

实验药理学及生物学进步 895

美国基因及细胞治疗协会

Toni Cathomen
Matthew Hirsch
Matthew Porteus 主编

陈元达 译

基因编辑

下一代基因疗法

 AMERICAN SOCIETY OF
GENE & CELL
THERAPY

 Springer

在实验药理学及生物学中的进步

895 卷

编辑人员：

IRUN R. COHEN, The Weizmann Institute of Science, Rehovot, Israel

ABEL LAJTHA, N.S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA

JOHN D. LAMBRIS, University of Pennsylvania, Philadelphia, PA, USA

RODOLFO PAOLETTI, University of Milan, Milan, Italy

更多的信息请关注：<http://www.springer.com/series/5584>

Toni Cathomen • Matthew Hirsch
Matthew Porteus
主编

基因编辑

下一代的基因疗法

注*本书只送不卖，不注册版权，只供学习交流使用！

 Springer

编辑

Toni Cathomen

药物中心-弗莱堡大学
弗莱堡, 德国

Matthew Hirsch

北卡罗来纳大学
教堂山 北卡罗莱纳州, 美国

Matthew Porteus

斯坦福药学院
帕洛阿尔托, 加州, 美国

ISSN 0065-2598

Advances in Experimental Medicine and Biology

ISBN 978-1-4939-3507-9

DOI 10.1007/978-1-4939-3509-3

ISSN 2214-8019 (electronic)

ISBN 978-1-4939-3509-3 (eBook)

Library of Congress Control Number: 2015958888

Springer New York Heidelberg Dordrecht London

© American Society of Gene and Cell Therapy 2016

本文禁止用于其他商业用途。所有版权归出版商所有。包括与内容相关的翻译、重印、插图、注释、广播、复刻等物理途径及传送、通知、存储、电子再版、电脑软件等其他领域。其他出版物中禁止使用与此出版物相关的一般描述性名称、注册名称、商标、服务商标等及暗示性词语。

出版商、作者及主编自发行之日起承担保证书中信息及建议的准确性及真实性的责任。但是无论是出版商还是作者或是主编都无法完全保证书中不会存在任何遗漏或谬误之处。

以上内容打印于无酸纸

Springer Science+Business Media LLC New York is part of Springer Science+Business Media
(www.springer.com)

谨以此书纪念 Carlos F. Barbas
III。在他传奇性的一生中，他由于
人工转录因子、锌指核苷酸、重组
酶中的锌指蛋白序列的设计在基因
编辑领域享有盛名。“我一直梦想
着能够引领药物研发的变革。”
(Carlos Barbas, 2008)

前言

在二十世纪早期，人类基因编辑还只是科幻作家、科学家以及一些哲学家玄奇的想象。早在100年前，伴随这个问题而来的多种伦理问题开始被人们所讨论：基因编辑的滥用是否会打破“自然”的进化？被进行编辑的生物是否会破坏我们的生活？基因蓝图的改变是否会使我们对于原有无害的病原体敏感？相关的担心无穷无尽。在今天，驱动我们进行基因编辑的动力来源于其潜在的治愈基因异常疾病的可能性。实际上，一些新兴的基因编辑技术效率很高而且易于操作，以至于非人类灵长类的基因也可被改变。基于前述观点，一些本来被用于治疗基因异常疾病的技术也可被用于改变人类种族使人获得更多的能力，如智慧或美貌。我们应该如何划定两者之间的红线呢？为了给予我们更多的时间进行公开的讨论或更好的理解如何保证我们现行的基因编辑工具在安全的范围之内，科学家和伦理学家已经叫停了人类基因组编辑的进程。

人类基因编辑是疾病疗法的最前沿，其应用基于多种具有创造性的观察，致使一个细胞具有更强的改造其DNA的能力。基于细胞作用方向的不同，无论是染色体的破坏还是原有位点的精确编辑在内源性或外源性的DNA上都是有可能的。为了提升我们进行基因突变的能力，我们花费了巨大的努力使核苷酸能够高效地进入人类细胞。现在的我们正处于基因时代的风口浪尖，现在正是时候将我们现有知识和二十世纪早期的相关理论进行结合。

我们非常高兴的发现，世界上的一些最为卓越的遗传学家、生物学家以及生物信息学家，都在刷新着我们对基因编辑的认知并赋予我们更强改造人类基因组的能力，也为这本书的最后成型做出贡献。为了使本卷能够囊括最大范围的优秀作者，每章的独立作者和编辑都尽力提供每个基因最为完善的背景信息。最终的成稿是囊括了人类基因编辑的多个方面的综合性成果。我们正在逐步接近一个没有疾病的时代。这太令人高兴了。

Freiburg, Germany
Chapel Hill, NC
Palo Alto, CA

Toni Cathomen
Matthew Louis Hirsch
Matthew Porteus

起航，基因码农的黄金时代！

2005年，人类基因组草图已经完成。随后，以一代基因测序、二代基因测序为基础的基因测序公司如雨后春笋般不断涌现。随着2015年“精准医疗”概念进入美国国情咨文，基因测序从基础研究实验室过渡到个人的独立医学实验室，并有部分相关检测产品进入临床。从来没有哪个行业像是基因检测行业一样，基于某种特定的技术成长为一个产业。

但是伴随着基因测序成本的不断降低，和大基因数据下人们后续解读知识的匮乏，今后的基因行业如何发展就成为了摆在所有基因相关企业面前的一道无法绕开的话题。有人说：“基因行业就像是趴在玻璃窗上的苍蝇，前途一片光明，就是没有出路。”到底坚守还是逃离，到底是进行产业升级还是扩大市场，如何辨别出基因和表型之间的不确定性，如何让我们已有的基因大数据库发挥出更大价值呢？

作为一名药物研发专业方向二年级的研究生，接触到基因编辑领域纯属偶然。本来就是在药物敏感性研究中的一个小问题，也想借着近些年来兴起的CRISPR/Cas9技术建几个模型，蹭蹭热度，结果一发不可收拾，一脚入坑往事不堪回首。经过近一年的课题的研究，对基因编辑的理解从原来扁平快的对CRISPR/Cas9粗浅的理解到TALENs到ZFNs再到归巢内切酶，从人类疾病细胞到正常细胞到鼠原细胞再到出芽嗜热菌，基因编辑发展的历史发展卷轴在面前缓缓摊开。突然感到小时候红警游戏中那种瞬间“秒变”半兽人的时代好像真的要来了。修仙小说中的那些所谓的“淬体筑基”、“逆天改命”也仿佛不再遥远。2014年，全球首款基因治疗药物“Glybera”在德国上市，用于治疗脂蛋白酶缺乏遗传病，由荷兰生物科技公司UniQure研发，一剂百万的高昂治疗价格也进一步的颠覆了传统医学的认知。今年前几个月，由于目标市场极端狭窄，疗效缺乏可控性，该药物宣布永久性撤市，但是带给业界的冲击缺远未结束。

挡在基因行业这只“小苍蝇”前面的玻璃正变得越来越薄，资本的风向下我们貌似越来越接近了行业财富“井喷”的窗口。虽然基因编辑脱靶的报道不断见诸报端，但是我们也明显感觉到，从初代ZFNs在有限的位点造成的数以千计的脱靶位点到TALENs的可被预测，再到前几个月中科院生化细胞所李劲松课题组利用全基因组测序才找到一百多个CRISPR脱靶位点，我们有理由相信随着计算机科学和合成生物学的发展，基因编辑技术会逐渐成为一种绝对安全有效的生物技术，基因编辑到临床应用的距离正在不断地缩短。而越来越多的人也开始相信，这一技术所带给我们的惊喜才刚刚开始。

随着阿里巴巴、京东、腾讯、百度、网易等几大互联网巨擎在一些互联网细分领域实现了碾压性的垄断优势后，有越来越多迹象表明互联网创业已经逐渐进入了疲软的“退潮期”。

2017年，Alpha Go打败了柯洁、量子计算机出现、日本用增强现实技术还原了邓丽君的演唱会，这些划时代技术的突破无不提示我们新的“高科技井喷”的创业时代已经拉开了帷幕。上个月，有幸和上海的一名互联网搜索引擎公司的老总谈话。该老总是上海本地人，高中学历，从98年开始涉足互联网搜索引擎行业，精通多种计算机算法语言。他谈到自己当年第一次接触互联网编程时，IT行业还被认为是一种没有前途的行业。“醒的比鸡早，睡的比小姐晚，活的比驴还累，挣得比民工还少。”是当时的业界常态。“现在呢？”我问。“现在？一个大中专生应届毕业生都敢和我说，一个月少于一万六不做。想想还是我自己敲代

码划算些。就是零六年的时候，突然经过了行业的拐点，一下子整个 IT 行业就都做起来了，带着我们这些半死不活的小企业也越来越好做，越来越赚钱了，这就把那些人养的眼高手低了。”

联想我们今天的生物技术，基因检测行业是何其的相似。随着技术壁垒的不断被打破，我们有理由相信属于“基因码农”的时代不会再遥远了，我不敢高调的宣布下一个时代就是属于“基因编辑”或是“生物技术”的时代，但是可以肯定的是“基因编辑”在下一波“科技创业”的大潮中必然是一股不容小觑的力量。

我不是英语专业毕业，本书也没有经过校正，翻译肯定会存在错误，而且错误还可能很多，有任何出入或常识性的错误请以原文为准。如果你也对“基因编辑”或此类这样有趣的“生物技术”感兴趣，如果你也恰巧读到了这一页并没有什么用的“作者闲话”，如果你也打算通过自己的力量拥有生物技术领域的一席之地，欢迎加一下这个并不怎么靠谱的小编的微信，或许我们可能存在一些共同语言呢？

最后如果本书能给你的实验或是认知一些小小的启迪，都将是本人莫大的荣幸。

祝好！

我的微信：Unisophy_para_chen

我的邮箱：1305206309@qq.com

陈元达
2017年7月4日
于中国科学院上海药物研究所

目录

近 20 年的基因编辑.....	1
Maria Jasin	
锌指核酸酶的发展和应用.....	14
Dana Carroll	
TAL 效应核酸酶的应用和发展.....	28
Alexandre Juillerat , Philippe Duchateau , Toni Cathomen , and Claudio Mussolino	
神经肌肉疾病的基因编辑.....	47
David G. Ousterout and Charles A. Gersbach	
噬菌体整合酶基因编辑.....	771
Michele P. Calos	
采用三链寡核苷酸和肽核酸进行精确基因组修饰.....	80
Raman Bahal , Anisha Gupta , and Peter M. Glazer	
通过适配体介导的基因靶向 (AGT)	95
Patrick Ruff and Francesca Storici	
通过双链断裂修复刺激 AAV 基因的编辑.....	108
Angela M. Mitchell , Rachel Moser , Richard Jude Samulski , and Matthew Louis Hirsch	
核苷酸酶编辑和三核苷酸重复序列疾病.....	119
John H. Wilson , Christopher Moye , and David Mittelman	
应用人工编辑核酸酶产生 HIV 抵抗细胞.....	138
George Nicholas Llewellyn , Colin M. Exline , Nathalia Holt , and Paula M. Cannon	
鉴定人工编辑核苷酸脱靶效应的策略.....	161
Eli J. Fine , Thomas James Cradick , and Gang Bao	
通过基因编辑核苷酸酶进行细胞编辑和疾病建模.....	190
Mark J. Osborn and Jakub Tolar	

参与编纂

Raman Bahal , Ph.D. Department of Therapeutic Radiology , Yale University School of Medicine , New Haven , CT , USA

Gang Bao , Ph.D. Department of Biomedical Engineering , Georgia Institute of Technology and Emory University , Atlanta , GA , USA

Michele P. Calos , Ph.D. Department of Genetics , Stanford University School of Medicine , Stanford , CA , USA

Paula M. Cannon , Ph.D. Department of Molecular Microbiology and Immunology , Keck School of Medicine, University of Southern California , Los Angeles , CA , USA

Dana Carroll , Ph.D. Department of Biochemistry , University of Utah School of Medicine , Salt Lake City , UT , USA

Toni Cathomen , Ph.D. Institute for Cell and Gene Therapy , Medical Center - University of Freiburg , Freiburg , Germany Thomas James Cradick Department of Biomedical Engineering , Georgia Institute of Technology and Emory University , Atlanta , GA , USA

Philippe Duchateau , Ph.D. Research Department , Paris , France Colin M. Exline , Ph.D. Department of Molecular Microbiology and Immunology , Keck School of Medicine, University of Southern California , Los Angeles , CA , USA

Eli J. Fine Department of Biomedical Engineering , Georgia Institute of Technology and Emory University , Atlanta , GA , USA

Charles A. Gersbach , Ph.D. Department of Biomedical Engineering , Duke University , Durham , NC , USA

Peter M. Glazer , M.D., Ph.D. Department of Therapeutic Radiology , Yale University , New Haven , CT , USA Department of Genetics, Yale University, New Haven, CT, USA

Anisha Gupta , Ph.D. Department of Therapeutic Radiology , Yale University School of Medicine , New Haven , CT , USA

Matthew Louis Hirsch , Ph.D. Gene Therapy Center , University of North Carolina at Chapel Hill , Chapel Hill , NC , USA Department of Ophthalmology , University of North Carolina at Chapel Hill , Chapel Hill , NC , USA

Nathalia Holt , Ph.D. Department of Molecular Microbiology and Immunology, Keck School of Medicine, University of Southern California, Los Angeles , CA , USA Maria Jasin , Ph.D. Developmental Biology Program , Memorial Sloan Kettering Cancer Center , New York , NY , USA

Alexandre Juillerat , Ph.D. Research Department , Paris , France George Nicholas Llewellyn , Ph.D. Department of Molecular Microbiology and Immunology , Keck School of Medicine, University of Southern California , Los Angeles , CA , USA

Angela M. Mitchell , Ph.D. Gene Therapy Center , University of North Carolina at Chapel Hill , Chapel Hill , NC , USA David Mittelman Department of Biological Sciences , Virginia Tech , Blacksburg , VA , USA Virginia Bioinformatics Institute ,

Blacksburg , VA , USA

Rachel Moser , B.A. Department of Ophthalmology , University of North Carolina at Chapel Hill , Chapel Hill , NC , USA

Christopher Moyer , Ph.D. Verna and Marrs McLean Biochemistry and Molecular Biology , Baylor College of Medicine , Houston , TX , USA

Claudio Mussolino , Ph.D. Center for Chronic Immunodeficiency , Medical Center -University of Freiburg , Freiburg , Germany

Mark J. Osborn , Ph.D. Department of Pediatrics , Division of Blood and Marrow Transplantation, Masonic Cancer Center, Stem Cell Institute, and Center for

Genome Engineering, University of Minnesota , Minneapolis , MN , USA

David G. Ousterout , Ph.D. Department of Biomedical Engineering , Duke University , Durham , NC , USA

Patrick Ruff , Ph.D. Department of Microbiology and Immunology , Columbia University Medical Center , New York , NY , USA

Richard Jude Samulski , Ph.D. Department of Pharmacology , Gene Therapy Center, University of North Carolina at Chapel Hill , Chapel Hill , NC , USA

Francesca Storici , Ph.D. School of Biology , Georgia Institute of Technology ,Atlanta , GA , USA

Jakub Tolar , M.D., Ph.D. Department of Pediatrics , Division of Blood and Marrow Transplantation, Stem Cell Institute, University of Minnesota , Minneapolis ,MN , USA

John H. Wilson , Ph.D. Verna and Marrs McLean Biochemistry and Molecular Biology , Baylor College of Medicine , Houston , TX , USA

关于编辑

Toni Cathomen, Ph.D 是德国弗莱堡医科大学细胞和基因疗法研究所的专家和编辑。研究所为医学中心提供血液和细胞制品和转移相关的诊断服务。Toni Cathomen 在瑞士苏黎世大学获得了博士学位。在他任职于弗莱堡大学之前，在美国圣迭戈索尔克研究所做博士后，作为柏林查利特医学院分子病毒学的助理教授以及汉诺威医学院实验血液学的副教授。Toni Cathomen 的主要研究课题包括：（1）进一步的提升基因编程工具（TALENs, CRISPR 等）在人类干细胞治疗中的安全性。（2）基于多功能诱导干细胞（iPSCs）构建疾病模型和细胞疗法以及（3）将细胞和基因疗法用于临床。

Matthew Hirsch 是教堂山北卡罗莱纳州大学眼科的助理教授。他在微生物学和免疫学、遗传学和分子生物学领域中有研究并在北卡罗来纳州大学的基因疗法中心也有任职。Dr. Hirsch 在西弗吉尼亚大学获得了他的博士学位，目前工作于西弗吉尼亚州摩根唐大肠杆菌和沙门杆菌基因公司。他在北卡罗来纳州研究中心与 Jude Samulski 共同完成了有关于应用腺病毒群（AAV）进行游离基因和染色体基因的博士后课题。Dr. Hirsch 目前还在继续进行 AAV 的基础研究并在盲人和肌肉萎缩症领域中有多种处于前临床评价中的药物。

Dr. Matthew Porteus 生于加利福尼亚，是甘恩高中的本地毕业生，在哈佛大学获得了 A.B. 历史和科学学位，著有论文“危险还是安全：重组 DNA 技术在不同社会阶层中的矛盾。”他随后回到自己的出生地并在斯坦福医学院继续攻读 M.D. 和 Ph.D，他的 Ph.D 课题致力于了解前脑发育的分子机制，他的 Ph.D 论文题目是“TES-1/DLX-2: 一种全新的在哺乳动物前脑发育中表达的同源异型基因的分离和鉴定。”在他完成了他的双学位后，他在波士顿儿童医院任职成为一名儿科的见习医师和住院医师，并随后在波士顿儿童医院/达纳法伯癌症研究所获得了儿科的血液学/肿瘤学奖学金。在他的奖学金项目和博士后研究中，他和计算机领域的 Dr. David Baltimore 合作在加利福尼亚理工学院开始了他在同源重组治疗干细胞点突变疾病的明确有效治疗儿童血液、尤其是镰刀症细胞手段的研究。在他和 Dr. Baltimore 合作的过程中，他在西南医学中心儿科和生化研究部获得了独立的教职职位，并在 2010 年以副教授的身份重返斯坦福。在这期间，他第一次通过研究发现基因纠正在人类细胞中的效率足以对患者进行治疗，并被认为是基因编辑领域的先驱和缔造者，在全世界范围内，这一领域中已经出现了数千个实验室并建立了多家公司。他的研究课题目前仍致力于通过基因重组手段治疗患有遗传疾病的儿童，并在不同人群的克隆动力学和采用基因编辑技术更好的了解儿童疾病——包括婴儿白血病和肌肉疾病。在临床上，Dr. Porteus 任职于露西尔帕卡德儿童医院，在那里他对于相关疾病的儿科患者进行不同的干细胞移植治疗。

近 20 年的基因编辑

Maria Jasin

摘要 直接的基因修饰对于了解基因的功能和基因疗法的应用至关重要。近二十年基因组中的双链损伤 (DSB) 被发现能够诱导高效的基因编辑, 方式包括诱导 DNA 进行同源重组, 如基因靶向或非精确的基因末端修复引起的突变形成。大范围核苷酸酶、ZFNs、TALENs 和 CRISPR/Cas9 技术的加速发展使得我们能够在双链损伤位点进行精确的修复, 该技术引起了生物学和生化学的变革。本章为这些基因编辑技术描绘了一幅全景, 并总结了在哺乳动物细胞中的 DSB 修复机制。

关键词 基因靶向 • 基因编辑 • 基因组编辑 • 同源重组 • 非同源末端重组 • NHEJ • Cas9 • 双链损伤 • 基因重组

M. Jasin, Ph.D. (*)

Developmental Biology Program, Memorial Sloan Kettering Cancer Center,
1275 York Avenue, New York, NY 10065, USA

e-mail: m-jasin@ski.mskcc.org

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_1

介绍

直接进行基因修饰的技术对于了解基因功能非常重要且具有成为新的基因疗法的巨大潜力[1]。DNA 双链损伤（DSB）诱导高效的基因修饰的发现促进了生物科学领域和生化领域中定向诱导 DSBs 的技术的发展与变革。引起这些技术进步的发现在本章中将进行详细的总结。有关于 DSB 修复机制的总结在本章中也将提及。

基因修饰在酵母中简单但在哺乳动物细胞中困难

在 1978 年遗传学家发现能够通过同源重组（HR）将 DNA 整合进入酵母基因组中[2]。随后，基于 HR 的多种方法被研发出来[3,4]。结果，酵母基因组的高效修饰被用来进行基因功能的研究，使得酵母成为真核生物基因分子模型研究中的一个典型的选择模型。

与此相反，整合进入哺乳动物细胞中的 DNA 被发现处于非同源末端区域[5]。同源整合——也就是所谓的基因靶向，其效率非常低，需要进行广泛性的筛选[6]。基因编辑技术在药物靶标的选择/反选[7]以及非启动子选择性靶标[8]中的应用促进了基因靶向细胞的发现，类似的基因靶向技术在一些系统中早已司空见惯，如产生突变小鼠的鼠胚胎干细胞[9]。虽然基因靶向的手段有所进步，但是在实际应用中仍然存在多种限制性因子，例如低效率产生的筛选的需求，在单轮筛选中无法同时靶向二倍体细胞的全部等位基因以及对于模式小鼠胚胎细胞系的需求。

DSBs 是同源重组的起始事件

在哺乳动物细胞中质粒 DNA 的非同源重组延缓了 HR 相关技术的发展。然而，在酵母中的深入研究为这一问题的解决提供了突破口。在染色体同源序列中 DSB 的产生被认为能够提升基因靶向的效率[10]。其中产生了两类问题：一是 DSB 可以不通过质粒上的同源区段修复（图 1a）；因此，在 DSB 位点有可能发生质粒片段的缺失（例如一个缺口），质粒的修复有可能引起原始片段的复活。二是质粒有可能在修复过程中整合到基因组中（图 1b）。

在酵母（见[11]）、细菌及其噬菌体[12]中的研究通过 γ -射线诱导产生 DSBs 发现 DNA 末端能够重组进入酵母和真核细胞系统从而通过 HR 产生 DSB 修复模型[13]。在这个模型中，DSBs 是 HR 的起始事件，损伤 DNA（实验样本）通过未损伤的序列以及同源片段 DNA（供体质粒）进行修复。在这一过程中起到重要节点作用的是一种双向 Holliday 复制叉，能够决定重组是通过交联还是非交联的方式进行。质粒介导的交联 DSB 修复能够将质粒上的片断整合进入染色体中（图 1b），而非交联的修复则不存在质粒序列的整合（图 1a）。

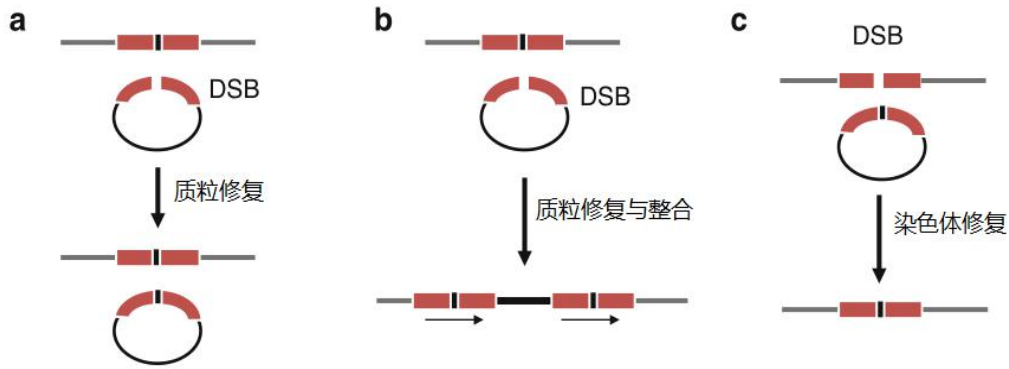


图1. DSB诱导的染色体和质粒DNA之间的同源重组 (HR)。染色体和质粒DNA之间的同源区段用红色表示。(a) 质粒DNA上的缺口或DSB可采用染色体作为模板进行简单的基因替换。(b) 质粒DNA上的缺口或DSB能够以染色体为模板进行基因的交叉修复,使质粒整合进入靶向的序列之间,在哺乳动物细胞中交叉整合被抑制。(c) 染色体上的DSB可通过质粒DNA为模板进行修复,并导致高效的基因整合。

在酵母中 DSB 修复的研究以及 DSB 模型的建立促使我们思考哺乳动物的 DSB 修复是否也通过 HR 途径[14]。最初的研究发现质粒上的 DSB (缺口) 可通过染色体修复 (图 1a), 并导致完整病毒的产生。这些实验发现 HR 途径的 DSB 修复效率很高, 有近 10% 的线性化质粒能够通过 HR 产生非交联作用。虽然在实验设计中整合基因并未被激活, 但是该实验为 HR 是主要的 DSB 修复机制提供了明确的证据, 而在该实验前这一说法经常被人怀疑。

为了对基因修饰进行检验, 实验设计选取了一些没有选择性的基因位点[8]。在实验中 DSB 修复能够引起质粒的整合 (图 1b), 基因组的修饰进一步的揭示了在哺乳动物细胞中 DSBs 重组整合的自然属性。DSB 进行同源重组的效率相较于其它机制高出 100 倍[8]。除了高效率之外, 重组频率 (交联) 被发现显著低于非重组频率 (非交联) [14]。这一不平衡的现象提示在哺乳动物细胞中有特异性的抑制交联作用的机制, 但这一观点目前还未获得充足的证明 (例如[15])。

开创性的实验: 基因组中的 DSBs 能够诱导靶向效率的提升

质粒中的 DSBs 能够诱导染色体的重组证明 DSB 同源重组修复能够被用于哺乳动物的基因编辑。然而这种技术还存在很多限制: 由于 DSB DNA 是遗传信息的载体[13], 因此模型中的 DSB 应该发生于基因组而非质粒上, 并直接引起基因组的改变 (图 1c)。由于在酵母中质粒 DNA 的同源整合易于实现, 因此基因组的破坏试验并未展开, 这一模型提示在哺乳动物中位点特异性的 DSB 的引入能够引起哺乳动物细胞中同源重组频率的提高。因而关键就在于如何在酵母中建立特异性的配套转换机制, 从而诱导两个染色体序列之间位点特异性的 DSB 并引起 HR 修复[16]。通过基因组特定位点的 DSB 的引入以及配套的 HO 核酸内切酶的表达, 染色体同源区域和人工导入同源序列之间的 HR 被激活[17]。在 *Drosophila* 中的转座子研究发现在同源序列之间的 HR 可通过 DSBs 诱导产生[18]。

为了验证 DSB 是否可以刺激基因靶向 (图 1c), 我们在基因组中诱导产生了 DSB 并提供了可用于进行 DSB 修复的同源片段[19]。尤其重要的是我们表达了 I-SceI 核酸内切酶 [20], 该酶与 HO 核酸内切酶相关, 能够引起酵母中位点特异性的 HR[21]。I-SceI 内切酶被用于本实验是由于起作用位点明确且具有足够的长度——18bp, 因此在染色体中该位点的表达并不能引起细胞的死亡[20]。

I-SceI 识别位点被整合进入哺乳动物染色体，并在 I-SceI 位点两侧引入 700bp 的同源片段（图 2a）。基因靶向的效率被提高了几个数量级，该现象证实染色体中 DSB 的引入相较于自然状态下基因靶向效率提升[23]。随后的有关于哺乳动物细胞基因靶向位点 DSBs 的重组整合的研究发现在胚胎干细胞和其他类型细胞中也存在类似的现象[24-28]。在 1994 年进行的一系列试验实现了 HR 途径的生物基因编辑，其结果将在以下内容中讨论。

在基因组中诱导突变发生的 DSBs

在 1994 年开展的实验证实诱导基因靶向的 DSBs 也能够引起基因突变（图 2b）。在这个案例中，由 I-SceI 产生的 DNA 末端并未引入同源末端或只引入几个 bp 的同源末端序列[19]。在这些实验中非精确性的末端重组（NHEJ）概率被认为是同源重组（HR）概率的二倍。在损伤处可见 NHEJ 引起的短片段删除或插入[29-31]。由 DSB 引起的长片段删除克隆也筛选得到[32]。这些实验为今天的 NHEJ 机制的基因编辑方式奠定了基础，其余信息将在以下内容中讨论。

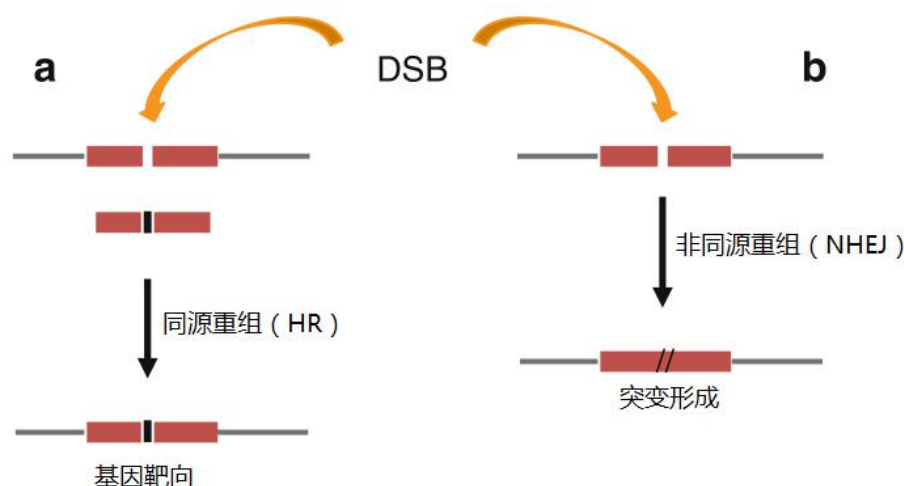


图2. 在染色体中由DSBs诱导产生的基因编辑。在染色体中的一个DSB能够通过HR途径引入同源DNA中，例如基因靶向，（a）或通过非精确的NHEJ引发突变（b）。最初在哺乳动物细胞中由I-SceI内切酶能够引起两种事件同时发生，其中NHEJ事件更为频繁[19]。采用 ZFNs, TALENs和CRISPR/Cas9的实验也被证实两种重组机制的并行发生。

