

申 报	系列：实验技术系列
	专业：生物物理学
	职称：实验师

业绩成果材料

（申报人的业绩成果材料包括论文、科研项目、获奖以及其他成果等）

单 位（二级单位） 生命科学学院

姓 名 刘 伟智

材料核对人：

单位盖章：

核对时间：

华南农业大学制

目 录

一、教学研究业绩

二、科研项目

- 1.主参：国家自然科学基金面上项目“水稻粒长基因 GL3n 的克隆和分子机制研究”项目计划书及批准通知书 3

三、论文、著作等

- 1.检索证明 18
- 2.以第一作者发表本专业论文情况
- 2.1. DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations 19
- 2.2. STI PCR: An efficient method for amplification and de novo synthesis of long DNA sequences 23

四、科研成果

五、其他业绩



项目批准号	32172010
申请代码	C1307
归口管理部门	
依托单位代码	51064208A0499-0932



国家自然科学基金 资助项目计划书 (预算制项目)

资助类别: 面上项目

亚类说明: _____

附注说明: _____

项目名称: 水稻粒长基因GL3n的克隆和分子机制研究

直接费用: 58万元 执行年限: 2022.01-2025.12

负责人: 谢先荣

通讯地址: 广东省广州市天河区华南农业大学科技楼502

邮政编码: 510642 电 话: 020-38297231

电子邮件: xie079799@163.com

依托单位: 华南农业大学

联系人: 唐家林 电 话: 020-85280070

填表日期: 2021年10月19日

国家自然科学基金委员会制

Version: 1.002.899



国家自然科学基金资助项目计划书填报说明 （预算制项目）

- 一、项目负责人收到《国家自然科学基金资助项目批准通知》（以下简称《批准通知》）后，请认真阅读本填报说明，参照国家自然科学基金相关项目管理办​​法和新修订的《国家自然科学基金资助项目资金管理办法》（以下简称《资金管理办法》，请查阅国家自然科学基金委员会官方网站首页“政策法规”栏目），按《批准通知》的要求认真填写和提交《国家自然科学基金资助项目计划书》（以下简称《计划书》）。
- 二、填写《计划书》时要科学严谨、实事求是、表述清晰、准确。《计划书》经国家自然科学基金委员会相关项目管理部门审核批准后，将作为项目研究计划执行、检查和验收的依据。
- 三、《计划书》各部分填写要求如下：
 - （一）简表：由系统自动生成。
 - （二）摘要及关键词：各类获资助项目都应当填写中、英文摘要及关键词。
 - （三）项目组主要成员：计划书中列出姓名的项目组主要成员由系统自动生成，与申请书原成员保持一致，不可随意调整。如果《批准通知》所附“项目评审意见及修改意见表”中“修改意见”栏目有调整项目组成员相关要求的，待项目开始执行后，按照项目成员变更程序另行办理。
 - （四）资金预算表：根据批准的项目资助额度，按规定调整项目预算，并按照《国家自然科学基金项目计划书预算表编制说明》填报资金预算表和预算说明书。
 - （五）正文：
 1. 面上项目、地区科学基金项目：如果《批准通知》所附“项目评审意见及修改意见表”中“修改意见”栏目没有修改要求的，只需选择“研究内容和研究目标按照申请书执行”即可；如果《批准通知》中上述栏目明确要求调整研究期限或研究内容等的，须选择“根据研究方案修改意见更改”并填报相关修改内容。
 2. 重点项目、重点国际（地区）合作研究项目、重大项目、国家重大科研仪器研制项目、原创探索计划项目：须选择“根据研究方案修改意见更改”，根据《批准通知》的要求填写研究（研制）内容，不得自行降低、更改研究目标（或仪器研制的技术性能与主要技术指标、验收技术指标等）或缩减研究（研制）内容。此外，还要突出以下几点：
 - （1）研究的难点和在实施过程中可能遇到的问题（或仪器研制风险），拟采用的研究（研制）方案和技术路线；
 - （2）项目主要参与者分工，合作研究单位（如有）之间的关系与分工，重大项目还需说明课题之间的关联；
 - （3）详细的年度研究（研制）计划。
 3. 创新研究群体项目：须选择“根据研究方案修改意见更改”，按下列提纲撰写：
 - （1）研究方向；



- （2）结合国内外研究现状，说明研究工作的学术思想和科学意义（限两个页面）；
 - （3）研究内容、研究方案及预期目标（限两个页面）；
 - （4）年度研究计划；
 - （5）研究队伍的组成情况。
- 4. 基础科学中心项目：须选择“根据研究方案修改意见更改”，根据《批准通知》的要求和现场考察专家组的意见和建议，进一步完善并细化研究计划，按下列提纲撰写：
 - （1）五年拟开展的研究工作（包括主要研究方向、关键科学问题与研究内容）；
 - （2）研究方案（包括骨干成员之间的分工及合作方式、学科交叉融合研究计划等）；
 - （3）年度研究计划；
 - （4）五年预期目标和可能取得的重大突破等；
 - （5）研究队伍的组成情况。
- 5. 对于其他类型项目，参照面上项目的方式进行选择和填写。



简表

项目负责人信息	姓 名	谢先荣	性 别	男	出生年月	1988年12月	民 族	汉族
	学 位	博士			职称	副研究员		
	是否在站博士后	否			电子邮件	xie079799@163.com		
	电 话	020-38297231			个人网页			
	工 作 单 位	华南农业大学						
	所 在 院 系 所	生命科学学院						
依托单位信息	名 称	华南农业大学					代码	51064208A0499
	联 系 人	唐家林		电子邮件	kycjkh@scau.edu.cn			
	电 话	020-85280070		网站地址	http://kjc.scau.edu.cn/			
合作单位信息	单 位 名 称							
项目基本信息	项 目 名 称	水稻粒长基因GL3n的克隆和分子机制研究						
	资 助 类 别	面上项目				亚 类 说 明		
	附 注 说 明							
	申 请 代 码	C1307:作物基因组及遗传学						
	基 地 类 别	亚热带农业生物资源保护与利用国家重点实验室						
	执 行 年 限	2022.01-2025.12						
	直 接 费 用	58万元						



项目摘要

中文摘要:

水稻粒型不仅是重要的产量性状，还与稻米品质有直接的关系。挖掘和研究更多的水稻粒型相关的基因及其机理，有利于揭开水稻粒型形成的调控机制，为培育高产优质的水稻新品种提供重要的基因资源和理论指导。本项目通过正向遗传学的方法，鉴定了一个新的控制水稻粒长的QTL基因GL3n，其增加籽粒的粒长并提高了稻米的食味品质。本项目拟在前期研究的基础上，围绕GL3n的克隆及其控制粒长的分子机制展开研究，主要包括：GL3n的图位克隆和功能验证；通过细胞学观察、表达分析、互作因子筛选、生理生化分析和进化研究，解析GL3n参与粒型调控的分子机制及其演化历程，并从中发现GL3n用于提高水稻品质和大田产量的潜力。

Abstract:

The shape of rice seeds affects not only the grain yield but also grain quality. Cloning more seed development related genes and studying their mechanisms will benefit the understanding of the formation of rice grain shape, and provide important genetic resource and theoretical guidance for breeding high-yield and superior-quality elite varieties. In this proposal, we identified a new QTL gene, GL3n, which controls the grain length in rice using forward genetics method. The gene is able to increase the grain length and also the quality. On the basis of our previous studies, we will focus on the cloning of GL3n and studying its molecular mechanism in controlling the grain length. The main contents include follows: the map-based cloning and validation of GL3n; elucidation of GL3n function in the development of grain shape through cytological observation, expression analysis, screening of interaction factors, molecular and biochemical analysis; further evaluation of the potential of GL3n for promoting the field yield and quality.

关键词(用分号分开): 水稻; 粒型; 基因定位; 产量性状; 分子机制

Keywords(用分号分开): Rice; Grain shape; Map-based cloning; Yield traits; Molecular mechanism



项目组主要成员

编号	姓名	出生年月	性别	职称	学位	单位名称	电话	证件号码	项目分工	每年工作 时间 (月)
1	谢先荣	1988.12	男	副研究员	博士	华南农业大学	020-38297231	360734198812031313	项目负责人	10
2	刘伟智	1984.11	男	实验师	硕士	华南农业大学	020-38297231	440106198411201838	组学和进化分析	8
3	谭习羽	1993.08	男	博士生	学士	华南农业大学	202-85288395	430381199308137114	互作因子筛选和验证	10
4	陈绍通	1992.10	女	博士生	硕士	华南农业大学	020-85288395	410725199210210560	生理生化分析	10
5	林展生	1998.02	男	硕士生	学士	华南农业大学	020-85288395	440582199802146998	基因定位和表型鉴定	10
6	李福权	1997.10	男	硕士生	学士	华南农业大学	020-85288395	440183199710287333	载体构建、转化和进化分析	10
总人数		高级		中级		初级		博士后	博士生	硕士生
6		1		1		0		0	2	2



国家自然科学基金预算制项目预算表

项目批准号：32172010

项目负责人：谢先荣

金额单位：万元

序号	科目名称	金额
1	一、基金资助项目直接费用合计	58.0000
2	1、设备费	1.4000
3	其中：设备购置费	1.4000
4	2、业务费	39.2000
5	3、劳务费	17.4000
6	二、其他来源资金	0.0000
7	三、合计	58.0000

注：请按照项目研究实际需要合理填写各科目预算金额。



预算说明书

（请按照《国家自然科学基金项目计划书预算表编制说明》等的有关要求，按照政策相符性、目标相关性和经济合理性原则，实事求是编制项目预算。填报时，直接费用应按设备费、业务费、劳务费三个类别填报，每个类别结合科研任务按支出用途进行说明。填报时，对单价 ≥ 50 万元的设备详细说明，对单价 < 50 万元的设备费用分类说明，对合作研究单位资质及资金外拨情况、自筹资金进行必要说明。）

本课题直接经费58万元。各项支出预算比例按照《国家自然科学基金资助项目资金管理办法》的相关规定，详细经费预算说明如下：

一、设备费（1.4万元，占总经费2.41%）：课题主持单位具有良好的科研条件，基本可保障课题科研工作的顺利实施。但需要对实验中常用且必需的老旧易损仪器设备进行更新。

1. 购置移液器1.4万元：课题缺少2 μ L以下量程和5mL量程的移液器，在进行定量分析和样本制备时，需要用到2 μ L和5mL移液器。因此，拟购买4支2 μ L移液器，2000元/支，共0.8万元；拟购买2支5mL移液器，3000元/支，共0.6万元。

