

面向大肠杆菌和酿酒酵母标准化构建、基因组整合与定量测试的工具包

谢荣辉^{a,1}, 王然^{a,1}, 詹依^{a,c}, 魏博妍^a, 徐思杨^a, 陈业^{a,b,*}, 刘莉莉^{a,d,*}

^a 中国科学院深圳先进技术研究院深圳合成生物研究所, 广东省深圳市南山区深圳大学城学苑大道 1068 号, 518055

^b 中国科学院大学, 北京市怀柔区雁栖湖东路 1 号, 北京 101408

^c 中国农业大学生物学院, 北京市海淀区圆明园西路 2 号, 100193

^d 深圳理工大学合成生物学院, 广东省深圳市光明区公常路 1 号, 518107

¹ 本文共同第一作者。

*通讯作者。

电子邮箱: liulili@suat-sz.edu.cn (刘莉莉); chenye.h.ai@gmail.com (陈业)

摘要

现有合成生物学工具包扩展了遗传元件库和组装能力。然而, 新使用者在构建体定义、层级转换、基因组整合路径和定量测试流程方面仍可能面临不明确之处。本研究提出 **Polaris**, 一套适用于大肠杆菌和酿酒酵母的 **Golden Gate** 工具包。**Polaris** 定义了两种宿主共享的三级层级体系: **Level 1** 储存可复用元件, **Level 2** 生成转录单元及其他中间构建体, **Level 3** 支持多基因组装以及测试骨架与整合骨架之间的切换。该工具包区分元件接口和转录单元接口, 在连续的组装阶段分别使用 **Bpil** 和 **Bsal**, 并提供 **ccdB** 和 **mScarlet13** 两种并行筛选方案。我们在两种宿主中建立了整合模块和定量表征框架, 并使用代表性遗传元件以及激活和抑制系统进行了验证。配套数字平台进一步编码了 **Level 1** 至 **Level 3** 的组装规划规则。**Polaris** 将层级化构建、稳定遗传实现和定量测试整合于统一执行框架中, 为数据积累、元件标准化和未来自动化提供了实用基础。

关键词:

合成生物学工具包; **Golden Gate** 组装; 遗传元件标准化; 层级化 **DNA** 组装; 基因组整合; 定量表征

Polaris: A toolkit for standardized construction, integration, and quantitative testing in *Escherichia coli* and *Saccharomyces cerevisiae*

Abstract

Existing synthetic biology toolkits have expanded part libraries and assembly capabilities. However, new users may still encounter ambiguity in construct definitions, hierarchical transitions, genomic integration routes, and quantitative testing procedures. Here, we present Polaris, a Golden Gate toolkit for *Escherichia coli* and *Saccharomyces cerevisiae*. Polaris defines a shared three-level hierarchy. Level 1 stores reusable parts, Level 2 generates transcription units and other intermediate constructs, and Level 3 supports multigene assembly and switching between testing and integration backbones. The toolkit distinguishes part interfaces from transcription unit interfaces, uses Bpil and Bsal in successive assembly stages, and provides parallel ccdB and mScarletI3 screening options. Integration modules and quantitative characterization frameworks were implemented in both hosts and demonstrated using representative genetic parts and activation and repression systems. A companion digital platform further encodes the Level 1 to Level 3 rules for assembly planning. Polaris connects hierarchical construction, stable genetic implementation, and quantitative testing within a common execution framework, providing a practical basis for data accumulation, part standardization, and future automation.

Keywords: Synthetic biology toolkit; Golden Gate assembly; Genetic part standardization; Hierarchical DNA assembly; Genomic integration; Quantitative characterization

1 引言

合成生物学的工程化实施依赖于遗传元件的标准化、模块化与可复用组装。基于 Type IIS 限制性内切酶的 **Golden Gate** 及其发展出的模块化克隆 (**MoClo**) 体系, 已成为当前标准件储存、层级装配与自动化执行的重要基础框架^[1-7]。该类方法能够将基础元件以统一格式入库, 再按预定义接口完成一锅式装配, 并已在酵母工具包、**CIDAR MoClo** 和 **EcoFlex** 等体系中证明其在不同宿主中的通用性^[3-5]。随着高通量构建和设计—构建—测试循环不断加快, 这类层级化方案也逐渐成为元件复用、自动化组装、结果积累、后续数据库化管理的共同技术底座^[2,8-11]。在底盘选择上, 大肠杆菌和酿酒酵母分别代表原核和真核中最成熟、最易获得、也最便于多数实验室稳定操作的模式系统^[3,5,7]。二者既覆盖了两类最基本的表达与调控框架, 也拥有完备的遗传操作与测试手段, 因此适合作为建立标准化构建与定量测试流程的共同基础, 并有利于不同实验室之间的方法对齐和结果比较。尤其对于需要快速建立工作流程的新实验室而言, 以这两个底盘为共同基础, 更容易形成可直接复制和逐步扩展的实验体系。

随着组装体系日益成熟, 工程问题已不再只是“能否拼接成功”, 而转向“对象是否定义明确、流程是否足够低歧义、结果是否能够直接比较”。对实际使用者而言, 真正消耗学习成本的往往不是组装反应本身, 而是元件属于哪一级、接口如何划分、不同酶和骨架分别承担什么职责、构建完成后如何进入稳定背景以及如何进入测试框架。已有工具包在元件扩展、宿主适配和复杂系统构建方面提供了丰富资源, 并通过标准化操作和最佳实践显著简化了部分宿主的工程化流程^[1-7]; 然而, 在跨宿主统一构建路径、整合入口和标准化测试对象定义方面, 仍缺少面向初始用户的统一实现。

标准化测试进一步强化了这一要求。可比较的对象并不只是某段序列本身, 而是处在固定接口、固定骨架、固定宿主与固定测量条件下的性能对象。**RPU** 等参考框架说明, 只有在统一参考元件和读出逻辑下, 定量数据才更具有可比性^[12-14]; 而关于遗传上下文和绝缘元件的研究也表明, 局部序列连接关系、宿主背景与环境条件都会持续影响系统行为^[15,16]。因此, 真正有工程价值的工具包不应只提供组装能力, 还应同时给出清楚的基因组整合路径和标准化测试入口^[11,17,18], 使构建对象能够自然过渡为可测、可比的性能对象。

基于这一需求, 本研究构建了一套同时适用于大肠杆菌和酿酒酵母的双宿主 **Golden Gate** 工具包。该体系采用统一的一级至三级层级架构, 将基础元件存储、转录单元组装、

多转录单元构建、基因组整合和标准化测试纳入同一工作流；通过区分元件接口与转录单元接口、分层使用 *Bpil* 和 *Bsal*^[2,5,6]，并结合 *ccdB* 反筛及其平行红色荧光版本^[19-21]，降低构建和判读的学习负担，同时提高不同实验室条件下的可执行性。本文重点说明这套体系如何使使用者能够直接进入元件表征、整合构建和系统测试，并为后续元件标准化、数据库建设和自动化衔接提供基础。

2 材料和方法

2.1 菌株和质粒

本研究用到的大肠杆菌克隆菌株为 DH10B，DH10B-Alt (*attB::lacI,tetR,araC*)，DH10B-12S^[22]，DB3.1。大肠杆菌系统测试菌株为 MG1655-Alt (*attB::lacI,tetR,araC*)，ECint034^[23]。酿酒酵母测试菌株为 BY4741 和 CYE72。测试质粒及菌株详细信息见附表 2 和 3。

2.2 培养基和试剂

LB 液体培养基 (Tryptone 10 g/L, Yeast extract 5 g/L, NaCl 10 g/L)；LB 固体培养基 (Tryptone 10 g/L, Yeast extract 5 g/L, NaCl 10 g/L, agar 20 g/L)；2×YT (Tryptone 16 g/L, Yeast extract 10 g/L, NaCl 5 g/L)；TSS 溶液 (85% 液体 LB 溶液，10% PEG (wt/vol, MW8000)，5% DMSO (vol/vol)，50 mM MgCl₂ (pH6.5)，以上 LB 液体、MgCl₂ 为高温灭菌，PEG、DMSO 为过滤灭菌)；5×KCM (0.5M KCl, 0.15M CaCl₂, 0.25M MgCl₂)；M9 培养基添加 0.4% 甘油、1% 的 100× 微量元素溶液和 0.2% Casamino acids; 100× trace element solution (1.0 g/L FeSO₄·7H₂O, 1.2 g/L MnCl₂·4H₂O, 0.8 g/L CoCl₂·6H₂O, 0.7 g/L ZnSO₄·7H₂O, 0.3 g/L CuCl₂·2H₂O, 20 mg/L H₃BO₃, 0.25 g/L (NH₄)₆Mo₇O₂₄·4H₂O, 5.0 g/L EDTA)；SD/-Leu(Coolaber, PM2201)；SD/-Ura(Coolaber, PM2271)；SD/-Ura-Leu (Coolaber, PM22991)；MSG-Ura(Coolaber, T2001)；SD-All 培养基(Coolaber, PM2101)；琼脂(Bacto, 214010)。抗生素工作浓度为：Kanamycin(50 µg/mL，阿拉丁，K103024)，Chloramphenicol(25 µg/mL，生工生物，A100230)，Gentamicin(10 µg/mL，BBI，A620217)，Carbenicillin (100 µg/mL，Macklin/麦克林，C805408-10g)，G418 (200 mg/L，Coolaber，CG5471)。测试使用到的诱导剂信息见附表 4。

2.3 质粒构建-Gibson 组装

骨架质粒 (Backbone) 均采用 Gibson 组装方式构建 (CU201，全式金生物)。线性化载体通过限制性内切酶或 PCR 扩增获得。各组装片段与线性载体存在 15-25 nt 的重叠区

域。反应体系为包括 5 μ L 2 $\times$ Basic Assembly Mix, 线性化载体 5-100 ng (推荐 0.03 pmol), 各插入片段为 0.03-0.06 pmol (按载体与插入片段摩尔比 1:1 至 1:2 加入), ddH₂O 补加至总体积 10 μ L, 轻轻混匀后, 置于 50 $^{\circ}$ C 反应 5-15 分钟。反应结束后产物可直接用于转化或 -20 $^{\circ}$ C 保存。

2.4 质粒构建-Golden Gate 组装

标准反应混合液包含一定体积 DNA 片段 (各片段加入量按比例调整, 插入片段与载体骨架的摩尔比控制在 2:1 左右, Backbone 使用约 0.05 pmol), Bpil (Thermo Scientific, ER1012) 或 Bsal-HF $\text{v}2$ (NEB, R3733L) 0.5 μ L, T4 DNA 连接酶 (NEB, M0202L) 0.5 μ L, 10 $\times$ T4 DNA 连接酶缓冲液 1.0 μ L, 其余用无核酸酶水 (dd H₂O) 补足到 10 μ L。反应按以下程序进行: (37 $^{\circ}$ C 2 min, 16 $^{\circ}$ C 2 min) 循环 25–35 次 (当片段数量超过 6 个或进行复杂构建时, 循环条件优化为: (37 $^{\circ}$ C 5 min, 16 $^{\circ}$ C 5 min) 循环 35 次); 循环结束后, 于 50 $^{\circ}$ C 孵育 15 min, 以促进残余未切或错误连接模板的再切割; 若不立即进行转化, 则 80 $^{\circ}$ C 孵育 20 min 使内切酶彻底失活。

2.5 大肠杆菌感受态制备

将菌株划线于 LB 平板, 37 $^{\circ}$ C 过夜培养。挑取单菌落于 2 mL LB 液体培养基中, 37 $^{\circ}$ C、220 rpm 振荡过夜培养。取 1 mL 菌液转接至 100 mL LB 培养基 (500 mL 锥形瓶), 37 $^{\circ}$ C、220 rpm 振荡培养 2–3 h 至 OD₆₀₀ 为 0.30-0.35。将锥形瓶冰浴 20 min, 4 $^{\circ}$ C、1000 g 离心 15 min 收集菌体。离心弃上清后, 用 4 mL 预冷 TSS 重悬; 再次离心并重复 TSS 重悬离心一次。最后, 用 4 mL 预冷的 TSS 再次轻轻重悬菌体, 以每管 100 μ L 感受态细胞分装于 0.2 mL 离心管中, -80 $^{\circ}$ C 保存。

2.6 大肠杆菌化学转化法

取出感受态细胞于冰上融化, 将适量感受态加入 1.5 mL EP 管中 (TSS 感受态需补加 5 $\times$ KCM 至终浓度 1 $\times$)。加入质粒或组装产物 (不超过感受态体积的 10%), 轻弹混匀 4-5 次, 冰浴 30 min。42 $^{\circ}$ C 水浴热激 30s, 立即冰浴 2 min。加入 5-10 倍感受态细胞体积的 2 $\times$ YT 培养基, 37 $^{\circ}$ C、220 rpm 培养 1 h。600-800 g 离心 5 min, 弃部分上清, 留约 100 μ L 重悬菌体, 涂布于含相应抗生素的 LB 平板, 37 $^{\circ}$ C 过夜培养。其中 Backbone 质粒需转化至 DB3.1 克隆底盘 (携带 gyrA462 基因), 非 Backbone 质粒转化至 DH10B 系列菌株。

2.7 大肠杆菌基因组整合

含 attP 的 Backbone 经 Golden Gate 组装后, 使用 Bsal (或 AvrII) 和 SpeI 双酶切去除 pSC101 复制子, 通过 0.8% 琼脂糖凝胶电泳分离并回收目的片段。将回收片段与 T4 DNA