两种 DSBs 能够诱导基因重组

这些实验也证实在相同染色体上的两个 DSBs 能够引起遗传信息的删除[19]。而且，在一些实验中还发生了由 DSBs 引发的其他类型的染色体异常。在肿瘤细胞中染色体的重排是常见的，包括周期性发生的、可产生倒置的易位[33]。I-SceI 核酸内切酶能够被用来在不同的染色体上诱导产生 I-SceI 位点的染色体易位[34,35]。染色体易位不只在染色体受到 DSB 损伤时可观测到[27]。两个位点的 DSBs 被发现能够通过 NHEJ 引起染色体易位[36]，但无法引起 HR，这一现象可能是由于交联抑制[34,35]。由于肿瘤两个染色体之间的染色体易位损伤结构并未显示出同源性，NHEJ 介导的易位系统可通过致癌性染色体易位的产生进行相关机制的探究[37]。

采用 I-SceI 核酸内切酶进行 HR 的研究

通过 HR 进行 DSB 的修复（有时也称为直接同源修复，或 HDR）的过程起始于 DNA 末端的

切除引起的 DNA 单链末端的产生（来自于一篇回顾性文章，见[38]及其参考文献）（图 3a）。这个单链为一种单链交换蛋白——真核细胞中的 RAD51 提供了底物，并形成核酸蛋白纤维插入到损伤的同源序列中。损伤 DNA 的修复合成起始于核酸蛋白纤维的插入位点并采用同源序列作为模板。在这一节点后可发生多种生物学过程，其中最为简单的一个过程是合成一条与外来片段不相关的新的 DNA 链（见[39]及其参考文献）。在哺乳动物细胞中的 HR 更易于姊妹染色单体之间发生，而发生于其它染色体或载体上的同源区段的同源重组处于一个更低的频率中[40]。

当同源序列接近于 DSB 损伤位点，在寡聚单链损伤末端（SSA）进行同源重组的方式可行（图 3b）。这一方式同样起始于 DNA 末端损伤，但是随后的寡聚单链形成通过切除后的退火而非起始单链的插入。在退火后，DNA 末端进行结扎前的修整。SSA 相关的 HR 频率取决于损伤位点和目的序列之间的距离。因此，虽然 SSA 已经被用于 HR 途径中相关蛋白的鉴定以及后续 HR 步骤的探索，但是 SSA 的作用仍是不清楚的。

在哺乳动物细胞中对于 HR 及相关因素的理解促进了 I-SceI 核酸内切酶的应用。由于目标基因无法指示 HR 事件的发生，染色体内的 HR 报告系统已经建成。其中最为常见的是基于 GFP 荧光蛋白的 DR-GFP 系统[41]（图 3c）。在这个报告系统中，在 DSB 位点的非损伤序列中存在一段简单的保守序列，具有同源序列的染色单体或姊妹染色单体能够引起细胞 GFP 的激活。DR-GFP 报告系统现在已被导入小鼠中研究在早期体细胞中的 HR 过程[42]。相关的 SA-GFP 报告系统也用于 SSA 事件的评估[43]（图 3d）。

随着 DR-GFP 和相关报告系统的产生，哺乳动物 HR 突变能够被最终确定。随后，与 RAD51 蛋白[41,44]和 BRCA1、BRCA2 相关的肿瘤抑制因子[45-47]被发现具有促进 HR 的作用。在哺乳动物细胞中比较 HR 和 SSA 的作用发现 BRCA1 和 BRCA2 分别作用于 HR 的不同途径：BRCA1 突变细胞被发现具有 HR 和 SSA 缺陷，提示该蛋白作用于早期、例如末端删除，而 BRCA2 被发现具有 HR 缺陷但是在 SSA 过程中表达增加[43]。BRCA1 在末端切除中的作用通过随后的实验被证实[48]，而 BRCA2 在随后的 RAD51 介导的后续单链置换中所起到的作用也通过后续的生物学实验得到证实[49]。在 BRCA2 突变细胞中，SSA 的增加是由于 HR 末端切除分子 DNA 链退火抑制的解除[43,50]。

除了在 DR-GFP 报告系统中发现的简单的 HR 机制外，长修复片段的合成也被发现，该过程能够引起大段序列的复制[51]（图 4a）。这些过程的调控存在一些不同的基因；例如，在 BRCA1 和其他的突变细胞中 HR 事件频率的提高（见[52]及其参考文献）。

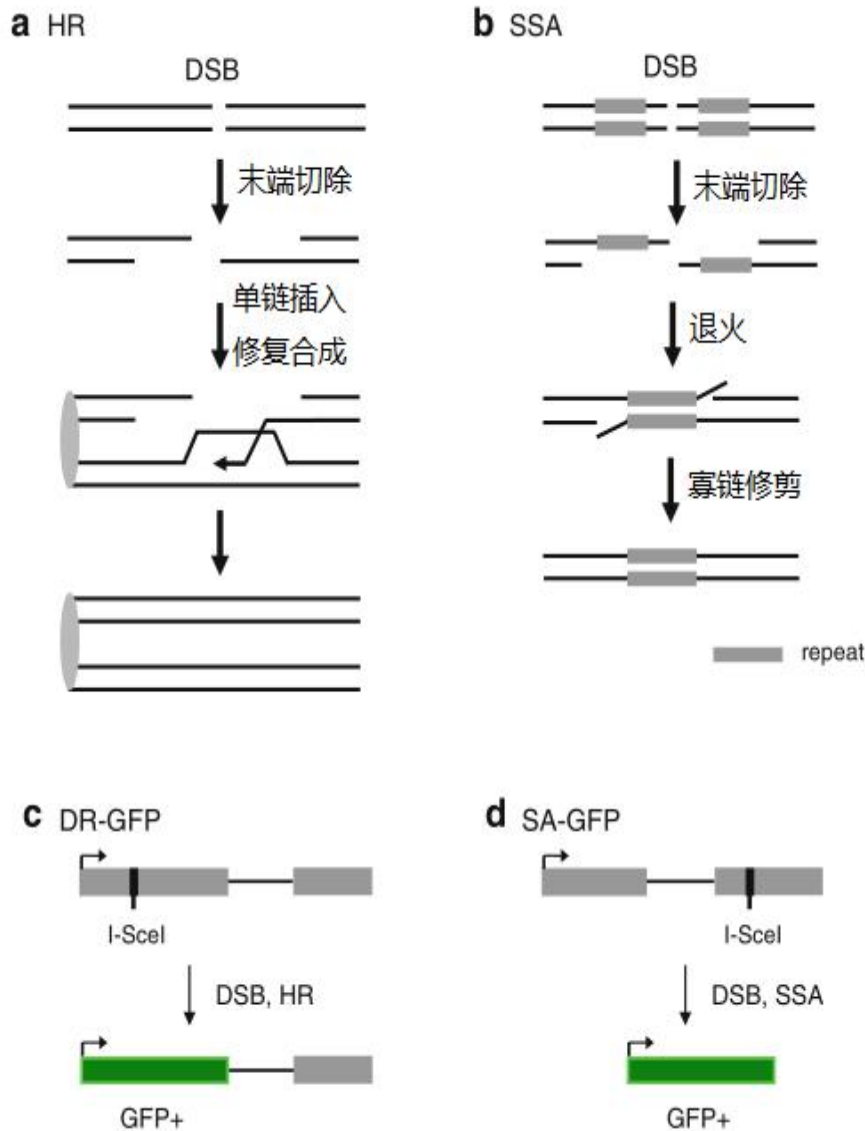


图.3 HR和寡聚单链退火 (SSA)。(a)简单的由HR引起的DSB修复。DNA末端的切除能够引起DNA末端与RAD51蛋白结合形成单纤维促使损伤的DNA单链插入到完整的同源DNA上，在正常情况下是姊妹同源染色体，如图所示。寡聚链的3'末端起始DNA的合成；姊妹染色体能够修复损伤DNA使其恢复到原有状态。在完成了HR过程后，新合成的链解聚连接至另一侧，或者也可能发生其他结果[38]。(b)由SSA产生的DSB，SSA发生于DSB一侧的重复序列。与HR相类似，SSA也起始于末端切除，但是在DSB的任何一侧的单链DNA都能进行互补配对，配对异常处可经核酸酶的修剪作用修复。(c)检测HR的DR-GFP报告系统。在DR-GFP受体实验中，通过HR机制产生的DSB修复作用可以是两个无功能的GFP基因重组形成具有功能的GFP基因，并可通过流式细胞术检验。左侧的GFP由于存在I-SceI位点而没有功能，右侧的GFP基因由于双侧被截短而失去作用。(d)检测SSA的SA-GFP系统。在SA-GFP实验中，SSA机制进行的DSB修复能够使两个失去功能的GFP基因重组形成一个具有功能的GFP基因。在这个案例中，左侧的GFP基因的3'端被截短，右侧GFP的5'端被截短并具有一个I-SceI位点。

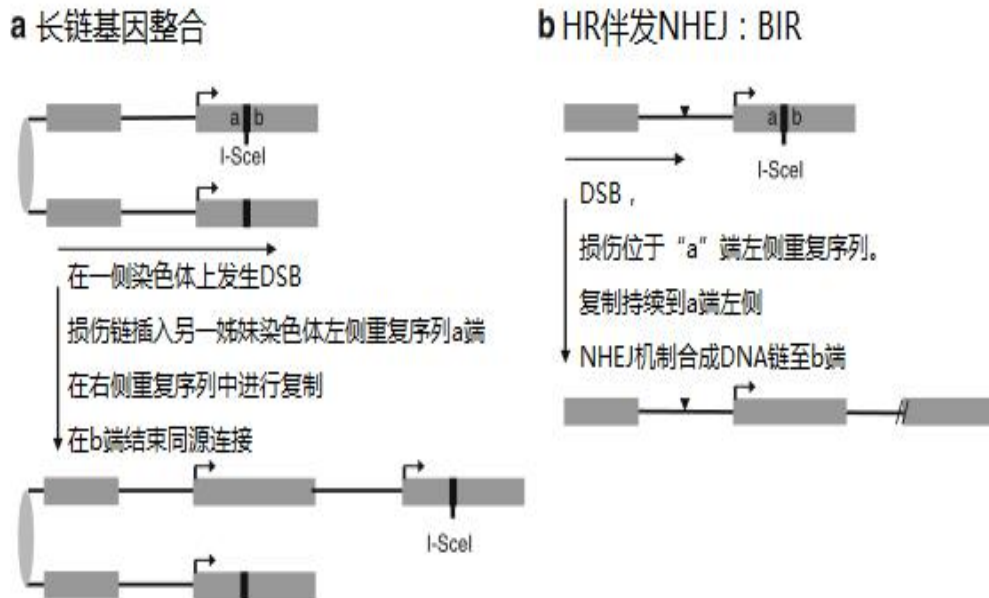


图4.在I-SceI位点发生的其它类型的HR。(a)长基因片段的整合是延展复制中的一类HR事件,通过该过程造成长基因片段的复制,如图所示该事件终止于同源序列处。(b)HR伴发NHEJ,也称作BIR。在这一过程中也会发生片段的复制,但是这一事件终止于NHEJ。

HR 和 NHEJ 之间的协同和竞争

被发现的多种 NHEJ 因子都被发现需要抗原受体的重组结合,由 DSBs 介导的一种特异性通路蛋白称为 RAG 蛋白[29-31]。NHEJ 机制能够介导多样性的非精确重组的产生,这对于免疫应答非常重要。在 I-SceI 引起 DSB 之后, NHEJ 突变能够提升 HR 和 SSA 之间的竞争[43,53],这说明 DSB 修复通路之间存在竞争关系[54]。最近的研究发现 NHEJ 相关蛋白的缺失能够造成 BRCA1 突变亚型,这一现象进一步证实了上述观点[48]。

HR 和 NHEJ 有时可能发生在同一 DSB 位点。在这一案例中,一个 DNA 损伤末端能够插入到其它染色体的同源序列中起始合成(HR),但是新合成的 DNA 链可通过 NHEJ 作用产生其它的 DNA 序列[51,55](图 4b)。这一类型的事件最近被定义为损伤诱导复制(BIR)[56],但是这一现象和在酵母中染色体末端序列的延展过程不同[57],但是这两个过程具有共同的复制因子[56]。BIR 与其它的同源重组事件相类似(图 4a),都会对存在的染色体序列进行复制(图 4b);有证据显示在癌细胞中这种模式的大片段基因复制概率增加[56]。

由多种核酸内切酶诱发的基因编辑

I-SceI 核酸内切酶起始的基因编辑研究证明运用罕见的核酸内切酶进行基因组的人工操作是可行的[23]。这一方法成功的关键在于在修饰位点附近存在酶切位点。寡核苷酸试剂由于其可以通过 Watson-Crick 配对靶向于基因的任何位点被认为具有前景[23],例如将寡聚核苷酸和细菌链交换蛋白 RecA 蛋白化学部分进行整合进行体外实验的基因操作[58]。如 I-SceI 这种的归巢内切酶也被称为大范围核酸酶,被要求具有很低的脱靶率,而这一特性可通过 I-SceI 和 DNA 之间的复合物晶体结构而得到证实[59]。而大范围核苷酸的特异性是在重新设计的过程中就被决定的(例如[60]或[61]中有关于 I-SceI 的部分)。虽然大范围核苷酸不大可能发展成为广义的裂解试剂,但是在一些情况下其可被直接应用。例如,仅由 235 个氨基酸和能够促进转染效率的单链所组成的小 I-SceI,可用于基因治疗。

锌指 DNA 连接结构域具有一个特异性识别区域的可变 DNA 连接结构[62-64]。限制酶 FokI 的双叶结构具有独特的 DNA 连接结构和裂解结构[65,66]，具有和锌指 DNA 识别区域相融合的独特分子裂解结构域[67]，现在被称为锌指核苷酸（ZFNs）。原则上讲，每个锌指结构可以和三个碱基相互作用，多个锌指结构能够在 Koz1 裂解区域排列融合并在基因组上产生独特的裂解位点[68]。最初采用 ZFNs 进行基因突变的报道用于 *Drosophila* 模型的 yellow 基因，靶向于锌指连接区域 GNN 位点[69,70]。在多基因靶向系统中，可应用 I-SceI 内切酶为离体基因组 DNA 的 HR 重组提供重组载体[70]。

最早的 DSB 诱导的人类基因重组是 IL-2R γ 基因[71]。采用供体质粒和 ZFN 进行基因靶向的效率即使是在干细胞中也非常高。而具有更高效率的是双等位基因靶向，这是通过传统的基因编辑方式无法完成的。目前已经具有多种 ZFNs 的基因编辑应用系统[68]。DSB 诱导突变形成的另一种强大的方式是 HIV 共受体 CCR5 基因[72]，这一方式目前被用于基因治疗[1]。

除了这些成功之处，具有 DNA 特异性的锌指模块的组装对于研究者来说仍具有一些挑战；例如，每个锌指单位很难不受其它单位的影响，而且也无法获得 64 种不同类型的锌指。由于植物病原体中 DNA 识别编码 TAL 效应分子的发现极大地促进了核酸的设计，在 TAL 模块中每两个氨基酸识别一个碱基[73-76]。和 ZFNs 相似的是，在 FokI 裂解结构域中具有 DNA 连接重复序列的融合从而能够产生 TAL 效应核苷酸酶（TALENs）[77]。

TALEN 技术的发展解决了以基因编辑为目的的 DNA 损伤的制造。然而，产生于细菌适应性免疫系统中的 RNA 介导核酸酶的发现——即 CRISPR/Cas9[78]，使这一过程更加简便，该系统由一个独立的核酸酶（Cas9）和一个可以通过 Watson-Crick 配对进行特异性靶向的 RNA 序列的复合体组成。研究者们很快地适应了采用 CRISPR/Cas9 的基因编辑技术[79,80]，该技术也成为了研究者们的一种常用的方法[81]。除了在构建的简便性上和 TALENs 技术相似之外，CRISPR/Cas9 相对于其他技术能适用于更加复杂的体系，因为在靶向于不同的基因位点时，只有 RNA 部分需要做出相应的调整。

在基因靶向和突变成过程中，人工设计的核酸酶被用于产生染色体的重组。例如，在多种人类细胞系中 ZFNs、TALENs 和 CRISPR/Cas9 都被用于肿瘤相关的染色体易位[37,82,83]。类似的方法也被用于在小鼠中诱导染色体产生具有肿瘤特征的易位或倒置，从而提高肺癌发生的概率[84]。

总结

在上个 20 年之内，基因编辑技术取得了飞速的发展。从人们最初开始进行基因编辑到学会运用 ZFNs 进行基因编辑用了近十年。在 ZFNs 应用的五年后，TALENs 技术的发展使得基础研究工作者具有几乎能够编辑所有基因的能力，而在这之后三年，CRISPR/Cas9 随后的应用通过对于核苷酸靶点设计的简化进一步地促进了基因编辑的进行。目前几乎所有领域的研究者们都能运用这些技术来解答不同的生物学问题。未来在这些技术中所取得的进步也值得期待，但是我们却在事先无法描绘出在该技术中到底还会出现什么样的突破从而足以引领整个生物科学的发展。而我们清楚的是由于基础研究的持续推进，来源于不同系统的多种基因编辑技术的发展将会给生化科学带来一场前所未有的革命。

致谢 感谢 Rohit Prakash 和 Matt Krawczyk 为本书的写作带来的灵感。这项工作由 NIH CA185660 和 GM054668 出资支持。

参考文献

1. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med.* 2014;370(10):901 – 10.
2. Hinnen A, Hicks JB, Fink GR. Transformation of yeast. *Proc Natl Acad Sci U S A.* 1978; 75(4): 1929 – 33.
3. Scherer S, Davis RW. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc Natl Acad Sci U S A.* 1979;76(10):4951 – 5.
4. Rothstein RJ. One-step gene disruption in yeast. *Methods Enzymol.* 1983;101:202 – 11.
5. Lacy E, Roberts S, Evans EP, Burtenshaw MD, Costantini FD. A foreign beta-globin gene in transgenic mice: integration at abnormal chromosomal positions and expression in inappropriate tissues. *Cell.* 1983;34(2):343 – 58.
6. Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS. Insertion of DNA sequences into the human chromosomal β -globin locus by homologous recombination. *Nature.* 1985;317:230 – 4.
7. Mansour SL, Thomas KR, Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature.* 1988;336(6197):348 – 52.
8. Jasin M, Berg P. Homologous integration in mammalian cells without target gene selection. *Genes Dev.* 1988;2:1353 – 63.
9. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet.* 2005;6(6):507 – 12.
10. Orr-Weaver TL, Szostak JW, Rothstein RJ. Yeast transformation: a model system for the study of recombination. *Proc Natl Acad Sci U S A.* 1981;78(10):6354 – 8.
11. Resnick MA. The repair of double-strand breaks in DNA; a model involving recombination. *J Theor Biol.* 1976;59(1):97 – 106.
12. Thaler DS, Stahl FW. DNA double-chain breaks in recombination of phage lambda and of yeast. *Annu Rev Genet.* 1988;22:169 – 97.
13. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. *Cell.* 1983;33(1):25 – 35.
14. Jasin M, deVilliers J, Weber F, Schaffner W. High frequency of homologous recombination in mammalian cells between endogenous and introduced SV40 genomes. *Cell.* 1985;43:695 – 703.
15. LaRocque JR, Stark JM, Oh J, Bojilova E, Yusa K, Horie K, et al. Interhomolog recombination and loss of heterozygosity in wild-type and Bloom syndrome helicase (BLM)-deficient mammalian cells. *Proc Natl Acad Sci U S A.* 2011;108(29):11971 – 6.
16. Strathern JN, Klar AJ, Hicks JB, Abraham JA, Ivy JM, Nasmyth KA, et al. Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus. *Cell.* 1982;31(1):183 – 92.
17. Nickoloff JA, Chen EY, Heffron F. A 24-base-pair DNA sequence from the MAT locus stimulates intergenic recombination in yeast. *Proc Natl Acad Sci U S A.* 1986;83(20):7831 – 5.
18. Gloor GB, Nassif NA, Johnson-Schlitz DM, Preston CR, Engels WR. Targeted gene replacement in *Drosophila* via P element-induced gap repair. *Science.* 1991;253(5024):1110 – 7.
19. Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol.* 1994;14:8096 – 106.

20. Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A.* 1994;91:6064 – 8.
21. Belfort M, Roberts RJ. Homing endonucleases: keeping the house in order. *Nucleic Acids Res.* 1997;25(17):3379 – 88.
22. Colleaux L, d'Auriol L, Gailbert F, Dujon B. Recognition and cleavage site of the intron-encoded omega transposase. *Proc Natl Acad Sci U S A.* 1988;85:6022 – 6.
23. Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet.* 1996;12(6):224 – 8.
24. Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res.* 1995;23(24):5012 – 9.
25. Choulika A, Perrin A, Dujon B, Nicolas JF. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol.* 1995;15(4):1968 – 73.
26. Liang F, Romanienko PJ, Weaver DT, Jeggo PA, Jasin M. Chromosomal double-strand break repair in Ku80-deficient cells. *Proc Natl Acad Sci U S A.* 1996;93(17):8929 – 33.
27. Richardson C, Moynahan ME, Jasin M. Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. *Genes Dev.* 1998;12(24):3831 – 42.
28. Donoho G, Jasin M, Berg P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol.* 1998;18(7):4070 – 8.
29. Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. *Adv Genet.* 2013;82:1 – 45.
30. Deriano L, Roth DB. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet.* 2013;47:433 – 55.
31. Pannunzio NR, Li S, Watanabe G, Lieber MR. Non-homologous end joining often uses microhomology: Implications for alternative end joining. *DNA Repair (Amst).* 2014;17:74 – 80.
32. Sargent RG, Brenneman MA, Wilson JH. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Mol Cell Biol.* 1997;17(1):267 – 77.
33. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer.* 2007;7(4):233 – 45.
34. Richardson C, Jasin M. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature.* 2000;405(6787):697 – 700.
35. Elliott B, Richardson C, Jasin M. Chromosomal translocation mechanisms at intronic alu elements in mammalian cells. *Mol Cell.* 2005;17(6):885 – 94.
36. Weinstock DM, Elliott B, Jasin M. A model of oncogenic rearrangements: differences between chromosomal translocation mechanisms and simple double-strand break repair. *Blood.* 2006;107(2):777 – 80.
37. Ghezraoui H, Piganeau M, Renouf B, Renaud JB, Sallmyr A, Ruis B, et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol Cell.* 2014;55(6):829 – 42.
38. Jasin M, Rothstein R. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol.* 2013;5(11).
39. Ferguson DO, Holloman WK. Recombinational repair of gaps in DNA is asymmetric in

- Ustilago maydis* and can be explained by a migrating D-loop model. *Proc Natl Acad Sci U S A*. 1996;93(11):5419 – 24.
40. Johnson RD, Jasin M. Double-strand-break-induced homologous recombination in mammalian cells. *Biochem Soc Trans*. 2001;29(Pt 2):196 – 201.
41. Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev*. 1999;13(20):2633 – 8.
42. Kass EM, Helgadottir HR, Chen CC, Barbera M, Wang R, Westermarck UK, et al. Double-strand break repair by homologous recombination in primary mouse somatic cells requires BRCA1 but not the ATM kinase. *Proc Natl Acad Sci U S A*. 2013;110(14):5564 – 9.
43. Stark JM, Pierce AJ, Oh J, Pastink A, Jasin M. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol*. 2004;24(21):9305 – 16.
44. Johnson RD, Liu N, Jasin M. Mammalian XRCC2 promotes the repair of DNA double-strand breaks by homologous recombination. *Nature*. 1999;401(6751):397 – 9.
45. Moynahan ME, Chiu JW, Koller BH, Jasin M. Brca1 controls homology-directed DNA repair. *Mol Cell*. 1999;4(4):511 – 8.
46. Moynahan ME, Cui TY, Jasin M. Homology-directed DNA repair, mitomycin-c resistance, and chromosome stability is restored with correction of a Brca1 mutation. *Cancer Res*. 2001;61(12):4842 – 50.
47. Moynahan ME, Pierce AJ, Jasin M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell*. 2001;7(2):263 – 72.
48. Bunting SF, Callen E, Wong N, Chen HT, Polato F, Gunn A, et al. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell*. 2010;141(2):243 – 54.
49. Jensen RB, Carreira A, Kowalczykowski SC. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*. 2010;467(7316):678 – 83.
50. Tutt A, Bertwistle D, Valentine J, Gabriel A, Swift S, Ross G, et al. Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *EMBO J*. 2001;20(17):4704 – 16.
51. Johnson RD, Jasin M. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J*. 2000;19(13):3398 – 407.
52. Chandramouly G, Kwok A, Huang B, Willis NA, Xie A, Scully R. BRCA1 and CtIP suppress long-tract gene conversion between sister chromatids. *Nat Commun*. 2013;4:2404.
53. Pierce AJ, Hu P, Han M, Ellis N, Jasin M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev*. 2001;15(24):3237 – 42.
54. Kass EM, Jasin M. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett*. 2010;584(17):3703 – 8.
55. Richardson C, Jasin M. Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells. *Mol Cell Biol*. 2000;20(23):9068 – 75.
56. Costantino L, Sotiriou SK, Rantala JK, Magin S, Mladenov E, Helleday T, et al. Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science*. 2014;343(6166):88 – 91.
57. Malkova A, Ira G. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev*. 2013;23(3):271 – 9.
58. Baliga R, Singleton JW, Dervan PB. RecA oligonucleotide filaments bind in the minor groove

- of double-stranded DNA. *Proc Natl Acad Sci U S A*. 1995;92(22):10393 – 7.
59. Moure CM, Gimble FS, Quijcho FA. The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J Mol Biol*. 2003;334(4):685 – 95.
60. Takeuchi R, Choi M, Stoddard BL. Redesign of extensive protein-DNA interfaces of meganucleases using iterative cycles of in vitro compartmentalization. *Proc Natl Acad Sci U S A*. 2014;111(11):4061 – 6.
61. Doyon JB, Pattanayak V, Meyer CB, Liu DR. Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc*. 2006;128(7):2477 – 84.
62. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus oocytes*. *EMBO J*. 1985;4(6):1609 – 14.
63. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*. 1991;252(5007):809 – 17.
64. Nardelli J, Gibson TJ, Vesque C, Charnay P. Base sequence discrimination by zinc finger DNA-binding domains. *Nature*. 1991;349(6305):175 – 8.
65. Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci U S A*. 1992;89(10):4275 – 9.
66. Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK. Structure of the multimodular endonuclease FokI bound to DNA. *Nature*. 1997;388(6637):97 – 100.
67. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A*. 1996;93(3):1156 – 60.
68. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*. 2010;11(9):636 – 46.
69. Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc finger nucleases. *Genetics*. 2002;161(3):1169 – 75.
70. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. *Science*. 2003;300(5620):764.
71. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc finger nucleases. *Nature*. 2005;435(7042):646 – 51.
72. Varela-Rohena A, Carpenito C, Perez EE, Richardson M, Parry RV, Milone M, et al. Genetic engineering of T cells for adoptive immunotherapy. *Immunol Res*. 2008;42(1-3):166 – 81.
73. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*. 2009;326(5959):1501.
74. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*. 2009;326(5959):1509 – 12.
75. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, et al. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science*. 2012;335(6069):720 – 3.
76. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*. 2012;335(6069):716 – 9.
77. Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ. TAL effectors: highly adaptable phyto-bacterial virulence factors and readily engineered DNA-targeting proteins. *Trends Cell Biol*. 2013;23(8):390 – 8.
78. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-

- RNA- guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096): 816 – 21.
79. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819 – 23.
80. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823 – 6.
81. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014;157(6):1262 – 78.
82. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, et al. Cancer translocations in human cells induced by zinc finger and TALE nucleases. *Genome Res*. 2013;23(7):1182 – 93.
83. Choi PS, Meyerson M. Targeted genomic rearrangements using CRISPR/Cas technology. *Nat Commun*. 2014;5:3728.
84. Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature*. 2014;516(7531):423 – 7.