二、业务费（39.2万元，占总经费67.59%）：包括项目实施过程中使用的试剂、辅助材料、测试化验加工、出版、会议和差旅等费用。

1. 分子生物学实验常用酶类试剂预算3.54万元。

高保真扩增酶 1400元 \times 10支=1.4万元

Taq酶混合液 60元 \times 50支=0.3万元

M-MLV逆转录酶 200元 \times 8支=0.16万元

限制性内切酶 300元 \times 6支=0.18万元

连接酶 200元 \times 5支=0.1万元

其他酶类2000元 \times 7支=1.4万元

2. 各类生物试剂盒预算4.58万元，如DNA提取、RNA提取、质粒提取、PCR产物纯化、反转录试剂盒、实时荧光定量PCR试剂盒等。

质粒提取试剂盒 400元 \times 6盒=0.24万元

DNA提取试剂盒 800元 \times 10盒=0.8万元

DNA纯化试剂盒 400元 \times 6盒=0.24万元

RNA提取试剂盒 700元 \times 6盒=0.42万元

RNA纯化试剂盒 350元 \times 8盒=0.28万元

荧光定量PCR试剂盒 1600元 \times 10盒=1.6万元



- 反转录试剂盒 2500元/盒×4盒=1.0万元
3. 各类常规生化 and 分子生物学试剂预算4.57万元。
- Yeast transformation system 3500元/瓶×5瓶=1.75万元
- Yeast maker carrier DNA 3300元/瓶×4瓶=1.32万元
- X-Gal 200元/瓶×5瓶=0.1万元
- X-gluc 1200元/瓶×2瓶=0.24万元
- IPTG 300元/瓶×2瓶=0.06万元
- 蛋白酶K 7000元/瓶×1瓶=0.7万元
- 蔗糖、丙酮、异戊醇20元/瓶×20瓶=0.4万元。
4. 用于DNA和蛋白质分子量检测Marker预算1.98万元。
- Trans15K DNA Marker 700元/支×8支=0.56万元
- Trans2K Plus DNA Marker 500元/支×10支=0.5万元
- Blue Plus IV Protein Marker 500元/支×5支=0.25万元
- PageRuler Prestained Protein Ladder 2500元/支×2支=0.5万元
- Lambda PFG Ladder 1700元/支×1支=0.17万元
5. 各类实验所需耗材2.93万元。
- 8联排PCR管配套管盖 150元/包×60包=0.9万元
- PCR用96孔板 40元/包×80包=0.32万元
- PCR用硅胶垫400元/包×5包=0.2万元
- 10、200、1000μl 吸头 75元/包×40包=0.3万元
- 500、1500μl离心管 120元/包×20包=0.24万元
- 单个PCR薄壁管 200元/包×20包=0.4万元
- 1.5ml、2ml离心管 50元/包×60包=0.3万元
- 100ml三角瓶、9cm培养皿、100ml密封瓶 20元/个×60个=0.12万元
- 500ml三角瓶、12cm培养皿、250ml密封瓶30元/个×50个=0.15万元
6. 用于水稻种植所需肥料耗材等农资预算1万元。
- 化肥250元/包×10包=0.25万元
- 农药、除草剂100元/包×30包=0.3万元
- 育秧盆 100元/个×20个=0.2万元
- 桶 50元/个×50个=0.25万元



7. DNA测序、引物合成、和遗传转化等测试化验加工预算16.6万元，测算依据如下：

引物合成：引物800条×25元/条=2万元

DNA常规测序：反应数600个×20元/个=1.2万元

籼稻遗传转化服务费：10个载体×4800元/个=4.8万元

粳稻遗传转化服务费：6个载体×3000元/个=1.8万元

基因组DNA重测序：6个样本×2500元/个=1.5万元

转录组测序：24个样本×1500元/个=3.6万元、

激光共聚焦观察：85小时×200元/小时=1.7万元

8. 差旅/会议/国际合作与交流费预算1.6万元

国内学术会议（如全国植物生物学大会等）2次，会议注册费和差旅费预算共0.8万元；

参加国际学术交流会议1次，预算共0.8万元。

9. 出版/文献/信息传播/知识产权事务费预算2.4万元

论文发表版面费等，以2篇为测算依据，平均每篇版面费10000元，共2.00万元；

申请专利1项，专利申请费0.4万元。

三、劳务费（17.4万元，占总经费30.00%）：用于参与课题研究的研究生、聘用的科研辅助人员和临时用工人员的劳务费用。

1. 研究生劳务费14.4万元

参加本课题的博士研究生2人，每人每月补助1200元，每人每年工作10个月，4年研究生劳务费合计9.6万元；参加课题硕士生2名，每人每月补助600元，每人每年工作10个月，4年研究生劳务费共4.8万元。

2. 临工劳务费3万元

聘请临工进行田间插秧、取样、收割等劳动，平均0.75万元/年，4年共3万元。



报告正文

研究内容和研究目标按照申请书执行。



国家自然科学基金项目负责人、依托单位承诺书

国家自然科学基金项目负责人承诺书

本人郑重承诺：我接受国家自然科学基金的资助，严格遵守中共中央办公厅、国务院办公厅《关于进一步加强科研诚信建设的若干意见》《关于进一步弘扬科学家精神加强作风和学风建设的意见》等规定，及国家自然科学基金委员会关于资助项目管理、项目资金管理等各项规章制度，在《计划书》填写及项目执行过程中：

（一）按照《批准通知》《国家自然科学基金资助项目计划书填报说明》的要求填写《计划书》，未自行降低、更改目标任务或约定要求，或缩减研究（研制）内容；

（二）树立“红线”意识，严格履行科研合同义务，按照《计划书》负责实施本项目（批准号：32172010），切实保证研究工作时间，按时报送有关材料，及时报告重大情况变动，不违规将科研任务转包、分包他人，不以项目实施周期外或不相关成果充抵交差；

（三）遵守科研诚信、科研伦理规范和学术道德，认真开展研究工作，对资助项目发表的论著和取得的研究成果按规定进行标注，反对无实质学术贡献者“挂名”，不在成果署名、知识产权归属等方面侵占他人合法权益，并如实报告本人及团队成员发生的违背科研诚信要求的任何行为；

（四）尊重科研规律，弘扬科学家精神，严谨求实，追求卓越，反对浮夸浮躁、投机取巧，不人为夸大学术或技术价值，不传播未经科学验证的现象和观点；

（五）将项目资金全部用于与本项目研究工作相关的支出，并结合科研活动需要，科学合理安排项目资金支出进度。

如违背上述承诺，本人愿接受国家自然科学基金委员会和相关部门做出的各项处理决定。

项目负责人（签字）：

年 月 日

依托单位科研管理部门：

负责人（签章）：

年 月 日

依托单位财务管理部门：

负责人（签章）：

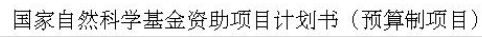
年 月 日

国家自然科学基金项目依托单位承诺书

我单位同意承担上述国家自然科学基金项目，将保证项目负责人及其研究队伍的稳定和研究项目实施所需的条件，严格遵守国家自然科学基金委员会有关资助项目管理、项目资金管理和科研诚信管理等各项规定，并督促实施。

依托单位（公章）

年 月 日



本栏目由自然科学基金委填写	科学处审查意见：	
	负责人（签章）： 年 月 日	
	科学部审查意见：	
	负责人（签章）： 年 月 日	

国家自然科学基金资助项目批准通知

(预算制项目)

谢先荣 先生/女士：

根据《国家自然科学基金条例》、相关项目管理办法规定和专家评审意见，国家自然科学基金委员会（以下简称自然科学基金委）决定资助您申请的项目。项目批准号：32172010，项目名称：水稻粒长基因GL3n的克隆和分子机制研究，直接费用：58.00万元，项目起止年月：2022年01月至2025年12月，有关项目的评审意见及修改意见附后。

请您尽快登录科学基金网络信息系统（<https://isisn.nsf.gov.cn>），**认真阅读《国家自然科学基金资助项目计划书填报说明》并按要求填写《国家自然科学基金资助项目计划书》（以下简称计划书）**。对于有修改意见的项目，请您按修改意见及时调整计划书相关内容；如您对修改意见有异议，须在电子版计划书报送截止日期前向相关科学处提出。

请您将电子版计划书通过科学基金网络信息系统（<https://isisn.nsf.gov.cn>）提交，由依托单位审核后提交至自然科学基金委。自然科学基金委审核未通过者，将退回的电子版计划书修改后再行提交；审核通过者，打印纸质版计划书（一式两份，双面打印）并在项目负责人承诺栏签字，由依托单位科研、财务管理等部门审核、签章并在承诺栏加盖依托单位公章，且将申请书纸质签字盖章页订在其中一份计划书之后，一并报送至自然科学基金委项目材料接收工作组。纸质版计划书应当保证与审核通过的电子版计划书内容一致。**自然科学基金委将对申请书纸质签字盖章页进行审核，对存在问题的，允许依托单位进行一次修改或补齐。**

向自然科学基金委提交电子版计划书、报送纸质版计划书并补交申请书纸质签字盖章页截止时间节点如下：

1. **2021年10月22日16点**：提交电子版计划书的截止时间（视为计划书正式提交时间）；
2. **2021年10月29日16点**：提交修改后电子版计划书的截止时间；
3. **2021年11月5日16点**：报送纸质版计划书（其中一份包含申请书纸质签字盖章页）的截止时间。

三、请评述申请人的研究基础及研究方案的创新性和可行性。
团队具备相应的实验平台，前期工作基础良好，研究方案可行。

四、其他建议

修改意见：

生命科学部

2021年10月12日



SCAUUB202519479

检索证明

根据委托人提供的论文材料，委托人华南农业大学生命科学学院 刘伟智 2 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区	引用
1	DSDecode: A Web Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations	MOLECULAR PLANT 出版年: 2015 出版日期: SEP 7 卷期: 8 9 页码: 1431-1433 文献类型: Letter	第一作者	A 类	华南农业大学 生命科学学院	SCI	IF2-year=7.142 IF5-year=6.885 (2015)	生物 2 区 Top 期刊: 否 (2015)	SCI 核心合集 总引: 222
2	STI PCR: An efficient method for amplification and de novo synthesis of long DNA sequences	MOLECULAR PLANT 出版年: 2022 出版日期: APR 4 卷期: 15 4 页码: 620-629 文献类型: Article	并列第一作者	T2 类	华南农业大学 生命科学学院	SCI	IF2-year=27.5 IF5-year=22.6 (2022)	生物学 1 区 Top 期刊: 是 (2022)	SCI 核心合集 总引: 16