连接酶于 25 °C 连接 2 h（或 16 °C 过夜），65 °C 灭活 15 min；随后加入 BsaI（或 AvrII）和 SpeI 于 37 °C 酶切 30 min，80 °C 灭活，得到环化产物。采用 TSS 法制备底盘菌株感受态细胞，转化温敏型质粒 pEcint01，于 30 °C 培养获得转化子。挑取单菌落接种于含羧苄青霉素的 LB 液体培养基中，30 °C 振荡培养过夜，按 1:100 稀释至新鲜 LB（含羧苄青霉素），30 °C 培养至 OD₆₀₀≈0.3-0.35，迅速转移至 37 °C 诱导 15 min。TSS 法再次制备感受态细胞，并转化上述环化产物，于 30 °C 培养获得转化子。菌落 PCR 验证转化子的基因组整合情况。挑取单个已验证的菌落，接种至无抗生素的 LB 培养基中，于 37 °C 下振荡过夜培养，以消除温度敏感型质粒 pEcint01

将成功整合的菌株制备感受态细胞，转化 pCP20 质粒（30 °C 培养）。挑取单菌落划线于羧苄青霉素平板，30 °C 培养 3-6 h，经 PCR 验证卡那霉素抗性基因（KanR）的消除。选取 KanR 阴性克隆接种于无抗 LB，37 °C 培养 2-3 h，分别划线于无抗和羧苄青霉素平板，37 °C 过夜培养。在羧苄青霉素平板上无法生长的克隆表明 pCP20 质粒已被成功消除。

2.8 大肠杆菌元件和转录调控系统测试

在 96 孔板（Costar 3799）中，每孔加入 200 μL LB 培养基（无抗生素或相应抗生素或诱导剂），挑取单克隆进行接种，置于恒温振荡器（Biosan, PST-60HL-4）中，37 °C、800 rpm 条件下过夜培养。随后，在另一 96 孔板中每孔加入 100 μL M9 培养基（无抗生素或添加相应抗生素；若实验体系需诱导表达转录因子，则含相应诱导剂诱导其表达），并取 1 μL 过夜培养菌液转移至其中，于 37 °C、800 rpm 培养 2.0 h（第一阶段培养）。之后，每孔补加 100 μL M9 培养基（无抗生素或添加相应抗生素，转录因子诱导条件同第一阶段；如转录因子需加对应配体分子，则在此阶段添加），继续于 37 °C、800 rpm 培养 3.0 h（第二阶段培养）。最后，使用酶标仪（Tecan Infinite 200 Pro）测定 96 孔板中各孔的 OD₆₀₀ 值及相应荧光值。

2.9 酿酒酵母转化

经典 HR 整合方法中，选取 5μg 质粒，用 BsaI 酶切消化后，将反应产物直接转化到制备好的感受态细胞，CRISPR/Cas9 整合方法中，除以上酶切消化产物外，还需要 Cas9 质粒（pCfB2312）与 sgRNA 质粒（sgRNA+pScCom02）各 200ng。CRISPR/Cas9 整合方法的酵母菌经验证插入正确后，需要经过消去，将 sgRNA 质粒去除。将过夜培养的酵母菌在 10000X 稀释涂布在 YPD+5-FoA（1mg/mL）（HARVEYBIO, 5-FoA）平板上，于 30°C 倒置培养 2 天得到转化子。挑取长出的菌落，分别在 YPD+G418（200μg/mL），液体培养基 SD/-Ura，及划线 SD/-Ura 板培养，2 天后无法长出则说明 Cas9 以及 sgRNA 质粒已经

丢失，保菌，后续实验待用。

使用经典酵母感受态制备与转化试剂盒（冻存型）（Coolaber, SK2400）进行酿酒酵母感受态细胞制备和酵母转化。将转化产物涂布于相应的缺陷型 SD 平板上，于 30 °C 倒置培养 2 天得到转化子。或使用 Frozen EZ Yeast Transformation Kit II 试剂盒（Zymo Research Corporation, T2001）进行酿酒酵母感受态细胞制备和酵母转化。将转化产物涂布于相应的缺陷型 SD 平板上，于 30 °C 倒置培养 2 天得到转化子。

2.10 酿酒酵母元件和转录调控系统测试

对于组成型以及诱导型启动子元件表征，挑取单菌落接种至含 500 μ L 选择性 SD 培养基的 96 孔深孔板（Yderbio, YDASQ01-S）中，于 30 °C、800 rpm 振荡培养 24 h。将培养物按 1:200 比例转接至添加诱导剂的新鲜 SD 培养基，于 30 °C、800 rpm 继续培养 16 h。取 20 μ L 菌液与 180 μ L 磷酸盐缓冲液（PBS, Proteintech, PR20014）加入 96 孔 U 底板中混匀，制备成流式待测样品。对于终止子元件表征，挑取待测转化子至含 10 mmol·L⁻¹ D-木糖的相应缺陷型培养基中，培养 24 h 后按 1:200 比例转接至含 10 mmol·L⁻¹ D-木糖的 SD 培养基，于 30 °C、800 rpm 培养 16 h 后取样用于流式分析，样品制备过程与启动子元件表征一致。对于转录因子定量测试，挑取单菌落接种至含 500 μ L 缺陷型 SD 培养基的 96 孔深孔板中，于 30 °C、800 rpm 振荡培养 24 h 后，按 1:200 转接至含梯度浓度 TF 表达诱导剂的新鲜 SD 培养基，并分别设置不添加或添加满诱导浓度 TF 效应物的非诱导组和诱导组；继续培养 16 h 后按上述方法进行流式检测。结合输入模块获得的 Input RPU 与报告模块的 Output RPU 绘制 Input RPU–Output RPU 曲线，确定两组输出差异最大的 TF 输入条件；随后固定该 TF 输入强度，设置 TF 效应物浓度梯度，按相同流程培养和检测。

2.11 酶标仪读值（适用于大肠杆菌系统的信号测定）

用酶标仪（Tecan Infinite 200 Pro）对每个样品孔的 OD₆₀₀ 和对应荧光通道进行读取，荧光需要选择合适增益（Gain），以获得合适数值。酶标仪参数按照：mTagBFP2: 420/455nm; sfCFP: 433/470 nm; sfGFP: 480/510 nm; sfYFP: 500/530 nm; mCherry: 580/610 nm 设定。每个样品的初始数值分析处理按照以下公式进行：

$$\text{Flu}/\text{OD}_{(\text{样品})} = \frac{\text{Flu}_{(\text{样品})} - \text{Flu}_{(\text{空白对照})}}{\text{OD}_{(\text{样品})} - \text{OD}_{(\text{培养基})}} \quad (\text{A.1})$$

Flu_(样品) 为测试样品的酶标仪荧光读值，Flu_(空白对照) 为空白对照细胞（无荧光对照菌株）的酶标仪荧光读值，OD_(样品) 为测试样品的酶标仪 OD₆₀₀ 读值，OD_(培养基) 为测试培养基的酶标仪 OD₆₀₀ 读值。

2.12 荧光蛋白扫描

荧光蛋白光谱特性使用 Infinite 200 PRO E Plex 型多功能微孔板检测仪 (Tecan) 进行表征。在采集荧光前立即测定各培养物的 OD₆₀₀ 以监测细胞密度, 随后在同一孔内进行光谱测量。在大肠杆菌的初步表征中, 对 mTagBFP2、sfCFP、sfGFP、sfYFP、mCherry 和 mScarletI3 进行了配对激发/发射波长扫描。激发波长在 400 – 650 nm 的蛋白适宜范围内以 5 或 10 nm 步长变化, 每个激发波长下的发射波长均设为激发波长加 30 nm ($E_m = E_x + 30 \text{ nm}$)。利用所得的信背比 (SBR) 生成配对激发/发射扫描曲线。为优化 sfCFP 和 mTagBFP2 的激发和发射设置, 还进行了额外的顺序扫描。对于 sfCFP, 先将激发波长固定在 430 nm, 发射波长从 460 nm 扫描至 550 nm, 步长为 5 或 10 nm; 随后将发射波长固定在 470 nm, 激发波长从 400 nm 扫描至 440 nm, 步长为 3、5 或 10 nm。在发射波长固定为 475 nm 的条件下, 以 5 或 10 nm 步长从 400 nm 到 445 nm 进行了第二次激发扫描。对于 mTagBFP2, 将激发波长固定在 420 nm, 发射波长从 450 nm 扫描至 520 nm, 步长为 5 或 10 nm; 随后将发射波长固定在 455 nm, 激发波长以 5 nm 步长从 400 nm 扫描至 425 nm。在酿酒酵母的光谱表征中, 对 mTagBFP2、mCitrine、mNeonGreen、mScarletI3 和 mRuby2 进行了配对激发/发射波长扫描, 波长范围覆盖各蛋白的特定区间 (255 – 700 nm), 步长为 5 或 10 nm。mTagBFP2 和 mRuby2 采用固定的 50 nm 激发-发射间隔, 而 mCitrine、mNeonGreen 和 mScarletI3 则采用固定的 30 nm 间隔。对于每个波长设置, 均使用相同的波长和增益设置测定对应的非荧光宿主对照。增益根据需要调整, 以使荧光读数保持在可测量范围内, 但在每个波长组合下, 荧光样品及其对应的背景对照使用完全相同的设置。利用样品信号和背景对照信号计算每个波长组合下的 SBR。

2.13 流式细胞术操作及分析

荧光测量使用配备高通量进样器 (HTS, BD Biosciences) 的 FACSCelesta 流式细胞仪进行。仪器设置针对酿酒酵母和大肠杆菌分别优化。对于酿酒酵母, 调整前向角散射 (FSC)、侧向角散射 (SSC) 和荧光检测器电压, 同时监测 FSC-A、SSC-A 及荧光信号, 使主细胞群体位于检测范围内的合适区域。对于大肠杆菌, 调整 SSC 和荧光检测器电压, 同时监测 SSC-H 和荧光信号, 使细菌群体位于合适的检测范围内。采集时为抑制低水平背景噪声, 对酿酒酵母和大肠杆菌均应用了 300 的 SSC 阈值。每个样品至少收集 10,000 个单细胞事

件。酿酒酵母的荧光信号记录于 BV421-A、BV510-A、FITC-A 和 PE-Texas Red-A 通道，大肠杆菌的荧光信号记录于 BV421-H、BV510-H、FITC-H 和 PE-Texas Red-H 通道。使用 FlowJo 软件（版本 10；BD Biosciences）分析数据，并提取每个设门细胞群体的中位荧光强度。经背景校正的荧光值以任意单位（a.u.）报告，计算方法为从测得的中位荧光强度中减去相应的背景自发荧光。

2.14 软件实现与源代码可用性

配套的 Polaris Kit Assembly 平台依据本工具包定义的 Level 1-3 规则开发，用于支持遗传元件和骨架选择、兼容性检查、层级化构建体生成、质粒图谱可视化以及组装路径管理。部署的网络平台可通过 <https://infrasynbio.siat.ac.cn/kitassembly> 访问。Polaris Kit Assembly 的源代码及相关资源可通过 <https://github.com/ChenYe-Lab/PolarisKitAssembly> 获取。

3 结果

3.1 面向标准化构建的工具包层级设计

本工具包沿用 Golden Gate 的层级化组装框架^[1-7]，但设计重点并不在于继续增加层级或扩展更多功能分支，而在于将双宿主中的构建对象、层级边界和筛选逻辑预先定义清楚，使使用者无需先开发一套定制化体系，便可直接开展元件测试、基因组整合和调控系统构建。为此，我们采用统一的一级至三级架构：一级用于存储可复用基础元件，二级作为构建中间态，用于形成单个转录单元（TU）或其他可继续升级的模块，三级则对应最终应用载体，可容纳多个转录单元，并可通过预定义接口在不同骨架之间高效切换，从而完成质粒测试载体、基因组整合载体以及不同宿主或复制子背景之间的转换。考虑到元件表征通常不需要经过完整多转录单元升级，我们还提供了面向标准测试的“元件直通三级”设计，以进一步降低测试构建的复杂度。

在接口设计上，本工具包沿用已有兼容的 scar 框架（表 1），同时进一步强化了元件上下游边界的定义。元件接口 scar 与转录单元接口 scar 被明确区分：前者服务于启动子、核糖体结合位点、编码序列和终止子等基础元件的顺序拼接，后者服务于多个转录单元在更高层级中的组合。关键连接位点不仅承担装配功能，也承担表达框架定义功能。例如，A2 位于启动子下游并贴近转录起始区，B 以 AATG 将起始密码子直接纳入连接语法^[2,4,6]，C 以 TAAA 将终止密码子 TAA 纳入标准边界，从而使元件对象的起始和终止位置在组装阶段即

被固定（补充说明 1）。由此，scar 不再只是连接片段的短序列，而成为元件对象和模块对象的显式边界定义。为降低上游模块对下游测试构型的影响，部分二级骨架还在转录单元接口与元件接口之间引入中性间隔序列或终止/隔离序列，从而减弱遗传上下文干扰^[15,16]（图 1）。

Scar	Sequence(5'→3')	Scar	Sequence(5'→3')	Scar	Sequence(5'→3')
A	GTGC	I	AGGT	R	AAAG
A2	ATCA	J	CACC	S	ACAA
B	AATG	K	TGTC	T	ACTC
C	TAAA	L	CGCT	U	AGAC
D	CCTC	M	CCGT	V	CAGG
E	GCTT	N	GCGA	W	CCAC
F	CTGA	O	GGGC	X	CCCA
G	TACG	P	AGTC	Y	CGAG
H	TTCC	Q	ATTG	Z	GTTA

表 1：大肠杆菌与酿酒酵母通用 Golden Gate 组装 scar 序列表。scar 分成两种类型，一类用于元件连接，如 Scar A、A2、B、C 和 D。另一类用于转录单元间连接，如 Scar E、F、GZ。

