。同时重复性不强,数据存储和可视化也是需要优化的问题。在算法优化、软件并行化、流程自动化、大规模数据存储、处理和高级分析等方面需要做大量工作。样品污染也是影响测序结果准确性的主要问题。

由于高通量测序技术的出现,这些序列的增长趋势[14][15]犹如潮水一般。一个实验室甚至可以每年发布 PT 级别的数据。如此大规模数据的高效存储、高效分析、共享循环利用是一项巨大的挑战,是高性能计算系统面临的重大挑战。目前,这些序列数据还不到冰山一角,只有少数完成了深入的分析。

据资料表明,高通量测序技术含有很多的优点,但高通量测序技术的局限性也不容忽视。一是测序速度提高了,但是后续对大量的测序数据的分析成为了一个非常困难的问题。其次,高通量测序技术不适合小型的测序。虽然高通量测序技术的价格正在逐渐下降,但每个反应的成本仍然在几千到几万元。普通客户会发现很难接受 PCR 产物和质粒等数十到数千个碱基的测序。目前,传统的 Sanger 测序法肯定是最佳选择。而且最新一代的测序仪器报价十分高,需要将近几百万,简单的小研究所难以承受此价格。因此,传统的 Sanger 测序法将与高通量测序技术一起成长、发展,短期内不会被淘汰。

4.2. 高通量测序发展动态

高通量测序技术的逐步发展促进了菌株进化工程在微生物工程领域的应用,在提高微生物菌株的环境耐受性、底物利用、代谢产物生产等方面取得了许多有意义的成就。随着生物学合成技术的发展,研究人员对菌种有了更加深切的认知,尤其是近年来人工合成酵母的出现,能够使研究人员能够更合理、更有效地进行菌株设计和大片段基因重组,以及获得更多样化的基因型和表型。这就是人工合成酵母的出现,它可以在未来的基础科学研究领域更好地研究基因结构的变异。这可以使研究人员更合理、更高效地实施菌种设计和大片段基因重组,产生更多样化的基因型和表型,是基础科学研究领域更好地研究基因结构变异与表型功能关系的重要模板。未来,为菌株的快速进化带来巨大机遇的同时,也为开发合适的高通量筛选平台提出了更大的挑战。尽管近年来高通量筛选方法的研究和应用取得了长足的进步,但它们一般只适用于筛选特定的代谢物、酶或代谢途径,缺乏普遍性和适用性。将来,由于生物检测信息的不断深化,在生物传感器领域,我们有更多的机会尝试开发更多对小分子具有特异性反应的生物传感器。同时,我们还需要探索更多的生物传感器机制,以使生物传感器的精度和准确度能够得到很大的提高。随着人工智能的发展和对人体微生物系统认识的不断加深,将极大地推动自动筛选的进步和自动数据分析的能力,为打造高通量测序技术提供强有力的技术支持。

高通量测序技术的成功与否是由筛选广度和精确度决定的,其实主要是由所选择的筛选模型决定。其中,最大程度地简化筛选步骤是最为重要的。这就好比用笼子捕鸟,笼子的大小对捕鸟能否顺利有决定性作用。如果笼子太大,就比较显眼,同时比较小的鸟也有可能逃脱掉;如果笼子太大,大一点的鸟就不会被逮到,有可能一天下来一只鸟都没有。总而言之,高通量 SrDNA 测序技术[12]现在已经成为生命科学研究项目中必不可少的工具。随着上一代测序技术和最新一代测序技术的快速发展,将高通量测序技术与筛选方法相结合必将加快菌种检测研究的发展。

基金项目

国家级大学生创新创业训练计划项目:202110449152;滨州学院实验技术项目:BZXYSYXM202008。

参考文献

- [1] 李霞, 赵晶, 鞠躬. 生物化学与分子生物学: 生命之窗-生命科学前沿纵览[M]. 西安: 第四军医大学出版社, 2014.
- [2] 吴林寰, 陆震鸣, 龚劲松, 等. 高通量测序技术在食品微生物研究中的应用[J]. 生物工程学报, 2016, 32(9): 1164-1174.
- [3] 杨祖明, 王颖, 姚明东, 等. 高通量筛选技术在菌种进化中的研究进展[J]. 化工进展, 2019, 38(5): 2402-2412.
- [4] 潘亭亭, 汪伟伟, 吉栩. 高通量自动化磁珠提取血凝块 DNA 的方法研究[J]. 中国卫生标准管理, 2020, 11(12): 132-136.
- [5] Inoue, D., Hinoura, T., Suzuki, N., *et al.* (2015) High-Throughput DNA Microarray Detection of Pathogenic Bacteria in Shallow Well Groundwater in the Kathmandu Valley, Nepal. *Current Microbiology*, **70**, 43-50.
<https://doi.org/10.1007/s00284-014-0681-x>
- [6] Lu, X.M. and Lu, P.Z. (2014) Characterization of Bacterial Communities in Sediments Receiving Various Wastewater Effluents with High-Throughput Sequencing Analysis. *Microbial Ecology*, **67**, 612-623.
<https://doi.org/10.1007/s00248-014-0370-0>
- [7] Maugeri, T.L., Gugliandolo, C. and Lentini, V. (2013) Diversity of Prokaryotes at a Shallow Submarine Vent of Panarea Island (Italy) by High-Throughput Sequencing. *Atti della Accademia Peloritana dei Pericolanti*, **91**, A1-1-A1-9.
- [8] 王鹏, 陈波, 张华. 基于高通量测序的鄱阳湖典型湿地土壤细菌群落特征分析[J]. 生态学报, 2017, 37(5): 1650-1658.
- [9] 徐瑛, 孙永明, 郑涛, 等. 高通量测序技术辅助筛选脱硫菌[J]. 化工学报, 2014(5): 1808-1814.
- [10] 何志刚, 刘慧屿, 刘艳, 等. 基于高通量测序技术筛选低温秸秆降解菌群的研究[J]. 山西农业大学学报(自然科学版), 2021, 41(6): 75-84.
- [11] Al Rwahnih, M., Daubert, S., Golino, D., *et al.* (2009) Deep Sequencing Analysis of RNAs from a Grapevine Showing Syrah Decline Symptoms Reveals a Multiple Virus Infection That Includes a Novel Virus. *Virology*, **387**, 395-401.
<https://doi.org/10.1016/j.virol.2009.02.028>
- [12] Kreuze, J.F., *et al.* (2009) Complete Viral Genome Sequence and Discovery of Novel Viruses by Deep Sequencing of Small RNAs: A Generic Method for Diagnosis, Discovery and Sequencing of Viruses. *Virology*, **388**, 1-7.
<https://doi.org/10.1016/j.virol.2009.03.024>
- [13] Adams, I.P., Glover, R.H., Monger, W.A., *et al.* (2009) Next-Generation Sequencing and Metagenomic Analysis: A Universal Diagnostic Tool in Plant Virology. *Molecular Plant Pathology*, **10**, 537-545.
<https://doi.org/10.1111/j.1364-3703.2009.00545.x>
- [14] 张文力. 高通量测序数据分析现状与挑战[J]. 集成技术, 2012, 1(3): 20-24.
- [15] 席建忠, 马明, 王干诚, 等. 高通量 siRNA 筛选技术发展展望[J]. 生命科学, 2014, 26(3): 276-286.