说明：论文等级和中科院大类分区按《华南农业大学学位论文评价方案（试行）》划分。

报告免责声明：如未盖章，报告无效



ISSN (PRINT) 1674-2052 ISSN (ONLINE) 1752-9967 CN 31-2013/Q 分子植物

Molecular Plant

Volume 8
Number 9

September 7, 2015

www.cell.com/molecular-plant
www.mplant.org

CSPB
IPPE, SIBS, CAS

CellPress

DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations

Dear Editor,

The transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 genome editing systems have greatly improved the efficiency for generating targeted mutations in various organisms including plants (Li et al., 2012; Cong et al., 2013; Li et al., 2013; Feng et al., 2014; Ma et al., 2015b; Zhang et al., 2014, 2015). In some plant species, the majority of mutations induced by TALENs and CRISPR/Cas9 systems are in uniform biallelic and heterozygous status in the first transgenic generation, although in some other plant species, chimeric mutations (with three or more allelic edited events within a single individual) may frequently occur (Li et al., 2013; Feng et al., 2014; Zhang et al., 2014, 2015; Ma et al., 2015b). In many cases, it is necessary to determine the mutated sequences of the targeted alleles. However, direct sequencing (with the Sanger method) of PCR amplicons containing such biallelic or heterozygous mutations results in superimposed sequencing peaks starting from the mutation sites. Therefore, cloning of the mutation-containing amplicons and sequencing of multiple clones for each target editing site are required to determine the mutated sequences of the targeted alleles, which is tedious, time consuming, and expensive. Aimed at this problem, we have recently developed a highly reliable Degenerate Sequence Decoding (DSD) method (Ma et al., 2015a) and applied it to decode hundreds of targeted mutation events in rice and *Arabidopsis* (Ma et al., 2015b). The DSD method decodes superimposed sequencing chromatograms in the following steps: (1) starting from the first overlapping-peak position on the chromatogram, manually generate a short degenerate sequence (DS) that is adjacent to the anchor sequence (AS), which sits upstream of the first overlapping-peak; (2) query the DS against the intact reference sequence twice with a sequence analysis program to find the matched sequence(s); and (3) link the AS with the query-matched sequences to generate the allele sequences or, if detecting only one matching hit, generate the second allele sequence by subtracting the allele 1 nucleotides from the degenerate bases. Even though simple and highly efficient, manual operation of this DSD method is still time consuming when decoding a large number of superimposed sequencing chromatograms.

Here, we describe a web-based tool, DSDecode, for automatic decoding of superimposed and regular sequencing chromatograms derived from PCR amplicons containing various types of mutations. Based on the DSD method, DSDecode can manipulate the whole decoding process mentioned above automatically, including reading and generating the AS and DS nucleotides from a corresponding sequencing chromatogram

file. DSDecode is available from SCAU/SCGene (<http://dsdecode.scgene.com/>).

The workflow of DSDecode is illustrated in Figure 1A. The first step in initiating a new task is to paste an intact reference sequence derived from sequencing of the non-edited parental line, which contains the target site and flanking sequences and completely covers the range of the sequencing chromatogram, into the text area. If the reference sequence is obtained from a public database of the species, this sequence must be identical to that of the parental line for genome editing. The second step is to upload a sequencing chromatogram file in .ab1 format for each decoding job. The parameters for the base numbers of AS (default 20 nt) and DS (default 15 nt) can be adjusted if necessary. The decoding process is started by clicking the "Decode" button (Figure 1B). Behind the scenes, the program carries out the following decoding steps automatically. (1) Read the sequencing digital information from the ab1 file and locate the boundary between regular single-peak signals and overlapping-peak signals. (2) Generate an AS and a DS from the sequencing information. (3) Query the DS against the reference sequence for the first hit. (4) Connect the AS and the first query-matched sequence to yield allele 1 sequence. (5) Subtract the nucleotides of allele 1 (in the degenerate part) from the degenerate bases; then join the resulting nucleotides with the AS to yield allele 2 sequence. (6) Align the decoded allele 1 and allele 2 sequences with the reference sequence. (7) Output the decoding results, including the AS and DS, decoded allele 1 and allele 2 sequences, and their alignments with the reference sequence (Figure 1C). In most cases the time for a decoding task, from data upload to display of the result, requires about 13–15 s. One unique ID is generated simultaneously for each decoding job and it can be used for reviewing the corresponding decoding results in the "QUERY" page within a week.

A search with 13–15 bases of a DS usually produces enough uniqueness, while a query with longer DS may increase the risk of interference with the decoding from the high-level noise signals present within the DS region. However, in some instances, one mutated allele may possess a relatively large deletion that is longer than the primary set DS length; in such cases, some base mutations, if any, present at the other allele may not be detected because the mutation site is not completely covered by the primary shorter DS (Supplemental Figure 1A and 1B). To resolve this problem, the program is designed to recognize this type of mutation and automatically reset a longer DS (the bases

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, SIBS, CAS.

Molecular Plant 8, 1431–1433, September 2015 © The Author 2015. 1431

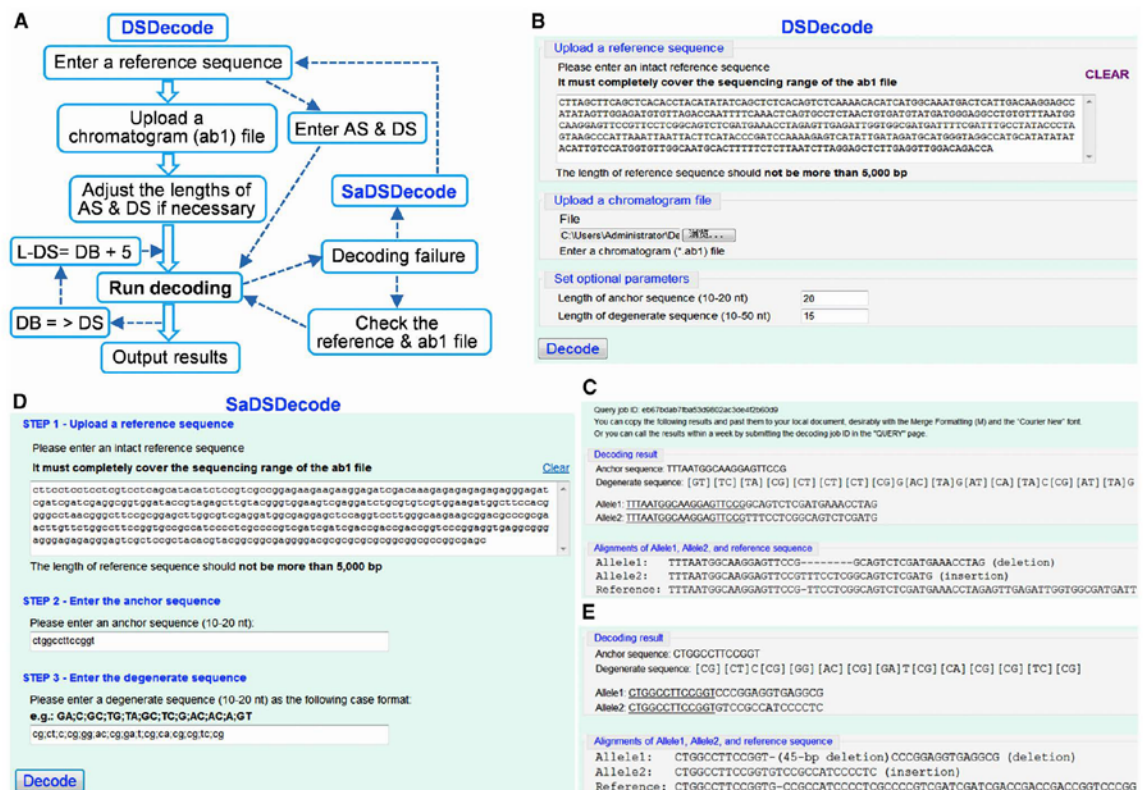


Figure 1. The Workflow of DSDecode/SaDSDecode and the Resulting Output.

(A) The main steps of the DSDecode/SaDSDecode workflow. AS, anchor sequence; DS, degenerate sequence; DB, deleted bases; L-DS, long DS. If a large deletion equal to or longer than the primary DS is detected in an allele, a new L-DS is reset by the program for decoding again. If automatic decoding fails, semi-automatic decoding (SaDSDecode) can be used.

(B and C) Data input areas of DSDecode and the resulting display. Additional “-” indicates a base deletion in the decoded allele(s) or the corresponding position of a base insertion in the reference sequence. In most cases, it is not necessary to change the default settings of 20 nt for AS and 15 nt for DS.

(D and E) Data input areas of SaDSDecode and the resulting display. This case was derived from a chromatogram shown in Supplemental Figure 2C (3).

of large deletion plus 5 nt, see Figure 1A), then run the decoding process again to produce a new result (Supplemental Figure 1C). If a decoding task fails, an inspection is displayed and links to a sub-program Semi-automatic DSDecode (see below).

Sequencing chromatograms with low or intermediate noise signals (Supplemental Figure 2A and 2B) can be effectively decoded by DSDecode. However, high-level noise signals in some low-quality sequencing chromatograms (Supplemental Figure 2C) could interfere with the identification of AS and DS, leading to decoding failure. To decode such low-quality sequencing chromatograms, a sub-program page, Semi-automatic DSDecode (SaDSDecode), is available (Figure 1D), which can be visited directly via <http://dsdecode.scgene.com/home/semidecode>. By inputting the manually generated AS and DS from a sequencing chromatogram as described (Ma et al., 2015a) into the text area, decoding results can be obtained with the sub-program (Figure 1E). To obtain high-quality sequencing chromatograms for DSDecode-based automatic decoding, we recommend that

targeted genomic sequences of about 400–700 bp, in which the targeted sites are located on the center regions, are PCR amplified and purified with a gel purification kit. Then the amplicons are sequenced with internal (nested) specific primers, with the binding positions preferably at 150–300 bp (no less than 100 bp and no more than 400 bp) upstream of the targeted sites. The primers used for the PCR amplification are not suitable for sequencing because they often produce high-level noise signals, which may interfere with the decoding by DSDecode. For multiplex targeting involving two or more closely arrayed sites, an amplicon containing these target sites can be obtained by a single PCR amplification, but independent sequencing for each target site (from both sides) is recommended, although fragmental deletions between two close target sites can be detected by DSDecode.

DSDecode and SaDSDecode can genotype various types of biallelic and heterozygous mutations in diploid organisms as shown previously (Ma et al., 2015a). In addition, the simple types of biallelic and homozygous mutations with single-base insertions

in both alleles and homozygous base deletions can also be analyzed with DSDecode (Supplemental Figure 3). However, chimeric mutations and the rare type, if any, of biallelic mutations with insertion or substitution of two or more bases in both alleles cannot be decoded by this program. To verify the efficiency and accuracy of DSDecode, we tested more than 150 sequencing chromatograms derived from various types of mutations in rice and *Arabidopsis* with DSDecode and further compared the results (Ma et al., 2015b) decoded with the manual DSD method (Ma et al., 2015a). The results showed that 125 chromatograms of high and ordinary quality were successfully and correctly decoded by DSDecode. The remaining chromatograms of low quality, while unable to be properly decoded with DSDecode, could be completely decoded by SaDSDecode or by the manual DSD method. The manual DSD method is also useful for verifying the results obtained with DSDecode or SaDSDecode.

