

数字化驱动合成生物学的未来： 从研发到生产的智能化路径（上）

王少鹏

柯尔柏医药科技

摘要：作为生物经济的核心驱动力，合成生物学正借助数字化技术的力量，从实验室研发阶段迅速向工业化生产迈进。本文聚焦于制造执行系统、工艺仿真模型、数字孪生模型和数据分析平台在合成生物学中的关键作用，探讨如何利用这些工具优化研发流程、简化技术转移过程、缩短产品上市周期、提高生产效率，并确保质量可追溯性和合规性。本文作者计划分上下两部分介绍数据如何助力合成生物学公司在商业化生产前后提高效率并降低风险。

关键词：数字化技术；合成生物学；智能化路径

当前，合成生物学企业面临诸多挑战，包括如何高效推进工艺开发、如何快速将实验室的创新技术转化为商业化生产、以及如何持续通过数据打通、整合与应用获取收益。这些挑战涉及繁杂的研发工作、复杂的工艺参数优化、大规模生产的可行性验证以及数据的一致性管理。此外，技术转移过程中还面临

着研发到生产阶段的数据衔接难题，以及高复杂性、长周期所带来的成本和时间压力。

在此背景下，数字化工具成为了解决这些难题的关键。制造执行系统、工艺仿真模型和数字孪生模型等数字化工具，通过连接研发与生产的各个环节，帮助企业优化技术转移流程，确保

关键工艺参数的稳定性，进而降低失败风险。这些工具的结合不仅能提升生产效率，还能实现从研发到生产的全面可追溯性，为合成生物学企业在快速发展的行业中保持竞争力提供有力支持。

1 合成生物学研发中的数据管理与工艺优化

合成生物学作为一个数据密集型的领域，涉及从结构设计到生产工艺开发的多重复杂流程。研发阶段需要高效整合实验数据，并通过反复迭代实现目标产品的高效表达。

为此，面对大量研发批次的数据，数据分析平台和工艺仿真模型发挥了重要作用。它们能够自动化地完成数据的收集、整理和分析工作，显著提高了研发效率，降低了传统人工操作的复杂性和错误风险，从而既提升了效率，又减少了时间和成本浪费。

1.1 自动化数据处理与分析

在合成生物学研发过程中，每个



图 1 数据协作管理平台可以助力合成生物学研发团队实现可视化分析

实验批次都会产生海量的数据，包括培养条件、代谢产物浓度、基因表达水平等。然而，这些数据往往分散在多个独立的系统中，如电子实验记录本(ELN)、实验室信息管理系统(LIMS)和 Excel 文件中，缺乏统一的标准化格式，导致团队难以高效地整合和分析这些数据。

为了解决这个问题，数据分析平台应运而生。这些平台能够自动收集实验数据，并通过预设的标准化流程实现数据的自动整理，将实验结果与各种变量清晰地关联起来。随后，平台会基于工艺模型对数据进行深入的统计学分析，帮助研发团队快速识别出影响目标产物产量和质量的关键因素。

以某酶产品的生产为例，它涉及多达 50 余种培养条件（如温度、pH 值、溶氧水平等）和数百种代谢产物变量的组合关系。传统的人工分析方式不仅费时费力，而且容易遗漏关键变量。在优化其生产工艺时，柯尔柏的用户研发团队利用数据平台 PAS-X Savvy 对数百种酶的动力学参数进行了全面的分析，成功筛选出了最佳的培养条件组合。这一举措不仅显著提升了目标产物的产量，还有效减少了副产物的生成。

1.2 数据标记与可视化分析

采用传统的分析方式，每轮实验的设计与评估均需要耗时数周，极大地拖慢了项目进度。为了显著提高分析效率，数据平台如 PAS-X Savvy 发挥了重要作用。该平台能够对不同实验数据进行清晰的标记，并以图表或图谱的形式直观地展示变量间的关系，即便是人工难以快速处理的多变量分析，也能通过计算机化系统迅速完成。这种数据标记与可视化功能使研发团队能够迅速识别变量与目标产物之间的相关性，从而为优化工艺设计提供了有力依据。