体系采用 Bpil/Bsal 分层：Bsal 用于最常见的二级组装，Bpil 用于一级进入和三级扩展。已有部分层级化工具包同样在不同层级间交换 Type IIS 酶的职责^[4-6]，因此“交换层级逻辑”本身并不是本体系的主要区别；真正关键的设计选择在于选用 Bpil 而非需要更高反应温度的替代酶。由于 T4 DNA 连接酶在较高温度下活性下降，而二级组装是发生频率最高的步骤，三级多转录单元组装又对整体切割与连接效率更加敏感，因此选择反应条件成熟、应用广泛的 Bsal，并选择最适反应温度为 37°C 的 Bpil 承担 L1 与 L3，从而使整个层级流程更容易维持在与 T4 DNA 连接酶兼容的统一温度窗口内。驯化后的元件首先通过 Bpil 导入一级骨架，正确插入后 Bpil 位点被移除而 Bsal 位点保留；随后由 Bsal 完成二级转录单元组装，并保留 Bpil 位点服务于三级扩展（图 1）。这样的分工使层级升级路径和酶学职责保持明确，也更适合后续自动化装配与元件管理流程的衔接。

在筛选策略上，所有骨架均提供 ccdB 反筛版本，并以 mScarletI3 完整替换 ccdB 构建平行的红色荧光版本（附图 1）。前者通过生长与否直接压低未切骨架、自连和错误连接带来的背景^[19,20]；后者则便于未配备 DB3.1 等 ccdB 传播菌株的实验室进行直观筛选^[21]。两类骨架遵循完全相同的层级关系和接口规则，只是在筛选方式上适配不同实验条件。我们

未采用 *lacZ* 蓝白斑体系，以避免 *X-gal* 使用不便及颜色判读带来的假阳性^[24,25]。综合来看，这些设计共同构成了一条对象定义明确、层级升级清楚、筛选判读直接且适合不同实验室条件的双宿主构建路径。

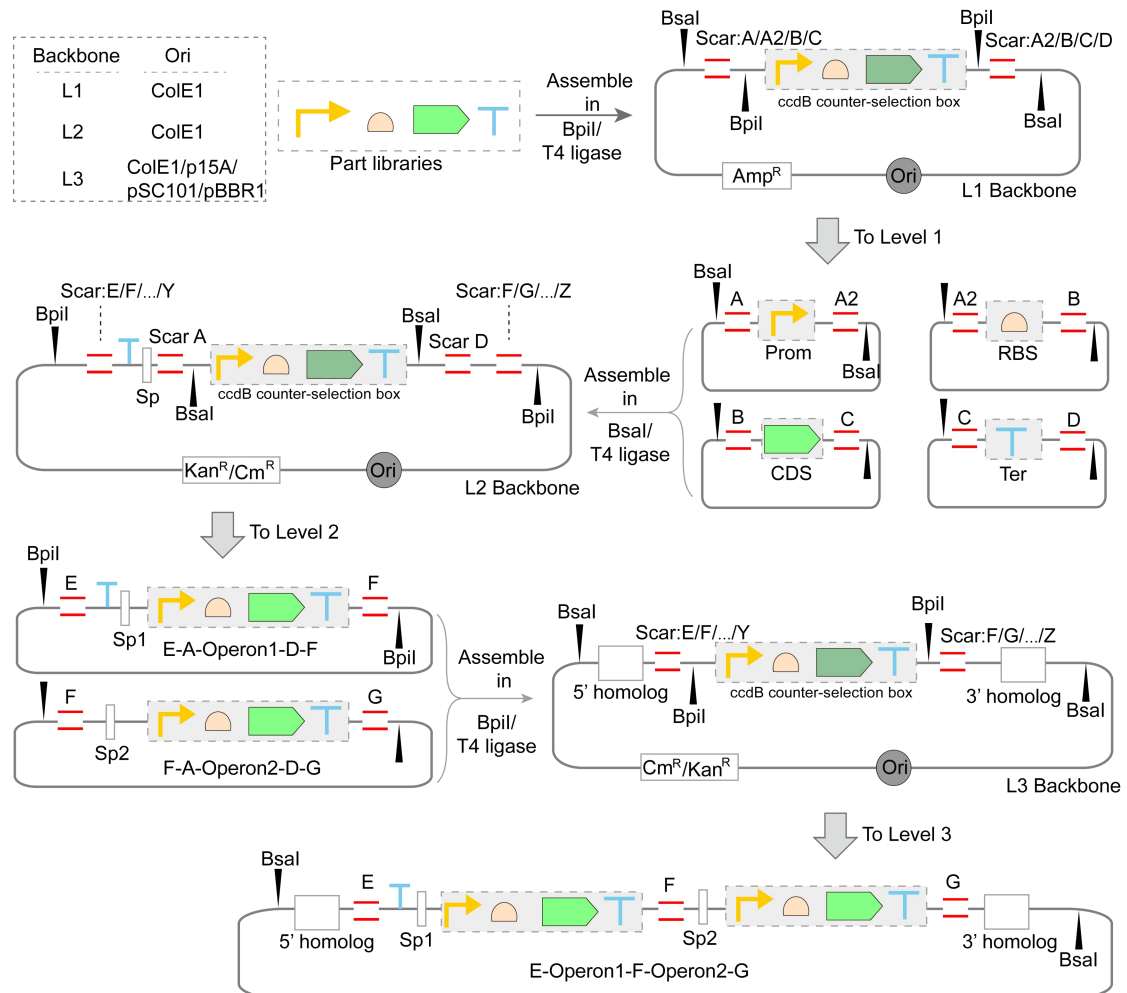


图 1. 层级 Golden Gate 组装。在标准 Golden Gate 工作流程中，首先以驯化的元件库为标准序列，通过 PCR 产生用于 Golden Gate 反应的线性 DNA 片段。随后，通过 *BpiI*/T4 连接酶反应将这些片段组装到相应的 L1 骨架中，生成不同类型的 Level 1 元件质粒。Level 1 元件质粒进一步通过 *BsaI*/T4 连接酶反应组装成 Level 2 质粒（单转录单元），多个 Level 2 质粒随后可通过 *BpiI*/T4 连接酶反应组装成 Level 3 质粒（多转录单元）。此处图示的 L3 骨架代表一个兼容整合的示例，因此含有侧翼同源臂。这些同源臂是骨架特异性的可选特征，而非所有 L3 组装的内在组成部分。相同的 L3 组装逻辑也适用于非整合型 L3 质粒（包括许多大肠杆菌 L3 骨架），只需使用不含同源臂的骨架即可。不同层级的骨架可携带不同的筛选标记和复制起点（补充说明 2 和 3）。为清晰起见，图中仅展示了使用基于 *ccdB* 反筛选的骨架质粒。相应的红色荧光筛选版本（其中 *ccdB* 被替换为 *mScarlet13*，补充图 1）遵循相同的层级架构和组装逻辑。每个层级的质粒均有其对应的骨架载体，不同层级的骨架携带不同的筛选标记或复制起点（补充说明 2、3）。Pro: promoter, Ter: Terminator。

3.2 面向稳定测试背景的双宿主基因组整合模块

为提高构建后表征的稳定性和可比性,本工具包将基因组整合作为双宿主体系中的关键环节。尤其在酿酒酵母中,单拷贝基因组整合有助于降低质粒拷贝数波动和筛选背景差异对定量结果的影响,从而形成更稳定的测试背景^[3];在大肠杆菌中,基因组整合同样有利于获得稳定表达背景,而更高通量的多位点方案则通常需要预先构建含预置整合位点(landing pad)的底盘菌株^[26,27]。基于这一考虑,三级骨架不仅承担多转录单元载体的组装功能,也被设计为构建对象进入稳定基因组背景的统一入口。

在大肠杆菌中,我们建立了基于 λ 噬菌体整合酶的位点特异性整合模块。携带 attP 位点的三级构建经处理后可作为供体 DNA 进入基因组 attB 位点,实现稳定的定点整合^[28]。实验结果显示,该系统在单次转化中产生了超过 10^4 个整合菌落,且所有 16 个受测菌落均显示出与预期整合连接处相符的条带模式(图 2)。除 λ 系统外,本工具包还包含三个正交的整合酶-位点系统(源文献中命名为 Int2、Int5 和 Int7)^[26,27],以及相应的 landing-pad 工具质粒^[29],用于未来的多位点基因组整合。这些扩展系统未在本研究中进行实验验证。此类系统需要先在底盘菌株中预置相应整合位点,再导入目标构建体。

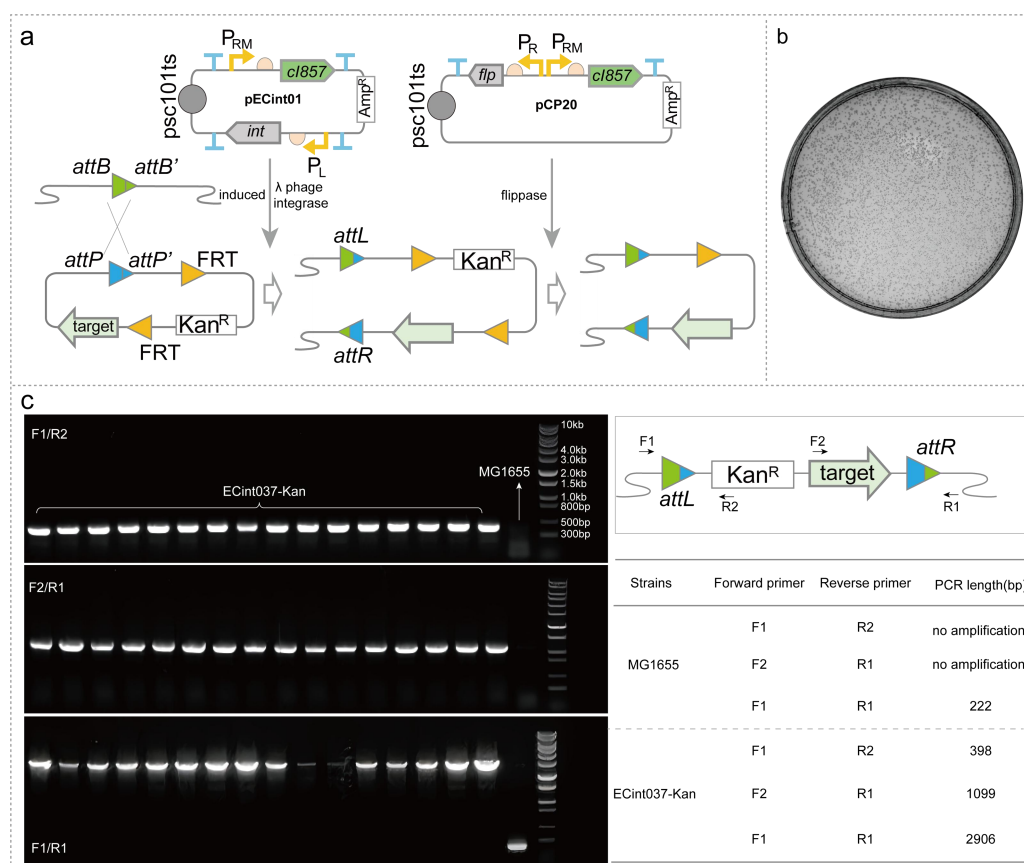


图 2. λ 噬菌体整合酶介导的大肠杆菌基因组整合。(a) λ 噬菌体位点特异性整合与标记去除流程示意图。携带 attP 位点和目标表达盒的供体环状 DNA 在 λ 整合酶作用下与染色体上的 attB 位点发生重组,生成 attL 和 attR 连接。卡那霉素抗性标记随后可通过使用 pCP20 的 FLP/FRT 介导重组被切除。(b) 将约 0.02 pmol 环化供体连接产物导入 50 μ L 大肠杆菌 MG1655 感受态细胞后,单次 λ 整合酶介导

整合实验所得的转化平板。在完整整合方案下，观察到超过 10^4 个抗生素抗性菌落。（c）从（b）所示平板上挑取的 16 个菌落的菌落 PCR 分析。所有 16 个受测菌落均显示出与预期整合连接处相符的条带模式。引物对 F1/R2 和 F2/R1 分别用于检测左、右整合连接，F1/R1 则用于区分野生型 attB 位点和已整合位点。所有 16 个受测菌落的条带模式均与预期的整合连接一致。

在酿酒酵母中，基因组整合模块同时覆盖经典同源重组和 CRISPR–Cas9 辅助整合两条路径^[3, 30, 31]。三级整合骨架预置对应基因组位点的同源臂和筛选标记，使组装完成的构建能够直接进入预定基因组背景；在需要提高位点特异性和整合效率时，同类构建又可与 Cas9/sgRNA 路线配合使用。工具包提供的整合骨架覆盖多个筛选标记和多个染色体位点，其中报告模块和转录因子模块可分别被固定到不同染色体背景中，为后续元件标定和系统测试提供稳定遗传环境（附表 1；补充说明 3）。两类整合路径均能够获得用于后续荧光验证和功能表征的整合株（图 3；附图 2），从而使酿酒酵母中的元件测试摆脱游离质粒带来的额外变异。