锌指核酸酶的发展和应⽤

Dana Carroll

摘要 锌指核酸酶 (ZFNs) 是最早被应⽤于基因编辑的靶向核酸酶。该酶被证实具有很⾼的效率,能够提⾼数个数量级的基因靶向频率。锌指 DNA 识别区域的模块性使得 ZFNs 能够在多种组织和细胞系中进行基因编辑。应⽤这一平台能够帮助采集核酸酶相关操作基因组的参数和信息。虽然近些年来这种技术已经逐步地让位于更易操作的 TALENs 和 CRISPR/Cas 核酸酶,但是 ZFNs 仍成功地在包括临床试验等多种领域中得到了应⽤。

关键词 锌指核酸酶 (ZFN) • 非同源末端修复 (NHEJ) • 同源重组修复 (HR) • 基因编辑

D. Carroll, Ph.D. (*)

Department of Biochemistry, University of Utah School of Medicine,
15 N. Medical Drive East, Room 4100, Salt Lake City, UT 84112-5650, USA
e-mail: dana@biochem.utah.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_2

介绍

如果你是一个基因工作者，你有两种方式在你感兴趣的基因上产生突变——前进或后退。通过传统方式（基因前进）你能够产生随机的突变，从而获得一个感兴趣的表型，然后尽力地对具有相应突变的基因特征进行描述，虽然这种特征可能并不是由于该突变造成的。随着基因提取和 DNA 测序技术的发展，该技术正朝向另一方向发展（反向基因技术），首先对于目标基因进行检测鉴定，然后通过特异性产生突变从而测试其表型。然而直到最近，这一工具的运用才受到限制。

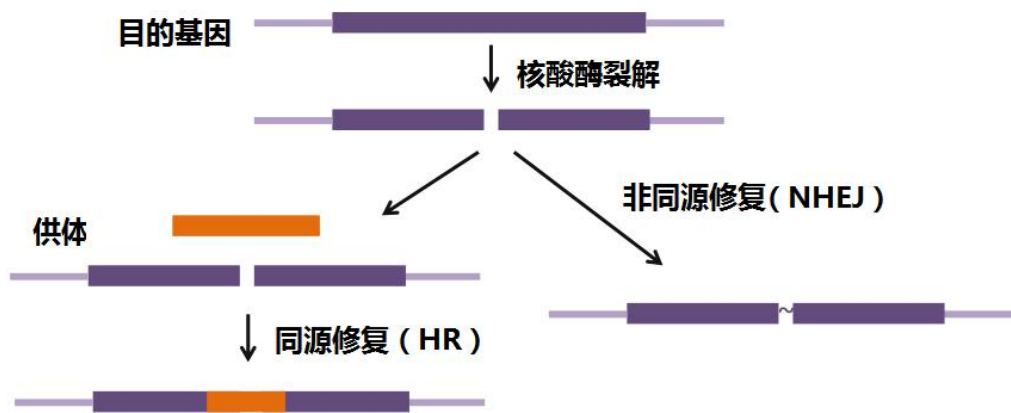


图1.在靶向核酸酶诱导损伤后的基因修饰途径。目的基因用紫色的长方形表示。在核酸酶对目的基因进行裂解后，在同源供体DNA片段存在的情况下，损伤通过同源重组的方式进行修复（HR），将供体片段整合到基因组上。另一种修复通路是非同源末端重组（NHEJ），能够造成损伤位点的短片段突变，如图中的波浪线。

在二十世纪七十年代到八十年代之间，研究者们基于外源性 DNA 模块和内源性基因靶点之间的同源重组（HR）在酵母[1,2]和小鼠[3,4]中建立了多种基因编辑的方法。即使是在酵母中，同源重组发生的频率也很低，但是强筛选使获得目的细胞成为了可能。由于技术上和生物学原理上的一些原因，这些方法很难扩展用于其它组织。在二十几九十年代初，全基因组被解析完毕，进行基因操作的需求逐渐提升。研究者们能够发现他们感兴趣的序列，但是很难去改变它们。

其中的一个很重要的原因是靶向基因的识别要依赖于细胞 DNA 修复酶的活性，然而在通常情况下目的基因是完整的，并不需要修复。在染色体上发生的双链损伤（DSB）是致死性损伤，必须得到修复[5]。而且，在包括减数分裂等很多情况下 DSBs 都能够刺激 HR 的进行。因此，提高基因编辑中同源修复的效率就需要对目的基因进行破坏或损伤。除了 HR 之外，在所有的组织中 DSBs 的修复还可以通过另一条易错途径——非同源末端修复（NHEJ）。在常规情况下，靶点的 DSBs 能够通过两种方式产生突变，一是非同源末端重组（NHEJ）以及在供体 DNA 存在的情况下对目的基因进行置换（图 1）。

已经有多种方法被采用来提高基因靶向的特异性，其中的大多数将在本章中进行讨论。其中最为主要的影响因素是识别和切割的模式。本书讨论内容包括锌指核酸酶（ZFNs）、转录激活样效应核酸酶（TALENs）以及 CRISPR/Cas RNA 介导的核酸酶（CRISPRs）。我的任务就是对 ZFNs 的发展和应用中的标志性成果进行回顾。更加详尽的回顾也可在其他地方获得 [6-9]。

ZFNs 的起源

ZFNs 并不是自然产生的蛋白质，但是它们来源于天然的成分。早在二十世纪九十年代，Chandrasegaran 及其同事发现 II 型限制酶——FokI 具有独立的 DNA 识别结构域和 DNA 切割结构域 [10]。这一发现启示了人们可以将没有特异性的切割结构域和人工改造过的 DNA 连接结构域相连使其获得全新的靶向特异性。这一想法首先通过 *Drosophila* 转录因子的同源异型框——Ubx 的融合成为了显示 [11]。

与此同时，可作为 DNA 连接转录因子的被称为锌指的重复性结构模块在多种真核细胞中被发现 [12,13]。这一结构由 Pavletich 和 Pabo 发现 [14]，识别模块中每三个锌指结构决定一段特异性的序列，在每个指状结构中锌原子的状态由一个半胱氨酸残基和两个组氨酸残基共同决定。通常情况下每个指状结构对应三个连续的 DNA 碱基对 [15]（图 2）。Chandrasegaran 在其第二种嵌合酶的 FokI 裂解结构域中融合了两种由其同事 Jeremy Berg 提供的不同的锌指结构，并使其具有了全新的特异性 [16]。这些嵌合酶就是最初的 ZFNs。

1996 年的研究发现多种 DNA 序列都可被锌指特异性靶向，也或许存在一种可以识别全序列的编码语言 [17,18]。但是后一种猜想并未实现 [19]，但是前一种发现已被证实。通过设计、分选和特征描述，研究者们已经研发出多种类型的锌指嵌合酶并可对多种不同的序列进行整合。而通用识别编码的建立则由于锌指结构在不同环境中活性的无法一致而失败 [20,21]。

ZFNs 的特征

Smith 等人 [22] 及 Bibikova 等人 [23] 通过溶液和细胞中的实验发现了 ZFN 进行裂解的影响因素。FokI 裂解结构域必须通过二聚化才能形成一个具有功能的核酸酶 [22,24]。很明显在锌指核酸酶嵌合的过程中自然状态下一些与二聚化相关的因子缺失了。因此，由于二聚化的作用变弱，蛋白只有在很高的浓度下才能发挥功能 [24]。为了通过 ZFNs 获得高效率的剪切，需要设计两组锌指结构，每一组锌指都与剪切模块相连并直接对目的 DNA 区域附近进行切割 [22,23]（图 2）。在高浓度下，二聚化和剪切更易发生。

ZFNs 能够对完整细胞的染色体进行切割并促进同源重组的发生 [23]。然而向非洲爪蟾卵母细胞中注射人工修饰的模块，却发现细菌的 FokI 核酸酶在该环境下无法正常识别这些序列。ZFNs 究竟是如何识别这些染色体序列的我们目前还不知道，但是我们也未发现有何因素能够导致其活性的丧失。在卵母细胞系统中，如果识别区域和切割区域之间的连接部分少于四个氨基酸，那么锌指连接结构之间的最佳距离是 6bp [23]。而这种连接方式也是目前最常采用的，在多项研究中也都采用 6bp 间隔的模式 [25,26]，但这不代表 5bp 甚至是 4bp 的间隔就肯定会造成一些不同。

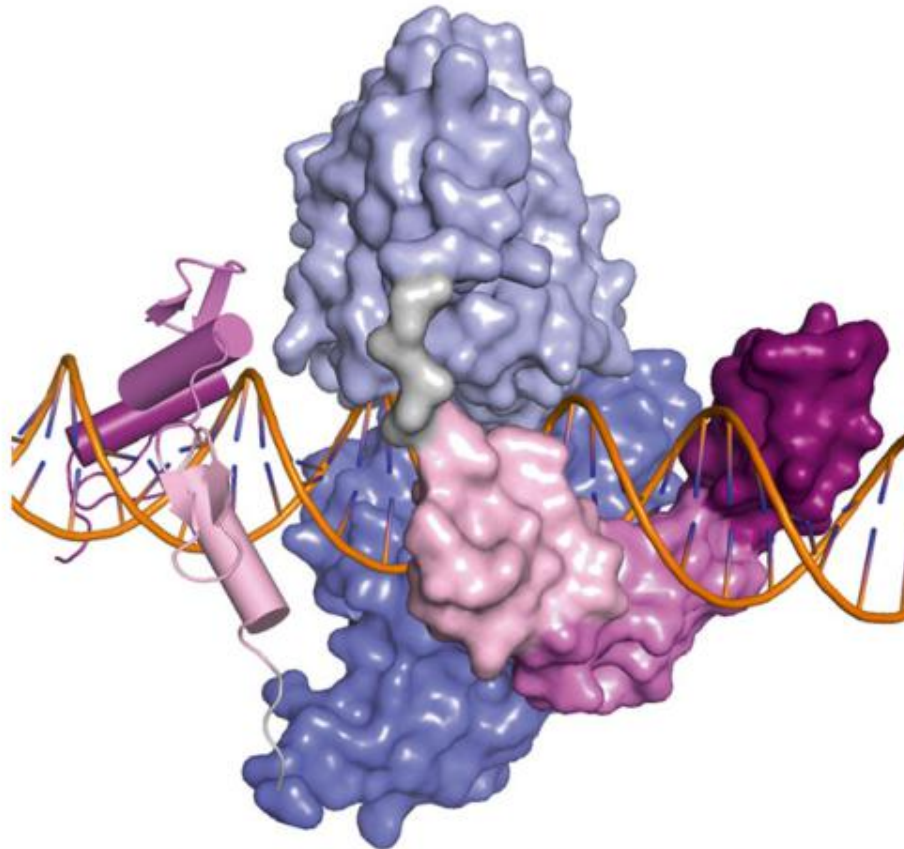


图2.三指ZFNs与DNA的连接模型。锌指结构是位于左侧的粉红带状结构和右侧的空间填充结构。FokI核酸酶是蓝色，之间的连接部分是灰色。DNA链水平插入指状结构中，其骨架为橙色。锌指结构之间相距6bp，模型的构建依托于 1MEY和2FOK蛋白数据库。

ZFNs 的设计

如上所述，锌指识别结构的模块性无法转化为简单的编码语言。有很多课题组也通过设计或筛选合成了新的指状结构或对其进行了新的组合。通过对已有的指状结构进行重新的排列组合在一些时候有效，但是并不可靠[21,27-29]。有很多研究者建立了筛选系统[30-33]，并将性能可靠的指状结构整合到一起[34,35]，并试图去探究究竟是什么因素调控着识别过程的进行[36]。在这个过程中多种特异性的 ZFN 对被研发出来（例如，见[8,36]），但是目前还没有足够简单的创造新结构的方法。

在多年之前，Klug 课题组发现在 DNA 识别序列中再连入两个锌指结构能够提高其效率[9,37]。具有更低脱靶效率的六指序列从此被建立，其中一个结构发生错配都会影响全部两个锌指单位的稳定性[37]。最全的锌指和锌指对的数据库由 Sangamo Bioscience 等整理，所有数据收录在 Sigma-Aldrich（<http://www.sigmaaldrich.com/life-science/zinc-finger-nuclease-technology.html>）数据库中。这些信息的价值在最近几年内大幅贬值，但仍然相对高昂。Sigma 的优势在于可以进行 ZFNs 全部的设计、检测和应用。

最早用于基因靶向的 ZFNs 在高表达时由于脱靶效应的存在具有显著的毒性[38,39]，而这一问题也同样困扰着其他的新的设计。在识别区域两侧增加锌指单位的数量被认为可以降低

低脱靶率，但这一策略并不是经常有效[40,41]。毒性经常是由锌指对中的一个 ZFN 引起的，而且由于其不完整的特异性，能够导致同源二聚化并引起非目标位点即脱靶位点的切割[39]。前述过程中的一个关键性步骤是介导切割结构域二聚化作用的替换，用必要的异源二聚化阻止同源二聚化的发生[42,43]。后一种的异源二聚体的产生[44]也能够修复由前一种同源二聚体引起的缺陷，而这些在通常情况下现在重要通过 ZFNs 和 TALENs 完成[45]。

在科研中的很多情况下，位于脱靶位点的切割和突变产生（通过 NHEJ）处于一个可接受的范围内，这一效应可以通过回交、补充或等位基因的独立表达使其降到最低（例如[46]）。然而在人源样本的应用中，组织的选择就显得更为重要。基于体外识别和切割的特异性，用来检测脱靶突变的方法已经得到了发展[9,47-49]。实验最后，还需对细胞和组织中的其他位点是否产生其它的二次突变进行检测[50]。由于细胞群落中的罕见突变在导入患者体内时会被选择且作用有可能被放大，因而对于全基因组测序的深度具有一定的要求。

在基因编辑中 ZFNs 的应用

ZFNs 作用于完整组织的首次实验是直接靶向于黑果蝇基因序列，通过 NHEJ 进行原位突变生成[38]或通过 HR 进行序列的置换[52]。随后类似的实验在人类细胞中开展，靶向于人工的[53]或自然的[41]序列，在一些研究中也作用于模型组织。到现在为止，ZFNs 已经在从酵母到蝴蝶到人类等 25 个不同的物种中进行成功的应用[7,8]。

该技术每一次在新的组织、细胞类型或终端进行应用时，都必须对核酸酶进行重新设计，包括在哪里需要 HR，如何设计供体 DNA 等。这些方面在其他的技術中也会涉及，人们从 ZFN 研究中获得的知识加速了 TALENs 和 CRISPRs 技术的发展。

由于 ZFN 造成的基因组的改变是无法逆转的，因此只需要瞬时转染表达。在体外的哺乳动物细胞中，研究者通过质粒[41,53]、病毒载体[54-56]、mRNA 转染和直接的添加蛋白的方式[57]进行 ZFNs 的转导。后一种方式能够起效是因为 ZFNs 具有很强的固有活性，虽然这种活性在不同类型的细胞之间存在差异，但是在仍能在很多类型的细胞中发挥功能。长片段双链供体 DNAs 能够被整合到质粒或病毒载体上，而短片段单链的供体则经常直接加入到培养基中[54,55,58,59]。

为了达到对全组织中的基因进行编辑的目的，必须实现在生殖细胞系中的基因编辑。在极为早期的胚胎中的生殖细胞系大多易于编辑。在这一时期进行 ZFN mRNAs 的注射在包括昆虫[60-63]、鱼[64-68]、蛙[69]、海胆[70,71]、小鼠[72-74]、大鼠[75-77]和兔子[78]等多个种系中均取得了成功。在 HR 发生的案例中[60,71,73-75]，供体 DNA 只是简单的混合于注射液中。在猪和牛中，基因编辑动物的产生是通过在体外诱导体细胞产生突变并将其核酸移植到卵母细胞中[79,80]。

植物对基因编辑试剂的传递提出了更高的挑战。在一些情况下，整个植物能够被解离成单个细胞或愈伤组织，而基因操作可在这些成份上进行。类似的工作包括马铃薯[81]和玉米[82]，ZFN 在其它植物中的递送主要借助农杆菌的 T-DNA 转运[83-86]。在技术的发展过程中也有病毒的运用[87]，但是还没有适用于所有植物的方法。

在植物中进行基因编辑的例子很好的说明了在不同的生物系统中需要采用的最佳方式不同且结果也可能不尽相同。有两项著名的实验揭示了这种技术存在的局限性。当 ZFNs 首次应用于线虫——秀丽隐杆线虫时，体细胞产生了正确的突变，然而生殖细胞系却没有[88]。在这个研究中核酸酶通过整合到转座子的外染色体序列进行递送，在生殖细胞系中很可能受到了强 RNA 干扰。当研究者们直接在发育中的性腺内注入 mRNA 时，ZFN 突变才被发现，但是突变频率很低[89]。随着更高效的 TALENs[89,90]，尤其是 CRISPRs[90-98]的出现，DNA、mRNA 和蛋白的注入效率均得到了提高。

另一个例子是斑马鱼。在 ZFN 的早期突变生成的实验中这一模型取得了成功[66,67,99]，

但是没有观察到 HR 产物。随着更有切割效率的 TALENs 的出现，DNA 寡聚核苷酸（oligos）和长片段、双链供体诱发的 HR 成为了现实[100,101]。有趣的是，很多寡聚产物只具有半同源性，半末端的插入[100]。这一现象说明斑马鱼胚胎更倾向于采用 NHEJ 的方式对 DSB 进行修复。

ZFN 的意义

有关于 ZFNs 和归巢内切酶（也称为大范围核酸酶）[102]的研究为基因编辑系统的优化提供了重要的信息。在上述内容中值得注意的是，这也包括在不同组织和细胞类型中核酸酶的递送。NHEJ 修复和 HR 修复之间的平衡也进行了讨论[60,103]，包括采用单链损伤（缺口）而不是 DSBs 更有可能引起 HR[104-107]。供体 DNA 的设计已经进行过研究[108]，而人工合成的单链寡聚供体的效率也被阐明[59,109]。目标序列的同源需求和目的基因的整合也被发现[108,110]。一种依赖于有限的同源序列进行基因整合的方法也被发现[111]。具有更加复杂的突变诱导能力——包括精确地删除和置换[59,112,113]、cDNA 的插入引起的基因纠正[56,114]以及染色体倒置[115,116]等的基因编辑方法已经建立。除此之外，在不进行筛选的条件下对核酸酶诱导产生的突变进行检测和定量的方法已经建立[42,117,118]。

对于很多的应用领域，编辑的自由性使得 TALENs 和 CRISPR/Cas9 核酸酶更受关注。而且 CRISPRs 具有蛋白单一的优势，而且由于其其在细胞基因库中的特异性由向导 RNAs 决定[119,120]或能靶向于同一细胞内的多个基因[121]。如上所述，通过异源二聚化作用能够提高 TALENs 的内源特异性[45,122]。CRISPR/Cas9 核酸酶的特异性虽然受到了质疑[123-127]，但是已经发展出了一些高效率的操作方式。例如通过缩短靶向序列的方式凸显错配的影响[128]，利用 Cas9 变体只对与向导 RNAs 结合的单链附近位点进行切割[126,129,130]，或者将无活性的 Cas9 和 FokI 切割结构域相融合，在 FokI 切割结构域上连有两个向导 RNAs 能够促进二聚化[131,132]。

在 ZFNs 被 TALENs 和 CRISPRs 完全取代之前，值得我们注意的是如果一个单链受到了反复的攻击，这与它所选用的平台没有任何关系。对于在人类治疗、牲畜和农作物的应用领域，切割试剂只需要很小的投入就能达到目的。有关于特异性和编辑的自由性就成为了需要主要考量的因素。在此背景下，小尺寸的 ZFNs 在一些情况中可能存在优势。最后，一些具有很强的特异性和效率的 ZFNs 核酸酶仍在使用中[49,114,133-135]。ZFNs 靶向于人类 CCR5 基因[49,136]的疗法已经进行了多年的临床试验[9,137]并被证实安全且在扩大的 I 期分析中被证实有效。另外还有其他的 ZFN 对被用于靶向人类疾病基因[7,8]，这在一些特殊的领域中已经被证实具有实用性并在不远的未来很可能得到进一步的发展。

致谢 我非常感谢已经和我共同奋战在基因编辑项目上多年的实验室的同事，以及所有在这一领域中加速了这项技术发展的研究者们，你们使这项研究更加有趣并具有更好的前景。我实验室的科研工作由 NIH 资助，最近也有 R01 GM078571 的帮助。

参考文献

1. Rothstein RJ. One-step gene disruption in yeast. *Methods Enzymol.* 1983;101:202 – 11.
2. Scherer S, Davis RW. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc Natl Acad Sci U S A.* 1979;76:4951 – 5.
3. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet.* 2005;6:507 – 12.
4. Thomas KR, Folger KR, Capecchi MR. High frequency targeting of genes to specific sites in the mammalian genome. *Cell.* 1986;44(3):419 – 28.
5. Chapman JR, Taylor MR, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell.* 2012;47(4):497 – 510.
6. Carroll D. Genome engineering with zinc-finger nucleases. *Genetics.* 2011;188(4):773 – 82.
7. Carroll D. Genome Engineering with Targetable Nucleases. *Ann Rev Biochem.* 2014;83:409 – 39.
8. Segal DJ, Meckler JF. Genome engineering at the dawn of the golden age. *Annu Rev Genomics Hum Genet.* 2013;14:135 – 58.
9. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet.* 2010;11(9):636 – 46.
10. Li L, Wu LP, Chandrasegaran S. Functional domains in FokI restriction endonuclease. *Proc Natl Acad Sci U S A.* 1992;89:4275 – 9.
11. Kim Y-G, Chandrasegaran S. Chimeric restriction endonuclease. *Proc Natl Acad Sci U S A.* 1994;91:883 – 7.
12. Klug A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem.* 2010;79:213 – 31.
13. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* 1985;4(6):1609 – 14.
14. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å resolution. *Science.* 1991;252:809 – 17.
15. Pabo CO, Peisach E, Grant RA. Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem.* 2001;70:313 – 40.
16. Kim Y-G, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. *Proc Natl Acad Sci U S A.* 1996;93:1156 – 60.
17. Choo Y, Klug A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc Natl Acad Sci U S A.* 1994;91:11163 – 7.
18. Desjarlais JR, Berg JM. Toward rules relating zinc finger protein sequences and DNA binding site preferences. *Proc Natl Acad Sci U S A.* 1992;89:7345 – 9.
19. Wolfe SA, Neklodova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct.* 2000;29:183 – 212.
20. Lam KN, van Bakel H, Cote AG, van der Ven A, Hughes TR. Sequence specificity is obtained from the majority of modular C2H2 zinc-finger arrays. *Nucleic Acids Res.* 2011;39:4680 – 90.
21. Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods.* 2008;5(5):374 – 5.

22. Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* 2000;28:3361 – 9.
23. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim Y-G, et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol.* 2001;21:289 – 97.
24. Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A.* 1998;95:10570 – 5.
25. Handel EM, Alwin S, Cathomen T. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol Ther.* 2009;17(1):104 – 11.
26. Shimizu Y, Bhakta MS, Segal DJ. Restricted spacer tolerance of a zinc finger nuclease with a six amino acid linker. *Bioorg Med Chem Lett.* 2009;19:3970 – 2.
27. Carroll D, Morton JJ, Beumer KJ, Segal DJ. Design, construction and in vitro testing of zinc finger nucleases. *Nat Protoc.* 2006;1(3):1329 – 41.
28. Kim JS, Lee HJ, Carroll D. Genome editing with modularly assembled zinc-finger nucleases. *Nat Methods.* 2010;7(2):91.
29. Segal DJ, Beerli RR, Blancafort P, Dreier B, Effertz K, Huber A, et al. Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. *Biochemistry.* 2003;42:2137 – 48.
30. Christensen RG, Gupta A, Zuo Z, Schriefer LA, Wolfe SA, Stormo GD. A modified bacterial one-hybrid system yields improved quantitative models of transcription factor specificity. *Nucleic Acids Res.* 2011;39(12), e83.
31. Maeder ML, Thibodeau-Beganny S, Osiaik A, Wright DA, Anthony RM, Eichinger M, et al. Rapid “Open-Source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell.* 2008;31:294 – 301.
32. Persikov AV, Rowland EF, Oakes BL, Singh M, Noyes MB. Deep sequencing of large library selections allows computational discovery of diverse sets of zinc fingers that bind common targets. *Nucleic Acids Res.* 2014;42:1497 – 508.
33. Segal DJ, Dreier B, Beerli RR, Barbas III CF. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5' -GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A.* 1999;96:2758 – 63.
34. Gupta A, Christensen RG, Rayla AL, Lakshmanan A, Stormo GD, Wolfe SA. An optimized two-finger archive for ZFN-mediated gene targeting. *Nat Methods.* 2012;9(6):588 – 90.
35. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods.* 2011;8(1):67 – 9.
36. Zhu C, Gupta A, Hall VL, Rayla AL, Christensen RG, Dake B, et al. Using defined finger interfaces as units of assembly for constructing zinc-finger nucleases. *Nucleic Acids Res.* 2013;41(4):2455 – 65.
37. Moore M, Klug A, Choo Y. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. *Proc Natl Acad Sci U S A.* 2001;98:1437 – 41.
38. Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics.* 2002;161:1169 – 75.

39. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics*. 2006;172(4):2391 – 403.
40. Shimizu Y, Sollu C, Meckler JF, Adriaenssens A, Zykovich A, Cathomen T, et al. Adding fingers to an engineered zinc finger nuclease can reduce activity. *Biochemistry*. 2011;50(22):5033 – 41.
41. Urnov FD, Miller JC, Lee Y-L, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous gene correction using designed zinc-finger nucleases. *Nature*. 2005;435:646 – 51.
42. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome cleavage. *Nat Biotechnol*. 2007;25:778 – 85.
43. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol*. 2007;25:786 – 93.
44. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimer architectures. *Nat Methods*. 2011;8:74 – 9.
45. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol*. 2013;14(1):49 – 55.
46. Liu JL, Wu Z, Deryusheva S, Rajendra TK, Beumer KJ, Gao H, et al. Coilin is essential for Cajal body organization in *Drosophila melanogaster*. *Mol Biol Cell*. 2009;20:1661 – 70.
47. Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA. Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Res*. 2011;39(1):381 – 92.
48. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods*. 2011;8(9):765 – 70.
49. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol*. 2008;26(7):808 – 16.
50. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaepffel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol*. 2011;29(9):816 – 23.
51. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*. 2003;302:415 – 9.
52. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. *Science*. 2003;300(5620):764.
53. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science*. 2003;300:763.
54. Ellis BL, Hirsch ML, Porter SN, Samulski RJ, Porteus MH. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. *Gene Ther*. 2013;20(1):35 – 42.
55. Handel EM, Gellhaus K, Khan K, Bednarski C, Cornu TI, Muller-Lerch F, et al. Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-

- associated viral vectors. *Hum Gene Ther.* 2012;23(3):321 – 9.
56. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol.* 2007;25(11):1298 – 306.
57. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas 3rd CF. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat Methods.* 2012;9(8):805 – 7.
58. Asuri P, Bartel MA, Vazin T, Jang JH, Wong TB, Schaffer DV. Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol Ther.* 2012;20(2):329 – 38.
59. Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods.* 2011;8(9):753 – 5.
60. Beumer KJ, Trautman JK, Bozas A, Liu JL, Rutter J, Gall JG, et al. Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc Natl Acad Sci U S A.* 2008;105(50):19821 – 6.
61. Merlin C, Beaver LE, Taylor OR, Wolfe SA, Reppert SM. Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases. *Genome Res.* 2013;23(1):159 – 68.
62. Takasu Y, Kobayashi I, Beumer K, Uchino K, Sezutsu H, Carroll D, et al. Targeted mutagenesis in the silkworm *Bombyx mori* using zinc-finger nuclease mRNA injections. *Insect Biochem Mol Biol.* 2010;40:759 – 65.
63. Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T, et al. Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. *Nat Commun.* 2012;3:1017.
64. Ansai S, Ochiai H, Kanie Y, Kamei Y, Gou Y, Kitano T, et al. Targeted disruption of exogenous EGFP gene in medaka using zinc-finger nucleases. *Dev Growth Differ.* 2012;54(5):546 – 56.
65. Dong Z, Ge J, Li K, Xu Z, Liang D, Li J, et al. Heritable targeted inactivation of myostatin gene in yellow catfish (*Pelteobagrus fulvidraco*) using engineered zinc finger nucleases. *PLoS One.* 2011;6(12), e28897.
66. Doyon Y, MaCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, et al. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol.* 2008;26:702 – 8.
67. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol.* 2008;26:695 – 701.
68. Yano A, Guyomard R, Nicol B, Jouanno E, Quillet E, Klopp C, et al. An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. *Curr Biol.* 2012;22(15):1423 – 8.
69. Young JJ, Cherone JM, Doyon Y, Ankoudinova I, Faraji FM, Lee AH, et al. Efficient targeted gene disruption in the soma and germ line of the frog *Xenopus tropicalis* using engineered zinc-finger nucleases. *Proc Natl Acad Sci U S A.* 2011;108(17):7052 – 7.
70. Ochiai H, Fujita K, Suzuki K, Nishikawa M, Shibata T, Sakamoto N, et al. Targeted mutagenesis in the sea urchin embryo using zinc-finger nucleases. *Genes Cells.* 2010;15(8):875 – 85.
71. Ochiai H, Sakamoto N, Fujita K, Nishikawa M, Suzuki K, Matsuura S, et al. Zinc-finger nuclease-mediated targeted insertion of reporter genes for quantitative imaging of gene