In summary, we present a versatile and user-friendly tool for genotyping of various types of uniform mutations by direct sequencing of PCR amplicons containing targeted sites. Although developed for genome targeting analysis, DSDecode can also be used for genotyping of other source-derived nucleotide variations at single sites of sequencing chromatograms. This tool, in combination with the TALEN and CRISPR/Cas9 technologies, will greatly facilitate basic and applied biological research.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

This work was supported by grants from the Ministry of Agriculture of the People's Republic of China (2014ZX08010-001, 2014ZX08009-002).

AUTHOR CONTRIBUTIONS

W.L. and X.X. wrote the program, X.M. tested the program and wrote the paper, J.L. integrated the program on the server computer, J.C. and Y.-G.L. designed the program, and Y.-G.L. wrote the paper.

ACKNOWLEDGMENTS

We are grateful to Dr. Letian Chen for comments and editing of the manuscript. No conflict of interest declared.

Received: April 28, 2015

Revised: May 20, 2015

Accepted: May 26, 2015

Published: May 29, 2015

Weizhi Liu^{1,2,3,5}, Xianrong Xie^{1,2,3,5},
Xingliang Ma^{1,2,3}, Jun Li⁴, Jiehu Chen⁴ and
Yao-Guang Liu^{1,2,3,*}

¹State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangzhou 510642, China

²Key Laboratory of Plant Functional Genomics and Biotechnology of Guangdong Provincial Higher Education Institutions, Guangzhou 510642, China

³College of Life Sciences, South China Agricultural University, Guangzhou 510642, China

⁴Science Corporation of Gene Ltd, Guangzhou 510540, China

⁵These authors contributed equally to this article.

*Correspondence: Yao-Guang Liu (ygliu@scau.edu.cn)

<http://dx.doi.org/10.1016/j.molp.2015.05.009>

REFERENCES

- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**:819–823.
- Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **111**:4632–4637.
- Li, T., Liu, B., Spalding, M.H., Weeks, D.P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **30**:390–392.
- Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **31**:688–691.
- Ma, X., Chen, L., Zhu, Q., and Liu, Y. (2015a). Rapid decoding of sequence-specific nuclease-induced heterozygous and biallelic mutations by direct sequencing of PCR products. *Mol. Plant* <http://dx.doi.org/10.1016/j.molp.2015.02.012>.
- Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y., et al. (2015b). A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* <http://dx.doi.org/10.1016/j.molp.2015.04.007>.
- Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N., et al. (2014). The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* **12**:797–807.
- Zhang, H., Gou, F., Zhang, J., Liu, W., Li, Q., Mao, Y., Botella, J.R., and Zhu, J.-K. (2015). TALEN-mediated targeted mutagenesis produces a large variety of heritable mutations in rice. *Plant Biotechnol. J.* <http://dx.doi.org/10.1111/pbi.12372>.

ISSN (PRINT) 1674-2052 ISSN (ONLINE) 1752-9967 CN 31-2013/Q 分子植物

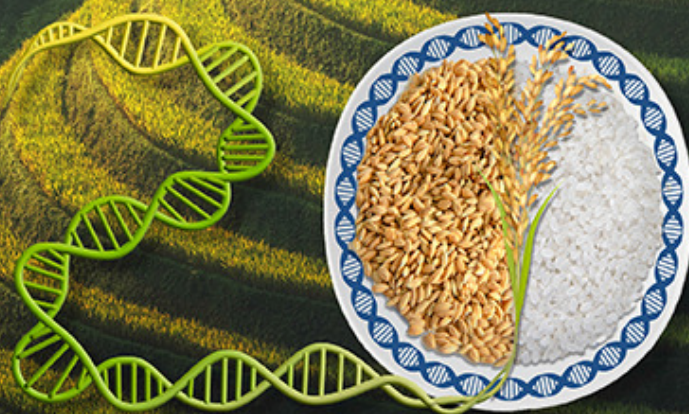
Molecular Plant

Volume 15
Number 4

April 4, 2022

www.cell.com/molecular-plant
www.molplant.org

Decoding Rice Genome



Rice 稻 米 基因组

CSPB
CEMPS, CAS

CellPress
Partner Journal

STI PCR: An efficient method for amplification and *de novo* synthesis of long DNA sequences

Zhe Zhao^{1,2,3}, Xianrong Xie^{1,2,3}, Weizhi Liu^{1,2,3}, Jingjing Huang^{1,2}, Jiantao Tan^{1,2}, Haixin Yu^{1,2}, Wubei Zong^{1,2}, Jintao Tang^{1,2}, Yanchang Zhao^{1,2}, Yang Xue^{1,2}, Zhizhan Chu^{1,2}, Letian Chen^{1,2,*} and Yao-Guang Liu^{1,2,*}

¹State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Life Sciences, South China Agricultural University, Guangzhou 510642, China

²Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou 510642, China

³These authors contributed equally to this article.

*Correspondence: Letian Chen (lotichen@scau.edu.cn), Yao-Guang Liu (ygliu@scau.edu.cn)

<https://doi.org/10.1016/j.molp.2021.12.018>

ABSTRACT

Despite continuous improvements, it is difficult to efficiently amplify large sequences from complex templates using current PCR methods. Here, we developed a suppression thermo-interlaced (STI) PCR method for the efficient and specific amplification of long DNA sequences from genomes and synthetic DNA pools. This method uses site-specific primers containing a common 5' tag to generate a stem-loop structure, thereby repressing the amplification of smaller non-specific products through PCR suppression (PS). However, large target products are less affected by PS and show enhanced amplification when the competitive amplification of non-specific products is suppressed. Furthermore, this method uses nested thermo-interlaced cycling with varied temperatures to optimize strand extension of long sequences with an uneven GC distribution. The combination of these two factors in STI PCR produces a multiplier effect, markedly increasing specificity and amplification capacity. We also developed a webtool, calGC, for analyzing the GC distribution of target DNA sequences and selecting suitable thermo-cycling programs for STI PCR. Using this method, we stably amplified very long genomic fragments (up to 38 kb) from plants and human and greatly increased the length of *de novo* DNA synthesis, which has many applications such as cloning, expression, and targeted genomic sequencing. Our method greatly extends PCR capacity and has great potential for use in biological fields.

Key words: STI PCR, PCR suppression, thermo-interlaced cycling, long genomic fragments, *de novo* DNA synthesis

Zhao Z., Xie X., Liu W., Huang J., Tan J., Yu H., Zong W., Tang J., Zhao Y., Xue Y., Chu Z., Chen L., and Liu Y.-G. (2022). STI PCR: An efficient method for amplification and *de novo* synthesis of long DNA sequences. *Mol. Plant*. 15, 620–629.

INTRODUCTION

PCR for DNA amplification is a foundational, powerful technique with widespread applications in molecular biology, biotechnology, synthetic biology, diagnostics, detection, identification, and forensic analysis (Moore, 2005). Since the invention of PCR (Mullis et al., 1986; Mullis and Faloona, 1987), numerous PCR-derived techniques have been developed. However, some challenges remain. In particular, the efficient and specific amplification of long DNA fragments from complex genomes and synthetic DNA pools, which is often required for analyzing the functions of large genes, multigene clusters, and genomic structural variations and for constructing large functional structures, has long been a vexing problem. In addition, synthetic biology is emerging as an important discipline with the potential to impact

many academic and industrial applications. The engineering of biological systems in synthetic biology often requires *de novo* synthesis of long DNA sequences (synthons) that comprise complete functional units, but current technologies are inefficient for building large DNA synthons (Hughes and Ellington, 2017).

A number of long PCR (or long-range PCR) approaches have been developed to improve the amplification of long DNA fragments, including the use of engineered high-fidelity/high-performance thermo-stable DNA polymerases and secondary

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

proofreading-competent polymerases, optimized buffer compositions, additives, decreased denaturation times, higher annealing temperatures, and longer strand extension times (Cheng et al., 1994a, 1994b; Villbrandt et al., 2000; Davidson et al., 2003; Wang et al., 2004; Hogrefe and Borns, 2011; Ignatov et al., 2014; Chua et al., 2017). Although some cases of the amplification of large fragments (~20 kb or larger) from the human genome have been reported, it is usually difficult to stably and specifically amplify genomic DNA fragments greater than ~10 kb in ordinary PCR experiments. This is especially true for the amplification of long genomic DNA in plants.

In our attempts to amplify long genomic DNA fragments using existing long PCR methods, which generally require more PCR cycles (35–40 cycles), we found that the PCRs often generated non-specific products, which were usually shorter than the target products. Because short non-specific products can be amplified more efficiently than longer target products, we reasoned that these products may strongly compete with the large target products for amplification, thus greatly reducing the amplification efficiency of the target. We therefore speculated that the development of a new long PCR method that avoids amplification of non-specific products could improve both the capacity and the specificity of amplifying long DNA sequences from genomes.

When a DNA fragment contains inverted terminal repeats (ITRs) of a certain length (which have a relatively higher melting temperature [T_m]), the self-complementary ends of the denatured single strands can pair during PCR to form duplex stems or panhandle-like structures, blocking primer binding and producing an effect termed PCR suppression (PS) (Lukyanov et al., 1995, 1999; Siebert et al., 1995; Diatchenko et al., 1996). Shorter fragment strands with ITRs of certain lengths are especially prone to forming such stem-loop structures, thus producing a stronger PS effect (Shagin et al., 1999). This PS effect has been utilized in some methods to construct subtracted or size-balanced cDNA libraries (Lukyanov et al., 1995; Diatchenko et al., 1996; Dai et al., 2007), isolate cDNA ends (Matz et al., 1999), and recover unknown flanking DNA sequences by thermal asymmetric interlaced PCR (Liu and Chen, 2007; Tan et al., 2019).

In most natural DNA sequences, the levels of G and C nucleotides (GC content) are unevenly distributed, especially for long sequences. Regions with low GC contents have low T_m values, and during the strand extension step of PCR at higher temperatures (e.g., 68°C–72°C), the base pairing in these regions is therefore unstable, decreasing the efficiency of strand synthesis. However, for high-GC regions, the activity of the DNA polymerase used and the strand extension are mostly optimal at higher temperatures (e.g., 72°C). A recently developed nested-loop PCR method improves the amplification efficiency of genomic sequences to some extent by changing the temperatures (58°C–68°C) for strand extension within PCR cycles (Long et al., 2019), but this method often produces non-specific products that impair the amplification of long fragments from complex genomes.

In this study, we developed a suppression thermo-interlaced (STI) PCR method that uses modified site-specific primers to generate the PS effect and a thermo-interlaced cycling (TIC) strategy with varied temperatures for strand extension to efficiently amplify

long DNA sequences with high specificity from complex genomes of various organisms for various applications. This method is also powerful for increasing the capacity for *in vitro* assembly and amplification of long *de novo*-synthesized DNA.