1.3 全球团队协作与结果共享

在跨国企业或跨国研发项目中，数据分析平台通过支持全球团队实时共享分析结果和模型仿真数据，显著提升了

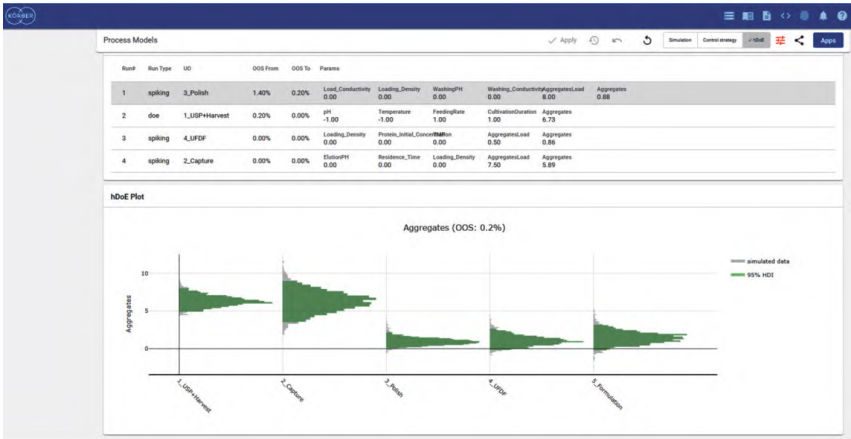


图 2 PAS-X Savvy 工艺模型的整体实验设计 (hDoE) 算法预测了最佳工艺，通过在需要的地方有效地分配资源显著减少生产结束后的 OOS 事件

团队协作效率。例如，当一组实验完成后，全球各地的分析团队能够同时访问这一统一的数据平台，基于相同的实验数据进行模型仿真和统计学分析。这样一来，团队便能快速评估优化策略的有效性。这种共享模式不仅缩短了实验决策的周期，还有效避免了低效的图片或 excel 邮件往来，以及重复实验和资源浪费的问题。

1.4 基于模型的仿真与工艺表征优化

数据分析平台与工艺仿真模型的结合，为评估不同参数对产品质量和产量的影响提供了直观的支持。在合成生物学的工艺优化过程中，参数调整往往会对后续流程产生链式影响，而常规方法难以全面量化这些复杂的关联关系。幸运的是，通过集成工艺模型和蒙特卡洛模拟，这些问题得到了系统化的解决。

集成工艺模型 (IPM) 通过整合细胞培养、纯化等多个环节的模型，实现了全流程的数字化仿真。例如，细胞培养过程中的溶氧水平不仅会影响代谢路径，还可能改变纯化阶段的杂质比例。通过模型的串联，这些链式影响能够被提前量化与预测。

作为平台的重要分析工具，蒙特卡洛模拟通过随机生成大量参数组合，帮助研发团队更全面地评估变量对整个流

程的系统性影响。例如，在某产品的生产过程中，蒙特卡洛模拟被用于分析温度变化对代谢速率的影响，以及这种变化如何逐步传递到纯化效率的变化中。通过大规模模拟不同参数组合，团队能够预测全流程的可能结果，并提前识别出潜在的风险。

此外，蒙特卡洛模拟还支持多目标优化，帮助研发团队在多个目标之间找到最佳的平衡点。例如，在调整溶氧水平和 pH 值的组合时，模拟会评估这些调整对提高产品产量和减少副产物生成的综合影响。这种基于数据驱动的优化方法不仅提升了合成生物学的研发效率，还大幅降低了工艺开发中的试错成本，为工艺设计提供了坚实的科学依据。

2 结论

在合成生物学的研发领域，使用合适的数字化工具，企业不仅能够有效应对复杂的工艺开发和高失败率的挑战，还能借助实时数据分析和模拟预测，实现产品性能的提升和资源利用效率的最大化。

在后续篇章中，作者将详细阐述合成生物学中技术转移的关键要素、商业化生产的合规性与实时优化策略，以及数据驱动的全生命周期管理方法，敬请读者持续关注。