- expression in sea urchin embryos. *Proc Natl Acad Sci U S A*. 2012;109(27):10915 – 20.
72. Carbery ID, Ji D, Harrington A, Brown V, Weinstein EJ, Liaw L, et al. Targeted genome modification in mice using zinc-finger nucleases. *Genetics*. 2010;186(2):451 – 9.
73. Meyer M, Hrabe de Angelis M, Wurst W, Kuhn R. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci U S A*. 2010;107:15022 – 6.
74. Meyer M, Ortiz O, Hrabe de Angelis M, Wurst W, Kuhn R. Modeling disease mutations by gene targeting in one-cell mouse embryos. *Proc Natl Acad Sci U S A*. 2012;109:9354 – 9.
75. Cui X, Ji D, Fisher DA, Wu Y, Briner DM, Weinstein EJ. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol*. 2011;29(1):64 – 7.
76. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*. 2009;325:433.
77. Mashimo T, Takizawa A, Voigt B, Yoshimi K, Hiai H, Kuramoto T, et al. Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS One*. 2010;5(1), e8870.
78. Flisikowska T, Thorey IS, Offner S, Ros F, Lifke V, Zeitler B, et al. Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. *PLoS One*. 2011;6(6), e21045.
79. Hauschild J, Petersen B, Santiago Y, Queisser AL, Carnwath JW, Lucas-Hahn A, et al. Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proc Natl Acad Sci U S A*. 2011;108(29):12013 – 7.
80. Yu S, Luo J, Song Z, Ding F, Dai Y, Li N. Highly efficient modification of beta-lactoglobulin (BLG) gene via zinc-finger nucleases in cattle. *Cell Res*. 2011;21(11):1638 – 40.
81. Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, et al. High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*. 2009;459:442 – 5.
82. Shukla VK, Doyon Y, Miller JC, DeKolver RC, Moehle EA, Worden SE, et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*. 2009;459(7245):437 – 41.
83. Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C, Baltes NJ, et al. Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. *Plant Physiol*. 2011;156(2):466 – 73.
84. Lloyd A, Plaisier CL, Carroll D, Drews GN. Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc Natl Acad Sci U S A*. 2005;102:2232 – 7.
85. Osakabe K, Osakabe Y, Toki S. Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. *Proc Natl Acad Sci U S A*. 2010;107(26):12034 – 9.
86. Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, et al. High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *Proc Natl Acad Sci U S A*. 2010;107(26):12028 – 33.
87. Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF. DNA replicons for plant genome engineering. *Plant Cell*. 2014;26(1):151 – 63.
88. Morton J, Davis MW, Jorgensen EM, Carroll D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. *Proc Natl Acad Sci U S A*. 2006;103(44):16370 – 5.
89. Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, et al. Targeted genome editing

- across species using ZFNs and TALENs. *Science*. 2011;333(6040):307.
90. Lo TW, Pickle CS, Lin S, Ralston EJ, Gurling M, Schartner CM, et al. Heritable Genome Editing Using TALENs and CRISPR/Cas9 to Engineer Precise Insertions and Deletions in Evolutionarily Diverse Nematode Species. *Genetics*. 2013;195:331 – 48.
 91. Chen C, Fenk LA, de Bono M. Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. *Nucleic Acids Res*. 2013;41, e193.
 92. Chiu H, Schwartz HT, Antoshechkin I, Sternberg PW. Transgene-Free Genome Editing in *Caenorhabditis elegans* Using CRISPR-Cas. *Genetics*. 2013;195:1167 – 71.
 93. Cho SW, Lee J, Carroll D, Kim JS, Lee J. Heritable Gene Knockout in *Caenorhabditis elegans* by Direct Injection of Cas9-sgRNA Ribonucleoproteins. *Genetics*. 2013;195:1177 – 80.
 94. Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods*. 2013;10:741 – 3.
 95. Frokjaer-Jensen C. Exciting Prospects for Precise Engineering of *Caenorhabditis elegans* Genomes with CRISPR/Cas9. *Genetics*. 2013;195(3):635 – 42.
 96. Katic I, Grosshans H. Targeted Heritable Mutation and Gene Conversion by Cas9-CRISPR in *Caenorhabditis elegans*. *Genetics*. 2013;195:1173 – 6.
 97. Tzur YB, Friedland AE, Nadarajan S, Church GM, Calarco JA, Colaiacovo MP. Heritable custom genomic modifications in *Caenorhabditis elegans* via a CRISPR-Cas9 System. *Genetics*. 2013;195:1181 – 5.
 98. Waaijers S, Portegijs V, Kerver J, Lemmens BB, Tijsterman M, van den Heuvel S, et al. CRISPR/Cas9-Targeted Mutagenesis in *Caenorhabditis elegans*. *Genetics*. 2013;195:1187 – 91.
 99. Foley JE, Yeh JR, Maeder ML, Reyon D, Sander JD, Peterson RT, et al. Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool Engineering (OPEN). *PLoS One*. 2009;4(2), e4348.
 100. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug 2nd RG, et al. In vivo genome editing using a high-efficiency TALEN system. *Nature*. 2012;491(7422):114 – 8.
 101. Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, et al. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods*. 2013;10(4):329 – 31.
 102. Stoddard BL. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modifications. *Structure*. 2011;19:7 – 15.
 103. Bozas A, Beumer KJ, Trautman JK, Carroll D. Genetic analysis of zinc-finger nuclease-induced gene targeting in *Drosophila*. *Genetics*. 2009;182(3):641 – 51.
 104. Kim E, Kim S, Kim DH, Choi BS, Choi IY, Kim JS. Precision genome engineering with programmable DNA-nicking enzymes. *Genome Res*. 2012;22(7):1327 – 33.
 105. McConnell Smith A, Takeuchi R, Pellenz S, Davis L, Maizels N, Monnat Jr RJ, et al. Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-Anil LAGLIDADG homing endonuclease. *Proc Natl Acad Sci U S A*. 2009;106(13):5099 – 104.
 106. Ramirez CL, Certo MT, Mussolino C, Goodwin MJ, Cradick TJ, McCaffrey AP, et al. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res*. 2012;40:5560 – 8.
 107. Wang J, Friedman G, Doyon Y, Wang NS, Li CJ, Miller JC, et al. Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res*. 2012;22:1316 – 26.
 108. Beumer KJ, Trautman JK, Mukherjee K, Donor CD, DNA. Utilization during gene targeting

- with zinc-finger nucleases. *G3 (Bethesda)*. 2013;3:657 – 64.
109. Radecke S, Radecke F, Cathomen T, Schwarz K. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. *Mol Ther*. 2010;18:743 – 53.
110. Elliott B, Richardson C, Winderbaum J, Nickoloff JA, Jasin M. Gene conversion tracts from double-strand break repair in mammalian cells. *Mol Cell Biol*. 1998;18(1):93 – 101.
111. Orlando SJ, Santiago Y, DeKolver RC, Freyvert Y, Boydston EA, Moehle EA, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res*. 2010;38(15), e152.
112. Lee HJ, Kim E, Kim JS. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res*. 2010;20(1):81 – 9.
113. Lee HJ, Kweon J, Kim E, Kim S, Kim JS. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. *Genome Res*. 2012;22(3):539 – 48.
114. Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature*. 2011;475(7355):217 – 21.
115. Brunet E, Simsek D, Tomishima M, DeKolver RC, Choi VM, Gregory P, et al. Chromosomal translocations induced as specified loci in human stem cells. *Proc Natl Acad Sci U S A*. 2009;106:10620 – 5.
116. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, et al. Cancer translocations in human cells induced by zinc finger and TALE nucleases. *Genome Res*. 2013;23(7):1182 – 93.
117. Dahlem TJ, Hoshijima K, Jurynek MJ, Gunther D, Starker CG, Locke AS, et al. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet*. 2012;8(8), e1002861.
118. Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res*. 2009;19(7):1279 – 88.
119. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, et al. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science*. 2014;343:84 – 7.
120. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic Screens in Human Cells Using the CRISPR/Cas9 System. *Science*. 2014;343:80 – 4.
121. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;153:910 – 8.
122. Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. *Science*. 2011;333(6051):1843 – 6.
123. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*. 2013;41:9584 – 92.
124. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013;31:822 – 6.
125. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31:827 – 32.
126. Mali P, Aach J, Stranges PB, Svelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome

- engineering. *Nat Biotechnol.* 2013;31:833 – 8.
127. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol.* 2013;31(9):839 – 43.
128. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol.* 2014;32:279 – 84.
129. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 2014;24:132 – 41.
130. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell.* 2013;154:1380 – 9.
131. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol.* 2014;32:577 – 82.
132. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol.* 2014;32:569 – 76.
133. Anguela XM, Sharma R, Doyon Y, Miller JC, Li H, Haurigot V, et al. Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood.* 2013;122(19):3283 – 7.
134. Li L, Krymskaya L, Wang J, Henley J, Rao A, Cao LF, et al. Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Mol Ther.* 2013;21(6):1259 – 69.
135. Yusa K, Rashid ST, Strick-Marchand H, Varela I, Liu PQ, Paschon DE, et al. Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. *Nature.* 2011;478(7369):391 – 4.
136. Maier DA, Brennan AL, Jiang S, Binder-Scholl GK, Lee G, Plesa G, et al. Efficient clinical scale gene modification via zinc finger nuclease-targeted disruption of the HIV co-receptor CCR5. *Hum Gene Ther.* 2013;24(3):245 – 58.
137. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med.* 2014;370:901 – 10.

TAL 效应核酸酶的应用和发展

Alexandre Juillerat , Philippe Duchateau , Toni Cathomen , Claudio Mussolino

摘要 在 2009 年植物遗传学家在植物病原体黄杆菌属的转录激活样效应分子 (TALEs) 的 DNA 识别区域发现了全新的 DNA 连接结构。在 DNA 的识别区域的具有独特的分子结构能够阻止其与 DNA 单链结合。该项技术的突破主要源于对每个阻滞区域中与决定 DNA 特异性的关键残基的解析以及在特异性识别调控中这些氨基酸残基所发生的特异性的改变的发现。通过将 TALE 来源的 DNA 识别区域与核酸内切酶转录调控分子或染色体重塑结构域进行融合设计得到的多种分子能够实现对于基因、转录或表观遗传的改变，这是该项技术里程碑式的发现。最近几年之内，酶学设计的发展随着该技术从基础领域到生物领域以及人类基因治疗领域的拓展取得了长足的进步。

关键词 基因编辑 • 基因敲除 • 基因组编辑 • 转录激活样效应核酸酶 • 锌指核酸酶 • TALE 编辑 • TAL 效应分子 • 金门

A. Juillerat , Ph.D. • P. Duchateau , Ph.D.

Research Department , Cellectis SA , Paris , France

T. Cathomen , Ph.D.

Institute for Cell and Gene Therapy , Medical Center - University of Freiburg ,
Hugstetter Str. 55 , 79106 Freiburg , Germany

C. Mussolino , Ph.D. (*)

Center for Chronic Immunodeficiency , Medical Center - University of Freiburg ,
Hugstetter Str. 55 , 79106 Freiburg , Germany

e-mail: claudio.mussolino@uniklinik-freiburg.de

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_3

转录激活样效应分子

转录激活样效应分子 (TALE) 是来源于黄杆菌属的蛋白, 是变形菌门中众多的植物病原体之一。在感染过程中, 包括 TALE 在内的蛋白混合物能够通过III型分泌系统侵入宿主植物细胞的细胞质中。在入核之后, TALEs 能够最小化真核转录因子的功能并整合连接到宿主基因组的顺势调控元件中并调控细胞通路, 并最终实现促进细菌增值的目的[1]。

TALEs 由 N-末端的分泌迁移信号肽、DNA 识别的中央结构域和一个 C 末端的酸式激活结构域、能够迁移入核的核算定位信号端构成[2] (图 1)。

最常见的蛋白编辑策略包括截短 N-和 C-末端, 以期于创造出人工的最小尺寸的具有活性的 TALE 来源的 DNA 识别结合结构域。在下一步部分我们将总结能够作用于多领域的不同的 TALE 脚手架的发展。

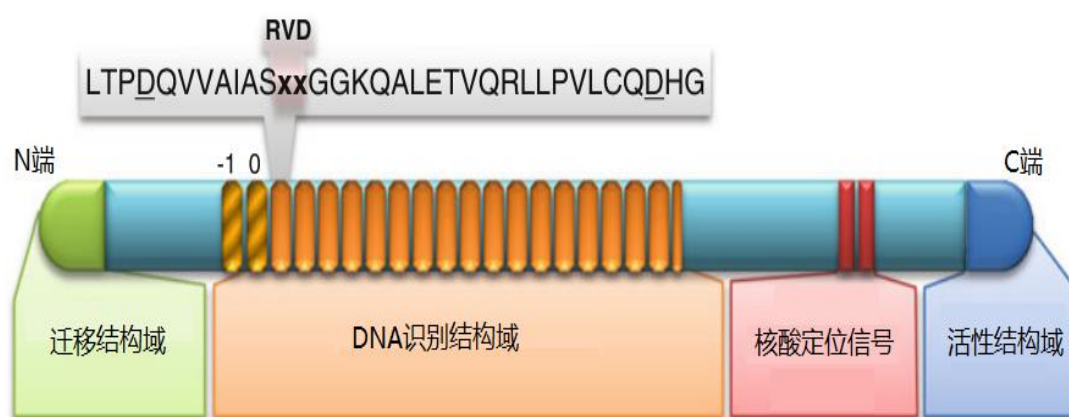


图1.TAL效应分子简图。TAL效应分子 (TALE) 包括一个N-末端迁移结构域和真核样核酸定位信号以及C-末端蛋白部分的活性结构域。蛋白的中央部分是DNA识别区域, 由多个具有DNA连接功能的模块所组成。除了额外的0和-1重复样结构, 蛋白质序列高度保守。在每个模块中的重复性多样化的二残基 (RVD) 具有DNA连接的特异性。

DNA 识别结构域

中央的 DNA 识别结构域是一条多样化的并行重复序列, 长度在 15.5 到 19.5 之间, 最末端的重复通常较短且被当作是“半个重复”。每个模块都由 33-35 个高度保守的氨基酸所组成的 DNA 识别结构域以及在 12 和 13 位点具有高度异质性的双残基 (RVDs) 所组成[2]。这两个氨基酸在核酸识别中发挥了重要的作用, 每个氨基酸对应一个编码即每个 RVD 对应一个单独的核苷酸。通过这种相互作用研究者能够根据 TALE 的蛋白序列发现未知的自然靶点, 或者进行逆向设计对 TALE 的搭档重复模块进行人工重组进而靶向于感兴趣的基因[3,4]。自然状态的 TALE 蛋白与其靶标 DNA 的晶体结构在 2912 年初获得解析, 并在此基础上加深了对 TALEs 进行 DNA 识别的理解[5,6]。每个模块都连有两个 α -螺旋和一个包含有 RVD 的环状结构。螺旋的亚结构为右手螺旋, RVD 残基为向内指向。这个蛋白缠绕在 DNA 双螺旋结构上, RVD 残基直接插入主要槽点, 但对于 DNA 的双螺旋结构没有大的影响。然而, RVD 的两个残基被认为具有不同的功能: 12 位的氨基酸残基用于稳定环状结构而 13 位的残基特异性的

与目的 DNA 的脱氧核苷酸产生特异性的作用。

表1 RVDs特异性

RVD	相应的核酸序列	参考文献
NI	A	[3, 4]
HD	C	[3, 4]
NN	G/A	[3, 4]
NK	G	[3, 4, 7-9]
NH	G	[8, 9]
NG	T	[3, 4]
NG	5mC	[12]
N*	5mC	[13]
H*	5mC	[13]

虽然 RVD-核酸相关编码在自然状态下具有 15 种不同的形式[3,4]，研究者们主要关注于来源于黄杆菌 TALEs 的四种主要的 RVDs：HD、NN、NI 和 NG，分别对应于 C、G 或 A、A、T（表 1）。这种直接的编码形式（例如四种不同的 RVDs 对应四种核苷酸）促使多种分子工具的产生并在多种组织中成功地进行了应用，使得 TALEs 成为 DNA 靶向平台中最具有潜力的形式之一。但是，一个关键性问题是 DNA 靶向工具的特异性。研究者们试图通过减少如 NH 或 NK 等 G 特异性的 RVDs 的频率提升靶向的特异性。但是除了在鸟嘌呤靶向中取得了预期结果外，在 TALE 背景的核酸酶设计和转录激活中，‘NK’ RVD（而不是 NN）的置换在最终的蛋白中会对最终蛋白的活性产生影响[7,8]。当‘NK’ RVDs 位于 N-末端序列时这一影响尤其显著[7,8]。高分辨率的 PthXo1 与靶向 DNA 相互作用的分子模型[6]揭示了 NN 包含结构相对于 NK 序列具有更强的亲和力[7]。另外潜在的环境依赖性的低频率 RVDs 也不容忽视。通过对 23 个自然 RVDs 的筛选，Zhang 及其同事发现在‘NH’ RVD 被替换成 NN 时，对鸟嘌呤具有更强的亲和力并可使生物学活性保持在原来的水平[9]。这一现象也同时被 Boch 及其同事在独立性研究中发现[8]。然而，由于 NH 相关序列样本数还远远未达到需求，建议科研工作者将‘NN’替换成‘NH’ RVDs 用于靶向于鸟嘌呤序列仍为时尚早。因此，虽然‘NN’ RVDs 对于鸟嘌呤和腺嘌呤具有双靶向效应，但是由于其对于所有嘌呤环具有总的高亲和力，‘NN’ RVDs 还是主要用于进行鸟嘌呤的靶向[10]。

最近，有关于改造后的 RVDs 的应用实验正在进行中。我们发现基于不同的 RVDs 的靶向特性进行人工改造后其靶向特异性有所增强[11]。Miller 及其同事研发了全新的 RVDs 试图提高已知 TALENs 的活性和特异性[12]。这项研究发现在 TALE-DNA 的识别区域全新的结构的靶点效率显著提高，并可靶向于全新位点的但互补的 RVDs 产生具有高活性和高特异性的 TALENs。

但是在神经干细胞中，以 TALE 为基础的转录激活却无法激活沉默的 Oct4 促进因子[13]。在 HEK293T 细胞的相同研究中，一个 TALE 相关转录激活因子在体外无法激活外源性转入的乙酰化受体。另一方面，采用 5-硫唑嘧啶-20-胞嘧啶对 DNA 乙酰化进行化学抑制能够使活性复苏。这些结果说明 TALE 为基础的分子工具具有影响 DNA 胞嘧啶甲基化的作用。T-特异性的 RVD ‘NG’ 的结构研究发现，NG 能够与 5-甲基胞嘧啶（5-mC）相互作用，说明 TALEs 能够被编辑识别 CpG 的甲基化[14]。Valton 及其同事[15]发现 TALE 序列中的‘N*’和‘H*’ RVDs（星号表示缺少 13 位残基）能够高效的靶向于 5-mC（表 1）。这项研究证实‘HD’ RVD

与 5-甲基胞嘧啶具有不兼容性，强调了‘N*’和‘H*’相对于‘NG’对于甲基化靶向的优势。因此，在甲基化胞嘧啶存在的情况下，规避 CpG 双核苷酸或对于靶向位点进行初筛能够提高 TALE 相关分子工具的成功率。

在不同植物中基因组的研究使得 TALE 样蛋白在黄枯病和洋葱伯霍尔德菌中的应用成为现实。博尔斯通菌蛋白与黄杆菌属的 TALE 具有很多的相似之处，包括它们的核酸定位和在 C-末端的酸性激活区域[16]。罗伯斯通菌 TALE 样效应分子（RTL 蛋白）的 DNA 识别区域在每个模块化的重复单元中都有 35 个从未在黄杆菌属 TALEs 中发现的含有特定的 RVDs 的保守残基。因此 RTLs 与底物之间具有与传统的 TALE 完全不同的亲和力和特异性。而且，相对于黄杆菌属的 TALEs，RTLs 优先与零位存在鸟嘌呤位点的序列结合。有报道称合成了特异性靶向 5'-胸苷的 TALE 序列[17]。来源于博克氏菌 TALE 样蛋白（Bat 蛋白）具有和黄杆菌属的 TALEs 一样的识别序列[18]。然而，Bat 蛋白通常短小并且具有独特的重复 DNA 识别序列。有趣的是，这些重复具有高度的异质性，只有 40%的已知序列，而且相对于同样的黄杆菌属 TALEs，其亲和力更低。我们已经报道了采用类似的洋葱伯霍尔德菌 TALE 样脚手架进行核酸酶的设计[19]，并创造出能够靶向于新的目标 DNA 序列的 TALE 样蛋白[20]。总的来说，自从 2009 年 TALEs 的‘一对一’编码被发现之后，很多技术的进步拓展了靶向的范围以及基于 TALE 平台的分子工具生物应用组合。

N-末端区域

TALE 效应的早期编辑工作发现 N-末端最初的 152 个氨基酸能够被删除且不影响蛋白质的活性，很可能该段区域在植物细胞中主要的功能是定位[21,22]。然而，人工进一步缩短 TALEs 的 N-末端能够导致其无法与 DNA 结合[23,24]而且 NΔ152 截断结构很快作为参考脚手架而出现。晶体结构显示第一个调控模块残基 152 的附近区域在两个重复样结构之间（通常称为重复 0 和-1），通过-1 重复结构的色氨酸和 DNA 靶向位点的 5'-胸苷相连，而这一特征在自然界的黄杆菌属 TALE 靶向位点中高度一致[2]。这一结构被证实是通过蛋白的这个位点介导 TALE 和 DNA 靶标相连，很容易形成高能键有助于靶向识别。

多种结构能够被整合到 TALE 的 N-末端结构并对最终分子的连接活性没有影响；在研究中，检测标记或纯化标签（如 FLAG，HA，S）或定位信号（核酸，线粒体）已经被成功地整合到天然的 TALEs 并连有不同长度的 N-末端区域。有趣的是，Yang 及其同时发现即使是一些功能性片段，如 FokI 催化片段也能被整合到 TALEs 的 N-末端并产生相应的核酸酶活性[25]。基于不同的 NΔ152，Beurdeley 及其同事融合了 I-TevI 归巢内切酶的催化区域和 TALE 的 N-末端区域。这一整合后的 TALE 核酸酶（cTALEN）同时具有了部分靶向催化可编程的 DNA 靶向区域的优势[26]。Barbas 及其同事进一步的研究了可变的 N-末端影响，运用了增幅的截断数据库筛选策略，发现 NΔ120 或 NΔ128 TALE 变体相对于其他的如嵌合 TALE 重组酶等人工设计的酶具有更大的优势[27]。

C-末端区域

很多 TALE 相关的人工蛋白的研究将关注点放在 C-末端的改造上，研究发现这一区域对于 DNA 的识别连接功能并不重要。诚然，TALE C-末端自然转录激活域和异源激活域能够产生无视内容截断的功能性蛋白[23,24,28,29]。然而，在人工的核酸酶中，C-末端 DNA 识别区域和 FokI 内切酶区域之间连接区域的长度对于两个酶之间的靶向活性和容错长度具有一定的影响。相对于自然长度的 TALENs 脚手架，含有 10-70 个氨基酸的 TALEN 脚手架具有更好的核酸酶活性[7,23,29]。另一方面，不同的 C-末端连接长度（>40 残基）能够引发更大范围的切割活性（12-30bp），具有更短‘连接结构’或完全缺少 C-末端的 TALEN 的 DNA 切割范围受限（13-16bp）[7,23,29]。因此，基于不同的 DNA 靶向位点的间隔长度，要选择相对应的合

适的 C-末端脚手架，通过缩短 C-末端间隔长度的容错率能够通过限制脱靶提高特异性。因此，TALE 的 DNA 连接脚手架的 C-末端能够与多种功能的效应分子连接，如转录抑制子[30]、激活片段[31]、染色体调控因子[32]和核酸酶，并成功地运用于哺乳动物和植物细胞的转录、表观遗传和基因操作中（图 2）。

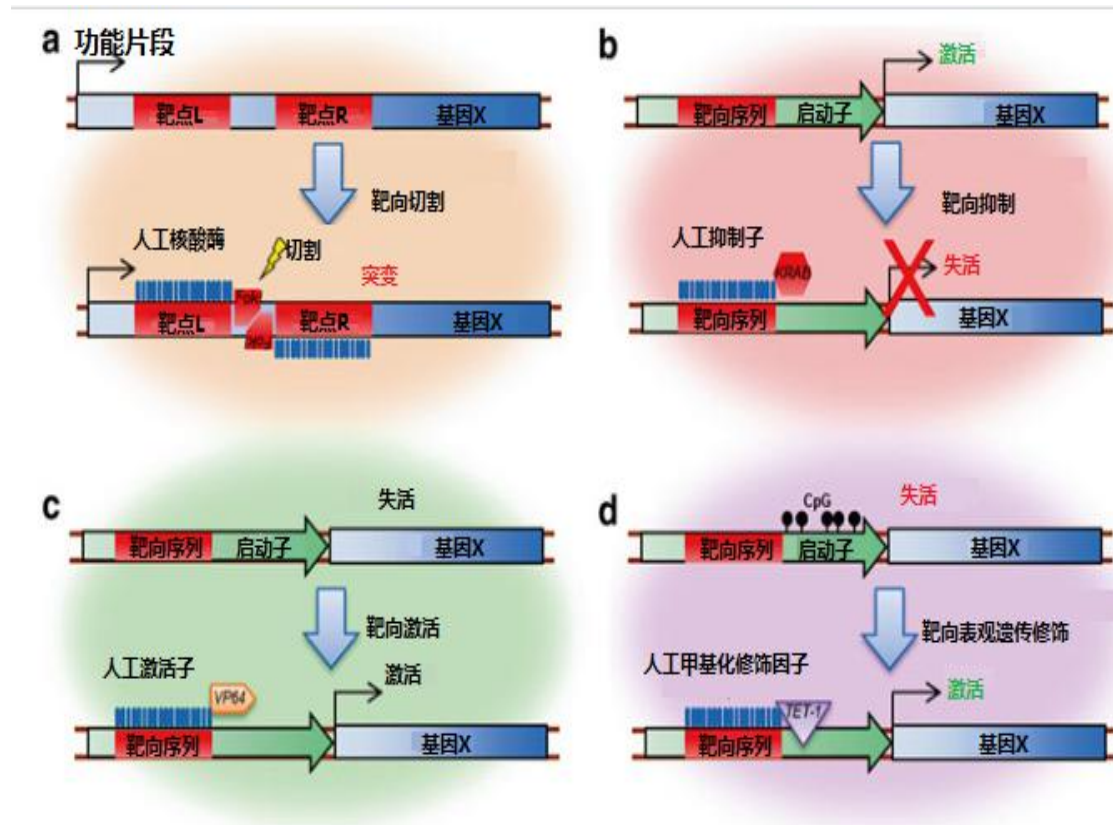


图2. TALE相关效应分子。多种效应分子和TALE DNA识别区域的融合形成了多种功能的融合分子。（a）用于产生DNA双链损伤的人工修饰核酸酶。（b，c）通过人工转录因子的修饰赋予TALEs分子调控转录激活和转录抑制的功能。（d）通过染色体修饰靶向表观调控。

TALE 序列的组装

虽然 TALE 的 DNA 识别结构能够被很容易的编辑靶向于特定的 DNA 目标序列，但是如果目标序列长度超过 100bp，对于传统的编辑策略仍然是个挑战。随着 TALE 相关分子工具的出现，很多平台都具有了快速建立靶向工具的能力[33-42]。这些平台在一些关键性参数上存在着差异，例如起始模块的数量和准备，序列最终的复杂程度及其生产量。在下面的章节我们将会讨论近些年出现的不同 TALE 序列并依据其组装方法将其归为四类(总结见图.3 和表.4)：

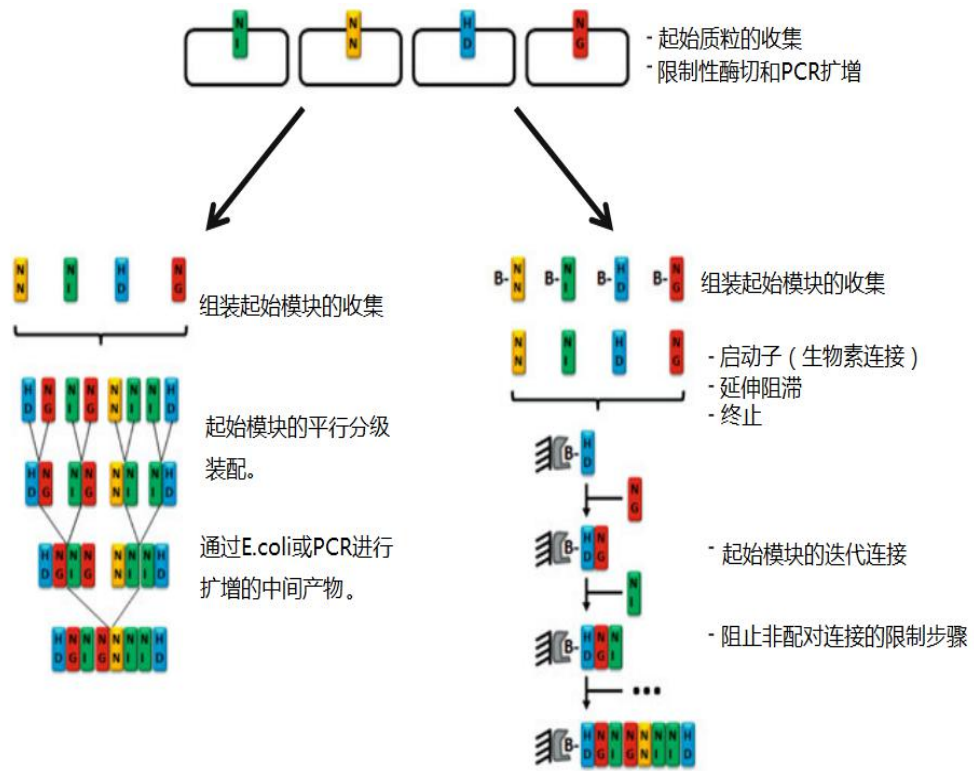
(1) 起始克隆，(2) 金门克隆，(3) 固定相组装，(4) 非结扎依赖性克隆。II S 类型限制酶的主要应用为在其识别位点的固定距离处对 DNA 进行切割，留下一个 4bp 的缺口。因此，如果 II S 识别位点位于每条 DNA 链的相反方向的 5'和 3'末端，采用一种限制酶能够产生四种不同的核苷酸缺口。无论是酶切还是修复过程，最终的结构都不会再包含原有的识别位点（“无痕克隆”）。