RESULTS

Design of STI PCR

The first principle underlying STI PCR is the use of the PS effect to suppress the amplification of non-specific PCR products, which are usually smaller than target products, to avoid their competitive amplification along with the specific products and thus to enhance the amplification of the target sequence. PCR amplicons with ITRs that have higher T_m values produce stronger PS (Lukyanov et al., 1995; Shagin et al., 1999). For this purpose, the forward and reverse site-specific primers for STI PCR, denoted PS primers, contain 3' regions that can specifically bind to genomic sites (with calculated T_m values of ~62°C–64°C for these initial binding sites) and an identical non-target 5' tail (5' tag) consisting of an arbitrary GC-rich sequence (20–25 nt) with a higher T_m of 68°C–72°C (Figure 1A). Examples include the 5' tags 5'-gcctggctccacgctccgagtg-3' with T_m = 68°C, 5'-gcctggctccacgctccgagtg-3' with T_m = 69.6°C, and 5'-gcctggctccacgctccgagtg-3' with T_m = 71.3°C used in this study (Supplemental Table 1). The higher T_m values ensure the stable pairing of ITRs at 68°C–72°C during the annealing and extension steps.

Primer concentration is another factor that affects PS (Lukyanov et al., 1995; Shagin et al., 1999). When primers are used at lower concentrations (such as 0.1–0.15 μ M used in this study compared with 0.3–0.5 μ M used in conventional PCR), the strands of non-specific products < ~3 kb easily form stem-loop structures, and their amplification is blocked or strongly suppressed. This occurs because intra-molecular annealing of the complementary ends of shorter product strands is kinetically favored over inter-molecular annealing between the primers and templates at lower primer concentrations. However, the PS effect is relatively weak for longer (e.g., >~4 kb) products because the longer distances between the complementary ends effectively decrease stem formation. Thus, the primers can bind to the template ends for priming and amplification (Figure 1B).

During STI PCR, non-specific products (including primer-dimers) may initially be generated by mispairing and priming, but most of these are smaller than the large target products. Although some large non-specific products may also be generated initially, they are amplified poorly because the efficiency of their initial off-target mispairing/priming is much lower than that of the on-target pairing/priming to target sequences. Therefore, these differential PS effects suppress the amplification of smaller non-specific products while simultaneously allowing efficient amplification of long target sequences. We reasoned that this PCR design could eliminate the competitive amplification of non-specific products, thus greatly promoting efficient, specific amplification of large target fragments.

The second principle is the use of nested TIC programs with step-wise- or ramping-varied temperature ranges (e.g., 60°C–70°C, 65°C–72°C, or 60°C–72°C) to optimize strand extension of target

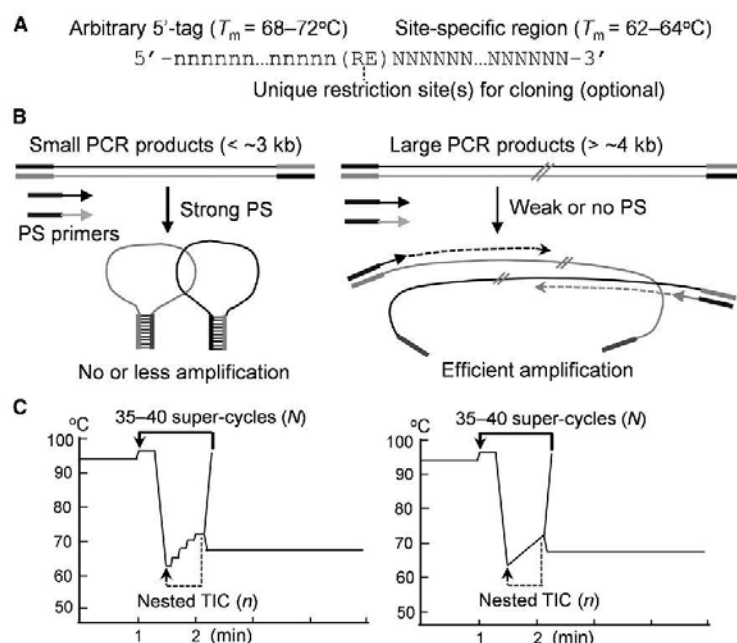


Figure 1. Diagrams illustrating STI PCR.

(A) Constitution of typical PCR suppression (PS) forward and reverse primers. The 5' tag sequence (GC rich) is identical in the forward and reverse PS primers.

(B) During the primer-binding and strand extension steps of STI PCR with PS primers (in relatively low concentrations, such as 0.1–0.15 μM), denatured single strands of smaller, non-specific products (<~3 kb, including primer-dimers) tend to form stable stem-loop structures with a strong PS effect, whereas those of large target products (>~4 kb) form stems much less efficiently. Thus, the amplification favors large target products.

(C) Representative thermo-cycling program patterns for STI PCR. An STI PCR program contains a total of 35–40 (N) super-cycles (complete thermo-cycles with denaturing, annealing and extension steps). In each super-cycle, the annealing and extension steps consist of repeated, multiple cycles (n) of nested thermo-interlaced cycling (TIC) with stepwise (left) or ramping (right) changes in temperature, thus optimizing the strand extension of target sequences with different GC nucleotide contents and distributions. The number n is mainly based on the sequence length.

sequences with unevenly distributed GC nucleotide content (Figure 1C). This approach works because relatively lower temperatures (e.g., 60°C – 65°C) promote strand extension of regions with lower GC content or AT-rich stretches, whereas higher temperatures (e.g., 68°C – 72°C), at which the selected DNA polymerase has higher activity, are optimal for the extension of regions with higher GC content. Importantly, the combination of both the PS effect and TIC in our STI PCR produces a multiplier effect that scales up both the specificity and amplification ability for long target sequences.

In amplifying some long genomic fragments (e.g., >~30 kb), it may still be difficult to stably amplify target products to high concentrations by one-round STI PCR. Thus, we performed a second PCR with nested site-specific primers that do not need to have the PS effect because of the limited number (18–23) of thermo-cycles used (using TIC for strand extension) to obtain long target products with high specificity at high concentrations.

STI PCR program conditions

To test the effectiveness of this method, we first designed several pairs of specific PS primers (Supplemental Table 1) to amplify genomic fragments (single copy in the genome) of various sizes from rice (*Oryza sativa* with genome size of 466 Mb). For comparison, we also prepared six sets of standard specific primers (with T_m values of $\sim 65^\circ\text{C}$) (Supplemental Table 1) with the same 3' site-specific regions as the above PS primers and used them to amplify the same 4.9–14.0 kb fragments.

To simplify the PCR program settings for target sequences with various GC content distributions, we designed four basic programs for STI PCR (Supplemental Figure 1). These programs have about 35–40 super-cycles (complete thermo-cycles with

denaturing, annealing, and extension steps), and in each super-cycle, the annealing/extension steps consist of specific numbers of nested thermo-interlaced cycles. Program I is for target sequences with a moderate GC content distribution (average 45%–55% GC with or without less GC-rich and/or AT-rich stretches). Program II is for target sequences in which some regions (≥ 200 bp) have higher GC contents ($\geq 70\%$ GC). A higher denaturing temperature (97°C , or 98°C if very GC-rich regions are present) is used in this program for complete DNA denaturing. Program III is for target sequences in which some regions (≥ 100 bp) have a lower GC content (<25% GC) or contain some long AT-rich stretches. Program IV is for target sequences that contain both high-GC regions (≥ 200 bp, $\geq 70\%$ GC) and low-GC regions (≥ 100 bp, <25% GC). We used stepwise changes in extension temperature in the nested TIC, although changes in extension temperature could also be ramped. A fifth program, program V, is also provided for two-round STI PCR based on the selection of one of the basic programs (I–IV) described above.

We used the high-fidelity/high-performance thermo-stable DNA polymerase KOD FX Neo (Toyobo, Japan) and ApexHF HS DNA Polymerase CL (Accurate Biotech, Hunan, China) kits to perform our PCRs. Because the use of high-quality genomic DNA (large size and high purity) as a template is important for long PCR, we prepared high-quality total DNA from rice, maize (*Zea mays*), and human cell lines. Most sequence fragments in the prepared DNA samples were larger than ~70 kb, as measured by pulsed-field gel electrophoresis (Supplemental Figure 2).

Development of a web-based tool, calGC

To facilitate analysis of GC content and distribution in target sequences, we developed the user-friendly web-based tool calGC (<http://skl.scau.edu.cn/calGC/>) in Python. After a target sequence

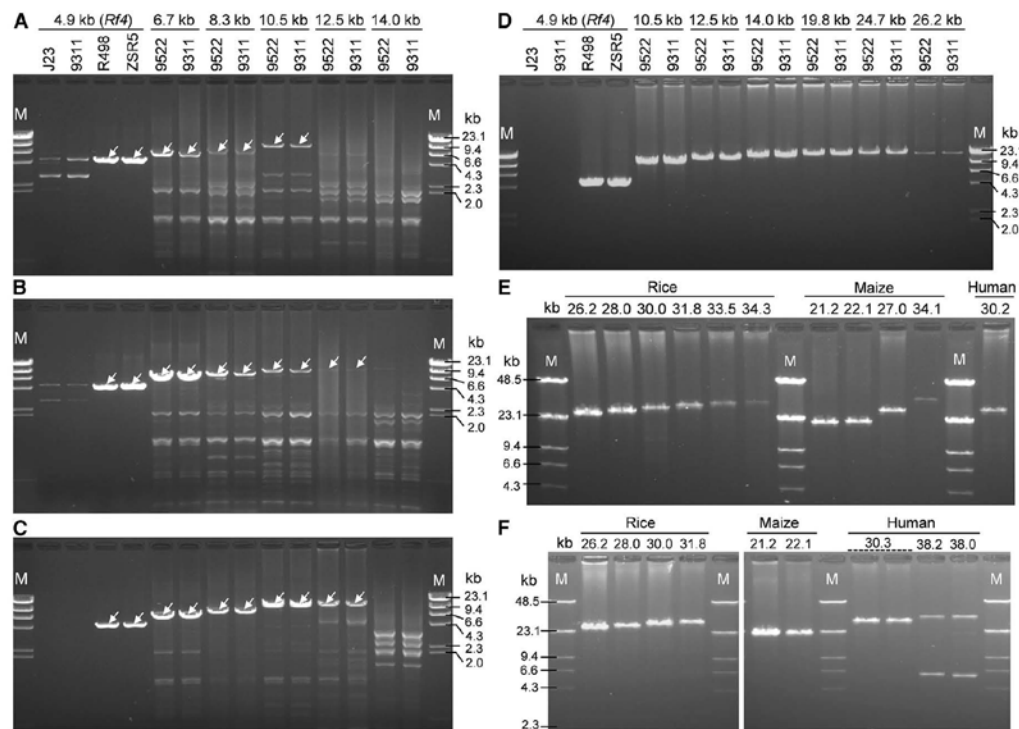


Figure 3. STI PCR enables specific amplification of long genomic DNA fragments.