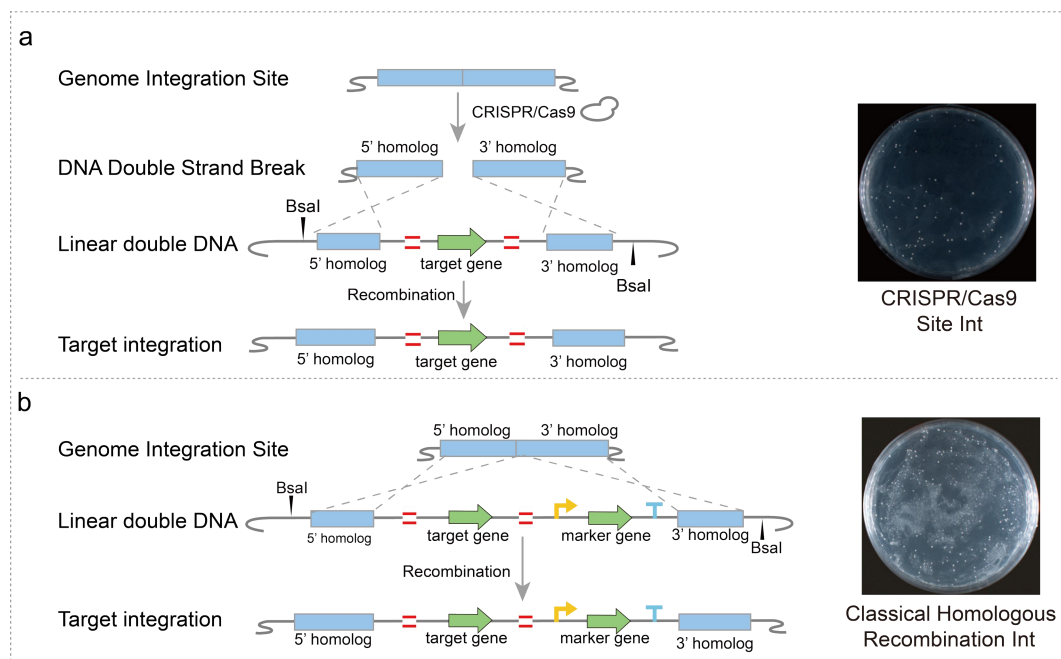


图 3：酿酒酵母基因组整合。（a）基于 Cas9 系统的重组方式的示意图。Cas9/sgRNA 在染色体 X:236336..237310 指定位点产生双链断裂，外源线性化 DNA 两侧的同源臂与基因组上的位点发生同源重组，实现目标片段的特异性整合。（b）常规酵母同源重组整合方式。外源线性化 DNA 两侧的同源臂与染色体 V 的 IIA 位点发生同源重组，将目标片段与 MAR4_CaURA3 筛选标记一并整合至该位点。

因此，在本工具包中，三级并不只是多转录单元组装的终点，也对应同一构建对象进入不同稳定基因组背景的切换节点。大肠杆菌和酿酒酵母虽然采用不同的宿主特异性整合机制，但在本体系中共享同一逻辑：先按统一规则完成构建，再通过切换合适的三级整合骨架进入定义好的测试背景，为后续标准化定量表征提供基础。

3.3 双宿主体系中的标准化测试框架

在完成层级构建和基因组整合设计后，我们进一步建立了双宿主统一背景下的元件与系统测试框架。这里的测试对象并不是孤立的序列本身，而是在固定骨架、固定宿主背景、固定参考元件和固定读出条件下被定义出来的性能对象。基于这一原则，工具包中的构建产物可直接进入标准化定量表征流程，而不需要为每一类元件或每一种调控系统重新设计测试条件。

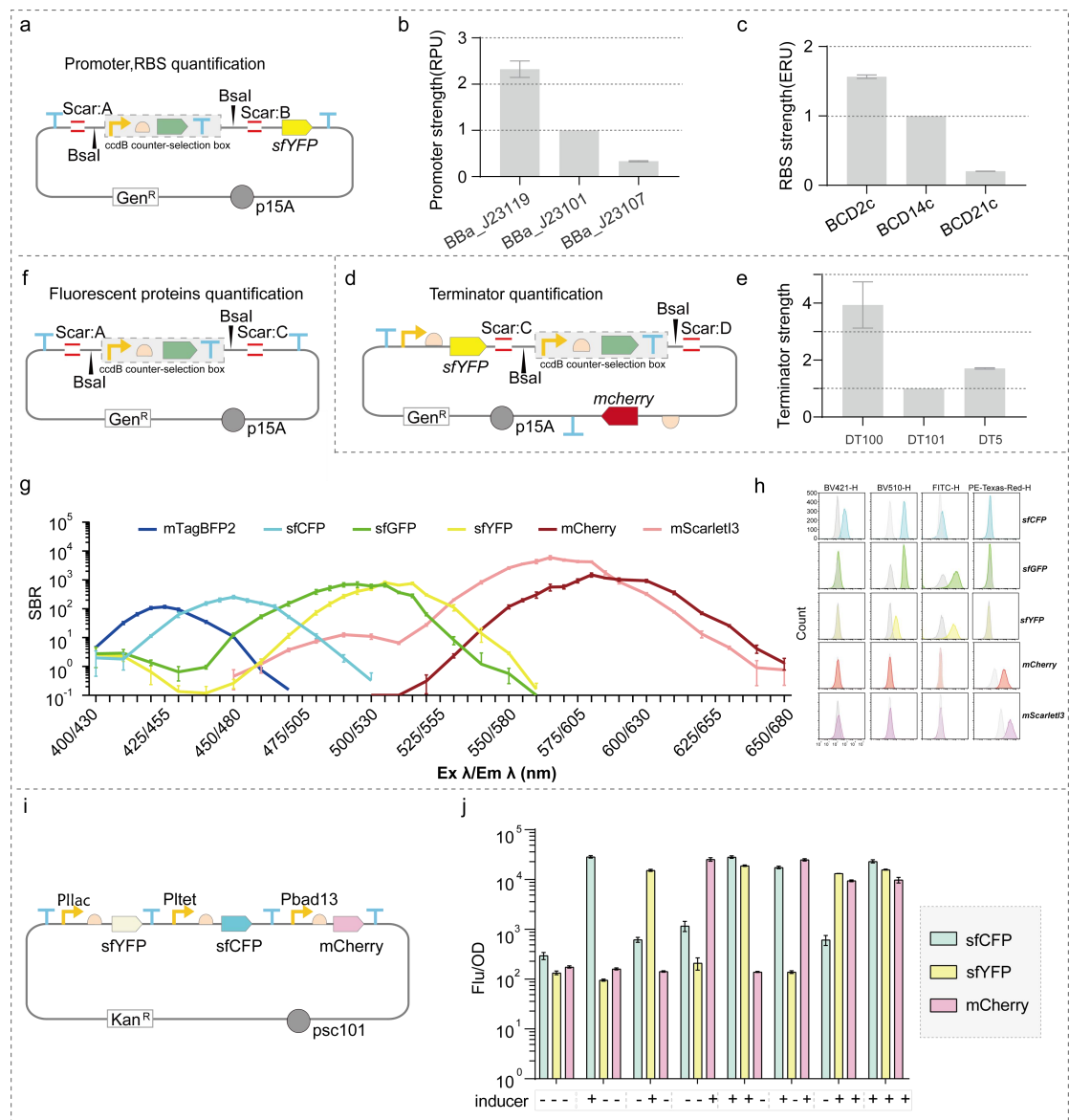


图 4: 大肠杆菌测试 Backbone 设计及元件表征。(a) Promoter 和 RBS 测试 Backbone, 该骨架设计成 AB scar 形式, 需要同时组装 Promoter 和 RBS; (b) Promoter 强度标定结果, 以 BBa_J23101 为参考元件, 计算各元件 RPU (补充说明 4.1); (c) RBS 强度标定结果, 以 BCD14c 为参考元件, 计算各 RBS 元件 REU (补充说明 4.2); (d) Terminator 测试 Backbone, 双荧光蛋白报告基因, 诱导表达上游 sfYFP, 同时测定下游 mCherry 通读水平; (e) Terminator 强度标定结果, 以 DT101 为参考元件,

计算各 Terminator 元件相对强度（补充说明 4.3）；（f）荧光蛋白测试 Backbone，该骨架设计成 AC scar 形式，需要同时组装 Promoter、RBS 和 CDS；（g）荧光蛋白的光谱扫描结果。激发扫描步长设为 5 nm 或 10 nm，激发光与检测发射光之间的波长间隔固定为 30 nm。Ex：激发；Em：发射；（h）sfCFP、sfGFP、sfYFP、mCherry 和 mScarlet3 在 BV421-H、BV510-H、FITC-H 和 PE-Texas Red-H 通道中的流式细胞信号分布。灰色曲线表示非诱导条件，彩色曲线表示相应荧光蛋白在诱导条件下的信号。未进行荧光补偿；（i）细胞多色共表达体系构建图谱；（j）多色表达系统的测试结果。sfCFP、sfYFP 和 mCherry 分别由 aTc、IPTG 和阿拉伯糖诱导。x 轴表示诱导剂的存在（+）或缺失（-），顺序为 aTc、IPTG、阿拉伯糖。信号使用酶标仪定量，以 Flu/OD₆₀₀ 表示（见材料与与方法）。数据以平均值 ± 标准差（SD）表示；每项实验均在不同日期进行 3 次独立重复（n = 3）。

在大肠杆菌中，启动子、核糖体结合位点和终止子的测试均采用统一逻辑的测试骨架（图 4a、d、f）。所有元件测试质粒均使用 p15A 复制子和庆大霉素抗性标记，以尽可能保持复制子背景和筛选条件一致。启动子测试以 BBa_J23101 为参考元件，完成了包括 BBa_J23119、BBa_J23101、BBa_J23107 等在内的代表性启动子标定，并以相对启动子单位（RPU）进行量化^[12]（图 4a、b）；核糖体结合位点测试以 BCD14c 为参考元件，对 BCD2c、BCD14c、BCD21c 等元件进行相对表达强度比较^[13]，并以相对表达单位（REU）表示（图 4a、c）；终止子测试采用上下游双报告读出框架^[14, 26]，以 DT101 为参考元件，标定了 DT101、DT100 和 DT5 等终止子的相对强度^[26]（图 4d、e；附图 3）。对于转录因子系统，我们进一步将测试拆分为两个对象：输入模块采用低拷贝 pSC101 骨架和卡那霉素抗性，用于稳定表达转录因子；输出模块采用 p15A 骨架和庆大霉素抗性，用于读取下游启动子的荧光响应。这种拆分使诱导剂浓度、转录因子表达量和输出信号之间的关系能够被分步骤量化，而不是混合在同一个构建中解释。

在酿酒酵母中，标准化测试框架进一步强调稳定遗传背景。启动子和终止子的元件标定主要建立在预定义的整合背景上：报告模块整合在染色体 V 的 IIA 位点，筛选标记为 MAR4_CaURA3^[32]；转录因子模块则整合在染色体 III 的 IIB 位点或染色体 XV 的 I 位点，并分别对应 MAR2_KILEU2 或 MAR7_clonNAT^[33]。启动子测试采用 pScbb23 骨架，仅替换待测启动子即可完成表征（图 5a、b）；终止子测试采用 pScbb46 和 pScbb47 骨架，在统一表达框架中替换待测终止子（图 5c、d；附图 4c、d）。与此同时，系统测试同样被组织成输入模块和输出模块：输入模块用于定义转录因子表达量，输出模块用于在固定报告构型下读取信号。以 pScLv305（SpacerD1-pPFY1-mCitrine-tENO1）为参考元件后，酿酒酵母中的组成型启动子、诱导型启动子、终止子以及转录因子系统都可以进入同一比较框架（图 5a、c、e；附图 4a、b、c）。

为减少读出层面的不确定性，我们还对两种宿主中代表性荧光报告蛋白的检测窗口和通

道干扰进行了系统校准。在大肠杆菌中，mTagBFP2、sfCFP、sfGFP、sfYFP、mCherry 和 mScarletI3 的激发扫描结果明确了推荐检测波长，并显示 sfGFP 与 sfYFP、mCherry 与 mScarletI3 之间存在显著光谱重叠（图 4f-g、附图 5a-e）；在酿酒酵母中，mCitrine、mNeonGreen、mScarletI3、mRuby2 和 mTagBFP2 的扫描和流式分析同样定义了推荐检测通道，并提示绿色/黄色和部分红色荧光蛋白之间存在可测的相互影响（图 5e-h、附图 6）。因此，荧光蛋白在本体系中不仅是报告器，也构成了标准化测试条件的一部分，为后续多色读出和系统测试提供了统一的检测窗口。