TALE 序列既可以在体外合成也可以从商业公司购买。

标准克隆组装

从原理上来讲这种方式是最简单的组装 TALE 序列的方法，而且通过 E.coli 质粒单链编码或多链构建的方式，以及标准化酶切/修复、扩增的步骤能够产生具有平行等级行为的中间序列（图.3a，左）。起始模块的设计能够影响同尾酶（如：单元组装）或 II S 限制酶（如：REAL，REAL-Fast）的整合。基于这个方法，起始质粒的数量的强度可以小于十[43]，或者几打[35,37,44]甚至是几百[30,34,38,45]。每个起始模块的尺寸整合单元的数量相关（在最大的质粒中有一到四个重复单位）。起始模块的准备能够影响 PCR 的扩增或直接降解外源质粒。在每一次的组装周期中，中间产物都能通过克隆 PCR 进行表征并成功完成限制性酶切的过程。基于不同的起始序列的设计，2-8[25]个独立性模块能够被整合到同一个克隆反应中。然而，只有很少的组装方式[37,38,43,45]能够为最终的序列长度和尺寸提供多于一个或两个的可能的灵活性。这些复杂的、大数量的分子生物学步骤（酶切、修复、DNA 质粒分离和 DNA 片段提纯）明显的限制了低产率方法的应用。想要将 TALE 序列组装成标准的基因编辑工具的大小（10-24 个重复单元），至少需要两周（取决于最终序列的长度）。然而，由于基础分子生物学技术已在多个实验室中成功运用，因此这种方法具有一定的优势。

a



b

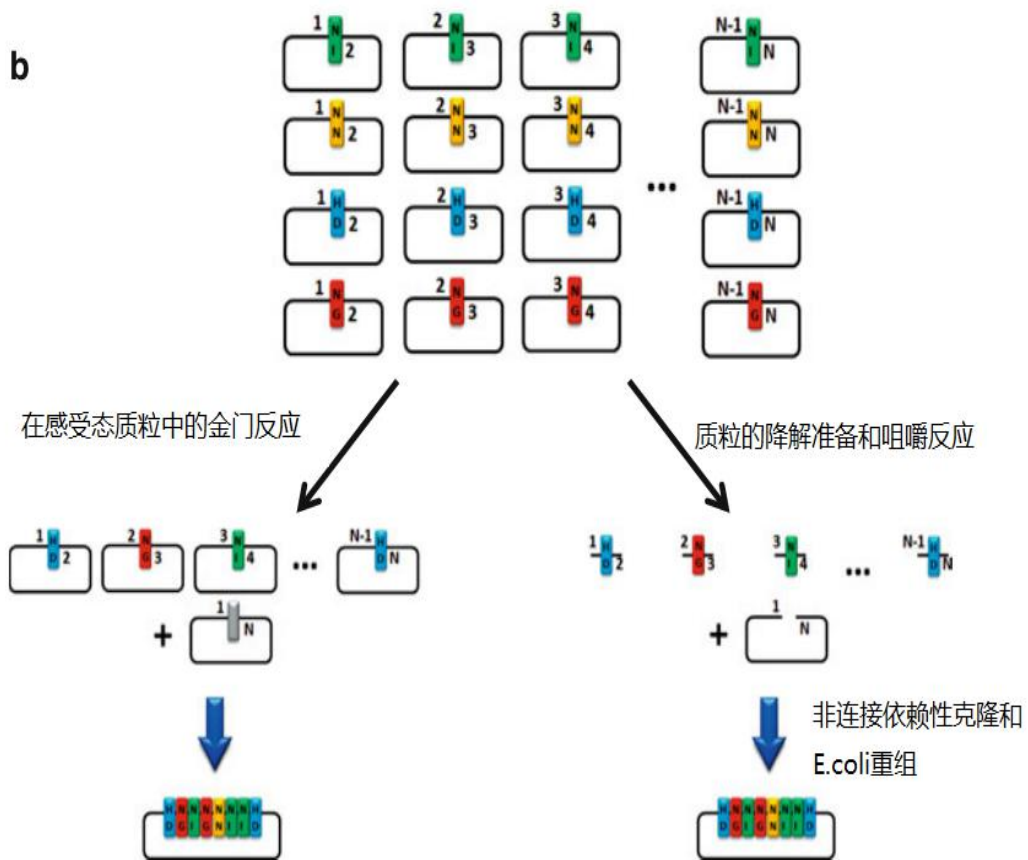


图3.TALE的装配方式。在起始质粒中的独立单元和多组重复单元片段。(a)采用平行分级反应(左)和迭代酶的固相装配反应(右)的简图。(b)金门“一孔一步”反应(左)和非连接依赖性克隆(LIC)(右)。

Table 2 TALE array assembly methods

	Building blocks ^a	Building block preparation	Intermediate amplification step	Possible array length ^b	Estimated time lines ^c	Terminal Half repeat	Throughput	Ref.
Standard cloning assembly	~32	Digestion	Several	various	1-2 weeks	In final plasmid	Low	[31]
	~376	Digestion	Several	various	1-2 weeks	In final plasmid	Low	[29]
	~20	Digestion	Several	various	1-2 weeks	In final plasmid	Low	[28]
	~100	Digestion	No	12.5	2-3 days	In building blocks	Low/medium	[33]
	~48	PCR	1 (<i>E. coli</i>)	16.5 and 24.5	~1 week	In final plasmid	Low	[30]
	~832	Digestion	No	14.5	2-3 days	In building blocks	Low/medium	[32]
	~67	Plasmid	1 (<i>E. coli</i>)	17.5	~1 week	In final plasmid	Low/medium	[35]
	~80	Plasmid	1 (<i>E. coli</i>)	17.5 and 20.5	~1 week	In final plasmid	Low/medium	[39]
	~41	Plasmid	1 (<i>E. coli</i>)	0.5-23.5	~1 week	In building blocks	Low/medium	[38]
	~68-84	Plasmid	1 (<i>E. coli</i>)	11.5-30.5	~1 week	In building blocks	Low/medium	[37]
Golden Gate based assembly	~78-94	Plasmid	1 (<i>E. coli</i>)	8.5-31.5	~1 week	In building blocks	Low/medium	[40]
	~48	PCR/digestion	1 (PCR)	12.5	2-3 days	In final plasmid	Low/medium	[19]
	~72	PCR/digestion	1 (PCR)	18.5-24.5	4 days	In final plasmid	Low/medium	[43]
	~64	PCR/USER digestion	no	Described for 14.5 and 17.5	2-3 days	In building blocks	Medium	[41]

Table 2 (continued)

	Building blocks ^a	Building block preparation	Intermediate amplification step	Possible array length ^b	Estimated time lines ^c	Terminal Half repeat	Throughput	Ref.
Solid phase assembly	~380	PCR or digestion	no	up to 19.5	3 days	In final plasmid	Medium/High	[46]
	unknown	unknown	no	16.5	2-3 days	In final plasmid?	Medium/High	[45]
	~30	PCR	no	up to 18.5	3 days	In building blocks	Medium/High	[44]
Ligation Independent cloning assembly	~64	Digestion and chew back reaction	1 (<i>E. coli</i>)	18.5 (9.5-18.5) ^d	3 days	In final plasmid	Medium/High	[47]
	~3072	Digestion and chew back reaction	1 (<i>E. coli</i>)	15.5 (9.5-18.5) ^d	3 days	In final plasmid	High	[47]

^aBuilding blocks collections include intermediate cloning plasmids and final backbones. Building block number is also dependent on the number of different repeat type implemented in the assembly strategy

^bTotal array length includes the last terminal half repeat that could be incorporated in the final receiving TALE-backbone containing plasmid

^cEstimated assembly time lines do not take in account preparation of the starting materials and sequencing of the final array in the final backbone

^dCombination of both LIC-based collection and implementation of single LIC-compatible single repeat extend synthesis flexibility

“金门” 组装

金门克隆技术最早被用于在已知序列中插入多条 DNA 片段酶切克隆[41,46]。这一方法的优势在于所有的克隆步骤（酶切和连接）都可以通过“一步一孔”的标准实验循环（例如温度）完成多片段的重组（例如 PCR 扩增或质粒）（图 3，左）。然而，克隆效率——即在 *E.coli* 转化涂板后，其阳性克隆在多于九条片段以上的克隆整合实验中显著降低[47]。因此，所产生的经典 TALE 序列具有超过 10 个重复单位的结构，需要进行多组平行的金门反应或 PCR[24] 或 *E.coli* 质粒的重组扩增[36,47-50]对于进一步扩增的亚序列在金门反应开始之前进行预先设置。由于原始的方法基于 II S 限制酶的运用，Yang 及其同事[42]开发了一种可变的 PCR 方法进行起始模块的组装。他们的策略依赖于一个单独反应中，PCR 产物进行 USER（尿嘧啶特异性的延伸试剂；[51]）降解后，利用含有尿嘧啶的引物对于起始模块进行扩增。Sanjana 及其同事[52]开发了另一种不同的金门策略，通过对中间序列进行环化并采用核酸外切酶对未反应或未环化的产物进行移除。环化序列随后通过 PCR 进行扩增，并随后结合到目标序列中。由于在 *E.coli* 的扩增中缺少了扩增步骤，减少了副产物生成的可能，并具有高产率的特征，但是 PCR 中间产物纯化的必要性掩盖了这些优势。

到今天为止所有用来进行 TALE 组装的金门策略都需要数十浓度（24-78）的质粒或无需特定仪器的能够在大多数实验室进行重复的 PCR 扩增引物。它们构建 TALE 所需的时间从一天到一周不等，与序列的长度（0.5-30.5）和中间步骤的类型（PCR 或 *E.coli* 质粒扩增）有关。由于这种‘一步一孔’组装策略的发展超出了经典分子生物技术的可能性，但是其所需要的中间步骤，例如涂板、克隆筛选、PCR 筛选和 DNA 分离等限制其作为一种高通量自动化技术的应用。

固相组装

TALE 重复单元的固相组装是一种高通量的替代性金门克隆策略(图 3a, 右)。和肽或 DNA/RNA 的固相化学合成相类似，这是一种基于磁性或包衣进行固相支撑的方法。这种方式能够很简单的移除过剩反应物或改变反应溶剂。因此这一序列能够与起始模块（或起始子）相连并进而在反应缓冲溶液中进行逐步地酶促延长。在没有降解作用发生时起始模块被认为具有保护作用，在酶切产生理想粘性末端的过程中这一模块被认为具有活化的作用。固相和底物之间的耦合作用通过生物素-链霉素亲和素之间的相互作用完成，由于商用的固相支持物的出现，使得生物素寡聚核酸的获得更加便利。

与所描述的前两种方式相一致，固态支撑底物依赖于 II S 限制酶[33,38,45]或同尾酶[38,40]。起始模块由生物素连接的（5'-末端编码序列）启动子，延伸模块和终止端组成。迭代限制装配法（ICA）由 Briggs 及其同事开发[33]，该方法能够通过屏蔽未反应产物（阻滞剂）阻止非互补连接的产生。通过组装前的多重复模块[38,40,45]或单片段 PCR 的扩增或质粒的设计[33]阻止延伸的发展。根据起始模块的尺寸和序列长度的不同，全序列的组装可以在 3 天之内完成。由于这一过程很容易通过流体工作站进行自动化而且由于中间步骤的缺少，这一方法可被用于中等剂量或高通量的生产。

非依赖性连接克隆（LIC）

由 Schmid-Burgk 及其同事[39]提出的策略是一种基础性的过程，不涉及起始模块的酶切或连接步骤（图 3b, 右）。非依赖性的连接克隆能够产生长的（30bp）非回文缺口，在只缺少四种 dNTP's 中的一种的情况下，能够最大的发挥出 T4 DNA 聚合酶的 3'-5'-外源性核酸酶的优势。为了提高组装过程的通量，采用了具有 5 组重复单位的 3072 个编码质粒替换原有质粒或与原有的双重复质粒合用。应用这一策略，一步 LIC 反应能够在 3 天之内将不同尺寸的

序列装配完成。而且，在有限稀释浓度下的细菌生长只在流体操作步骤中完成，所以进一步的 LIC 策略的通量能够进一步地提高。有趣的是，这一策略也可应用于上述的装配方法用来提高生产通量。

TALEs 的设计及其应用

TALE 的发现及其特殊的 DNA 识别方式对于生命科学的发展具有革命性的影响。TALE DNA 识别区域的模块性和 DNA 简单的互动编码推进了 DNA 识别编码从基础研究到人类基因治疗的各种不同领域的应用。自从 2009 年开始，当 TALE DNA 识别编码被发现开始[3,4]，有关于 TALE 研究的文章、技术创新或者应用研究的数量在 2013 年达到 231 件，比 2009 年提高十倍。就如前所提到的，多种不同的工具被用于进行 TALE 分子的装配以对于特定的 DNA 序列进行靶向。自然状态下的 TALEs 的识别区域具有 1.5 到 33.5 个重复单元，17.5 个中位模块数[2]；然而，有报道称最少有 6.5-10.5 个重复单元具有可测量的生物学活性[3]。一旦进行装配后，DNA 识别区域尾端能够与多种效应分子相融合从而产生不同功能的分子。最成功的一类人工酶具有人工设计的核酸酶区域。这些酶具有特异性的识别区域和切割区域，DNA 识别区域和核酸酶区域进行了融合，核酸酶区域通常来源于 FokI 限制酶。在已知位点的正确二聚化能够引起目的区域的双链损伤（DSB）。基因编辑领域得益于核酸酶设计的，能够介导靶向进行基因组的修饰。一旦基因组 DNA 收到了自然或人工的损伤后，细胞能够通过保守修复机制进行序列修复从而避免凋亡。在哺乳动物细胞中，最主要的两条 DNA 修复通路是通过 DSB 进行的 DNA 整合：（1）非同源性末端重组（NHEJ）或（2）基于同源重组（HR）的同源末端修复。NHEJ 在细胞周期中活性很强也是可以最快修复 DSB 的途径；然而，其引发的易错机制能够引起损伤位点小片段的插入/删除，如果 DSB 发生于基因编码区也能够引起基因功能的缺失。与此相反，HR 介导的修复机制能够引起 DSB 损伤位点的精确修复，在自然状态下，DSB 精确修复的信息来源于姊妹染色体，修复过程发生于 S/G₂ 期。HR 介导的 DNA 修复在哺乳动物细胞中十分罕见，发生率在每 10⁴-10⁷ 个细胞之间[53]。然而，在 Jasin 实验室[54,55]的指向性研究中发现，通过调整基因靶向位点的位置或外源性导入具有位点同源性的供体质粒能够将 HR 的效率提高几个数量级。在这些条件下，供体 DNA 上的遗传信息将会被整合到靶向位点上，使得精确的基因重组得以发生。因此，通过 NHEJ 或 HR 修复的机制，能够实现基因的敲除、修复和插入。通过手头的这些工具，科学家可以通过在不同组织中进行反向遗传的手段提升他们对于基因功能的理解[53]。除了基础性的研究之外，基因编辑技术还被发现可用于其他领域，包括生物技术、系统生物学或者人类基因治疗，目前这一技术已经被应用于编辑创造新的玉米、模拟人类疾病的离体细胞系、以及进行了缺陷修复的人类细胞系[56]。十五多年来，Cys₂-His₂ 锌指 DNA 识别区域在用于产生人工核酸酶的领域取得了极大的成功。锌指核算内切酶（ZFNs）是基因编辑领域的一个里程碑，使得在新物种的基因功能研究中进行基因操作[57-62]以及在人类异常基因缺陷中的基因修复[63-65]成为了可能。运用优化后的 ZFNs 进行 HIV 的治疗的 I 期临床试验取得了显著的成功[66]，该试验证实了基因破坏能够被用于进行 HIV 感染的抵抗[67,68]。然而，受限于锌指序列中 DNA 识别的内容以及无法避免的脱靶，使得其在多范围内的应用很难展开[69-72]。

TALE 相关 DNA 识别区域的发现使得可定制化的核酸酶设计平台成为了可能。TALE 相关的核酸酶（TALENs）与成熟的 ZFN 对相比具有更强的特异性和更加多元化的功能[23,73]。具有全新特异性的 TALEN 能够在相对合理的时间内设计完成[48]，每个 DNA 碱基对可以对平均 3 个 TALEN 对[74]。ZFNs 的靶向范围大多低于每 50-500bp[75,76]。因此不难理解为何很多实验室都引入了这种新技术并使得 TALENs 这项技术的应用范围显著性地扩大。

最早的 TALE 相关人工核酸酶的研究可追溯至 2010 年[28]，随后 TALEN 的脚手架及其效率都获得了改良[23,29,77]，这项技术也被应用于多种组织和细胞系中，包括斑马鱼[78]、小

鼠[79]、大鼠[73]、非人灵长类[80]以及人工诱导纯化的干细胞 (iPSCs) [81]中。TALENs 已经成功的被用于人类基因治疗模型中, 包括人类细胞中镰刀贫血症的突变修复[82]、慢性肉芽肿患者中 iPSC 诱导的粒细胞白细胞中 gp91phox 表达基因的修复[83]、杜氏肌营养不良患者细胞中抗萎缩肌蛋白表达的修复[84]以及隐性营养不良性大疱性表皮松解的治疗[85]。有趣的是, 人工核酸酶不但能切割基因组 DNA, 当采用合适的定位信号时, 也可被用于破坏线粒体 DNA[86,87]。虽然只是个开始, 这一方法为线粒体异常的治疗开启了新的途径。

除了核酸酶之外, 其他的效应分子片段也能和 DNA 识别区域末端相融合从而扩大该技术在基因编辑或是转录或是表观遗传中的应用。采用人工转录因子在转录水平调控分子基因的表达成功的解决了锌指相关 DNA 识别区域的不足。在人类疾病鼠模型中, 内源性基因的表达水平可以被高效调控, 提示定制化的转录因子可被用于进行新的疗法的开发[88,89]。随着 TALEs 技术应用的深入, 定制化 DNA 识别区域的优化使得人工转录激活因子和调控内源性基因表达的抑制因子的使用更加频繁[24,30]。采用高通量方法产生大量 TALE 相关转录因子[90]能够进一步扩大系统生物学在全通路的转录调控和新基因网络模拟中的应用。为了进一步扩大人工酶的应用, 定制化的 DNA 识别区域能够和组蛋白去乙酰基转移酶 (HDACs) 和甲基化转移酶 (HMTs) 相互融合从而实现目标位点的表观修饰[91,92]。

定制化 TALE 相关效应分子的特异性

对于未来的研究, 科研人员计划通过对效应分子的设计完成永久性的基因组编辑、表观遗传修饰护着转录调控。因此, 有警告称 TALE 相关的设计酶具有无法避免的脱靶性。而由特异性不强引起的无法预知的脱靶可能会导致意想不到的副作用。在第一代二聚 ZFNs 出现的时候这一问题是无法避免的主要问题[93], 而这一问题随后通过对 FokI 二聚化表面进行从新设计后避免了同源二聚化的发生的方法被解决[94,95]。很多用于评价 TALE 相关酶特异性的方法都集中于在 TALE 阻遏蛋白进行递送后的微阵列分析[30]或对 ZFNs 特异性进行评价的高通量方法。体外对于切割数据库的筛选或对于慢病毒载体整合能力的评估已经实现, 例如对于 CCR5 特异性的 ZFNs 的评估[70,72]。所有的实验都指出这些 ZFNs 具有无法忽视的脱靶活性[71]。其他的人工核酸酶平台的产生, 如 RNA 介导的核酸酶 (RGNs) 和 TALENs, 为基因编辑领域注入了新的力量。然而, 由于 RGNs 具有高度的脱靶效应[96-98], 研究中已经开始采用相对于二代、三代 ZFNs[23,102]毒性更小的二代 TALENs 脚手架[99]和经人工改造后的 CRISPR/Cas9[100,101]。更为重要的是, 我们最近发现高特异性和低毒性直接相关[103]。除了用于临床转化的方法之外, 这些结果都清晰的显示出如 TALENs 一样的高特异性核酸酶平台的重要性。

有证据显示一些 TALE 单位并非只具有唯一的核酸识别位点。虽然 NG、NI 和 HD 都倾向于靶向唯一的核苷酸序列[9], 最常见的 G 特异性的 ‘NN’ 模块也能与腺苷酸连接。如上所述, 有关于全新的、特异性更强的 TALE 模块的系统性研发工作已经展开, 而这将会进一步地提高 TALENs 的特异性[104]。有一个开放性问题曾经提到, 是否可以通过加强切割识别的特异性从而避免基因毒性的风险。基于这一假设, 有研究将 TALE 相关 DNA 识别序列和 I-TevI 归巢内切酶的切割位点相互融合产生了独立的 TALEN (cTALEN) [26]。在这种条件下, I-TevI 切割区域提供了第二重保障, 因为其只有在切割末端为 CNNNG 的情况下才会被激活[105]。由于序列中 I-TevI 的存在使得基因组中的潜在位点数下降, 因而 cTALEN 确实通过克服每个位点的二聚化作用简化了具有催化活性的 TALENs 的产生。而且, 最近有报道称 Lin 及其同事[106]发现 TALE 相关 DNA 识别区域能够能够和重组的大范围内切酶项链从而特异性的靶向于人类基因组。通过这种方式, 重编程的 TALE-I-SceI 融合蛋白像是传统的 TALEN 一样靶向于 β -珠蛋白基因并诱发 HDR 重组, 而且其毒性极大降低。与 ZFNs 相类似, 异源二聚化的 FokI 切割区域具有额外的特异性[107], 经人工改造过的高活性的 FokI 片段的应用有助于提高人

工 TALENs 的特异性和效果[108,109]。除此之外，靶点的理性化选择能够避免对于高度相似片段的错误结合从而最小化脱靶效应[104]，目前有很多联网工具能够帮助研究者完成这一工作[110]。

结论

采用人工设计的核酸酶造成基因永久性改变的实验数量正在以指数级增高。自从最初被认为是荒诞不经的 ZFNs 的定制化限制酶的设想开始[111]，在人类基因治疗中其进展飞速。随着 TALENs 技术的发展，随着部分特征如编辑的自由度、酶活性以及作为基因编程工具的特异性的变化，这类普适性酶的稳定性得以提高。随着具有更高效的 DBSs 诱导活性以及相对于 TALENs 更易编辑的 RGNs 的到来，进一步的加速了这一趋势[112]。虽然在早期 T 细胞中的基因编辑效率已经足够令人深刻[113]，研究者的一个限制步骤为在早期细胞中人工核酸酶的尺寸问题。TALENs 的蛋白尺寸更为巨大，其递送目前仍是一种挑战，尤其是在选用病毒载体的时候。他们的尺寸和高度重复的结构是在病毒载体系统中的限制因素，例如腺病毒载体和逆转录病毒载体。然而，Holkers 及其同事最近发现腺病毒能够将完整的 TALEN DNA 转运到人类细胞中[114]，然而 Yang 及其同事则是通过多样化处理 TALE 重复模块的核苷酸序列从而将 TALEN 基因打包连入慢病毒载体中[115]。由于重新编码的 TALENs 的效率下降，与此同时多种改进方法正在研究中。如 mRNA 模块转染的完整 TALE 编码的非病毒递呈方式被证实能够意外的诱导早期 T 细胞的敲除效率[99,113]，因此不难预测，在不久的将来 TALEN 相关的基因编辑将在多种临床相关细胞类型中取得突破性进展。

参考文献

1. Buttner D, Bonas U. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiol Rev.* 2010;34(2):107 – 33.
2. Boch J, Bonas U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol.* 2010;48:419 – 36.
3. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science.* 2009;326(5959):1509 – 12.
4. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science.* 2009;326(5959):1501.
5. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, et al. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science.* 2012;335(6069):720 – 3.
6. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science.* 2012;335(6069):716 – 9.
7. Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, et al. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One.* 2012;7(9), e45383.
8. Streubel J, Blucher C, Landgraf A, Boch J. TAL effector RVD specificities and efficiencies. *Nat Biotechnol.* 2012;30(7):593 – 5.
9. Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun.* 2012;3:968.
10. Meckler JF, Bhakta MS, Kim MS, Ovidia R, Habrian CH, Zykovich A, et al. Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res.* 2013;41(7):4118 – 28.

11. Juillerat A, Pessereau C, Dubois G, Guyot V, Marechal A, Valton J, et al. Optimized tuning of TALEN specificity using non-conventional RVDs. *Sci Rep.* 2015;5:8150.
12. Miller JC, Zhang L, Xia DF, Campo JJ, Ankoudinova IV, Guschin DY, et al. Improved specificity of TALE-based genome editing using an expanded RVD repertoire. *Nat Methods.* 2015;12(5):465 – 71.
13. Bultmann S, Morbitzer R, Schmidt CS, Thanisch K, Spada F, Elsaesser J, et al. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res.* 2012;40(12):5368 – 77.
14. Deng D, Yin P, Yan C, Pan X, Gong X, Qi S, et al. Recognition of methylated DNA by TAL effectors. *Cell Res.* 2012;22(10):1502 – 4.
15. Valton J, Dupuy A, Daboussi F, Thomas S, Marechal A, Macmaster R, et al. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J Biol Chem.* 2012;287(46):38427 – 32.
16. de Lange O, Schreiber T, Schandry N, Radeck J, Braun KH, Koszinowski J, et al. Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol.* 2013;199(3):773 – 86.
17. Lamb BM, Mercer AC, Barbas 3rd CF. Directed evolution of the TALE N-terminal domain for recognition of all 5c bases. *Nucleic Acids Res.* 2013;41(21):9779 – 85.
18. de Lange O, Wolf C, Dietze J, Elsaesser J, Morbitzer R, Lahaye T. Programmable DNA-binding proteins from *Burkholderia* provide a fresh perspective on the TALE-like repeat domain. *Nucleic Acids Res.* 2014;42(11):7436 – 49.
19. Juillerat A, Bertonati C, Dubois G, Guyot V, Thomas S, Valton J, et al. BurrH: a new modular DNA binding protein for genome engineering. *Sci Rep.* 2014;4:3831.
20. Stella S, Molina R, Lopez-Mendez B, Juillerat A, Bertonati C, Daboussi F, et al. BuD, a helix-loop-helix DNA-binding domain for genome modification. *Acta Crystallogr D Biol Crystallogr.* 2014;70(Pt 7):2042 – 52.
21. Gurlebeck D, Szurek B, Bonas U. Dimerization of the bacterial effector protein AvrBs3 in the plant cell cytoplasm prior to nuclear import. *Plant J.* 2005;42(2):175 – 87.
22. Szurek B, Rossier O, Hause G, Bonas U. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Mol Microbiol.* 2002;46(1):13 – 23.
23. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 2011;39(21):9283 – 93.
24. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol.* 2011;29(2):149 – 53.
25. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* 2011;39(1):359 – 72.
26. Beurdeley M, Bietz F, Li J, Thomas S, Stoddard T, Juillerat A, et al. Compact designer TALENs for efficient genome engineering. *Nat Commun.* 2013;4:1762.
27. Mercer AC, Gaj T, Fuller RP, Barbas 3rd CF. Chimeric TALE, recombinases with programmable DNA sequence specificity. *Nucleic Acids Res.* 2012;40(21):11163 – 72.

28. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. 2010;186(2):757 – 61.
29. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*. 2011;29(2):143 – 8.
30. Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, et al. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Mol Biol*. 2012;78(3):311 – 21.
31. Gao X, Yang J, Tsang JC, Ooi J, Wu D, Liu P. Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. *Stem Cell Reports*. 2013;1(2):183 – 97.
32. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol*. 2013;31(12):1137 – 42.
33. Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res*. 2012;40(15), e117.
34. Li L, Piatek MJ, Atef A, Piatek A, Wibowo A, Fang X, et al. Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. *Plant Mol Biol*. 2012;78(4-5):407 – 16.
35. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, et al. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res*. 2011;39(14):6315 – 25.
36. Morbitzer R, Elsaesser J, Hausner J, Lahaye T. Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res*. 2011;39(13):5790 – 9.
37. Reyon D, Khayter C, Regan MR, Joung JK, Sander JD. Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. *Curr Protoc Mol Biol*. 2012;Chapter 12:Unit 12.15.
38. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 2012;30(5):460 – 5.
39. Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotechnol*. 2013;31(1):76 – 81.
40. Wang Z, Li J, Huang H, Wang G, Jiang M, Yin S, et al. An integrated chip for the high-throughput synthesis of transcription activator-like effectors. *Angew Chem Int Ed Engl*. 2012;51(34):8505 – 8.
41. Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S. Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One*. 2011;6(5), e19722.
42. Yang J, Yuan P, Wen D, Sheng Y, Zhu S, Yu Y, et al. ULtiMATE system for rapid assembly of customized TAL effectors. *PLoS One*. 2013;8(9), e75649.
43. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol*. 2011;29(8):699 – 700.
44. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, et al. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol*. 2011;29(8):697 – 8.
45. Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, et al. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*. 2013;

12(2):238 – 51.

46. Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden gate shuffling: a one-pot DNA shuffling method based on type II restriction enzymes. *PLoS One*. 2009;4(5), e5553.
47. Thieme F, Engler C, Kandzia R, Marillonnet S. Quick and clean cloning: a ligation-independent cloning strategy for selective cloning of specific PCR products from non-specific mixes. *PLoS One*. 2011;6(6), e20556.
48. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011;39(12), e82.
49. Geissler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens SE, et al. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS One*. 2011;6(5), e19509.
50. Sakuma T, Hosoi S, Woltjen K, Suzuki K, Kashiwagi K, Wada H, et al. Efficient TALEN construction and evaluation methods for human cell and animal applications. *Genes Cells*. 2013;18(4):315 – 26.
51. Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res*. 2007;35(7), e55.
52. Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc*. 2012;7(1):171 – 92.
53. Carroll D. Genome engineering with zinc-finger nucleases. *Genetics*. 2011;188(4):773 – 82.
54. Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A*. 1994;91(13):6064 – 8.
55. Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res*. 1995;23(24):5012 – 9.
56. Perez-Pinera P, Ousterout DG, Gersbach CA. Advances in targeted genome editing. *Curr Opin Chem Biol*. 2012;16(3-4):268 – 77.
57. Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, et al. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol*. 2008;26(6):702 – 8.
58. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*. 2009;325(5939):433.
59. Merlin C, Beaver LE, Taylor OR, Wolfe SA, Reppert SM. Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases. *Genome Res*. 2013;23(1):159 – 68.
60. Young JJ, Cherone JM, Doyon Y, Ankoudinova I, Faraji FM, Lee AH, et al. Efficient targeted gene disruption in the soma and germ line of the frog *Xenopus tropicalis* using engineered zinc-finger nucleases. *Proc Natl Acad Sci U S A*. 2011;108(17):7052 – 7.
61. Shukla VK, Doyon Y, Miller JC, DeKaveler RC, Moehle EA, Worden SE, et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*. 2009;459(7245):437 – 41.
62. Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, et al. High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *Proc Natl Acad Sci U S A*. 2010;107(26):12028 – 33.
63. Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, et al. In vivo genome editing

- restores haemostasis in a mouse model of haemophilia. *Nature*. 2011;475(7355):217 – 21.
64. Soldner F, Laganieri J, Cheng AW, Hockemeyer D, Gao Q, Alagappan R, et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*. 2011;146(2):318 – 31.
65. Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood*. 2011;118(17):4599 – 608.
66. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med*. 2014;370(10):901 – 10.
67. Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol*. 2010;28(8):839 – 47.
68. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol*. 2008;26(7):808 – 16.
69. Cathomen T, Joung JK. Zinc-finger nucleases: the next generation emerges. *Mol Ther*. 2008;16(7):1200 – 7.
70. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppl C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol*. 2011;29(9):816 – 23.
71. Mussolino C, Cathomen T. On target? Tracing zinc-finger-nuclease specificity. *Nat Methods*. 2011;8(9):725 – 6.
72. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods*. 2011;8(9):765 – 70.
73. Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S, et al. Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol*. 2011;29(8):695 – 6.
74. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol*. 2013;14(1):49 – 55.
75. Bhakta MS, Henry IM, Ousterout DG, Das KT, Lockwood SH, Meckler JF, et al. Highly active zinc-finger nucleases by extended modular assembly. *Genome Res*. 2013;23(3):530 – 8.
76. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods*. 2011;8(1):67 – 9.
77. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug 2nd RG, et al. In vivo genome editing using a high-efficiency TALEN system. *Nature*. 2012;491(7422):114 – 8.
78. Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, et al. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods*. 2013;10(4):329 – 31.
79. Sung YH, Baek IJ, Kim DH, Jeon J, Lee J, Lee K, et al. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol*. 2013;31(1):23 – 4.
80. Liu H, Chen Y, Niu Y, Zhang K, Kang Y, Ge W, et al. TALEN-mediated gene mutagenesis in rhesus and cynomolgus monkeys. *Cell Stem Cell*. 2014;14(3):323 – 8.
81. Menon T, Firth AL, Scripture-Adams DD, Galic Z, Qualls SJ, Gilmore WB, et al. Lymphoid regeneration from gene-corrected SCID-X1 subject-derived iPSCs. *Cell Stem Cell*. 2015;16(4):367 – 72.
82. Sun N, Liang J, Abil Z, Zhao H. Optimized TAL effector nucleases (TALENs) for use in

treatment of sickle cell disease. *Mol Biosyst.* 2012;8(4):1255 – 63.

83. Dreyer A-K, Hoffmann D, Lachmann N, Ackermann M, Steinemann D, Timm B, Siler U, et al. TALEN-mediated functional correction of X-linked chronic granulomatous disease in patient-derived induced pluripotent stem cells. *Biomaterials* 2015;69:191 – 200.

84. Ousterout DG, Perez-Pinera P, Thakore PI, Kabadi AM, Brown MT, Qin X, et al. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther.* 2013;21(9):1718 – 26.

85. Osborn MJ, Starker CG, McElroy AN, Webber BR, Riddle MJ, Xia L, et al. TALEN-based gene correction for epidermolysis bullosa. *Mol Ther.* 2013;21(6):1151 – 9.

86. Bacman SR, Williams SL, Pinto M, Peralta S, Moraes CT. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat Med.* 2013;19(9):1111 – 3.

87. Minczuk M, Papworth MA, Miller JC, Murphy MP, Klug A. Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. *Nucleic Acids Res.* 2008;36(12):3926 – 38.

88. Mussolino C, Sanges D, Marrocco E, Bonetti C, Di Vicino U, Marigo V, et al. Zinc-finger-based transcriptional repression of rhodopsin in a model of dominant retinitis pigmentosa. *EMBO Mol Med.* 2011;3(3):118 – 28.

89. Rebar EJ, Huang Y, Hickey R, Nath AK, Meoli D, Nath S, et al. Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat Med.* 2002;8(12):1427 – 32.

90. Maeder ML, Linder SJ, Reyon D, Angstman JF, Fu Y, Sander JD, et al. Robust, synergistic regulation of human gene expression using TALE activators. *Nat Methods.* 2013;10(3):243 – 5.

91. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, et al. Optical control of mammalian endogenous transcription and epigenetic states. *Nature.* 2013;500(7463):472 – 6.

92. Bernstein DL, Le Lay JE, Ruano EG, Kaestner KH. TALE-mediated epigenetic suppression of CDKN2A increases replication in human fibroblasts. *J Clin Invest.* 2015;125(5):1998 – 2006.

93. Alwin S, Gere MB, Guhl E, Effertz K, Barbas 3rd CF, Segal DJ, et al. Custom zinc-finger nucleases for use in human cells. *Mol Ther.* 2005;12(4):610 – 7.

94. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol.* 2007;25(7):778 – 85.

95. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol.* 2007;25(7):786 – 93.

96. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* 2013;41(20):9584 – 92.

97. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol.* 2013;31(9):822 – 6.

98. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol.* 2013;31(9):839 – 43.

99. Mock U, Machowicz R, Hauber I, Horn S, Abramowski P, Berdien B, et al. mRNA transfection of a novel TAL effector nuclease (TALEN) facilitates efficient knockout of HIV co-receptor CCR5. *Nucleic Acids Res.* 2015;43(11):5560 – 71.
100. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol.* 2014;32(3):279 – 84.
101. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol.* 2014; 32(6):569 – 76.
102. Bogdanove AJ, Schornack S, Lahaye T. TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol.* 2010;13(4):394 – 401.
103. Mussolino C, Alzubi J, Fine EJ, Morbitzer R, Cradick TJ, Lahaye T, et al. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 2014;42(10):6762 – 73.
104. Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, et al. A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol.* 2013;31(3):251 – 8.
105. Dean AB, Stanger MJ, Dansereau JT, Van Roey P, Derbyshire V, Belfort M. Zinc finger as distance determinant in the flexible linker of intron endonuclease I-TevI. *Proc Natl Acad Sci U S A.* 2002;99(13):8554 – 61.
106. Lin J, Chen H, Luo L, Lai Y, Xie W, Kee K. Creating a monomeric endonuclease TALE-I-SceI with high specificity and low genotoxicity in human cells. *Nucleic Acids Res.* 2015;43(2):1112 – 22.
107. Cade L, Reyon D, Hwang WY, Tsai SQ, Patel S, Khayter C, et al. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.* 2012;40(16):8001 – 10.
108. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods.* 2011;8(1):74 – 9.
109. Guo J, Gaj T, Barbas 3rd CF. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J Mol Biol.* 2010;400(1):96 – 107.
110. Booher NJ, Bogdanove AJ. Tools for TAL effector design and target prediction. *Methods.* 2014;69(2):121 – 7.
111. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. *Proc Natl Acad Sci U S A.* 1996;93(3):1156 – 60.
112. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014;346(6213):1258096.
113. Valton J, Guyot V, Marechal A, Filhol JM, Juillerat A, Duclert A, et al. A multidrug resistant engineered CAR T cell for allogeneic combination immunotherapy. *Mol Ther.* 2015; 23(9):1507 – 18.
114. Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C, et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* 2013;41(5), e63.
115. Yang L, Guell M, Byrne S, Yang JL, De Los AA, Mali P, et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* 2013;41(19):9049 – 61.

神经肌肉疾病的基因编辑

David G. Ousterout and Charles A. Gersbach

摘要 神经肌肉疾病范围很广，包括肌肉或神经功能异常引起的肌肉功能紊乱。遗传性的神经肌肉疾病是多种基因突变的共同结果，包括点突变、基因插入和删除、染色体重构、表观遗传异常以及重复单位的扩张或收缩。靶向基因编辑是在各种异常情况下进行突变修复的具有潜力的治疗方法。在过去的几十年中，在靶向 DNA 识别蛋白和复杂基因组的编辑领域取得了极大的成功。这些飞速发展的基因编辑工具能够实现基因的敲入、互换、或对于人类基因组中异常基因的移除。在本章中，将对几种遗传性的肌肉疾病和神经疾病进行讨论。对于基因编辑工具引起的基因修饰进行了回顾。具有前景的细胞和基因靶向疗法将在神经肌肉异常和基因编辑治疗中进行介绍。最后，将为大家介绍一些基因编辑用于基因修复和特定的神经肌肉异常治疗的例子。

关键词 基因编辑 • 锌指核酸酶 • TALENs • CRISPR/Cas9-肌肉萎缩症 • 运动神经元疾病

D. G. Ousterout , Ph.D. • C. A. Gersbach , Ph.D. (*)
Department of Biomedical Engineering , Duke University ,
Room 136 Hudson Hall , Box 90281 , Durham , NC 27708-0281 , USA
e-mail: charles.gersbach@duke.edu

© American Society of Gene and Cell Therapy 2016
T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_4

缩略语

AAV	腺病毒
BMD	贝克尔型肌营养不良
CRISPR	集束调控性间隔短片段回文重复单位
DM	强直性肌营养不良
DMD	杜氏肌营养不良
DSB	双链损伤
FSHD	肩胛肱骨筋膜营养不良
gRNA	导向 RNA
HDR	同源直接修复
HT	亨廷顿疾病
iRSC	诱导性多功能干细胞
LGMD2B	2B 型肢带型肌营养不良
MGN	大范围核酸内切酶
MM	三好氏远端肌肉病变
NHEJ	非同源性末端重组
RVD	多变二残基单位
SMA	脊髓性肌萎缩
ssODN	单链寡聚脱氧核苷酸
TALEN	转录激活样效应核酸酶
ZFN	锌指核酸酶

介绍

神经肌肉疾病是指能够引起肌肉功能异常的肌肉和神经疾病，这些异常能够导致多种症状，包括从运动障碍到全身瘫痪直至死亡。其中有一些疾病是由于遗传性基因突变引起的重要的基因功能的损伤。重要的是，这些疾病是由于多个大范围的基因突变引起的。包括点突变、基因插入和删除、染色体重构、表观异常和重复单位的扩增或减少。在过去的几十年中，基因编辑技术在效率和编辑简易度上的突破使得该类疾病有了新的可能的治疗途径。这些技术能够实现研究者对于目的基因的插入、改变和置换。在本章中，将会对处于在研的神经肌肉的编辑方式进行总的回顾，也会对这种全新模式疗法的应用以及尚处于基础研究的工具进行讨论。

常见神经肌肉异常的回顾

肌肉萎缩是肌肉疾病的一种，主要特征是进行性肌肉纤维的降解。这一类型的疾病是由于大量的基因突变所引起的，目前已知最少需要 30 种基因的突变[1]。运动神经元异常是由于控制肌肉运动的中枢神经系统的退化导致的肌肉功能的损伤。在传统理念中，这类疾病重要通过表型进行分类，包括发病、疾病进行、肌肉群损伤、遗传模式或其他的非肌肉表型改变[2,3]。然而，很明显独立基因的突变能够引起相似表型的多种异常、在同一基因上的突变也能够引起完全不同的表型。这是由于肌肉内多种蛋白之间相互作用的复杂机制，因此通过分子机制或对常见的表型进行再定义进而重新分类疾病就显得尤为重要，例如 *dysferlin* 肌病和肢带型肌肉萎缩。这一部分将着重介绍五种主要类型的肌肉营养失调和三种运动神经疾病，并在每一部分都介绍一种能够引起基因破坏的不同的方法。

杜氏和贝克尔型肌营养不良

杜氏肌营养不良（DMD）是一种最为常见的单基因遗传异常、在每 3500 个男性婴儿中就有一例发病[4]。这是由 X 染色体携带的隐性异常基因。缺陷蛋白的主要功能是为细胞骨架和细胞外基质提供一个机械连接，从而在肌肉收缩的过程中保护肌纤维膜免受剪切力的影响。突变能够造成肌萎缩蛋白的截短，进而破坏这种机械连接，并最终导致进行性肌损伤并在三十岁的时候造成死亡。目前 DMD 的标准疗法为姑息治疗，主要侧重于采用类固醇和 ACE 阻滞剂平缓呼吸道和延缓心力衰竭。与 DMD 相类似，贝克尔型肌营养不良（BMD）也是一个由肌营养不良基因突变导致的单基因骨骼肌退化疾病。然而在这个疾病中，由肌营养不良基因突变导致的内源性遗传片段的删除能够引起功能缺陷的肌营养不良蛋白的表达。通常，由于 BMD 的致病蛋白仍具有部分功能，因而临床症状较之 DMD 轻。BMD 的症状和病程相较于 DMD 更为复杂，但是具有相似的肌肉退行性模式，只是进展程度相对较慢[4]。因此，随着有效的康复手段和疾病管理的进行，很多 BMD 患者都能独立的生活并于正常人具有大致相同的预期寿命。

有很多手段都被开发出来进行肌营养不良基因的纠正。进行功能性的肌营养不良基因拷贝的递送是一个巨大的挑战，主要的挑战源于它过大的尺寸（>11kb 的编码序列）。小型化的肌营养不良基因载体能够被整合到人工截短的腺病毒载体中，并被传送到具有肌肉问题的组织中[5-7]。这些小型的肌营养不良基因足以改善 DMD 的症状[5]，最小化基因不良的研究仍在进行中[8]。相较于这些方法，在肌营养不良基因原发位点进行基因纠正能够通过修复蛋白功能的方式提升临床效果[9]。在基因和分子治疗领域还有一些相对较新的领域，比如说“外显子跳出”的策略，能够选择性地排除损伤位点的功能，被编辑的基因能够将目标外显子在 mRNA、阅读框中移除出去[10]。直到最近，阅读框跳出策略的总效率仍在人类临床评估中[11-13]。

虽然仍存在一些无法突破的技术瓶颈，但是肌营养不良基因仍是基因治疗的一个主要靶点。有三种截然不同的基因编辑策略被认为具有纠正肌营养不良蛋白表达的良好前景，其中包括点突变的纠正[14]，在保证内源性肌营养不良基因缺陷表达的情况下靶点阅读框的转换和删除[15-18]，在靶位点插入缺失的外显子从而克服患者特异性的突变[18,19]，或者在一个理想基因位点引入一个具有功能的及营养不良基因表达盒子[20]。

2B 型肢带肌营养不良

dysferlin 异常疾病由功能性的 dysferlin 蛋白表达缺陷引起的常染色体隐性异常，是 2B 型肢带肌营养不良（LGMD2B）和三好氏远端肌肉病变（MM）的分子致病基础[21]。dysferlin 是一种钙离子敏感蛋白，肌纤维的纤维膜损伤修复中发挥作用[22]。虽然 LGMD2B 和 MM 的分子机制相同，但是它们的表型由于未知的原因而不同。LGMD2B 的病理表现起始于肩和膝盖的无力（肢带），随后进展为近肢端的无力[23]。由 dysferlin 异常引起的三好氏远端肌肉病变表现为小腿、前臂、手足的无力[23]。这两种疾病通常在三十岁之前发病，虽然患有这两种疾病的患者的运动功能具有障碍，但是并不影响患者的预期寿命。这两种异常性疾病都是由 dysferlin 基因的单点突变引起的蛋白功能异常。对于功能性的 dysferlin 基因的表达修复是有望治愈这类疾病的一种极具前景的疗法[24]。对于 dysferlin 异常疾病，可能的基因编辑策略与 DMD 相类似，包括靶向阅读框的转变即在保留阅读框功能的前提下尽可能删除非必需的基因片段，在患者特异性突变出进行外显子敲除、以及在理想的基因靶位点导入具有功能的 dysferlin 基因盒。

强直性肌营养不良

强直性肌营养不良是由不同的常染色体异常引起的显性突变，包括 1 类强直性肌营养不良（DM1）和 2 型强直性肌营养不良（DM2）。这些疾病具有进行性的多系统障碍的特征，包括渐进式肌肉紧张、肌强直、白内障、内分泌重构和其他的缺陷[25]。这些疾病的起因各有不同，发病时间也从童年到成年不同。通常情况下，DM1 的严重程度高于 DM2。DM1 的遗传机制在于强直性肌营养不良蛋白激酶（DMPK）基因的 3'非编码区三核苷酸序列胞嘧啶-胸腺嘧啶-鸟嘌呤（CTG）重复单位的扩增[26]。通常情况下 CTG 重复单位的数量大约在 5 个到 37 个之间，当单位数量扩增超过 50 个后就会造成 DMPK 基因的不稳定性从而产生疾病[27]。而且，重复单位的尺寸能够直接影响疾病的产生和病程的进展。有趣的是，CTG 单位的扩增能够引起基因位点的染色体压缩和基因沉默[25]，但是基因沉默所引起的后果目前还未知。第二种强制性肌营养不良——DM2，与 ZNF9 基因的内在片段的四核苷酸序列胞嘧啶-胞嘧啶-胸腺嘧啶-鸟嘌呤重复单位的扩增有关[25]。与 DM1 不同的是，DM2 中 CCTG 重复片段的尺寸和疾病的严重程度并没有相关性。

DM1 和 DM2 中重复序列扩增的致病机理还不清楚，有推测认为和 DMPK 或 ZNF9 基因功能的缺失有关，也可能与突变的 RNAs 的毒性有关[25,28]。最近，有研究正是单位扩增引发毒性的机理很可能和 RNA 连接蛋白的破坏引起的肌强直、胰岛素抵抗、和强心功能的多个基因的剪切错配有关[1]。研究显示重复扩增可能引起和肌强直不相关的副作用，例如睡眠失调、智力缺陷和其它的中枢神经异常[1]。通过进行基因编辑的方式删除目标基因或截短重复片段能够有效缓解疾病并保持基因功能，也可将基因完全删除或进行精确修复从而降低毒性。

肩胛骨筋膜营养不良

肩胛骨筋膜营养不良（FSHD）是在肌肉异常疾病中流行度排名第三的常染色体疾病，发生率在每 7000 个新生儿一例[1]。造成这一疾病的遗传机制非常复杂，主要源于 4 号染色体上 D4Z4 基因中异常片段的聚集和表观遗传的改变，这也是该病的病理学特征之一[1]。4 号染色体的 D4Z4 序列具有 3.3kb 的重复序列，每个重复单位具有一个编码 DUX4 蛋白的开放阅读框。有两个事件被认为和 FSHD 的发生相关[29]。第一，D4Z4 序列缩短。这一序列的缩短能够引起 D4Z4 重复单位中相邻 DUX4 阅读框的内源表达增强和聚腺苷酸化信号通路的激活并稳定 DUX4 mRNA 的转录。第二，随着 D4Z4 序列的缩短，周围序列的表观抑制将会被解除并引起该位点 DUX4 基因的表达。DUX4 蛋白随后能够异常活化多个原本只在早期发育中会被激活的基因，从而引起成熟肌肉组织中的细胞毒性[30]。然而，DUX4 的单独表达并不能完全解释疾病的发生，因为不是在所有的 FSHD 患者中都会出现 DUX4 的表达，而且在一些正常人中也会有该基因的表达。这提示除了 DUX4 的表达之外还存在其他的机制能够引起疾病的发生，而这一机制很可能和 4p 染色体低甲基化有关。基因编辑工具很可能通过编辑片段的 DUX4 序列的长短和存在形式或对染色体 4p 的表观遗传形式进行改造从而创造出全新的研究领域和治疗策略。

脊髓性肌萎缩（SMA）

脊髓性肌萎缩是一种常染色体隐性疾病，发生率在美国接近为每 11000 个新生儿一例，并且是导致新生儿死亡的遗传原因之一[31]。目前针对于该疾病几乎没有治疗建议，严重的 SMA 患者通常在童年就引起死亡。在 5 号染色体上的两个同源基因——SMN1 和 SMN2，编码运动神经元蛋白（SMN），在疾病的发生中具有重要的功能。在正常个体中，SMN2 和 SMN1 及其相似，除了在 SMN2 基因中存在单核苷酸突变（C840T）[31]并可引起可变剪切，使得

SMN 的全长编码蛋白中有 10%-20%的 SMN2 转录和另一种可变的 7 号外显子缺失的 SMN (SMN Δ 7) 的产生。重要的是, 7 号外显子的缺失能够引起表达的 SMN 由于寡聚区域的缺失并被快速的降解。SMA 的产生是由于 SMN1 基因的缺失, 由于 SMN2 基因引起的缺陷性的 SMN 蛋白的产生。有趣的是, 由于产生 SMN 全产物的 SMN2 之间也存在差异, 减轻了疾病的严重程度。在严重程度不同的 SMA 病人中的全转录分析发现 SMA 全长转录表达达到 20-25%能够缓解 SMA 的疾病症状[31]。目前有多种在研的治疗方法试图通过 SMN1 基因的修复重塑 SMN 的功能性表达、增加 SMN2 的转录, 或者破坏 SMN2 提高 SMN 全长的表达[32]。通过基因编辑技术导入完整的 SMN 基因表达功能域是一种在根本上解决疾病的一种具有吸引力的方式, 或者通过对 SMN2 的编辑恢复外显子 7 的功能。

亨廷顿疾病

亨廷顿疾病 (HD) 是一种常染色体疾病, 由于亨廷顿 (HTT) 基因的三碱基重复单位的扩增引起。这种疾病是一种典型的遗传性疾病, 由于 HTT 基因 1 号外显子上的 CAG 三重复单位的扩增引起, 这种不稳定的三重复单位能够引起谷氨酰胺多聚突变型亨廷顿蛋白的表达并引起亨廷顿蛋白的聚集和异常退化。这些降解产物随后泛素化并被转运至溶酶体, 诱发凋亡的产生[33]。通常情况下, 36-39 个突变单位能够引起亨廷顿蛋白的降解[34], 然而 40 个或者更多的重复单位能够引起该蛋白完全的降解并在 65 岁时加重疾病的症状[35]。这种疾病通常发病于中年时期, 绝大多数患者发病于 35 岁到 45 岁左右。这种疾病的准确发病机制还不清楚, 减少或剔除扩增的重复单位有可能改善疾病症状。因此, 基因编辑技术有可能通过删除或减少 CAG 重复单位的数量纠正 HTT 基因, 成为一种具有吸引力的疗法。事实上, 已经有基因技术被发展用于这些重复序列的靶向[36]。

基因编辑工具

在近年来大量实验的进行极大地促进了多种平台的基因编辑技术在人类细胞中点突变效率的进步。值得关注的进步是对于病毒载体的编辑、定制化的 DNA 连接酶或者是寡核苷酸。每种进步都提高了基因编辑技术的特异性、活性、简易性和效率。这些进步催生了多种基因编辑工具的产生[37,38], 包括锌指核酸酶 (ZFNs) [39]、转录激活样核酸酶 (TALENs) [40,41] 以及近年来的 RNA 介导的 CRISPR/Cas9 系统[48]、三聚化寡核苷酸酶 (TFO) 复合物[49,50]、整合酶[51,52]和重组编辑后的锌指核酸酶[53-55]或者 TALE DNA 连接功能域[56]。这一部分将会介绍这些基因编辑技术中的进步。

基因编辑策略的介绍

基因编辑能够通过多种策略完成。在通常情况下, 通过在编辑位点产生特异性的双链损伤 (DSBs) 刺激内源性 DNA 修复途径的产生 (图 1)。有两种完全不同的 DNA 修复过程能够被用于进行靶向序列的编辑——非同源末端重组 (NHEJ) 和同源修复重组 (HDR), 以上见章节“基因编辑: 近 20 年双链损伤诱导的基因靶向和突变”中更为详细的介绍。简单来说, NHEJ 是一种随机的、易错的修复过程, 能够被用于介导随机性的小片段的插入和 DNA 损伤位点片段的删除。NHEJ 介导的基因编辑技术能够被用于进行哺乳动物细胞中基因的破坏[38,57]、同源片段的删除[58,59]、以及在双链损伤位点线性化 DNA 片段的整合[60], 以及对异常阅读框的修复[15,61,62]。HDR 采用同源修复 DNA 模板进行双链损伤的修复或在基因组中产生特异性的片段改变, 包括在全基因组的靶向片段插入。HDR 能够以很高的效率 (6%) 在人类细胞的基因功能域插入多达 8kb 的片段[63], 即使有抗生素筛选被用于多种类型的细胞编辑, 但是发生 HDR 修复的水平仍然很低[64-66]。

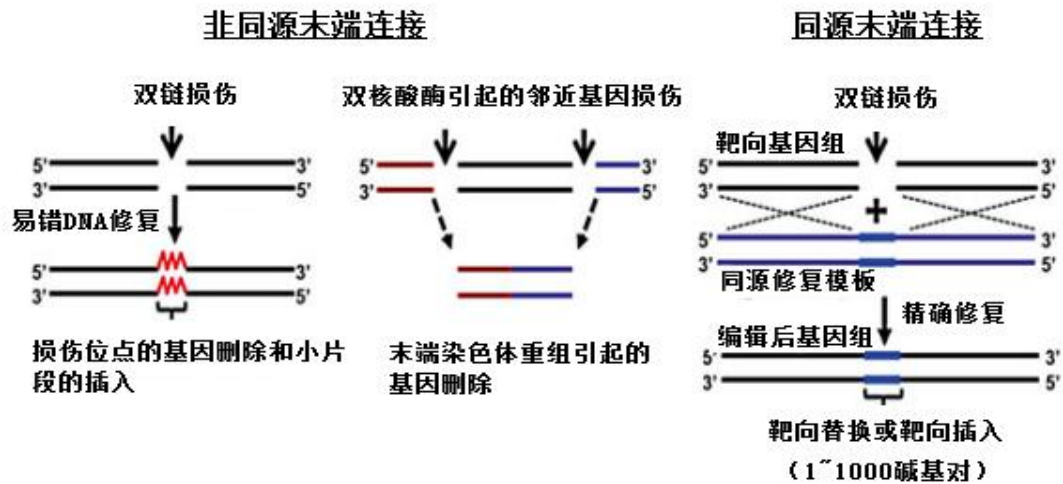


图1. 由编辑后的核酸酶产生的双链损伤诱发的DNA修复机制

靶向基因的模式工具

基因编辑的进步源于新型核酸酶的发现，例如 ZFNs 和 TALEN，以及对于自然条件下的 DNA 连接酶进行重新编辑，包括 MGNs 和最近发现的 CRISPR/Cas9 系统。所有的这些核酸酶都是通过靶向产生 DNA 特异性区段的双链损伤刺激细胞进行 HDR 或 NHEJ 修复（图 1，见“基因编辑策略介绍”部分上面）。与之类似的基因编辑或 mRNA 水平的基因修复技术，如单链寡聚核苷酸技术和腺病毒相关载体也在这一部分进行介绍。

归巢核酸酶

MGNs 具有游离的 DNA 连接和催化区域，能够刺激靶向 DNA 链的识别和裂解[67]。MGNs 在基因编辑技术中最为常见，例如 I-SceI 和 I-CreI，源于以常见多肽基序进行命名的 LAGLIDADG 家族。MGNs 通常以同源二聚化的模式或以单链蛋白的模式进行 DNA 的识别或裂解。相互依赖的 DNA 连接和 DNA 催化活性具有很强的特异性，但是这也使得靶向于全新序列的 MGNs 的重编程变得更为复杂。虽然存在挑战，但是已有 MGNs 被编辑用于全新基因位点的编辑[68]，包括人类细胞的基因相关编辑[19,48,69,70]。

基于 FokI 结构域的嵌合核酸酶

ZFNs 和 TALENs 作为一种嵌合核酸酶，可将独立性的、可编辑的 DNA 连接区域与 FokI 核酸内切酶的非特异性催化区域进行融合[71]。当两种独立的核酸酶与特异性的 DNA 片段识别结合后能够产生点特异性的双链损伤，并由此引发 FokI 的二聚化及靶向 DNA 序列的裂解（图 2a,b）。ZFNs 的 DNA 识别区域基于 Cys2-His2 锌指结构域，也是人类蛋白质组中最为常见的 DNA 连接模式。人工合成的锌指 DNA 连接区域能够通过直接点突变或理性设计或对相关大数据库的筛选提高 DNA 位点的识别效率[38]。这一工作使得 ZFNs 成为早期的被广泛使用的一种基因编辑工具[72-76]。TALENs 的 DNA 识别区域。TALENs 的 DNA 连接区域来源于一种植物病原体蛋白并具有重复性的不同结构的双残基（RVD）模块序列，每一模块能够特异性的识别一种 DNA 碱基对[77,78]。RVD 模块能够进行人工的序列设计用于识别不同的序列[77,78]。而 TALEN 技术则由于其显著地提高了基因编辑的成功率和效率而被广泛采用[40, 77-82]。