(A–C) (A) PCRs using standard primers and conventional thermo-cycling conditions. The target fragments in (A–C) are indicated by arrows, and the others are non-specific products. The standard (and PS) primers for the 4.9-kb fragment were designed to amplify the *Rf4* restorer gene in restorer lines R498 and ZSR5 but not the non-functional *rf4* allele in non-restorer lines J23 and 9311. The bands in J23 and 9311 (A and B) therefore represent non-specific products. 9522 and 9311 are *japonica* and *indica* rice lines, respectively, which are the same in (B and C). These PCR products (each 3–4 μ l) were separated by conventional electrophoresis on 0.8% agarose gel (A–D). M, molecular weight marker λ -Hind III. (B) PCRs using standard primers and TIC. (C) PCRs using PS primers and conventional thermo-cycling. (D) STI PCRs using PS primers and TIC. The fragments of 4.9, 10.5, 12.5, and 14.0 kb are the same target sequences used in (A–C). The sizes of all amplified target fragments in (A–D) fit the known sequences in the rice genome (note that DNA fragments of >20 kb could not be effectively separated by conventional agarose gel electrophoresis). The PCRs in (A–D) were performed using KOD FX Neo. (E) Additional one-round STI PCR (using ApexHF HS CL) for amplification of very long genomic sequences from rice, maize, and a human cell line. These PCR products (3–4 μ l each) were separated by pulsed-field gel (1.0% agarose) electrophoresis. M, mixed size markers of intact λ (48.5 kb) and λ -Hind III. (F) Two-round STI PCR for amplification of very long genomic fragments with high specificity. After the first-round STI PCR (using KOD FX Neo), a small amount (0.3–0.5 μ l) of the reaction was used as a template for the second PCR using nested specific primers. These can contain an additional 5' tag for cloning or barcoded sequencing. The second-round PCR products were separated by pulsed-field gel electrophoresis. An internal repeat sequence (~5.8 or ~6.0 kb) that was present in the 38.0- and 38.2-kb target sequences (at the *CYP1A1/CYP1A2* locus) of the human cell lines could be efficiently amplified by the nested primers.

Although one-round STI PCR can amplify long genomic sequences, we also carried out a two-round PCR procedure consisting of a first-round STI PCR (32–33 super-cycles) to greatly enrich the long target sequences and a second PCR (using ~0.3–0.5 μ l first-round PCR product as the template) with nested specific primers (without the PS tag and with T_m values of 64°C–66°C for the specific binding sites). In the second-round nested PCR with such a limited cycle number (18–23 super-cycles), only specific sequences can be amplified to high concentrations. This approach could amplify long genomic fragments (up to 38.2 kb) from rice, maize (with a genome size of ~2.3 Gb), and human cell lines (with a genome size of ~3.0 Gb) (Figure 3F). In the second PCR, the primers can also be modified with 5' tags for various purposes, such as cloning (e.g., introducing Gibson

Assembly [Gibson et al., 2009] tags or restriction enzyme sites) and sequencing (e.g., introducing adaptors/barcodes for pooled single-molecule sequencing).

To confirm the specificity of the amplified fragments, we selected 10 fragments from one-round and two-round STI PCR for DNA restriction analysis. The restriction patterns of all analyzed fragments matched the patterns expected based on their known sequences (Supplemental Figure 3).

DNA sequences with relatively high or low GC contents or that contain internal forward and/or reverse repeats may show impaired PCR amplification. Sequence analysis showed that our amplified genomic sequences included high-GC and/or

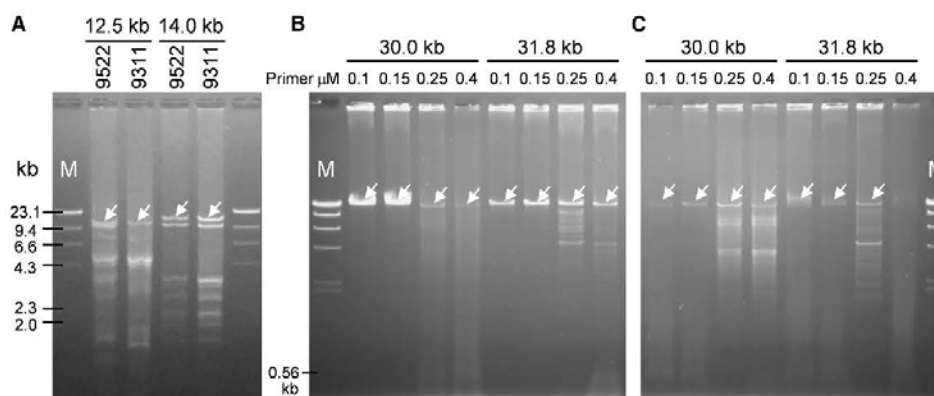


Figure 4. Factors that affect STI PCR performance.

(A) STI PCR using a forward primer (S58-Fm) modified from the PS primer S58-F by changing seven bases (G-to-C or C-to-G) in the 5' tag without changing the total primer T_m . The primer pairs S58-Fm/S58-12.5R and S58-Fm/S58-14R (0.15 μM each) (Supplemental Table 1) were used to amplify the same target fragments (12.5 and 14.0 kb) shown in Figure 3D that used the PS primer pairs S58-F/S58-12.5R and S58-F/S58-14R, respectively. These reactions produced weaker target fragment bands (arrowed), along with many non-specific products.

(B) One-round STI PCR using different concentrations of PS primers (and ApexHF HS CL) to amplify two genomic sequences of 9522 (arrowed, the same as those shown in Figure 3E). Weaker target bands and many non-specific products (in smear or bands) were produced in the reactions that used 0.25 and 0.4 μM primers.

(C) One-round STI PCR using modified PS primers (and ApexHF HS CL) with a lower T_m (57°C) in a shortened 5' tag (16 nt, Supplemental Table 1) to amplify the same genomic sequences shown in (B). Weaker or no target bands and many non-specific products (in smear or bands) were produced in these reactions.

low-GC regions (Supplemental Figure 4), and some contained internal forward and/or reverse repeats of various sizes (Supplemental Figure 5). Therefore, STI PCR is capable of amplifying PCR-recalcitrant genomic sequences.

Factors affecting the performance of STI PCR

To confirm that the high performance of STI PCR could be attributed to the PS effect rather than simply to the increased primer T_m of the longer primer with the additional 5' tag, we changed seven G and C nucleotides in the 5' tag of the forward primer S58-F to generate S58-Fm (Supplemental Table 1) while retaining the same total T_m value. The strand ends of PCR products obtained using this primer pair cannot form a stable stem structure. Two PCR tests using S58-Fm (in combination with the reverse primers S58-12.5R and S58-14R, primarily used for STI PCR) and TIC conditions produced much lower amounts of the target products (12.5 and 14.0 kb) and many non-specific products (Figure 4A) compared with the same pure target products (Figure 3D) amplified using the PS primer pairs S58-F/S58-12.5R and S58-F/S58-14R.

To determine whether primer concentration affects PS, amplification specificity, and amplification ability in STI PCR, we amplified two rice genomic fragments (30.0 and 31.8 kb) using different primer concentrations. Compared with the good performance obtained with PS primer concentrations of 0.1 or 0.15 μM, many non-specific products of various sizes and much weaker target sequence bands were obtained with primer concentrations of 0.25 or 0.4 μM (Figure 4B).

We next modified two PS primers with a lower T_m (57°C) in a shortened 5' tag (16 nt, Supplemental Table 1) for amplification of the same 30- and 31.8-kb sequences. STI PCR with these modified primers (at different concentrations) also produced

more non-specific products and weaker (or no) target bands (Figure 4C). These tests confirmed that PS strength is negatively related to primer concentration and demonstrated the important role of strong PS for enhancing the specific amplification of long DNA sequences.

STI PCR markedly increases the capacity for *de novo* DNA synthesis

Current *de novo* DNA synthesis technologies mainly include steps for design and synthesis of DNA oligonucleotides that cover target sequences, *in vitro* assembly of oligonucleotides into primary templates of specific lengths, PCR amplification of the fragments (synthons), cloning and sequencing verification of the synthons, and, if necessary, further assembly (*in vitro* or *in vivo*) of the synthons into larger assemblies or functional devices (Hughes and Ellington, 2017). Because of current technical limitations on the efficiency of synthesis of long DNA oligonucleotides (<200 nt, usually much shorter for quality assurance), oligonucleotide assembly, and amplification of assembled templates from the multi-oligonucleotide reaction pool, the length of synthons built by such one-round *in vitro* synthesis/assembly is usually limited to approximately 3 kb (Hughes and Ellington, 2017).

To test whether STI PCR can be used to overcome these limitations, we designed three synthons of 3.4, 4.2, and 7.0 kb with uneven GC distributions. Based on polymerase cycling assembly (PCA, also called overlapping PCR) (Stemmer et al., 1995), the most widely used and cost-effective method for *in vitro* assembly of DNA synthons (Hughes and Ellington, 2017), we designed 32, 38, and 58 oligonucleotides (each 110–150 nt in length with overlapped ends of 16–18 nt) (Figure 5) and synthesized them by the chemical method for assembly of 3.4-, 4.2-, and 7.0-kb

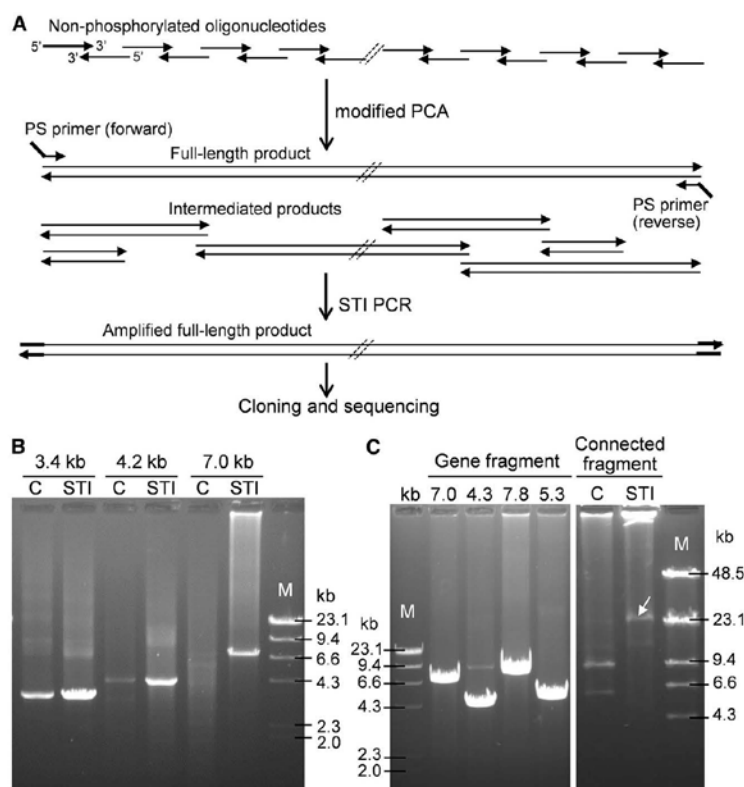


Figure 5. The use of modified polymerase cycling assembly and STI PCR markedly increases the capacity to perform *de novo* DNA synthesis.