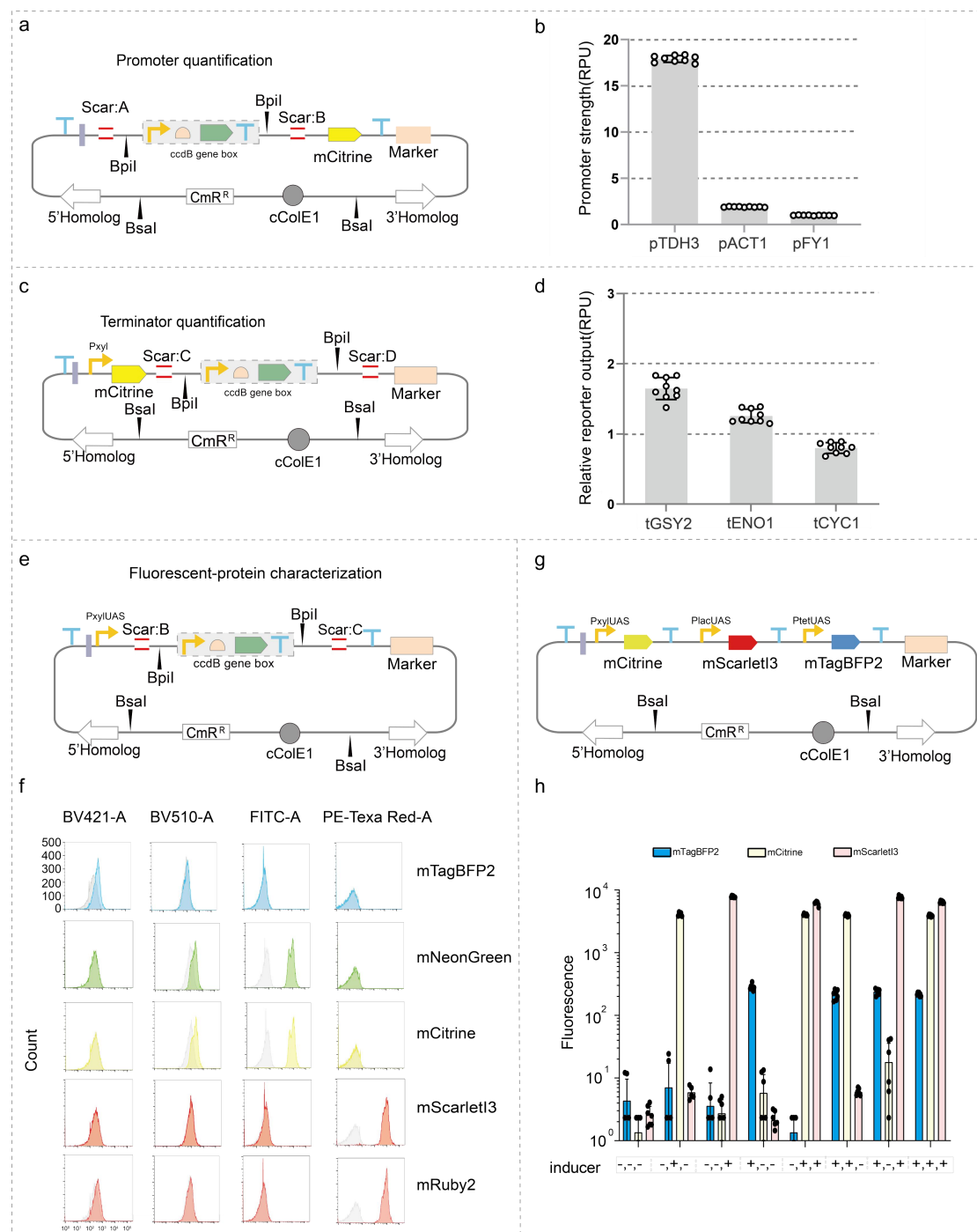


图 5: 酵母元件表征: (a) Promoter 测试 Backbone; (b) 启动子表征结果, 以携带整合的 pScLv305 (SpacerD1-pPFY1-mCitrine-tENO1; 酵母 RPU 参考构建) 的酵母菌株为参照, 计算各待测启动子的 RPU; (c) Terminator 测试 Backbone; (d) 基于相对报告输出(RPU)的终止子表征结果, 以携带整合的 pScLv305 (SpacerD1-pPFY1-mCitrine-tENO1; 酵母 RPU 参考构建) 的酵母菌株为参照进行计算; (e) 荧光蛋白扫描测试 Backbone; (f) mTagBFP2, mNeonGreen, mCitrine, mScarletl3 和 mRuby2 在 BV421-A, BV510-A, FITC-A, PE-Texas Red-A 通道的流式细胞信号分布结果, 灰色曲线代表非诱导条件, 带对应荧光颜色的曲线为诱导条件下测得; (g) 细胞多色共表达体系构建图谱; (h) 细胞多色共表达体系测试结果, 分别测定了三种诱导剂 8 种组合下的荧光值, 结果以 BV421-A, FITC-A 和 PE-Texas Red-A 通道的流式图中位数信号强度作为结果计算。mCitrine、mScarletl3 和 mTagBFP2 分别由木糖 (20 mmol/L)、IPTG (20 mmol/L) 和 aTc (100 ng/mL) 诱导表达。数据以平均值 $\pm$ 标准差 (SD) 表示; 每项实验均在不同日期进行 3 次独立重复 (n = 3)。

综上, 本工具包定义的并不仅是标准化组装规则, 也包括标准化测试规则: 在大肠杆菌中, 通过统一骨架、统一参考和统一输入/输出逻辑建立可比较的元件与系统测试对象; 在酿酒酵母中, 通过统一整合背景、统一参考构型和统一读出条件建立稳定的定量表征框架。这样形成的双宿主测试体系, 为后续元件积累、系统比较和跨实验条件校准提供了可直接复用的实验入口。

3.4 代表性调控线路验证了标准化工作流的可执行性

为验证上述标准化构建、基因组整合和测试框架能够在真实细胞背景中支持功能性对象, 我们进一步在两种宿主中构建并表征了代表性调控系统。这里的重点并不在于展示尽可能复杂的线路规模, 而在于检验在已定义的骨架、宿主背景和读出条件下, 工具包是否能够稳定地把标准元件组织成具有可解释输入—输出关系的功能模块。

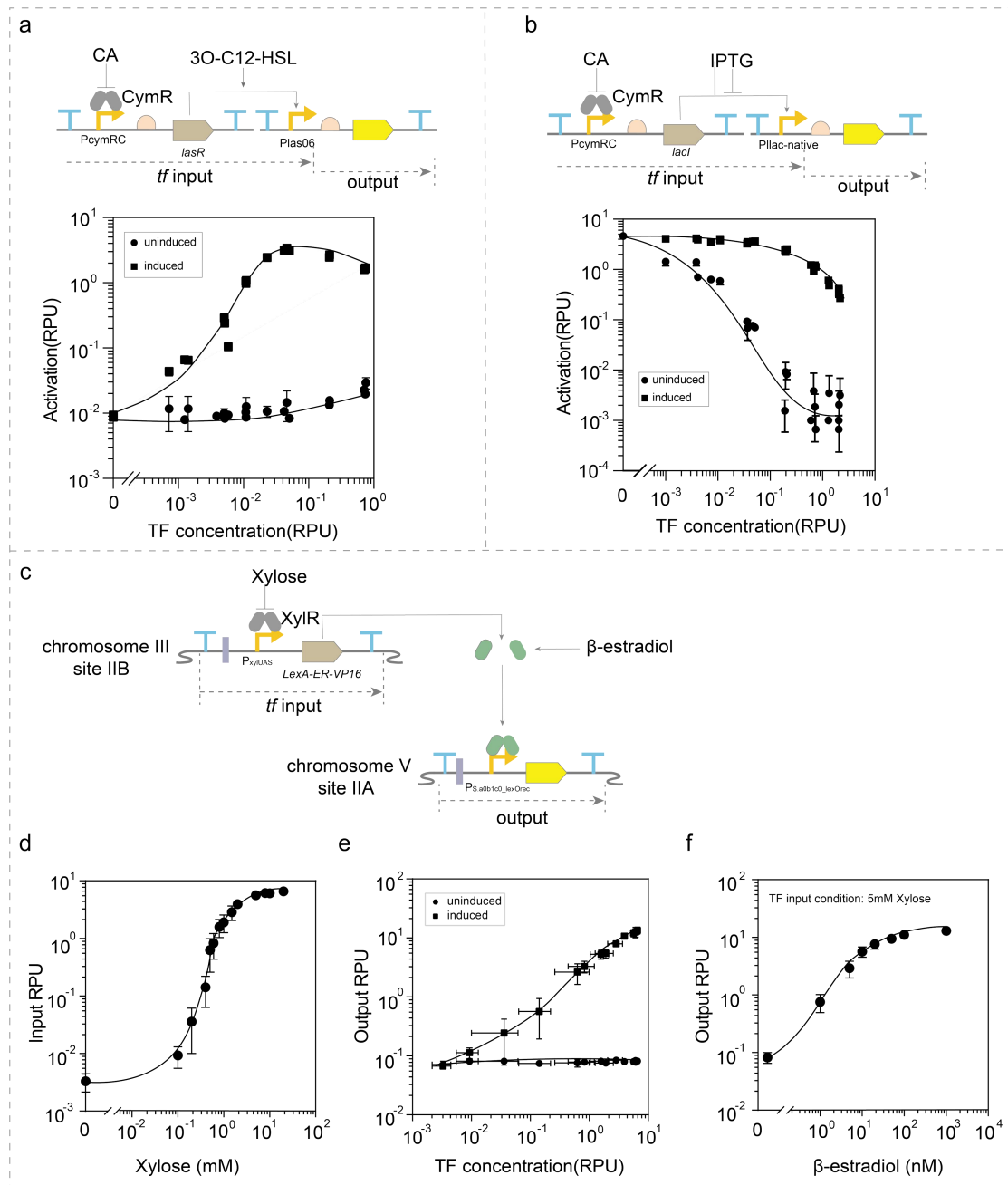


图 6: 大肠杆菌与酿酒酵母基因线路表征。(a) LasR 激活型基因线路在大肠杆菌中的示意图及系统输出。横轴表示转录因子 (TF) 表达水平 (以 RPU 表示), 受 P_{cymRC}-CymR 系统调控; 纵轴表示报告基因输出 (以 RPU 表示), 显示了在无配体或存在 1 μ M 3O-C12-HSL 条件下, 不同 TF 浓度梯度时的输出强度。(b) LacI 抑制型基因线路在大肠杆菌中的示意图及系统输出。横轴同 (a), 纵轴表示报告基因输出 (以 RPU 表示), 显示了在无配体或存在 1 mM IPTG 条件下, 不同 TF 浓度梯度时的输出强度。对于 LacI 系统, 过夜培养时添加了 100 μ M Cuminic acid (CA)。(c) 酿酒酵母 ER 激活型基因线路示意图。(d) 在携带整合的 pScLv013 (诱导启动子驱动的 mCitrine 输入表征报告构建) 的 CYE72 (预装 LacI、TetR 和 XylR 调控模块的工程酵母底盘) 中测得的 TF 输入滴定曲线。TF 输入水平以匹配的启动子报告构建的活性表示, 而非直接测量 TF 蛋白浓度。(e) 在无配体和存在 1 μ M β -雌二醇条件下, 不同 TF 输入水平对应的系统报告输出。(f) 在最适 TF 表达条件下 (输入诱导剂木糖浓度为 5 mM), ER 系统对 β -雌二醇的诱导响应曲线。数据以平均值 $\pm$ 标准差 (SD) 表示; 每项实验均在不同日期进行 3 次独立重复 (n = 3)。

在大肠杆菌中, 我们构建了 LasR 激活型和 LacI 抑制型两套调控系统。两类系统均

被拆分为输入模块和输出模块：输入模块用于定义转录因子表达水平，输出模块用于读取目标启动子的荧光响应。基于 **PcymRC** 的输入滴定首先确定了转录因子的工作窗口，其中输入曲线在高浓度 **Cuminic acid (CA)** 条件下达到上限；随后在选定的输入条件下测定系统对相应配体的响应（附图 7a）。结果显示，**LasR** 系统对 **3O-C12-HSL** 呈现预期的激活输出（图 6a；附图 7b），**LacI** 系统在 **IPTG** 存在时产生解除抑制后的信号恢复（图 6b；附图 7c）；两类系统在不同输入水平下均保持清晰的响应趋势，说明该框架能够将“输入定义”和“输出读取”分开组织，并支持可重复的定量测试。

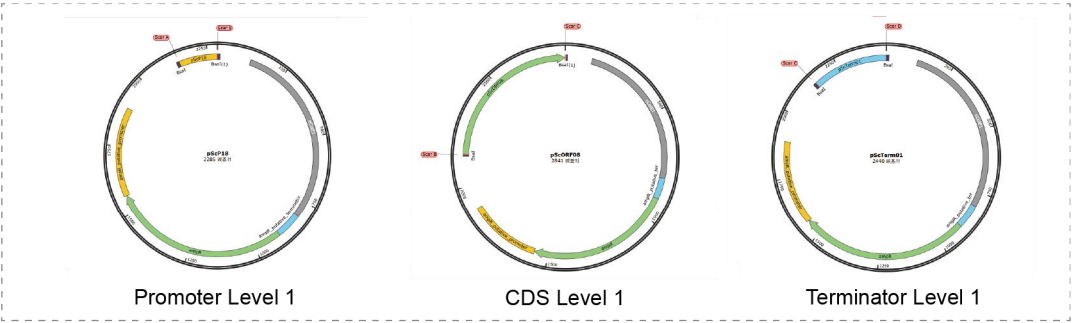
在酿酒酵母中，我们进一步构建并测试了 **XylR-PxylUAS**、**XylR/XlnR-Pxln.2b-xyl** 和 **LacI-PlacUAS** 等代表性调控系统（附图 8a-c）。各系统在统一整合背景和统一读出条件下均能够对相应诱导剂产生预期方向的输出变化，表明该工具包不仅适用于基础元件表征，也能够真核宿主中支撑激活型和抑制型调控对象的标准化构建与测试。进一步地，基于 **LexA-ER-VP16** 的模块化系统同样获得了可解释的转录因子输入曲线和 β -雌二醇响应曲线（图 6c-f），说明在酵母中，输入模块、信号转换模块和输出模块也可以在同一框架下被连续定义。

这些结果表明，本工具包建立的双宿主标准化工作流已足以支撑代表性调控线路从组装到测试的完整实现。更重要的是，线路结果的解释并不依赖于临时更换骨架、宿主或测试条件，而是建立在前述统一背景、统一参考和统一读出逻辑之上。因此，该体系验证的不是单个线路本身，而是一种可复用的标准化构建与测试入口：同一逻辑既可用于原核底盘，也可用于真核底盘，并可为后续元件积累、线路比较和跨实验条件校准提供直接基础。

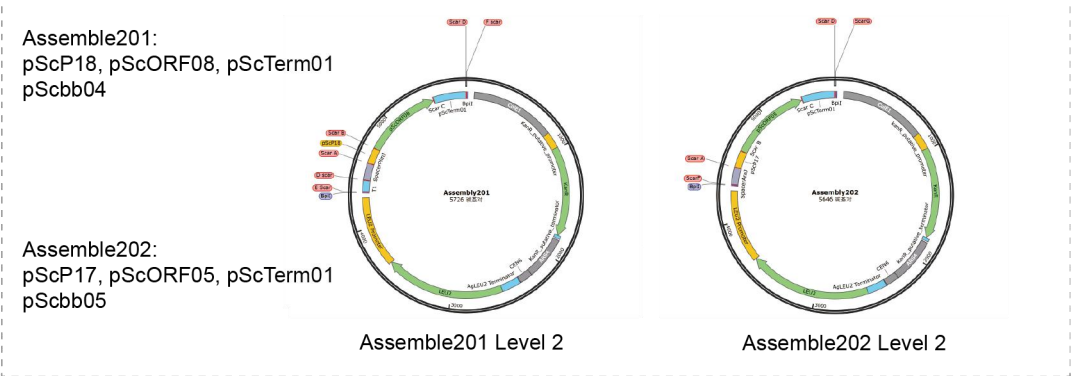
3.5 自动化质粒组装平台支持层级流程执行

本工具包的层级定义和接口规则不仅便于手工克隆，也适合数字化表达。基于相同的一级至三级对象定义、**scar** 规则和骨架切换关系，我们开发了配套的自动化质粒组装平台，用于质粒图谱可视化、层级路径生成和组装管理。平台当前主要支持从标准件调用到多级构建生成的数字化执行，而不直接承担遗传线路逻辑设计本身^[8,9]。由于元件类型、层级转换和骨架选择在工具包中已被预先定义，平台可以直接根据输入的元件和目标构建生成相应组装路径，使同一套标准化规则同时适用于实验执行和数字化管理。

Level 1 assembly



Level 2 assembly



Level 3 assembly

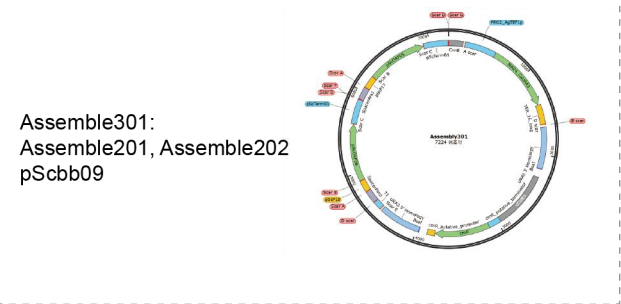


图 7：自动化平台介导的多层级质粒组装流程。以酿酒酵母工具包中的 6 个 Level 1 元件输入为起始模块，包括 2 个启动子质粒（pScP18 和 pScP17）、2 个 CDS 质粒（pScORF08 和 pScORF05），以及在两个转录单元中重复使用的终止子质粒 pScTerm01，因此共涉及 5 种不同的 Level 1 元件质粒。每组 Promoter–CDS–Terminator 被组装为一个含单转录单元的 Level 2 质粒；随后，2 个 Level 2 质粒进一步组装形成 1 个 Level 3 多转录单元质粒。该流程体现了自动化平台对标准化元件选择、单转录单元构建及多基因模块组合的连续组装能力。

为验证平台与工具包层级体系的兼容性，我们以酿酒酵母系统为例，设置 6 个一级元件输入（由 5 种不同质粒构成，其中 pScTerm01 重复使用），由平台自动组合为 2 个二级转录单元，并进一步生成 1 个三级多转录单元质粒（图 7）。该实例表明，前述层级设计、接口定义和骨架切换关系已经足够明确，可以被直接编码为可执行的数字化规则，从而支持从标准件调用到多级构建生成的连续执行。

因此，自动化平台在本研究中的意义，不在于替代遗传设计本身，而在于证明这套双宿

主工具包的规则已足够明确，能够被稳定转化为可执行的数字化组装流程。它进一步提升了工具包在标准件管理、层级构建和后续自动化衔接中的平台价值。

4 讨论

Polaris 工具包不仅提供质粒骨架，还将遗传元件存储、层级化构建、基因组整合和定量表征连接为统一工作流程。该框架预先定义了基础元件进入构建体系、升级为转录单元和多基因构建、转移至指定遗传背景并开展定量测试的过程，旨在减少使用者在常规实验开始前独立协调骨架选择、组装层级、筛选策略、整合途径和测试架构的工作。

这一框架由多项互补设计共同构成。**Level 1** 至 **Level 3** 层级明确了可复用元件、中间转录单元和最终应用构建的身份及组装路径；元件级接口与转录单元级接口的分离区分了不同层级的模块边界；**Bpil** 和 **Bsal** 的分层使用明确了各组装阶段的酶切规则。基于 **ccdB** 和 **mScarletI3** 的平行骨架为不同菌株条件和筛选需求提供了选择。**Level 3** 骨架进一步将组装产物连接至质粒测试体系或宿主特异性基因组整合途径。两种宿主共享层级逻辑，但保留各自的元件结构、骨架设计、整合机制和测试流程。

标准化测试是该工具包的另一核心组成。遗传元件不仅由 **DNA** 序列定义，其测得性能还与组装接口、载体背景、宿主、参考构建、培养条件和测量方法有关。固定的测试架构和匹配的参考构建，使启动子、核糖体结合位点、终止子、荧光蛋白及调控系统的测量结果能够在明确的实验背景中解释。将转录因子输入与报告输出分开表征，也使调控系统能够通过逐级输入—输出关系描述，而非仅以诱导剂浓度描述。

本研究的定量结果依赖所采用的宿主菌株、载体或整合背景、培养条件、参考构建和测量平台，不应视为可直接迁移的普通常数。不同实验室或平台之间转移定量值时，仍需进行参考测量和重新校准。本工具包的作用是明确可比较对象及其测量条件，而非消除全部生物学和实验变异。

明确的对象定义和层级转换规则也为数字化实现提供了基础。配套平台编码了元件类型、骨架选择、接口兼容性和 **Level 1** 至 **Level 3** 转换规则，支持元件检索、层级组装路径生成及质粒设计可视化。当前功能限于组装规划和构建体管理，不包括从头线路逻辑设计或自动化实验执行。已部署的平台和公开源代码为后续连接数据库、构建体管理及机器可读的设计—构建—测试流程提供了基础。

本研究仍有明确的适用边界。双宿主共享层级逻辑不意味着两种宿主的遗传元件可以在序列层面直接互换。**Int2**、**Int5** 和 **Int7** 整合系统作为扩展资源纳入工具包，但未在本研究中进行实验验证。已表征的元件和调控系统是工作流程的代表性示例，而非对所有元件或线路

架构的穷尽性验证。

后续工作可从三个方面推进：扩充已表征元件、整合骨架、宿主背景和调控系统等标准参考对象；建立关联序列、遗传背景、实验条件、参考测量和定量性能的结构化数据层；进一步完善构建体管理、数据回传及面向自动化执行的数字平台功能。这些扩展将支持可比较遗传对象和可复用实验流程的持续积累。

5. 结论

本研究建立了双宿主 **Golden Gate** 工具包 **Polaris**，将大肠杆菌和酿酒酵母中的层级化 DNA 组装、宿主特异性基因组整合与定量测试连接起来。该工具包定义了共享的三级层级，区分元件级与转录单元级接口，并结合 **BpiI/BsaI** 分层组装以及基于 **ccdB** 和 **mScarletI3** 的平行筛选方案。宿主特异性整合模块和预定义测试架构，使组装产物能够进入元件表征及调控系统分析；本研究通过两种宿主中的代表性遗传元件、激活系统和抑制系统验证了这一流程。配套数字平台编码了 **Level 1** 至 **Level 3** 的元件选择、兼容性检查、质粒可视化和多层级组装规划规则。**Polaris** 的贡献不仅在于组装方法，而在于将构建与测试整合为统一工作框架，为后续元件扩展、定量数据积累、数据库建设及与自动化兼容的合成生物学流程提供基础。

作者贡献声明

谢荣辉：数据整理、形式分析、实验研究、方法学、资源提供、验证、可视化、初稿撰写。

王然：数据整理、形式分析、实验研究、方法学、资源提供、验证、可视化、初稿撰写。

詹依：实验研究、方法学、资源提供、验证。

魏博妍：方法学、资源提供、软件开发。

徐思杨：实验研究、资源提供。

陈业：研究构思、经费获取、形式分析、实验研究、方法学、项目管理、监督指导、可视化、初稿撰写、审阅与编辑。

刘莉莉：研究构思、方法学、形式分析、验证、资源提供、项目管理、监督指导、初稿撰写、审阅与编辑。

利益冲突声明

作者声明无利益冲突。

生成式人工智能使用声明

在本稿准备过程中，作者使用 **ChatGPT (OpenAI)** 进行英文语言润色。使用该工具后，作者对相关内容进行了审阅和编辑，并对论文内容承担全部责任。

致谢

本研究得到中国科学院战略性先导科技专项（XDA0510200）的资助。感谢本实验室历届和现任成员提供遗传元件。

补充材料

本文相关补充材料以单独文件提供，包括补充信息、质粒和菌株清单、遗传元件序列，以及大肠杆菌和酿酒酵母的定量表征数据集。

数据获取

Polaris Kit Assembly 的源代码和全部 240 个质粒图谱（Kits_plasmids_map.zip）可从 <https://github.com/ChenYe-Lab/PolarisKitAssembly> 获取。处理后的定量数据、元件序列和质粒清单见补充材料。在线平台地址为 <https://infrasybio.siat.ac.cn/kitassembly>。

参考文献

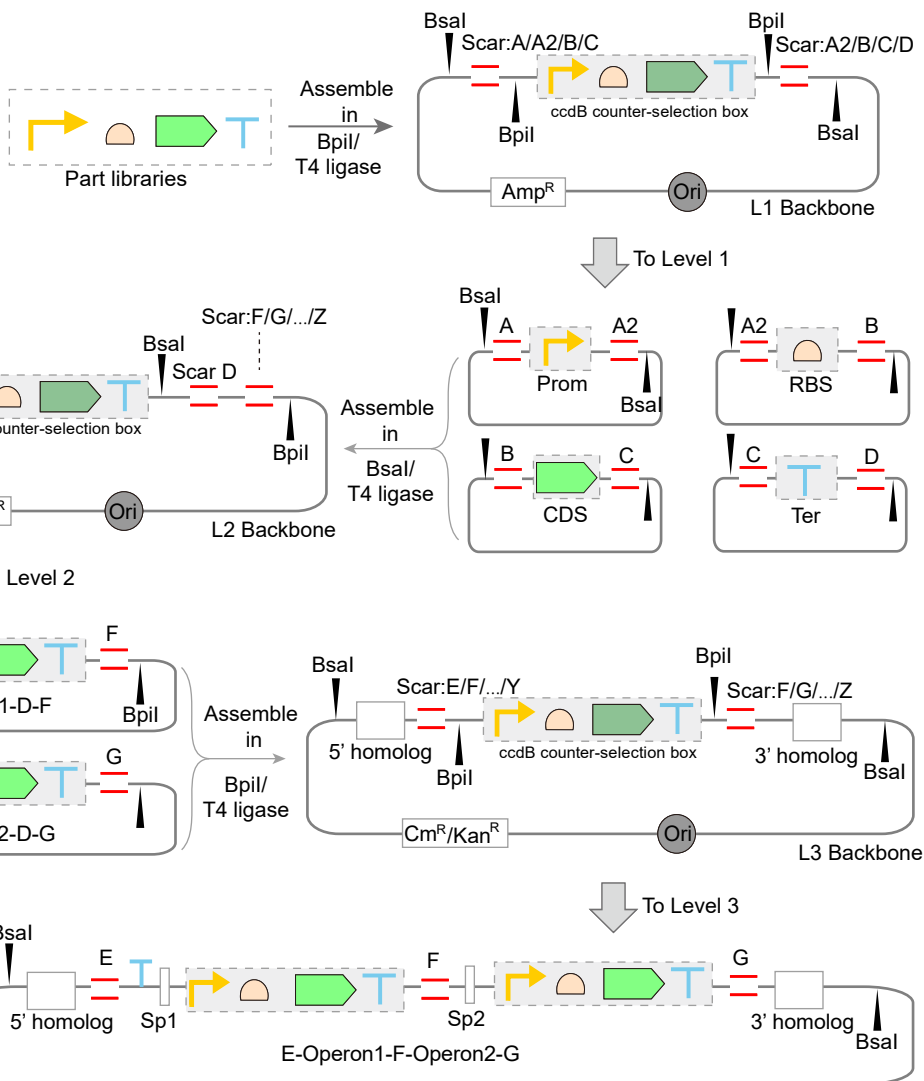
- [1] Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS One 2008;3:e3647.
<https://doi.org/10.1371/journal.pone.0003647>.
- [2] Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS One 2011;6:e16765.
<https://doi.org/10.1371/journal.pone.0016765>.
- [3] Lee ME, DeLoache WC, Cervantes B, Dueber JE. A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth Biol 2015;4:975-86.
<https://doi.org/10.1021/sb500366v>.
- [4] Iverson SV, Haddock TL, Beal J, Densmore DM. CIDAR MoClo: improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. ACS Synth Biol 2016;5:99-103.
<https://doi.org/10.1021/acssynbio.5b00124>.
- [5] Moore SJ, Lai HE, Kelwick RJR, Chee SM, Bell DJ, Polizzi KM, et al. EcoFlex: a multifunctional MoClo kit for *E. coli* synthetic biology. ACS Synth Biol 2016;5:1059-69.
<https://doi.org/10.1021/acssynbio.6b00031>.
- [6] Blázquez B, San León D, Torres-Bacete J, Gómez-Luengo A, Kniewel R, Martínez I, et al. Golden Standard: a complete standard, portable, and interoperative MoClo tool for model and non-model proteobacteria. Nucleic Acids Res 2023;51:e98.
<https://doi.org/10.1093/nar/gkad758>.