(A) Diagram of *de novo* synthesis of a long target DNA (synthon) from multiple overlapping oligonucleotides using modified polymerase cycling assembly (mPCA) and STI PCR.

(B) For *de novo* synthesis of 3.4-, 4.2-, and 7.0-kb DNA sequences, 32, 38, and 58 oligonucleotides (each 110–150 nt long with overlapped ends of 16–18 nt) were used for mPCA reactions (using ApexHF HS CL) with nested TIC (ranging from 62°C to 72°C) for strand extension and conventional PCA reactions with strand extension at 72°C to serve as controls. The mPCA and conventional PCA products (1.5–2 µl each) were then used for STI PCR (using PS primers) and conventional (C) PCR (using standard primers), respectively, to amplify the synthons.

(C) Assembly of four DNA fragments into a large fragment by mPCA and STI PCR. Left, the PCR-amplified DNA fragments with overlaps of 25–30 bp. Right, after mPCA and conventional PCA of the four fragments, the templates were amplified using STI PCR (STI) and conventional PCR (C), respectively. The target 24.4-kb sequence (arrowed) was produced only by mPCA/STI PCR.

synthons, respectively. We carried out modified PCA (mPCA) reactions (with ApexHF HS CL and the pooled oligonucleotides) using nested TIC (ranging from 62°C to 72°C) for strand extension (Figure 5A), and we performed conventional PCA reactions with strand extension at 72°C for comparison. Small amounts of the mPCA and conventional PCA products were then used for STI PCR (using PS primers) and conventional PCR (using standard primers), respectively, to amplify the synthons. Using the improved PCA and STI PCR methods, the 3.4-, 4.2-, and 7.0-kb synthons were generated and specifically amplified to high concentrations, whereas the conventional PCA and PCR methods produced only the 3.4- and 4.2-kb synthons, the latter at very low concentrations and accompanied with large amounts of intermediate products (Figure 5B). We cloned and sequenced the synthons and confirmed that all the oligonucleotides were correctly assembled into the synthetic sequences as designed.

To assess the usefulness of STI PCR for amplification of large DNA sequences assembled from multiple fragments, we PCR amplified four DNA fragments (7.0, 4.3, 7.8, and 5.3 kb) with overlaps of 25–30 bp (Figure 5C), linked them by mPCA, and then amplified the templates using STI PCR (and conventional PCR for comparison). STI PCR could amplify the assembled target sequence (24.4 kb), but conventional PCR could not (Figure 5C). These tests demonstrate that STI PCR (and the TIC used in mPCA) is very useful for markedly increasing the capacity for DNA synthesis and assembly.

Other applications of STI PCR

In functional genomics and biotechnology studies, it is often necessary to clone large genomic fragments in vectors for transformation, expression, and functional analyses.

Gibson Assembly is an efficient method for DNA fragment linking and cloning without the need for restriction enzyme digestion to generate cohesive ends in insert sequences (Gibson et al., 2009). Therefore, we prepared nested specific primers that contained additional unique sequences (20 and 25 bases) at the 5' tags (Supplemental Table 1) for Gibson Assembly cloning. We amplified two rice genomic sequences of 9.8 kb (containing one gene) and 22.9 kb (containing four functionally related genes) by two-round STI PCR (Supplemental Figure 6A). Using Gibson Assembly, we cloned these fragments into a binary vector (modified from pCAMBIA-1300) and a transformation-competent artificial chromosome (TAC)-based vector (Zhu et al., 2017), respectively (Supplemental Figure 6B), and we transferred these constructs into the rice genome for functional studies.

Current targeted genomic sequencing systems typically use synthetic oligonucleotides fixed on chips as capture reagents to enrich the genomic sequences of target regions. Another strategy for targeted genomic sequencing is based on PCR amplification of target genomic regions, with amplified fragments usually <10 kb. Here, we used one-round STI PCR (with KOD FX Neo) to amplify a target genomic region of ~64 kb in a rice mutant line (and a wild-type line) into four overlapping fragments (19.8, 20.8, 12.8, and 12.7 kb in the wild-type genome), and we detected a 5.8-kb insertion in this region of the mutant (Supplemental Figure 7A). We pooled these fragments from the mutant to construct an Illumina sequencing library for deep

sequencing. Based on the sequencing data, we estimated that the nucleotide mutation rate of the amplified products was 0.00012, about 1/12 of that obtained using the Taq DNA polymerase (Supplemental Figure 7B), thus demonstrating the high fidelity of DNA synthesis during the PCR.

DISCUSSION

Although numerous PCR-derived techniques have been developed and widely used in many fields of biology and biotechnology, amplification and specificity limitations of current PCR methods hinder the potential wider uses of related techniques. In this study, we utilized the principles of PCR suppression and TIC to develop a simple, robust, and powerful method, STI PCR, that enables efficient and specific amplification of long sequences from complex genomes and synthetic DNA pools.

In long PCR methods, a number of factors affect the ability to specifically amplify long DNA sequences from complex templates, and many efforts have been made to improve PCR performance (Cheng et al., 1994a, 1994b; Villbrandt et al., 2000; Davidson et al., 2003; Wang et al., 2004; Hogrefe and Boms, 2011; Ignatov et al., 2014; Chua et al., 2017). Nonetheless, the potential for long PCR has not been fully realized, and in practice, it is still often difficult to amplify large DNA in many applications. We assumed that the generation of non-specific products and their competitive amplification during PCR might substantially decrease the amplification efficiency of long target sequences and that blocking this competitive amplification would enhance the amplification of target sequences. In this study, we utilized a strong PS effect (through the use of PS primers with high ITR T_m and relatively low primer concentrations) to effectively suppress non-specific products, eliminating their competition with target products. In thermal asymmetric interlaced PCR, TIC is used for differential annealing of long, specific primers and short, arbitrary primers that have high and low T_m values to target and non-target sequences, respectively (Liu and Whittier, 1995). In nested-loop PCR (Long et al., 2019) and STI PCR, a similar TIC principle is used for optimal strand extension. We found that the use of either nested TIC or the PS effect alone improved the amplification efficiency and/or specificity to some extent (Figure 3B and 3C), but these individual effects were not sufficient to greatly scale up the amplification performance. However, the combination of these two factors in our STI PCR produces an obvious multiplier effect, enhancing both the amplification ability and specificity for long target sequences. For PCR amplification of DNA sequences of several kb using standard primers, we recommend using nested TIC to improve amplification effectiveness.

The limited capacity for amplification of long genomic DNA by current long PCR methods is generally attributed to the limited DNA synthesis activities of the DNA polymerases used. It is true that different thermo-stable DNA polymerases may have different performances in long PCR. We tested several thermo-stable high-fidelity DNA polymerases and found that KOD FX Neo and ApexHF HS CL had relatively higher amplification performance in STI PCR, but other thermo-stable high-fidelity DNA polymerases that we did not test may also be suitable for this method. However, our results showed that these same DNA polymerases (KOD FX Neo and ApexHF HS CL) could not efficiently amplify long DNA sequences when conventionally de-

signed primers and normal thermo-cycling conditions were used (Figures 3A–3C, 5B, and 5C).

Our STI PCR overcomes the limitations of current PCR methods in terms of amplicon size and specificity. Importantly, this method can also greatly scale up the capacity for *de novo* DNA synthesis. Our one-round *in vitro* synthesis tests successfully created large *de novo* DNA synthons up to 7.0 kb, but this may not be the size limit of the synthons, as this 7.0-kb sequence was produced at a high concentration with high specificity (Figure 5B). This method can also assemble multiple DNA fragments into large ones, as shown in Figure 5C, and is therefore useful for the preparation of multigene stacking constructs for genetic engineering of complex pathways and multiple traits. The large products amplified by STI PCR can be used for cloning, functional analysis, and direct sequencing (Sanger sequencing, next-generation sequencing, and single-molecule sequencing), and they may also enable other new applications. The associated webtool, calGC, also facilitates the analysis of DNA GC content and distribution features in routine molecular biology studies. Together with STI PCR, it represents an important technical advance that will greatly promote the development of synthetic biology, molecular biology, and biotechnology.

METHODS

Preparation of genomic DNA

High-quality genomic DNA was prepared from fresh leaves of ~30-day-old rice and maize plants using the CTAB method (Mazars et al., 1991; Zhou et al., 2016). High-quality genomic DNA was isolated from human cell lines (QSG-7701 and HEK293T) using a Universal Genomic DNA Kit (CW2298M, CoWin Biosciences, Beijing, China) according to the manufacturer's instructions.

Sequence analysis and primer design

The specificity of the primers to the corresponding reference genomes was examined by nucleotide BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi/>) using the algorithm "Somewhat similar sequences (blastn)" or by Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). T_m values of the primers were calculated according to the formula $T_m = 69.3 + 41 \times \text{GC\%} - 650/L$ (L = primer length) (Mazars et al., 1991). The primers used in this study (Supplemental Table 1) were synthesized and purified (with desalting) by Tsingke Biotechnology (China).

Development of calGC

The web-based tool calGC was developed in Python 3.8. The interface of calGC was implemented using Django as a backend program. The distribution of GC content along a given sequence is plotted by JavaScript of ECharts.

PCR conditions

High-fidelity/high-performance DNA polymerase kits of KOD FX Neo (Toyobo, Japan) and ApexHF HS DNA Polymerase CL (Accurate Biotech, Hunan, China), include buffers, were used for PCR tests with an enzyme concentration of 1 unit per 100 μl reaction for <20-kb target sequences (such as 0.4 U in 40 μl reaction/tube) or 1 unit per 90 μl reaction for >20-kb target sequences (such as 0.45 U in a 40 μl reaction/tube). It is notable that these reduced enzyme concentrations for STI PCR produced better amplification efficiency than that obtained using the standard concentration (1 unit per 50 μl reaction) recommended by the manufacturers. Therefore, for STI PCR using other DNA polymerases, different enzyme concentrations should be tested. One-round PCR (30 or 40 μl per reaction in 1 \times buffer with 200 μM each of dNTPs, DNA polymerase as above, and

Molecular Plant

~40 ng genomic DNA but without the primers) was run on a Biometra TAdvanced 96 S/SG thermocycler (Biometra, Germany) with a suitable thermo-cycling program. When the temperature reached ~90°C, 2.5 µl (or 3.0 µl) of the primer mix (2.0 µM each) was added (~0.1–0.15 µM final concentration) for manual hot start. Preliminary tests using different concentrations of PS primers (such as 0.1, 0.15, and 2.0 µM) may be carried out. For two-round PCR, the first-round STI PCR (20 µl per reaction) containing ~20–30 ng genomic DNA and the primers (0.1–0.15 µM each) was performed (using hot start) for 32–33 super-cycles. Then the first-round PCR product (0.3–0.5 µl) was used as the template for second-round PCR (30–40 µl) with nested site-specific primers (0.2 µM each, using hot start) for 18–23 super-cycles. The PCR products (each 3–4 µl) were run on 0.8% agarose gels for conventional electrophoresis or on 1.0% agarose gels for pulsed-field gel electrophoresis using a clamped homogeneous electric fields device (Bio-Rad, USA).

In vitro DNA synthesis and assembly

Oligonucleotides (110–150 nt each with overlapped ends of 16–18 nt and $T_m = 56^\circ\text{C}$ – 58°C) were designed (Figure 5) and synthesized by the chemical method (using de-salt purification) by Vazyme Biotech. The first-round oligonucleotide assembly reactions (20 µl each) by mPCA and conventional PCA were prepared using the ApexHF HS CL kit and contained sets of oligonucleotides (0.01–0.015 µM each). The mPCA reactions for the assembly of the 3.4-, 4.2-, and 7.0-kb synthons were performed using 22, 25, and 32 super-cycles of basic program IV. Each super-cycle included annealing at 57°C for 15 s and TIC numbers of 4, 6, and 10, respectively. The corresponding control PCA reactions were performed using 22, 25, and 32 cycles with annealing at 57°C for 15 s and extension at 72°C for 2, 3, and 5 min, respectively. The mPCA and control PCA products (1.5–2.0 µl each) were then used for STI PCR and conventional PCR (40 µl each using the ApexHF HS CL kit), respectively, to amplify the synthons. The STI PCRs (using 0.15 µM PS primers, Supplemental Table 1) were performed using 26 (for 3.4- and 4.2-kb synthons) or 30 (for the 7.0-kb synthon) super-cycles as described above for mPCA (basic program IV). The control PCRs (using 0.15 µM standard primers, Supplemental Table 1) were performed using 26 (for 3.4- and 4.2-kb synthons) or 30 (for the 7.0-kb synthon) cycles with annealing at 62°C for 15 s and extension at 72°C for 2, 3, and 5 min, respectively. For assembly of the four PCR-amplified fragments, about 30 ng of the fragments (each 6–8 ng) were added in 15 µl mPCA (15 super-cycles of program IV) or conventional PCA reactions (15 cycles with annealing at 65°C for 1 min and extension at 72°C for 18 min). Then 1.5 µl of the template was used for amplification by STI PCR (33 super-cycles of program IV) or conventional PCR (33 cycles with annealing at 65°C for 1 min and extension at 72°C for 22 min).

Agarose gel electrophoresis

The PCR products (3–4 µl each) were separated on 0.8% agarose gels by standard electrophoresis or on 1.0% agarose gels (for some products >20 kb) by pulsed-field gel electrophoresis using a clamped homogeneous electric fields device (Bio-Rad).

Cloning of long fragments

The rice genomic fragments amplified by two-round STI PCR were cloned into the modified pCambia-1300 binary vector (vector 1) or the TAC vector (vector 2) (Zhu et al., 2017) by the Gibson Assembly method using our homemade Gibson reaction as described previously (Gibson et al., 2009; Jiang et al., 2013).

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

FUNDING

This work was supported by the National Natural Science Foundation of China (31991222, 32030080 and 31760300), the Key Research Program of the Guangzhou Science, Technology and Innovation Commission

Amplification and *de novo* synthesis of long DNA sequences

(201904020030), the Major Program of Guangdong Basic and Applied Basic Research (2019B030302006), and the China Postdoctoral Science Foundation (2020M682730).

AUTHOR CONTRIBUTIONS

Y.-G.L. designed the research. Z.Z., X.X., W.L., J.H., J. Tan, H.Y., W.Z., J. Tang, Y.Z., Y.X., and Z.C. performed the experiments. X.X. developed the webtool. Y.-G.L. and L.C. wrote the paper.

ACKNOWLEDGMENTS

We thank Dr Jikai Wen for providing DNA samples from human cell lines. The authors have submitted a patent application (202111051247.2) based on the results reported in this paper, which is under review by the National Intellectual Property Administration, China. The authors declare no competing interests.

Received: October 19, 2021

Revised: December 6, 2021

Accepted: December 25, 2021

Published: December 27, 2021

REFERENCES

- Cheng, S., Chang, S.Y., Gravitt, P., and Respass, R. (1994a). Long PCR. *Nature* **369**:684–685.
- Cheng, S., Fockler, C., Barnes, W.M., and Higuchi, R. (1994b). Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci. U S A* **91**:5695–5699.
- Chua, E.W., Maggo, S., and Kennedy, M.A. (2017). Long fragment polymerase chain reaction. *Methods Mol. Biol.* **1620**:65–74.
- Dai, Z.M., Zhu, X.J., Chen, Q., and Yang, W.J. (2007). PCR-suppression effect: kinetic analysis and application to representative or long-molecule biased PCR-based amplification of complex samples. *J. Biotechnol.* **128**:435–443.
- Davidson, J.F., Fox, R., Harris, D.D., Lyons-Abbott, S., and Loeb, L.A. (2003). Insertion of the T3 DNA polymerase thioredoxin binding domain enhances the processivity and fidelity of Taq DNA polymerase. *Nucleic Acids Res.* **31**:4702–4709.
- Diatchenko, L., Chris Lau, Y.F., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., et al. (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. U S A* **93**:6025–6030.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**:343–345.
- Hogrefe, H.H., and Borns, M.C. (2011). Long-range PCR with a DNA polymerase fusion. *Methods Mol. Biol.* **687**:17–23.
- Hughes, R.A., and Ellington, A.D. (2017). Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* **9**:a023812.
- Ignatov, K.B., Barsova, E.V., Fradkov, A.F., Blagodatskikh, K.A., Kramarova, T.V., and Kramarov, V.M. (2014). A strong strand displacement activity of thermostable DNA polymerase markedly improves the results of DNA amplification. *Biotechniques* **57**:81–87.
- Jiang, Y., Xie, M., Zhu, Q., Ma, X., Li, X., Liu, Y.-G., and Zhang, Q. (2013). One-step cloning of intron-containing hairpin RNA constructs for RNA interference via isothermal *in vitro* recombination system. *Planta* **238**:325–330.
- Liu, Y.-G., and Chen, Y. (2007). High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques* **43**:649–656.

- Liu, Y.-G., and Whittier, R.F. (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* **25**:674–681.
- Long, Y., Sun, Q., Burnham, K., Steuernagel, B., Cai, X., Faris, J.D., Harris, M.O., Wulff, B.B.H., Lagudah, E.S., and Xu, S.S. (2019). Nested loop PCR (NL-PCR) for amplification of large DNA fragments with complex structures in wheat. In *The First International Wheat Congress (1st IWC)*, July 21–29, Saskatoon, Canada.
- Lukyanov, K.A., Launer, G.A., Tarabykin, V.S., Zaralsky, A.G., and Lukyanov, S.A. (1995). Inverted terminal repeats permit the average length of amplified DNA fragments to be regulated during preparation of cDNA libraries by polymerase chain reaction. *Anal. Biochem.* **229**:198–202.
- Lukyanov, K.A., Gurskaya, N.G., Bogdanova, E.A., and Lukyanov, S.A. (1999). Selective suppression of polymerase chain reaction. *Russ. J. Bioorg. Chem.* **25**:141–147.
- Matz, M., Shagin, D., Bogdanova, E., Britanova, O., Lukyanov, S., Diatchenko, L., and Chenchik, A. (1999). Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Res.* **27**:1558–1560.
- Mazars, G.R., Moyret, C., Jeanteur, P., and Theillet, C.G. (1991). Directing sequencing by thermal asymmetric PCR. *Nucleic Acids Res.* **19**:4783.
- Moore, P. (2005). PCR: replicating success. *Nature* **435**:235–238.
- Mullis, K.B., and Faloona, F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**:335–350.
- Mullis, K.B., Faloona, F.A., Scharf, S.J., Saiki, R.K., and Erlich, H.A. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Sym.* **51**:263–273.
- Shagin, D.A., Lukyanov, K.A., Vagner, L.L., and Matz, M.V. (1999). Regulation of average length of complex PCR product. *Nucleic Acids Res.* **27**:e23.
- Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995). An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23**:1087–1088.
- Stemmer, W.P., Cramer, A., Ha, K.D., Brennan, T.M., and Heyneker, H.L. (1995). Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxynucleotides. *Gene* **164**:49–53.
- Tan, J., Gong, Q., Yu, S., Hou, Y., Zeng, D., Zhu, Q., and Liu, Y.-G. (2019). A modified high-efficiency thermal asymmetric interlaced PCR method for amplifying long unknown flanking sequences. *J. Genet. Genomics* **46**:363–366.
- Tang, H., Luo, D., Zhou, D., Zhang, Q., Tian, D., Zheng, X., Chen, L., and Liu, Y.-G. (2014). The rice restorer *Rf4* for wild-abortive cytoplasmic male sterility encodes a mitochondrial-localized PPR protein that functions in reduction of *WA352* transcripts. *Mol. Plant* **7**:1497–1500.
- Villbrandt, B., Sobek, H., Frey, B., and Schomburg, D. (2000). Domain exchange: chimeras of *Thermus aquaticus* DNA polymerase, *Escherichia coli* DNA polymerase I and *Thermotoga neapolitana* DNA polymerase. *Protein Eng.* **13**:645–654.
- Wang, Y., Prosen, D.E., Mei, L., Sullivan, J.C., Finney, M., and Vander Horn, P.B. (2004). A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance *in vitro*. *Nucleic Acids Res.* **32**:1197–1207.
- Zhou, L., Shen, R., Ma, X., Li, H., Li, G., and Liu, Y.-G. (2016). Preparation of rice plant genomic DNA for various applications. *Curr. Protoc. Plant Biol.* **1**:29–42.
- Zhu, Q., Yu, S., Zeng, D., Liu, H., Wang, H., Yang, Z., Xie, X., Shen, R., Tan, J., Li, H., et al. (2017). Development of “Purple Endosperm Rice” by engineering anthocyanin biosynthesis in the endosperm with a high-efficiency transgene stacking system. *Mol. Plant* **10**:918–929.