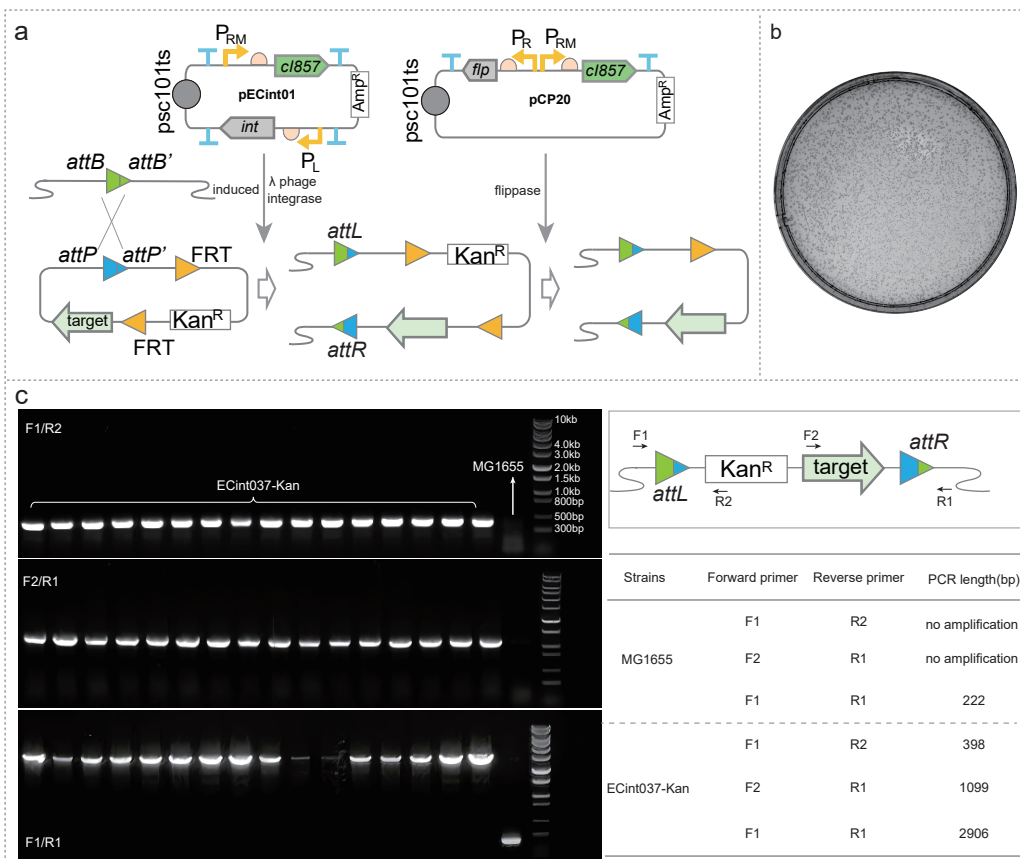
- [7] Mejlsted J, Kubaczka E, Wirth S, Koeppl H. The Coli Toolkit (CTK): an extension of the modular yeast toolkit for use in *E. coli*. ACS Synth Biol 2026;15:462-71.
<https://doi.org/10.1021/acssynbio.5c00489>.
- [8] Nielsen AAK, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, et al. Genetic circuit design automation. Science 2016;352:aac7341.
<https://doi.org/10.1126/science.aac7341>.
- [9] Chen Y, Zhang S, Young EM, Jones TS, Densmore D, Voigt CA. Genetic circuit design automation for yeast. Nat Microbiol 2020;5:1349-60.
<https://doi.org/10.1038/s41564-020-0757-2>.
- [10] Ham TS, Dmytriv Z, Plahar H, Chen J, Hillson NJ, Keasling JD. Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. Nucleic Acids Res 2012;40:e141.
<https://doi.org/10.1093/nar/gks531>.
- [11] McLaughlin JA, Beal J, Mısırlı G, Grünberg R, Bartley BA, Scott-Brown J, et al. The Synthetic Biology Open Language (SBOL) version 3: simplified data exchange for bioengineering. Front Bioeng Biotechnol 2020;8:1009.
<https://doi.org/10.3389/fbioe.2020.01009>.
- [12] Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, Czar MJ, et al. Measuring the activity of BioBrick promoters using an *in vivo* reference standard. J Biol Eng 2009;3:4. <https://doi.org/10.1186/1754-1611-3-4>.
- [13] Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai QA, et al. Precise and reliable gene expression via standard transcription and translation initiation elements. Nat Methods 2013;10:354-60.
<https://doi.org/10.1038/nmeth.2404>.
- [14] Chen YJ, Liu P, Nielsen AAK, Brophy JAN, Clancy K, Peterson T, et al. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nat Methods 2013;10:659-64.
<https://doi.org/10.1038/nmeth.2515>.
- [15] Cardinale S, Arkin AP. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. Biotechnol J 2012;7:856-66.
<https://doi.org/10.1002/biot.201200085>.
- [16] Lou C, Stanton B, Chen YJ, Munsky B, Voigt CA. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. Nat Biotechnol 2012;30:1137-42.
<https://doi.org/10.1038/nbt.2401>.

- [17] Bayer CN, Rennig M, Ehrmann AK, Nørholm MHH. A standardized genome architecture for bacterial synthetic biology (SEGA). *Nat Commun* 2021;12:5876. <https://doi.org/10.1038/s41467-021-26155-5>.
- [18] Lux MW, Strychalski EA, Vora GJ. Advancing reproducibility can ease the ‘hard truths’ of synthetic biology. *Synth Biol (Oxf)* 2023;8:ysad014. <https://doi.org/10.1093/synbio/ysad014>.
- [19] Lund BA, Leiros HKS, Bjerga GK. A high-throughput, restriction-free cloning and screening strategy based on ccdB-gene replacement. *Microb Cell Fact* 2014;13:38. <https://doi.org/10.1186/1475-2859-13-38>.
- [20] Wang H, Bian X, Xia L, Ding X, Müller R, Zhang Y, et al. Improved seamless mutagenesis by recombineering using ccdB for counterselection. *Nucleic Acids Res* 2014;42:e37. <https://doi.org/10.1093/nar/gkt1339>.
- [21] Zhang X, Teng C, Lyu K, Lyu S, Fan Y. ‘Two in One’ cloning vector applied for blunt-end and T-A cloning with one-step digestion-ligation and screening of positive recombinants by unaided eyes. *Curr Issues Mol Biol* 2025;47:17. <https://doi.org/10.3390/cimb47010017>.
- [22] Meyer AJ, Segall-Shapiro TH, Glassey E, Zhang J, Voigt CA. *Escherichia coli* “Marionette” strains with 12 highly optimized small-molecule sensors. *Nat Chem Biol* 2019;15:196-204. <https://doi.org/10.1038/s41589-018-0168-3>.
- [23] Wang T, Xie R, Ma Z, Chen Y. A unified computational framework for quantitative design and optimization of transcriptional regulation across bacterial species. *Nucleic Acids Res* 2026;54:gkaf1472. <https://doi.org/10.1093/nar/gkaf1472>.
- [24] Tafoya-Ramírez M, Padilla-Vaca F, Ramírez-Saldaña A, Mora-Garduño J, Rangel-Serrano A, Vargas-Maya N, et al. Replacing standard reporters from molecular cloning plasmids with chromoproteins for positive clone selection. *Molecules* 2018;23:1328. <https://doi.org/10.3390/molecules23061328>.
- [25] Wu P, Li X, Yang M, Huang Z, Mo H, Li T, et al. High-throughput, one-step screening, cloning and expression based on the lethality of DpnI in *Escherichia coli*. *Biochem Biophys Res Commun* 2018;504:177-83. <https://doi.org/10.1016/j.bbrc.2018.08.151>.
- [26] Park Y, Espah Borujeni A, Gorochowski TE, Shin J, Voigt CA. Precision design of stable genetic circuits carried in highly-insulated *E. coli* genomic landing pads. *Mol Syst Biol* 2020;16:e9584. <https://doi.org/10.15252/msb.20209584>.

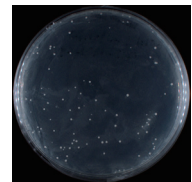
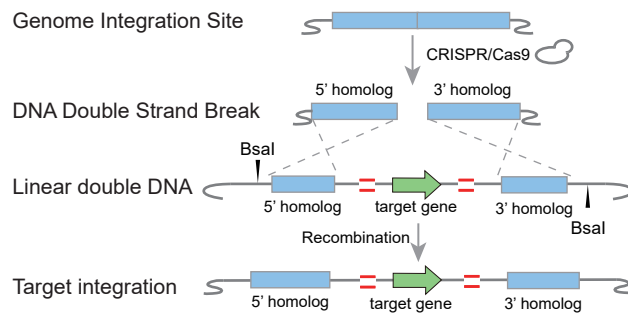
- [27] Yang L, Nielsen AAK, Fernandez-Rodriguez J, McClune CJ, Laub MT, Lu TK, et al. Permanent genetic memory with >1-byte capacity. *Nat Methods* 2014;11:1261-6. <https://doi.org/10.1038/nmeth.3147>.
- [28] Landy A. Dynamic, structural, and regulatory aspects of λ site-specific recombination. *Annu Rev Biochem* 1989;58:913-41. <https://doi.org/10.1146/annurev.bi.58.070189.004405>.
- [29] Li Q, Sun B, Chen J, Zhang Y, Jiang Y, Yang S. A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. *Acta Biochim Biophys Sin (Shanghai)* 2021;53:620-7. <https://doi.org/10.1093/abbs/gmab036>.
- [30] Reider Apel A, d'Espaux L, Wehrs M, Sachs D, Li RA, Tong GJ, et al. A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2017;45:496-508. <https://doi.org/10.1093/nar/gkw1023>.
- [31] Meng J, Qiu Y, Zhang Y, Zhao H, Shi S. CMI: CRISPR/Cas9 based efficient multiplexed integration in *Saccharomyces cerevisiae*. *ACS Synth Biol* 2023;12:1408-14. <https://doi.org/10.1021/acssynbio.2c00591>.
- [32] Goldstein AL, Pan X, McCusker JH. Heterologous URA3MX cassettes for gene replacement in *Saccharomyces cerevisiae*. *Yeast* 1999;15:507-11. [https://doi.org/10.1002/\(SICI\)1097-0061\(199904\)15:6<507::AID-YEA369>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1097-0061(199904)15:6<507::AID-YEA369>3.0.CO;2-P).
- [33] Gueldener U, Heinisch J, Koehler GJ, Voss D, Hegemann JH. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* 2002;30:e23. <https://doi.org/10.1093/nar/30.6.e23>.

Backbone	Ori
L1	ColE1
L2	ColE1
L3	ColE1/p15A/ pSC101/pBBR1



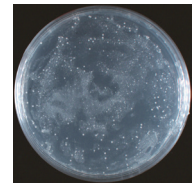
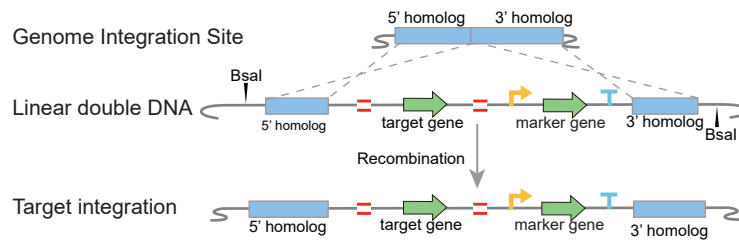


a

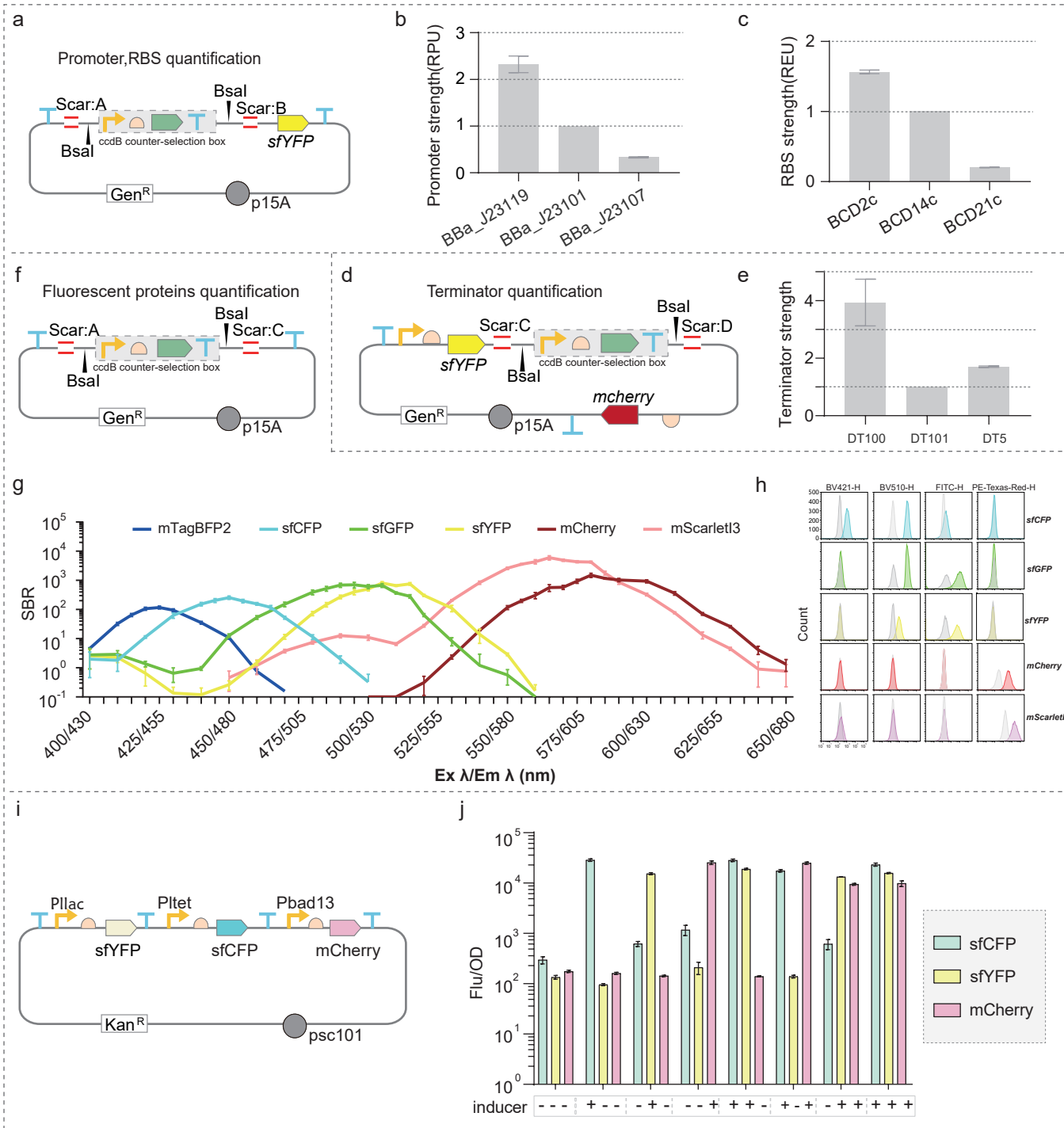


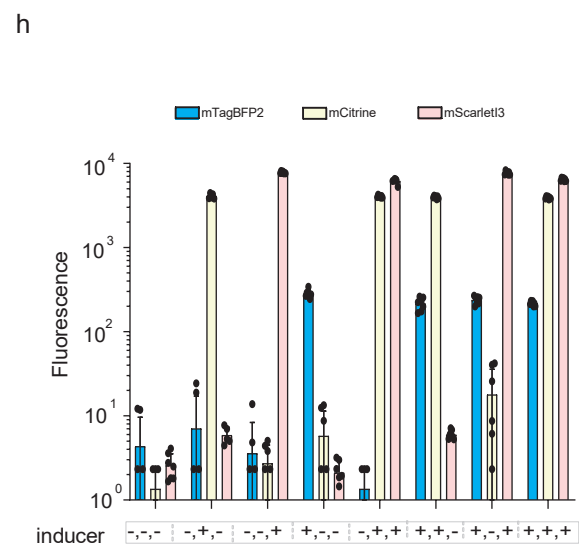
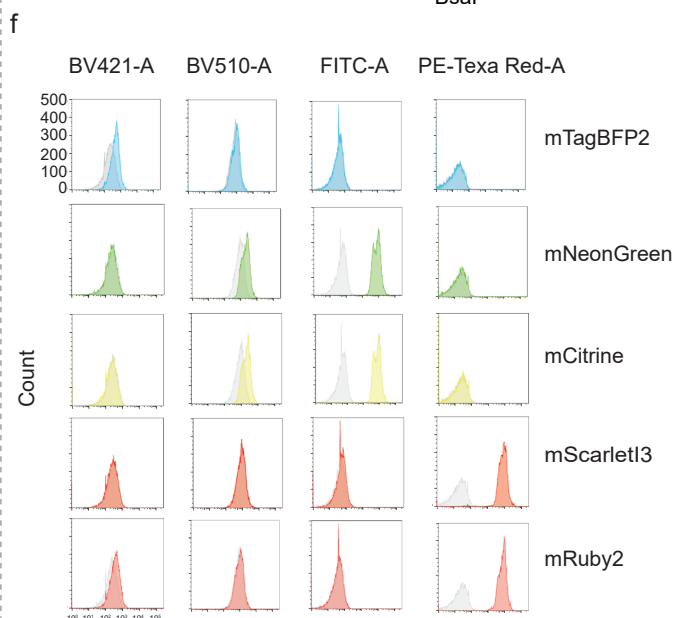
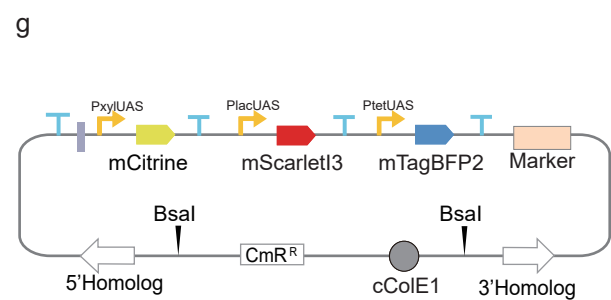
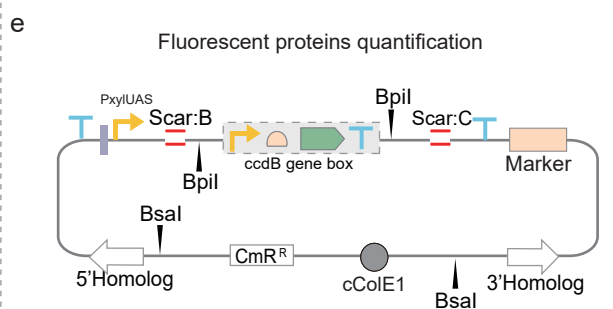
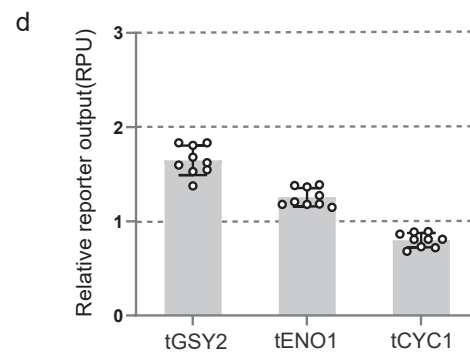
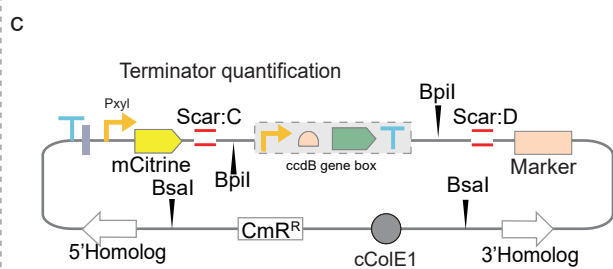
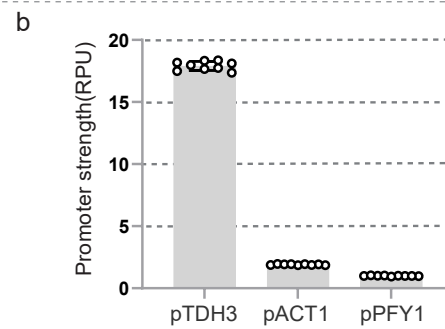
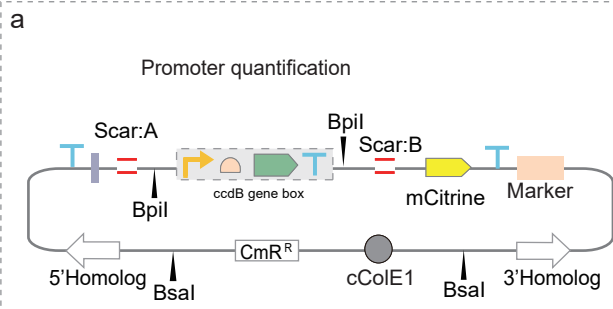
CRISPR/Cas9
Site Int

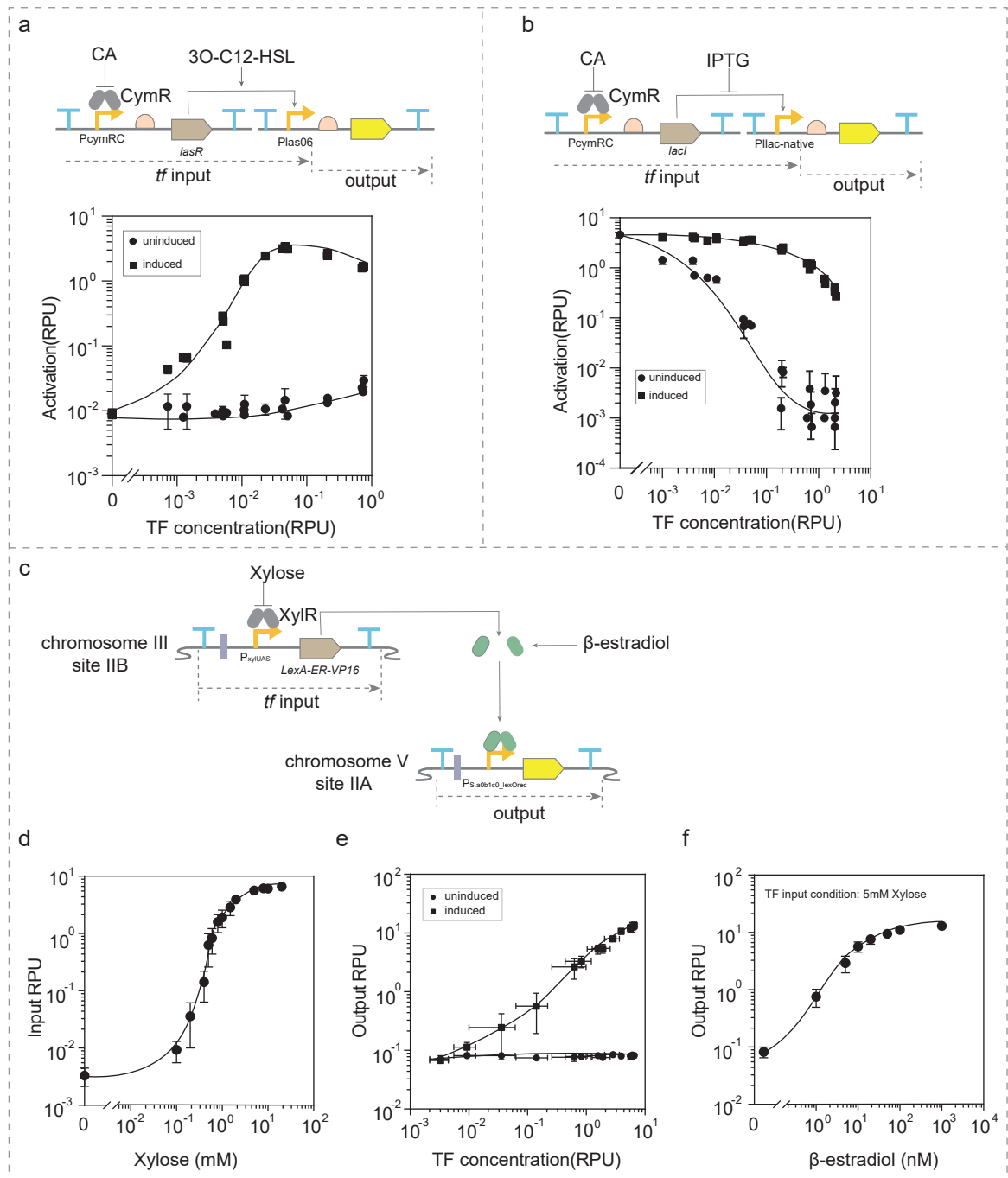
b



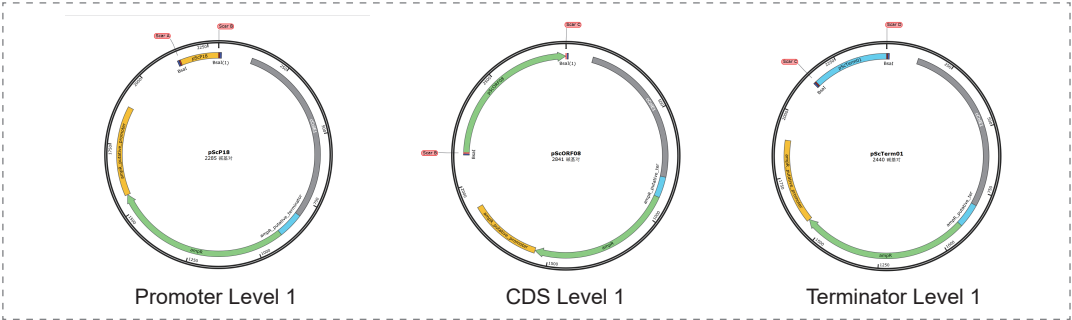
Classical Homologous
Recombination Int



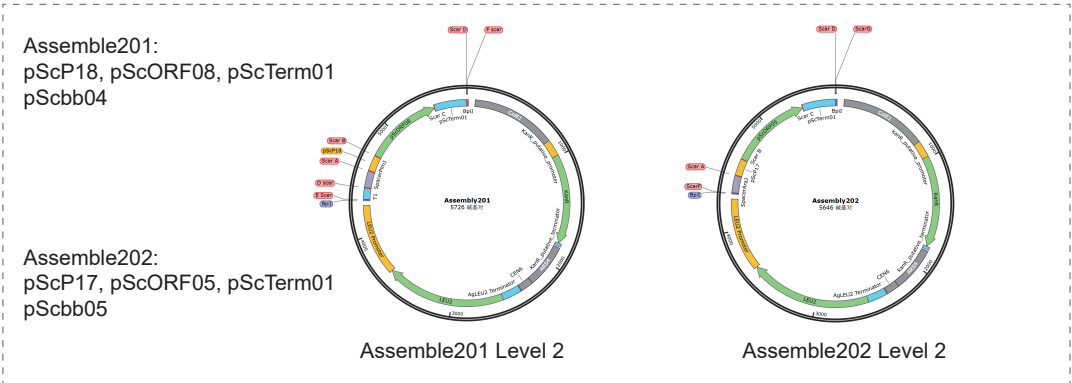




Level 1 assembly



Level 2 assembly



Level 3 assembly