CRISPR/Cas 系统

最近，一类来源于细菌内源性免疫系统的新的 DNA 编辑酶被用于进行哺乳动物细胞编辑 [45,47]。与 MGNs、ZFNs 或 TALENs 不同，CRISPR/Cas9 通过向导 RNA (sgRNA) 引起的 RNA-DNA 之间的相互作用识别靶向的 DNA 序列引导嵌合的 Cas9 核酸酶对于互补的 DNA 序列进行裂解 (图 2c)。CRISPR/Cas 系统已经成功的通过 Cas9 酶和向导 RNA (gRNA) 的共表达通过 gRNA 的 5' 端和基因组 DNA 进行的互补配对被用于哺乳动物细胞特定基因位点的编辑 [42-44,46,83]。而这一同源性靶位点必须邻近于一段特殊的前空白序列基序 (PAM)，一段用于 Cas9 核酸内切酶识别的短片段。多种研究探讨了 *S. pyogenes* 中 CRISPR/Cas 系统的特异性前空白区位置依赖性的错配特异性 [84-89]。其余的具有更小 Cas 基因片段的 CRISPR 系统被用于具有尺寸限制的 AAV 载体中 [90]。CRISPR/Cas 系统的一个独特的能力是能够同时性的靶向于不同的基因位点，也就是所谓的多路复用，通过单独的 Cas9 蛋白和两种以上的 gRNA 进行表达 [42,91-93]。

整合酶、重组酶和转座酶

自从核酸酶被直接用于通过 DNA 修复途径的 DNA 改造，能够直接进行靶向 DNA 序列催化改造的酶学编辑研究得到了极大的重视。整合酶通过 DNA 的点特异性的重组在基因组的特定区域插入大片段的基因功能域。ΦC31 整合酶能够被用于人类和小鼠基因组中有限位点的整合，并能够介导具有治疗作用的基因整合到功能已知的基因位点 [51,52]。程序性基因重组也被用于和独特的 DNA 连接功能域共同作用，例如锌指 [53,55,94,95] 或者 TALEs [56]，进行目标基因位点的点特异性重组。转座酶可以和 ZFPs [96]、TALEs [97] 进行融合形成新复合体介导基因表达功能域点特异性整合。最后，逆转录病毒整合酶能够与 ZFPs 在病毒基因位点的特定区域进行直接连接 [98,99]。这些技术的后续发展可能对高效的、可重复的及点特异性的肌肉异常疾病的治疗基因整合具有重要的价值。

腺病毒介导的基因编辑

腺相关病毒 (AAV) 是一种非致病性病毒，具有介导整合最多 4.7kb 长度基因的能力 [100]。(见“基因编辑工具直接介导体内基因修复”部分)。AAV 为基础的递呈系统与其他的技术相比在相同浓度下通过简便的同源重组提高基因靶向的效率的转染技术 [100-108]。这些基因的靶向率根据细胞的不同从 0.1% 到 1% 不等，而且能够造成基因的插入或删除、替换或整个功能域的插入 [100-108]。AAV 为基础的基因靶向具有多种优势，包括基因靶向载体的设计自由度、与其它转染技术相比高效的整合率以及毒性的降低。虽然这些载体引起的总基因靶向率仍然是个障碍，需要更高效率的基因修复，而这一效率也可通过核酸酶介导的双链损伤效率的增强而得到进一步的提高 [100,103-108]。AAV 载体也被可能发生脱靶引发非特异性的基因改变，以及载体的随机性重组 [107,109-111]。

由寡聚核苷酸复合体介导的基因编辑

永久性的基因编辑能够通过多种类型的寡聚核苷酸产生，包括 RNA-DNA 复合体 (chimeric) [49] 和化学合成的单链寡聚核苷酸 (ssODNs) [50,112]。这些寡聚核苷酸能够直接与靶向的 DNA 序列作用并刺激 DNA 修复的发生。通常情况下，ssODNs 对于靶向序列具有很高的敏感性，偏向作用于编码链并能诱导序列发生改变。虽然寡聚核苷酸片段诱导基因编辑的机制还在研究中，但是目前已有研究结果显示化学修饰的 ssODNs 能够提高这一途径的效率 [50,112]。ssODNs 被发现能够介导转染细胞 SMA [113] 和 DMD [114] 疾病相关突变的修复。

体内基因修复的介导

通过在体内环境中进行基因突变的修复从而缓解疾病症状是神经肌肉疾病进行全新基因疗法的终极目标。基因编辑在目的靶点的应用具有很多挑战，尤其是在效率和特异性方面。这一部分将会介绍在体内介导基因修复的多种关键性方法，包括以细胞基础疗法或通过病毒或者非病毒的转染方式直接进行基因编辑酶的导入。

在体内介导基因修复的细胞基础疗法

细胞基础疗法的目的是在病人组织中通过外源性基因修复细胞的移植实现具有治疗作用基因的功能性表达。这些移植细胞能够作为治疗基因或功能替代基因表达的储库。目前有多种类型的细胞正在被研究用于肌肉萎缩的细胞治疗并在将来用可能被用于进行神经肌肉异常的治疗，包括多功能干细胞（iPSCs）[118,119]、骨髓祖细胞[120]、骨骼肌祖细胞[121]、胚层成血管细胞[122]、CD133+细胞[123]和真皮纤维母细胞[61,124]。除此之外，在人类永生化肌肉细胞中所取得的进步极大的简化了基因修复肌肉细胞克隆的获取[125,126]。这一类的细胞能够在体外利用基因编辑技术进行修饰，并通过化学转染或电转的方式将质粒或 mRNA 导入细胞、或利用整合缺陷慢病毒进行转染[127]，或采用细胞可渗透性核酸酶进行处理[128,129]。

杜氏肌营养不良是能够采用基因治疗联合细胞治疗解决的原型疾病。因为在肌肉退化中肌组织更加易于整合外源性的移植祖细胞。最初，DMD 的细胞基础疗法只是通过注射和免疫抑制剂联用的方法将异源性的健康肌肉祖细胞移植到 DMD 患者的肌肉组织中[130]。一些挑战存在于基因编辑技术和细胞基础疗法的联用中，包括对于移植细胞的免疫应答、外源治疗基因的表达、移植细胞数的不足以及移植位点细胞迁移能力变弱。解决这些方法的一个有趣的方案是利用患者自体产生的 iPSCs，在体外纠正缺陷基因并诱导其分化形成肌原组织再将细胞回输到患者退化的肌组织中[15,20,118,119,122,131]。这一方法能够减少移植细胞潜在的免疫损伤并维持修复后细胞的多样性。

通过体内直接性基因编辑工具的导入进行基因修复

传统上来说，基因疗法的目的在于进行病灶部位突变基因的替换。外源性基因可以通过病毒或非病毒介导的方式利用局部注射或全身注射的途径直接进行基因导入。基因编辑疗法的一个重要特点是能够引起基因的永久性改变。因此，编辑工具的瞬时暴露或具有特定基因编辑密码的编辑酶的直接瞬时性表达是最为合适的处理方式，并且减少基因编辑核酸酶暴露的时间能够降低脱靶效率[128,132]。

病毒基因的导入具有病毒遗传性组织趋向的优势。一种较为广泛在体内基因治疗中采用的病毒基因导入载体称为腺相关病毒，是一种小型化的非致病性的非依赖病毒属，最大具有约 4.7kb 容纳装载量的单链 DNA 基因组[133]。重要的是，AAV 载体能够高效的转染非分裂类型细胞，例如肌肉纤维和神经元。AAV 非整合性的自然属性能够降低外源基因整合到基因组中诱发肿瘤或重要基因破坏的风险。对于基因编辑的目的，这一特征也是具有优势的，由于基因编辑核酸酶的表达是瞬时性的，因而也就降低了长时程的毒性。利用 AAV 为基础的外源编辑模板以及基因编辑核酸酶的基因靶向效率被提高[100,103-108,134]。而且，利用 AAV 进行 ZFNs 和外源修复模板的导入在体内环境下被证实具有更高的效率[60,138]。以 AAV2 假病毒为基础的 AAV 载体具有不同的肌肉趋向性 AAV 外壳，例如 AAV2/1、AAV2/6、AAV2/7、AAV2/8、AAV2/9 以及最近的 AAV2.5、AAV/SASTG 和 AAV2i8G9，被证实能够通过有丝分裂旺盛部位数个月的持续性全身静注或局部注射的途径高效导入到多种骨骼肌组织中[7,139-149]。有多种 AAV 载体被证实能够高效的导入神经元、神经胶质细胞和星形胶质细胞

中[150-152]。最近，首项以 AAV 为基础的肌肉注射途径的基因治疗方法在欧洲被获准应用[153]。

非病毒介导方法也是高效的、组织特异性的、具有吸引力的方式。此种方法的一个主要优势使能够对治疗暴露量进行高度精确的控制[154]。纳米颗粒能够被改造连接靶向肽段用于特定组织的传导，可用于多种基因编辑疗法的转染介导。而且，这些复合体除了能够被用于进行 DNA 的传导，也可用于治疗性肽段和 RNA 的转导。值得注意的是，有证据显示纳米颗粒能够同时介导基因编辑工具和合适的修复模板进行体内的基因编辑[155]。因此，非病毒的转导方式也很有可能成为小尺寸基因编辑药物的给药途径。

基因编辑手段在神经肌肉疾病中的应用

基因编辑工具能够被用于进行基因组 DNA 的操作从而进行神经肌肉疾病的治疗与研究。这一部分将要讨论在这一类的疾病中三种主要的基因编辑策略的应用：原位基因的修复、靶向基因的替换以及疾病相关基因的破坏。

原位点基因的修复

基因编辑的一项强大的应用是直接在患者原有的基因位点进行直接性的、高效率的疾病相关基因的修复。这种方法具有很多的优势，包括永久性的基因修复和在自然状态基因调控的作用下调控适当的生理性和组织特异性的表达。这一部分重要采用基因编辑工具进行大片段基因组的删除、小片段的插入、阅读框移码或利用 DNA 外源修复模板插入遗失的外显子（见表 1 和 2）。

通过靶向基因编辑进行阅读框的修复

基因编辑核酸酶能够高效地用于介导靶向基因片段的插入和删除并在肌营养不良基因的非必要区段附近进行阅读框的修复[15,18]。重要的是，在不借助于外源性修复模板的情况下，仅仅通过核酸酶的活性就能够启动靶向阅读框的 NHEJ 易错修复途径从而产生小范围基因的随机插入和删除从而造成移码突变。然而，由于插入或删除的片段长度不等，在修复蛋白的靶向位点会产生不同的序列。这对于修复蛋白的总功能具有决定性的影响，而这一过程也会引起免疫原性的获得进而有可能触发免疫反应和监控调控。

另一种方法能够利用靶向碱基的替换[50]，片段插入或者是删除扰乱早期 mRNA 的形成。有一种更为复杂的方式是共表达两种核酸酶切割相邻的两端基因组位点并高效的介导基因组特定区段的删除[58,157,158]，或对于基因组的异常片段进行有选择性的切除[16-18]。这两种方法均能够导致 mRNA 序列可控的改变因而可用于缺陷蛋白的修复。

表 1.基因编辑技术在神经肌肉疾病创新疗法中的应用策略

基因修复策略	突变诱导类型	基因编辑技术的应用	相关疾病	案例
外显子突变诱导	异常外显子序列	单链寡聚核苷酸, 核酸酶, AAV	SMA	序列突变产生的基因不稳定或致病性外显子删除
内含子基因编辑	终止子失活, 移码突变	核酸酶, AAV, 重组酶/整合酶/转座酶	DMD, BMD, LGMD2B	介导修复阅读框内部片段的小范围插入和删除
非必须外显子的删除	终止子失活, 移码突变	单链寡聚核苷酸, 核酸酶, AAV, 重组酶/整合酶/转座酶	DMD, BMD, LGMD2B	拼接位点破坏或删除整段非必需外显子修复内部缺陷基因
在原基因位点进行缺失位点的插入	移码突变, 重要编码区域的删除	核酸酶, AAV, 重组酶/整合酶/转座酶	DMD, BMD, LGMD2B	在合适的基因区域插入重要的缺失外显子
靶向敲入替代基因	所有编码区突变	核酸酶, AAV, 重组酶/整合酶/转座酶	DMD, BMD, LGMD2B, SMA	在原有启动子整合替换原有基因或在其它染色体区段诱导转录激活
靶向删除	重复扩增序列	核酸酶, 重组酶	HD, FSHD, DM1/DM2	在重复区域的靶向大片段删除
互补基因敲除	疾病相关表型模拟	核酸酶, AAV, 重组酶/整合酶/转座子	DMD, BMD	在疾病相关区域进行基因破坏

表 2.利用基因编辑工具进行遗传性神经肌肉异常的疾病基因修复的部分研究

疾病	靶向基因	修复方法	采用工具	参考文献
杜氏肌营养不良	DMD	在外显子内部进行小片段的插入和删除从而实现阅读框的修复	TALENs, CRISPR	[15,17,18]
	DMD	删除完整外显子从而修复阅读框	ZFNs, TALENs, CRISPR	[16-18]
	DMD	通过同源修复纠正点突变	CRISPR	[14]
	DMD	非必需外显子的瞬时删除从而纠正异常外显子表达	单链寡聚核苷酸	[10],[11-13]
	DMD	靶向缺失外显子的敲入	整合酶, TALENs, CRISPR	[18,19]

	DMD	靶向肌营养不良基因的敲入	ZFNs	[20]
	MSTN	破坏肌生成抑制蛋白从而产生代偿性表型	TALENs	[171]
肢带型肌营养不良	DYSF	瞬时性删除非必需外显子从而修复异常阅读框	单链寡核苷酸	[117]
脊髓性肌萎缩	SMN2	阻止 SNP 7 号外显子剪切位点的识别从而对该片段进行修复	单链寡核苷酸	[115] [116,156]
	SMN2	通过将 SNP 7 号外显子的置换为野生型从而将 SMN2 替换成 SMN1	单链寡核苷酸	[113]

通过在基因突变处敲入缺陷的必须基因片段从而完成基因修复

与靶向移码突变不同,完整的野生型基因产物的产生能够通过外源性的带有同源修复区段的完整外显子片段的导入进行介导整合[18]。例如,归巢核酸酶能够被用于刺激 DMD 患者细胞缺陷的 45-52 位外显子的进行同源重组插入[19]。或者,如上所述,单独利用 AAV 载体相较于传统的质粒介导的方式能够提高基因靶向的效率,虽然在一些事例中刺激同源修复的效率并没有显著性的提高。重要的是,相较于现行的核酸酶技术,AAV 载体显著性的降低了脱靶的风险。必须基因片段的敲入在 DMD、LGMD2B 异常代谢、SMA 的 SMN1 异常以及其他的肌营养不良疾病的治疗中非常有效。

点突变的修复

终止子的不成熟、单碱基的异常替换以及 SNP 的剪切点突变是常见的遗传发病原因。在“常见神经肌肉异常概述”中所描述的多种基因编辑工具在这类遗传突变疾病的治疗中都非常有效。例如,多肽核苷酸(PNAs)在体内具有很高的点突变诱导率并能够修复 DMD 动物模型中异常代谢基因的表达[50]。AAV 介导的基因靶向提供了体内单点突变的相似模式[159],包括在多种鼠模型中实现了基因修复,例如遗传性酪氨酸血症[160]。值得注意的是,这些方法的效率很低,除非在一些案例中,疾病相关突变能够使得目标细胞获得选择性的生长优势,例如在酪氨酸血症中。利用核酸酶介导的同源重组修复基因编辑在很多领域中都具有很高的效率。在这种方式下,核酸酶和同源修复模板都具有诱导细胞发生突变的目标序列。这种方法被用于在 mdx 小鼠 DMD 模型中进行肌营养不良点突变修复[14]。在体内环境下,由 AAV 进行 ZFNs 和修复模板的介导的方法能够高效、稳定的诱导基因发生目标突变并能够改善模型小鼠的疾病表型[60,138]。值得注意的是,这些方法造成永久性基因突变后在诱导肝组织再生方面显示出了持续性的作用。这种方法在体内转录中的早期挑战是外源性修复模

板的高效整合，而这与细胞状态和细胞类型有关。然而，这种方法在体外已经成功应用在多种类型的人类单点基因突变实验中得到了应用，包括严重的免疫缺陷并发症[73]、镰状细胞贫血[64,65]、 α -1 抗胰蛋白酶[66]和其它。

多点基因突变替换

基因编辑工具能够被用于介导完整的基因表达功能单位高效地替换原有的损伤内源性基因[127,161]。在这种方法中，置换基因被整合到基因组中替代原有损伤基因发挥作用。外源性插入的功能性片段包括一段目标基因和人工合成的能够持续性诱导基因在合理范围内进行表达的启动子。DNA 修复模板包括一段基因替换序列和人工合成的两侧同源序列，从而能够将靶向序列整合到目标位点。在引入核酸酶产生 DSBs 直接导入点特异性整合酶的情况下这一过程的效率能够被提高多个数量级。

促进驱动性转基因的表达

大多数被编辑基因都包括一个能够驱动目的基因表达的强启动子，并可应用多个已知启动子实现组织特异性或全基因总表达量的提高。一种常见的长时程启动子以巨噬细胞病毒(CMV)启动子为基础，能够介导高水平的表达，但是有时很难实现其在特定细胞类型中的沉默[162]。其它的长时程启动子能够介导多种水平的转录，包括猿猴病毒 40 (SV40) 早期启动子，人源泛素化 C(hUbc)启动子，人源性延长因子 1- α (EF1- α)和磷酸甘油酸激酶 1(PGK-1)启动子。通常来讲，长时程启动子能够驱动固定水平的转录，但是在不同细胞类型中的转录水平不同。

和全基因组表达不同，在目标组织中限制治疗转基因的表达是被需要的。诱导型启动系统，例如 tet-on 或 tet-off 系统，能够通过外源性的小分子——例如具有剂量依赖性表达特征的强力四环素能控制基因的表达。虽然 tet-诱导型系统可能无法直接导入患者，但是有相似的化学诱导基因表达系统正处于临床试验中[163]。在基因编辑应用中组织特异性表达十分必要，能够减少潜在的脱靶率。在肌肉组织中特异性的内源启动子能够驱动外来基因生理水平或高水平的表达，包括肌间线蛋白启动肌酸激酶启动子和肌红蛋白启动子[164,165]。相似的是，神经元特异性的基因转录能够被人工合成的启动子驱动，包括突触蛋白 1 (SYN1)和神经元特异性烯醇酶 (NSE/RU5) [166,167]。

不同的是，人工设计的基因靶向能够被用于整合目的基因下游内源性启动子，并利用其对启动子进行调控。除此之外，由于同源修复模板中不包括启动子，引入强启动子这一方法能够由于减少同源模板的尺寸并降低由于随机整合引起的肿瘤发生的潜在风险。

基因整合的“安全位点”

整合位点的选择对于治疗基因进行合适的表达和确保对邻近基因不产生影响具有重要意义。在外源引入的基因表达片段中具有多个被称为“安全位点”的整合位点。在理想条件下，安全基因不表达必须基因，并且相对于已知的肿瘤发生或破坏基因距离较远，并能够诱导组织中能力依赖性和高效的转基因的表达[168]。在人类基因组 AAVS1 中具有一个这样的“安全位点” [169]，能够诱导多种病毒在几乎所有的组织中介导转录性激活，而这种激活对于大多数的正常细胞生理功能没有显著影响。AAVS1 也因此由于其固有的启动激活属性成为一个能够驱动基因表达的具有吸引力的安全位点[20,127,161]。在修复原发位点进行直接性的基因整合从而高效替换编码区突变也具有巨大的研发潜力[60,138]。值得注意的是，AAV 基因转导和基因编辑技术联用能够高效地将功能性基因整合到处于分裂期和非分裂期的细胞并介导目标基因发生同源重组修复或非同源末端修复[60.108.134-137]。

通过代偿性基因靶向进行疾病模拟

进行疾病相关基因编辑从而模拟疾病表型的一个潜在性选择是通过对其它的相关代偿性基因的操作实现疾病的模拟。肌生成抑制蛋白基因及其通路通常情况下作为代偿性的基因进行靶向，因为该基因能够调控肌肉肥大。多项研究发现肌生成抑制蛋白的敲低和敲除[170,171]或者肌生成抑制蛋白受体 ACVR1B 的拮抗[172]，能够导致肌肉肥大并改善骨骼肌营养不良的症状，尤其 DMD 疾病。基因编辑技术被发现能够通过破坏肌形成抑制蛋白编码基因有效的用于该条通路的永久性敲除[171]。免疫应答大多具有神经肌肉异常的常见特征，并常被用于进行代偿性基因的靶向从而用于增强神经肌肉异常的细胞疗法效果或用于缓和免疫应答的相关症状。相似的是，疾病相关生长因子、受体或是其它的靶向基因的破坏能够为多种神经肌肉异常疾病提供新的治疗思路。

讨论和结论

基因编辑包括一系列的强大工具，能够被用于进行多种神经肌肉异常疾病的基因修复。基因编辑技术的迅猛发展持续、高效、多样化的拓宽着这项技术应用的领域。尤其是，最近新兴的 CRISPR/Cas 系统就是能够应用于人类细胞的、简易诱导突变发生的具有潜力的技术。在该系统中能够通过 Cas9 蛋白和 gRNA 结构进行编辑提高其特异性。而且，在点特异性整合酶的设计、靶向和特异性中的研究进展催生的第二代基因编辑技术有助于直接对目标 DNA 进行操作设计核酸酶。今后很可能会产生效率更高的基因编辑技术用于大片段的插入以及目前难以进行同源修复的细胞类型的基因编辑。最后，最近有研究发现经编程过的 DNA 酶能够产生表观调控修饰[173-180]。这在复杂基因异常的研究、处理和表观调控具有重要作用，例如 FSHD。

将修复基因或基因编辑酶导入到病灶组织能够诱发多种主要的焦点疗法基因转录的进行。腺相关病毒是一种能够促进基因转录的载体，能够更加高效安全地将外源基因导入到肌肉骨骼组织中。在靶向病毒或非病毒的转染方式能够更为高效安全的介导细胞中有限大小片段的基因编辑。而且，基因编程为神经肌肉异常的自体细胞治疗提供了一种令人兴奋的途径。通过在实验室体外环境下对自体干细胞进行操作产生具有更为广泛特征的下一代并将其诱导分化成特定的细胞类型从而重塑损伤组织。进一步的研究需要集中在如何对基因编辑工具和其转染策略进行优化用于特定的神经肌肉异常的治疗。

综上所述，基因编辑的最终目标是阻止疾病的进程或从根本上阻止疾病发生。在过去的五年中人们进行了非凡的努力发展了全新的分子工具，能够在复杂的人类基因组中编辑基因的序列和表观遗传状态。这些工具能够用于人类细胞在体外环境中的快速操作以及小动物模型的体内研究和潜在疗法的开发。这些基因编辑技术随着基因导入手段的进步和干细胞疗法的发展也得到了加强，并能够作为神经肌肉异常疾病治疗中的一种可行的方式。

参考文献

1. Leung DG, Wagner KR. Therapeutic advances in muscular dystrophy. *Ann Neurol*. 2013;74(3):404 – 11.
2. Peeters K, Chamova T, Jordanova A. Clinical and genetic diversity of SMN1-negative proximal spinal muscular atrophies. *Brain*. 2014;137(Pt 11):2879 – 96.
3. Farrar MA, Kiernan MC. The genetics of spinal muscular atrophy: progress and challenges. *Neurotherapeutics*. 2015;12(2):290 – 302.
4. Konieczny P, Swiderski K, Chamberlain JS. Gene and cell-mediated therapies for muscular dystrophy. *Muscle Nerve*. 2013;47(5):649 – 63.
5. Wang B, Li J, Xiao X. Adeno-associated virus vector carrying human minidystrophin genes

- effectively ameliorates muscular dystrophy in mdx mouse model. *Proc Natl Acad Sci U S A*. 2000;97(25):13714 – 9.
6. Pichavant C, Aartsma-Rus A, Clemens PR, Davies KE, Dickson G, Takeda S, et al. Current status of pharmaceutical and genetic therapeutic approaches to treat DMD. *Mol Ther*. 2011;19(5):830 – 40.
7. Qiao C, Koo T, Li J, Xiao X, Dickson JG. Gene therapy in skeletal muscle mediated by adeno-associated virus vectors. *Methods Mol Biol*. 2011;807:119 – 40.
8. Harper SQ, Hauser MA, DelloRusso C, Duan D, Crawford RW, Phelps SF, et al. Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nat Med*. 2002;8(3):253 – 61.
9. Ruszczak C, Mirza A, Menhart N. Differential stabilities of alternative exon-skipped rod motifs of dystrophin. *Biochim Biophys Acta*. 2009;1794(6):921 – 8.
10. Lu QL, Yokota T, Takeda S, Garcia L, Muntoni F, Partridge T. The status of exon skipping as a therapeutic approach to Duchenne muscular dystrophy. *Mol Ther*. 2010;19(1):9 – 15.
11. Cirak S, Arechavala-Gomez V, Guglieri M, Feng L, Torelli S, Anthony K, et al. Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet*. 2011;378(9791):595 – 605.
12. Goemans NM, Tulinius M, van den Akker JT, Burm BE, Ekhart PF, Heuvelmans N, et al. Systemic administration of PRO051 in Duchenne's muscular dystrophy. *N Engl J Med*. 2011;364(16):1513 – 22.
13. van Deutekom JC, Janson AA, Ginjaar IB, Frankhuizen WS, Aartsma-Rus A, Bremmer-Bout M, et al. Local dystrophin restoration with antisense oligonucleotide PRO051. *N Engl J Med*. 2007;357(26):2677 – 86.
14. Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science*. 2014;345(6201):1184 – 8.
15. Ousterout DG, Perez-Pinera P, Thakore PI, Kabadi AM, Brown MT, Qin X, et al. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther*. 2013;21(9):1718 – 26.
16. Ousterout DG, Kabadi AM, Thakore PI, Perez-Pinera P, Brown MT, Majoros WH, et al. Correction of dystrophin expression in cells from duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. *Mol Ther*. 2015;23(3):523 – 32.
17. Ousterout DG, Kabadi AM, Thakore PI, Majoros WH, Reddy TE, Gersbach CA. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat Commun*. 2015;6:6244.
18. Li HL, Fujimoto N, Sasakawa N, Shirai S, Ohkame T, Sakuma T, et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports*. 2015;4(1):143 – 54.
19. Popplewell L, Koo T, Leclerc X, Duclert A, Mamchaoui K, Gouble A, et al. Gene correction of a duchenne muscular dystrophy mutation by meganuclease-enhanced exon knock-in. *Hum Gene Ther*. 2013;24(7):692 – 701.
20. Benabdallah BF, Duval A, Rousseau J, Chapdelaine P, Holmes MC, Haddad E, et al. Targeted gene addition of microdystrophin in mice skeletal muscle via human myoblast transplantation.

Mol Ther Nucleic Acids. 2013;2, e68.

21. Barthelemy F, Wein N, Krahn M, Levy N, Bartoli M. Translational research and therapeutic perspectives in dysferlinopathies. *Mol Med*. 2011;17(9-10):875 – 82.
22. Bansal D, Miyake K, Vogel SS, Groh S, Chen CC, Williamson R, et al. Defective membrane repair in dysferlin-deficient muscular dystrophy. *Nature*. 2003;423(6936):168 – 72.
23. Rahimov F, Kunkel LM. The cell biology of disease: cellular and molecular mechanisms underlying muscular dystrophy. *J Cell Biol*. 2013;201(4):499 – 510.
24. Aartsma-Rus A, Singh KH, Fokkema IF, Ginjaar IB, van Ommen GJ, den Dunnen JT, et al. Therapeutic exon skipping for dysferlinopathies? *Eur J Hum Genet*. 2010;18(8):889 – 94.
25. Nelson DL, Orr HT, Warren ST. The unstable repeats—three evolving faces of neurological disease. *Neuron*. 2013;77(5):825 – 43.
26. Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, et al. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science*. 1992;255(5049):1253 – 5.
27. Turner C, Hilton-Jones D. The myotonic dystrophies: diagnosis and management. *J Neurol Neurosurg Psychiatry*. 2010;81(4):358 – 67.
28. Kumar A, Agarwal S, Agarwal D, Phadke SR. Myotonic dystrophy type 1 (DM1): a triplet repeat expansion disorder. *Gene*. 2013;522(2):226 – 30.
29. Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camano P, Dauwerse JG, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*. 2010;329(5999):1650 – 3.
30. Snider L, Geng LN, Lemmers RJ, Kyba M, Ware CB, Nelson AM, et al. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet*. 2010;6(10), e1001181.
31. Arnold WD, Burghes AH. Spinal muscular atrophy: development and implementation of potential treatments. *Ann Neurol*. 2013;74(3):348 – 62.
32. Seo J, Howell MD, Singh NN, Singh RN. Spinal muscular atrophy: an update on therapeutic progress. *Biochim Biophys Acta*. 2013;1832(12):2180 – 90.
33. Jonson I, Ougland R, Larsen E. DNA repair mechanisms in Huntington's disease. *Mol Neurobiol*. 2013;47(3):1093 – 102.
34. Rubinsztein DC, Leggo J, Coles R, Almqvist E, Biancalana V, Cassiman JJ, et al. Phenotypic characterization of individuals with 30-40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. *Am J Hum Genet*. 1996;59(1):16 – 22.
35. Ross CA, Tabrizi SJ. Huntington's disease: from molecular pathogenesis to clinical treatment. *Lancet Neurol*. 2011;10(1):83 – 98.
36. Mittelman D, Moye C, Morton J, Sykoudis K, Lin Y, Carroll D, et al. Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells. *Proc Natl Acad Sci U S A*. 2009;106(24):9607 – 12.
37. Gaj T, Gersbach CA, Barbas 3rd CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013;31(7):397 – 405.
38. Perez-Pinera P, Ousterout DG, Gersbach CA. Advances in targeted genome editing. *Curr Opin Chem Biol*. 2012;16(3-4):268 – 77.

39. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet.* 2010;11(9):636 – 46.
40. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol.* 2013;14(1):49 – 55.
41. Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol.* 2012;23(5):644 – 50.
42. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339(6121):819 – 23.
43. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science.* 2013;339(6121):823 – 6.
44. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. *Elife.* 2013;2, e00471.
45. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014;346(6213):1258096.
46. Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol.* 2013;31(3):230 – 2.
47. Cox DB, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. *Nat Med.* 2015;21(2):121 – 31.
48. Redondo P, Prieto J, Munoz IG, Alibes A, Stricher F, Serrano L, et al. Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature.* 2008;456(7218):107 – 11.
49. Bertoni C, Rando TA. Dystrophin gene repair in mdx muscle precursor cells in vitro and in vivo mediated by RNA-DNA chimeric oligonucleotides. *Hum Gene Ther.* 2002;13(6):707 – 18.
50. Kayali R, Bury F, Ballard M, Bertoni C. Site-directed gene repair of the dystrophin gene mediated by PNA-ssODNs. *Hum Mol Genet.* 2010;19(16):3266 – 81.
51. Quenneville SP, Chapdelaine P, Rousseau J, Beaulieu J, Caron NJ, Skuk D, et al. Nucleofection of muscle-derived stem cells and myoblasts with phiC31 integrase: stable expression of a full-length-dystrophin fusion gene by human myoblasts. *Mol Ther.* 2004;10(4):679 – 87
52. Bertoni C, Jarrachian S, Wheeler TM, Li Y, Olivares EC, Calos MP, et al. Enhancement of plasmid-mediated gene therapy for muscular dystrophy by directed plasmid integration. *Proc Natl Acad Sci U S A.* 2006;103(2):419 – 24.
53. Gaj T, Mercer AC, Gersbach CA, Gordley RM, Barbas 3rd CF. Structure-guided reprogramming of serine recombinase DNA sequence specificity. *Proc Natl Acad Sci U S A.* 2011;108(2):498 – 503.
54. Gersbach CA, Gaj T, Gordley RM, Barbas 3rd CF. Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res.* 2010;38(12):4198 – 206.
55. Gersbach CA, Gaj T, Gordley RM, Mercer AC, Barbas 3rd CF. Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. *Nucleic Acids Res.* 2011;39(17):7868 – 78.
56. Mercer AC, Gaj T, Fuller RP, Barbas 3rd CF. Chimeric TALE, recombinases with programmable DNA sequence specificity. *Nucleic Acids Res.* 2012;40(21):11163 – 72.
57. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1

resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol.* 2008;26(7):808 – 16.

58. Sollu C, Pars K, Cornu TI, Thibodeau-Beganny S, Maeder ML, Joung JK, et al. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids Res.* 2010;38(22):8269 – 76.

59. Lee HJ, Kim E, Kim JS. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* 2010;20(1):81 – 9.

60. Anguela XM, Sharma R, Doyon Y, Miller JC, Li H, Haurigot V, et al. Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood.* 2013;122(19):3283 – 7.

61. Peault B, Rudnicki M, Torrente Y, Cossu G, Tremblay JP, Partridge T, et al. Stem and progenitor cells in skeletal muscle development, maintenance, and therapy. *Mol Ther.* 2007;15(5):867 – 77.

62. Chapdelaine P, Pichavant C, Rousseau J, Paques F, Tremblay JP. Meganucleases can restore the reading frame of a mutated dystrophin. *Gene Ther.* 2010;17(7):846 – 58.

63. Moehle EA, Rock JM, Lee Y-L, Jouvenot Y, DeKolver RC, Gregory PD, et al. Targeted gene addition into a specific location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A.* 2007;104(9):3055 – 60.

64. Sebastiano V, Maeder ML, Angstman JF, Haddad B, Khayter C, Yeo DT, et al. In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells.* 2011;29(11):1717 – 26.

65. Zou J, Mali P, Huang X, Doney SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood.* 2011;118(17):4599 – 608.

66. Yusa K, Rashid ST, Strick-Marchand H, Varela I, Liu PQ, Paschon DE, et al. Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. *Nature.* 2011;478(7369):391 – 4.

67. Arnould S, Delenda C, Grizot S, Desseaux C, Paques F, Silva GH, et al. The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. *Protein Eng Des Sel.* 2011;24(1-2):27 – 31.

68. Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, et al. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther.* 2011;11(1):11 – 27.

69. Arnould S, Perez C, Cabaniols JP, Smith J, Gouble A, Grizot S, et al. Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J Mol Biol.* 2007;371(1):49 – 65.

70. Smith J, Grizot S, Arnould S, Duclert A, Epinat JC, Chames P, et al. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* 2006;34(22), e149.

71. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol.* 2011;29(2):143 – 8.

72. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science.* 2003;300(5620):763.

73. Urnov F, Miller J, Lee Y, Beausejour C, Rock J, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005;435:2.
74. Alwin S, Gere MB, Guhl E, Effertz K, Barbas 3rd CF, Segal DJ, et al. Custom zinc-finger nucleases for use in human cells. *Mol Ther*. 2005;12(4):610 – 7.
75. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. *Science*. 2003;300(5620):764.
76. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichinger M, et al. Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell*. 2008;31(2):294 – 301.
77. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*. 2009;326(5959):1501.
78. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*. 2009;326(5959):1509 – 12.
79. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol*. 2011;29(8):731 – 4.
80. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res*. 2011;39(21):9283 – 93.
81. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 2012;30(5):460 – 5.
82. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011;39(12), e82.
83. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31(3):227 – 9.
84. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*. 2013;41(20):9584 – 92.
85. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013;31(9):822 – 6.
86. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31(9):827 – 32.
87. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013;31(9):833 – 8.
88. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013;31(9):839 – 43.
89. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154(6):1380 – 9.
90. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. 2015;520(7546):186 – 91.
91. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation

- of endogenous human genes. *Nat Methods*. 2013;10(10):977 – 9.
92. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, et al. RNAguided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods*. 2013;10(10):973 – 6.
93. Yang L, Guell M, Byrne S, Yang JL, De Los AA, Mali P, et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res*. 2013;41(19):9049 – 61.
94. Gordley RM, Gersbach CA, Barbas 3rd CF. Synthesis of programmable integrases. *Proc Natl Acad Sci U S A*. 2009;106(13):5053 – 8.
95. Gaj T, Mercer AC, Sirk SJ, Smith HL, Barbas 3rd CF. A comprehensive approach to zinc finger recombinase customization enables genomic targeting in human cells. *Nucleic Acids Res*. 2013;41(6):3937 – 46.
96. Kettlun C, Galvan DL, George Jr AL, Kaja A, Wilson MH. Manipulating piggyBac transposon chromosomal integration site selection in human cells. *Mol Ther*. 2011;19(9):1636 – 44.
97. Owens JB, Mauro D, Stoytchev I, Bhakta MS, Kim MS, Segal DJ, et al. Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. *Nucleic Acids Res*. 2013;41(19):9197 – 207.
98. Tan W, Dong Z, Wilkinson TA, Barbas 3rd CF, Chow SA. Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *J Virol*. 2006;80(4):1939 – 48.
99. Lim KI, Klimczak R, Yu JH, Schaffer DV. Specific insertions of zinc finger domains into Gag-Pol yield engineered retroviral vectors with selective integration properties. *Proc Natl Acad Sci U S A*. 2010;107(28):12475 – 80.
100. Deyle DR, Russell DW. Adeno-associated virus vector integration. *Curr Opin Mol Ther*. 2009;11(4):442 – 7.
101. Inoue N, Dong R, Hirata RK, Russell DW. Introduction of single base substitutions at homologous chromosomal sequences by adeno-associated virus vectors. *Mol Ther*. 2001;3(4):526 – 30.
102. Hirata R, Chamberlain J, Dong R, Russell DW. Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. *Nat Biotechnol*. 2002;20(7):735 – 8.
103. Miller DG, Petek LM, Russell DW. Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks. *Mol Cell Biol*. 2003;23(10):3550 – 7.
104. Miller DG, Petek LM, Russell DW. Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat Genet*. 2004;36(7):767 – 73.
105. Porteus MH, Cathomen T, Weitzman MD, Baltimore D. Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Mol Cell Biol*. 2003;23(10):3558 – 65.
106. Hirsch ML, Green L, Porteus MH, Samulski RJ. Self-complementary AAV mediates gene targeting and enhances endonuclease delivery for double-strand break repair. *Gene Ther*. 2010;17(9):1175 – 80.
107. Gellhaus K, Cornu TI, Heilbronn R, Cathomen T. Fate of recombinant adeno-associated viral vector genomes during DNA double-strand break-induced gene targeting in human cells. *Hum Gene Ther*. 2010;21(5):543 – 53.
108. Asuri P, Bartel MA, Vazin T, Jang JH, Wong TB, Schaffer DV. Directed evolution of

adenoassociated

virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol Ther.* 2012;20(2):329 – 38.

109. Miller DG, Rutledge EA, Russell DW. Chromosomal effects of adeno-associated virus vector integration. *Nat Genet.* 2002;30(2):147 – 8.

110. Nakai H, Montini E, Fuess S, Storm TA, Grompe M, Kay MA. AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nat Genet.* 2003;34(3):297 – 302.

111. Donsante A, Miller DG, Li Y, Vogler C, Brunt EM, Russell DW, et al. AAV vector integration sites in mouse hepatocellular carcinoma. *Science.* 2007;317(5837):477.

112. Bertoni C, Rustagi A, Rando TA. Enhanced gene repair mediated by methyl-CpG-modified single-stranded oligonucleotides. *Nucleic Acids Res.* 2009;37(22):7468 – 82.

113. DiMatteo D, Callahan S, Kmiec EB. Genetic conversion of an SMN2 gene to SMN1: a novel approach to the treatment of spinal muscular atrophy. *Exp Cell Res.* 2008;314(4):878 – 86.

114. Maguire K, Suzuki T, DiMatteo D, Parekh-Olmedo H, Kmiec E. Genetic correction of splice site mutation in purified and enriched myoblasts isolated from mdx5cv mice. *BMC Mol Biol.* 2009;10:15.

115. Douglas AG, Wood MJ. Splicing therapy for neuromuscular disease. *Mol Cell Neurosci.* 2013;56:169 – 85.

116. Passini MA, Bu J, Richards AM, Kinnecom C, Sardi SP, Stanek LM, et al. Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Sci Transl Med.* 2011;3(72):72ra18.

117. Wein N, Avril A, Bartoli M, Beley C, Chaouch S, Laforet P, et al. Efficient bypass of mutations in dysferlin deficient patient cells by antisense-induced exon skipping. *Hum Mutat.* 2010;31(2):136 – 42.

118. Darabi R, Arpke RW, Irion S, Dimos JT, Grskovic M, Kyba M, et al. Human ES- and iPSC-derived myogenic progenitors restore dystrophin and improve contractility upon transplantation in dystrophic mice. *Cell Stem Cell.* 2012;10(5):610 – 9.

119. Tedesco FS, Gerli MF, Perani L, Benedetti S, Ungaro F, Cassano M, et al. Transplantation of genetically corrected human iPSC-derived progenitors in mice with limb-girdle muscular dystrophy. *Sci Transl Med.* 2012;4(140):140ra89.

120. Dezawa M, Ishikawa H, Itokazu Y, Yoshihara T, Hoshino M, Takeda S, et al. Bone marrow stromal cells generate muscle cells and repair muscle degeneration. *Science.* 2005;309(5732):314 – 7.

121. Cerletti M, Jurga S, Witczak CA, Hirshman MF, Shadrach JL, Goodyear LJ, et al. Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell.* 2008;134(1):37 – 47.

122. Tedesco FS, Hoshiya H, D'Antona G, Gerli MF, Messina G, Antonini S, et al. Stem cell-mediated transfer of a human artificial chromosome ameliorates muscular dystrophy. *Sci Transl Med.* 2011;3(96):96ra78.

123. Negroni E, Riederer I, Chaouch S, Belicchi M, Razini P, Di Santo J, et al. In vivo myogenic potential of human CD133+ muscle-derived stem cells: a quantitative study. *Mol Ther.*

2009;17(10):1771 – 8.

124. Kimura E, Han JJ, Li S, Fall B, Ra J, Haraguchi M, et al. Cell-lineage regulated myogenesis for dystrophin replacement: a novel therapeutic approach for treatment of muscular dystrophy.

Hum Mol Genet. 2008;17(16):2507 – 17.

125. Zhu CH, Mouly V, Cooper RN, Mamchaoui K, Bigot A, Shay JW, et al. Cellular senescence in human myoblasts is overcome by human telomerase reverse transcriptase and cyclindependent

kinase 4: consequences in aging muscle and therapeutic strategies for muscular dystrophies. Aging Cell. 2007;6(4):515 – 23.

126. Berghella L, De Angelis L, Coletta M, Berarducci B, Sonnino C, Salvatori G, et al. Reversible immortalization of human myogenic cells by site-specific excision of a retrovirally transferred oncogene. Hum Gene Ther. 1999;10(10):1607 – 17.

127. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 2007;25(11):1298 – 306.

128. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas 3rd CF. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods. 2012;9(8):805 – 7.

129. Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas lii CF. Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One. 2014;9(1), e85755.

130. Palmieri B, Tremblay JP. Myoblast transplantation: a possible surgical treatment for a severe pediatric disease. Surg Today. 2010;40(10):902 – 8.

131. Filareto A, Parker S, Darabi R, Borges L, Iacovino M, Schaaf T, et al. An ex vivo gene therapy approach to treat muscular dystrophy using inducible pluripotent stem cells. Nat Commun. 2013;4:1549.

132. Pruettt-Miller SM, Reading DW, Porter SN, Porteus MH. Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. PLoS Genet. 2009;5(2), e1000376.

133. Seto JT, Ramos JN, Muir L, Chamberlain JS, Odom GL. Gene replacement therapies for duchenne muscular dystrophy using adeno-associated viral vectors. Curr Gene Ther. 2012;12(3):139 – 51.

134. Ellis BL, Hirsch ML, Porter SN, Samulski RJ, Porteus MH. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. Gene Ther. 2013;20(1):35 – 42.

135. Hirsch ML, Green L, Porteus MH, Samulski RJ. Self-complementary AAV mediates gene targeting and enhances endonuclease delivery for double-strand break repair. Gene Ther. 2010;17(9):1175 – 80.

136. Handel EM, Gellhaus K, Khan K, Bednarski C, Cornu TI, Muller-Lerch F, et al. Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adenoassociated viral vectors. Hum Gene Ther. 2012;23(3):321 – 9.

137. Rahman SH, Bobis-Wozowicz S, Chatterjee D, Gellhaus K, Pars K, Heilbronn R, et al. The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. Hum Gene Ther. 2013;24(1):67 – 77.

138. Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. Nature. 2011;475(7355):217 – 21.

139. Asokan A, Conway JC, Phillips JL, Li C, Hegge J, Sinnott R, et al. Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. *Nat Biotechnol.* 2010;28(1):79 – 82.
140. Messina E, Nienaber J, Daneshmand M, Villamizar N, Samulski RJ, Milano CA, et al. Adenoassociated viral vectors based on serotype 3b use components of the fibroblast growth factor receptor signaling complex for efficient transduction. *Hum Gene Ther.* 2012;23(10):1031 – 42.
141. Bowles DE, McPhee SWJ, Li C, Gray SJ, Samulski JJ, Camp AS, et al. Phase 1 gene therapy for duchenne muscular dystrophy using a translational optimized AAV vector. *Mol Ther.* 2012;20(2):443 – 55.
142. Qiao C, Li J, Zheng H, Bogan J, Li J, Yuan Z, et al. Hydrodynamic limb vein injection of adeno-associated virus serotype 8 vector carrying canine myostatin propeptide gene into normal dogs enhances muscle growth. *Hum Gene Ther.* 2009;20(1):1 – 10.
143. Wang Z, Zhu T, Qiao C, Zhou L, Wang B, Zhang J, et al. Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. *Nat Biotechnol.* 2005;23(3):321 – 8.
144. Watchko J, O' Day T, Wang B, Zhou L, Tang Y, Li J, et al. Adeno-associated virus vectormediated minidystrophin gene therapy improves dystrophic muscle contractile function in mdx mice. *Hum Gene Ther.* 2002;13(12):1451 – 60.
145. Yang L, Jiang J, Drouin LM, Agbandje-McKenna M, Chen C, Qiao C, et al. A myocardium tropic adeno-associated virus (AAV) evolved by DNA shuffling and in vivo selection. *Proc Natl Acad Sci U S A.* 2009;106(10):3946 – 51.
146. Liu M, Yue Y, Harper SQ, Grange RW, Chamberlain JS, Duan D. Adeno-associated virusmediated microdystrophin expression protects young mdx muscle from contraction-induced injury. *Mol Ther.* 2005;11(2):245 – 56.
147. Odom GL, Gregorevic P, Allen JM, Finn E, Chamberlain JS. Microdystrophin delivery through rAAV6 increases lifespan and improves muscle function in dystrophic dystrophin/utrophindeficient mice. *Mol Ther.* 2008;16(9):1539 – 45.
148. Wang Z, Kuhr CS, Allen JM, Blankinship M, Gregorevic P, Chamberlain JS, et al. Sustained AAV-mediated dystrophin expression in a canine model of Duchenne muscular dystrophy with a brief course of immunosuppression. *Mol Ther.* 2007;15(6):1160 – 6.
149. Shen S, Horowitz ED, Troupes AN, Brown SM, Pulicherla N, Samulski RJ, et al. Engraftment of a galactose receptor footprint onto adeno-associated viral capsids improves transduction efficiency. *J Biol Chem.* 2013;288(40):28814 – 23.
150. Gray SJ. Gene therapy and neurodevelopmental disorders. *Neuropharmacology.* 2013; 68:136 – 42.
151. Hester ME, Foust KD, Kaspar RW, Kaspar BK. AAV as a gene transfer vector for the treatment of neurological disorders: novel treatment thoughts for ALS. *Curr Gene Ther.* 2009;9(5):428 – 33.
152. Ojala DS, Amara DP, Schaffer DV. Adeno-associated virus vectors and neurological gene therapy. *Neuroscientist.* 2015;21(1):84 – 98.
153. Gruber K. Europe gives gene therapy the green light. *Lancet.* 2012;380(9855), e10.
154. Lorden ER, Levinson HM, Leong KW. Integration of drug, protein, and gene delivery systems with regenerative medicine. *Drug Deliv Transl Res.* 2015;5(2):168 – 86.
155. McNeer NA, Schleifman EB, Cuthbert A, Brehm M, Jackson A, Cheng C, et al. Systemic

- delivery of triplex-forming PNA and donor DNA by nanoparticles mediates site-specific genome editing of human hematopoietic cells in vivo. *Gene Ther.* 2013;20(6):658 – 69.
156. Porensky PN, Mitrpant C, McGovern VL, Bevan AK, Foust KD, Kaspar BK, et al. A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. *Hum Mol Genet.* 2012;21(7):1625 – 38.
157. Lee HJ, Kweon J, Kim E, Kim S, Kim JS. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. *Genome Res.* 2012;22(3):539 – 48.
158. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, et al. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* 2013;41(14), e141.
159. Miller DG, Wang PR, Petek LM, Hirata RK, Sands MS, Russell DW. Gene targeting in vivo by adeno-associated virus vectors. *Nat Biotechnol.* 2006;24(8):1022 – 6.
160. Paulk NK, Wurstthorn K, Wang Z, Finegold MJ, Kay MA, Grompe M. Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo. *Hepatology.* 2010;51(4):1200 – 8.
161. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, et al. Sitespecific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods.* 2011;8(10):861 – 9.
162. Brooks AR, Harkins RN, Wang P, Qian HS, Liu P, Rubanyi GM. Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *J Gene Med.* 2004;6(4):395 – 404.
163. Huang C, Ramakrishnan R, Trkulja M, Ren X, Gabrilovich DI. Therapeutic effect of intratumoral administration of DCs with conditional expression of combination of different cytokines. *Cancer Immunol Immunother.* 2012;61(4):573 – 9.
164. Himeda CL, Tai PW, Hauschka SD. Analysis of muscle gene transcription in cultured skeletal muscle cells. *Methods Mol Biol.* 2012;798:425 – 43.
165. Salva MZ, Himeda CL, Tai PW, Nishiuchi E, Gregorevic P, Allen JM, et al. Design of tissuespecific regulatory cassettes for high-level rAAV-mediated expression in skeletal and cardiac muscle. *Mol Ther.* 2007;15(2):320 – 9.
166. Yaguchi M, Ohashi Y, Tsubota T, Sato A, Koyano KW, Wang N, et al. Characterization of the properties of seven promoters in the motor cortex of rats and monkeys after lentiviral vectormediated gene transfer. *Hum Gene Ther Methods.* 2013;24(6):333 – 44.
167. Delzor A, Dufour N, Petit F, Guillermier M, Houitte D, Auregan G, et al. Restricted transgene expression in the brain with cell-type specific neuronal promoters. *Hum Gene Ther Methods.* 2012.
168. Papapetrou EP, Lee G, Malani N, Setty M, Riviere I, Tirunagari LM, et al. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. *Nat Biotechnol.* 2011;29(1):73 – 8.
169. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKever RC, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol.* 2009;27(9):851 – 7.
170. McPherron AC, Lawler AM, Lee SJ. Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member. *Nature.* 1997;387(6628):83 – 90.
171. Xu L, Zhao P, Mariano A, Han R. Targeted myostatin gene editing in multiple mammalian

- species directed by a single pair of TALE nucleases. *Mol Ther Nucleic acids*. 2013;2, e112.
172. Lee SJ, Reed LA, Davies MV, Girgenrath S, Goad ME, Tomkinson KN, et al. Regulation of muscle growth by multiple ligands signaling through activin type II receptors. *Proc Natl Acad Sci U S A*. 2005;102(50):18117 – 22.
173. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol*. 2013;31(12):1137 – 42.
174. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, et al. Optical control of mammalian endogenous transcription and epigenetic states. *Nature*. 2013; 500(7463):472 – 6.
175. Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK, et al. Locus-specific editing of histone modifications at endogenous enhancers. *Nat Biotechnol*. 2013;31(12):1133 – 6.
176. Siddique AN, Nunna S, Rajavelu A, Zhang Y, Jurkowska RZ, Reinhardt R, et al. Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J Mol Biol*. 2013;425(3):479 – 91.
177. de Groote ML, Verschure PJ, Rots MG. Epigenetic Editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. *Nucleic Acids Res*. 2012;40(21):10596 – 613.
178. Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, et al. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat Methods*. 2015;12(5):401 – 3.
179. Hilton IB, D' Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol*. 2015;33(5):510 – 7.
180. Polstein L, Perez-Pinera P, Kocak D, Vockley C, Bledsoe P, Song L, et al. Genome-wide specificity of DNA-binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Res*. 2015;25(8):1158 – 69.

噬菌体整合酶基因编辑

Michele P. Calos

摘要 噬菌体整合酶是一种原核点特异性的重组酶，能够在 attB 和 attP 识别位点之间实现精确的切割和插入。这类酶的应用范围包括从细菌细胞到哺乳动物细胞，是一种非常有效的基因编辑工具。PhiC31 是第一种在哺乳动物细胞中进行应用的噬菌体整合酶。这种整合酶能够高效整合其自身的 attB 和 attP 位点。而且，phiC31 整合酶能够对大型基因组中与 att 的相似位点即伪 att 位点进行整合，造成插入基因的脱靶。PhiC31 整合酶也能够与其他的噬菌体整合酶进行联用，例如具有不同识别序列且脱靶率较低的 Bxb1。这些噬菌体整合酶的多种性能决定了其具有广泛的应用前景，总结如下：产生转基因的生物体并实现体内基因治疗，进行细胞的重编程或通过功能域的置换实现精确的基因编辑。最后介绍的系统——双整合酶功能域置换（DICE），联用 phiC31 和 Bxb1 对 attP 序列进行精确靶向并诱导基因组发生同源重组，该方式尤其是在多功能干细胞的迭代基因编辑中非常具有应用价值。

关键词 attB 位点 • attP 位点 • Bxb1 整合酶 • 功能域置换 • 胚胎干细胞 • 同源重组修复 • 多功能干细胞诱导 • phiC31 整合酶 • 重编程 • TALEN

M. P. Calos, Ph.D. (*)

Department of Genetics, Stanford University School of Medicine,
300 Pasteur Drive, Alway Bldg., Room M334, Stanford, CA 94305-5120, USA
e-mail: calos@stanford.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_5

噬菌体整合酶的应用

在1998年噬菌体 ϕ C31 整合酶被报道能够在大肠杆菌和缓冲液环境下而非原有的链霉素菌属细胞环境下对 $attB$ 和 $attP$ 位点进行整合[1]。这些数据说明该种酶具有独立完整的功能，并不需要宿主特异性的共表达因子支持其活性的产生。在这一基础上，我们首先假设了 ϕ C31 能够在外源性的细胞环境下发挥功能，例如需要进行位点整合的哺乳动物细胞。而且，在 att 位点多种不同的重组行为研究中发现重组反应是单向性的[1]，这一特征对于高效介导哺乳动物细胞基因组非常有利。

这些推测在我们2000年发表的研究中得到了证实， ϕ C31 整合酶确实能够在哺乳动物细胞中整合野生型的 $attB$ 和 $attP$ 位点[2]。 att 位点的长度也在那次研究中第一次得到了鉴定，大约在34bp或略小。如图1所示，整合酶能够通过介导 $attP$ 和 $attB$ 位点之间的重组，产生无法再进行整合的 $attL$ 和 $attR$ 位点，从而完成一次单向性的反应。目前已经发现每一个 att 位点都能够结合一组二聚化的整合酶分子，并介导随后剪切插入反应的发生[3]。

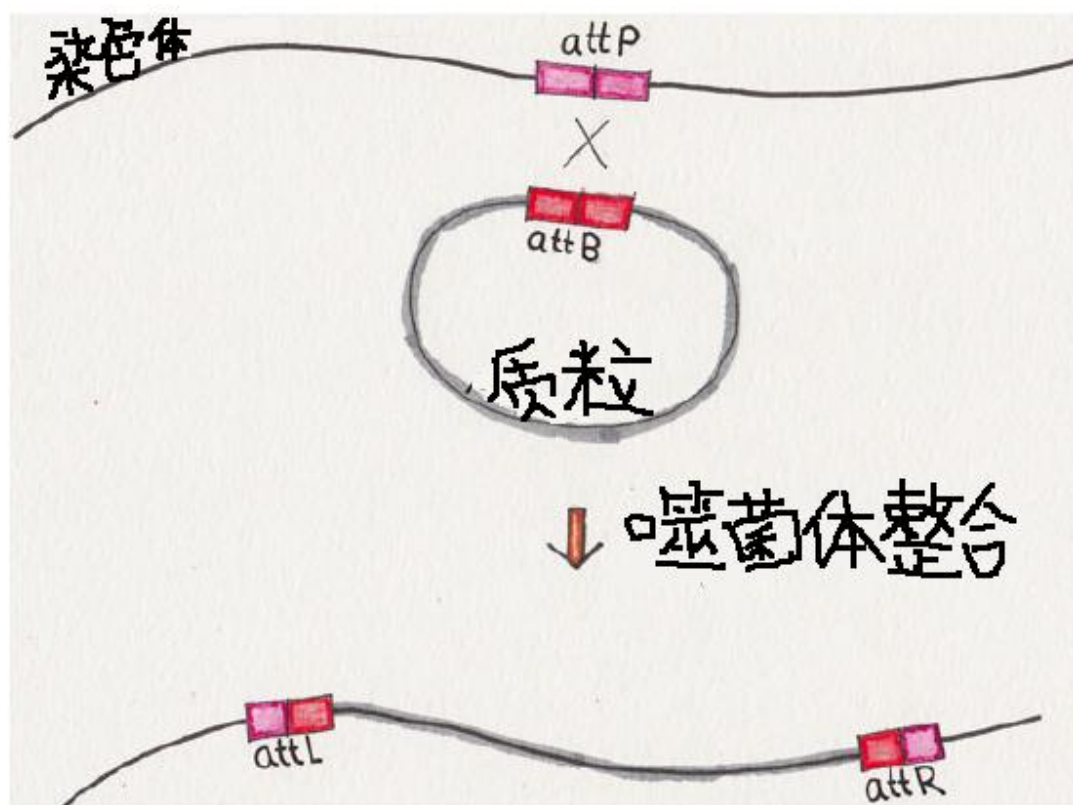


图1.基础噬菌体整合反应。噬菌体整合酶分别在宿主细菌和噬菌体基因组的 $attB$ 和 $attP$ 位点进行整合反应，其中大约有不到34个bp的回文序列。二聚化的噬菌体整合酶分别位于不同的 att 位点并进行整合（如上图所示），并通过切割插入反应在 att 位点之间永久性导入外源片段。通过这种途径，能够将具有 $attB$ 位点的质粒上序列导入到具有同源 $attP$ 位点的基因组中。整合后的 att 位点具有一半的 $attB$ 和 $attP$ ，不能够继续被整合，因此是一种单向性的反应。

attP 位点整合中伪位点的干扰

当 att 位点接近 30bpd 的时候，就可能存在具有相似特征序列或部分一致的序列也有可能被 phiC31 整合酶催化重组。因此不建议利用类似 30bp 大小的互补识别序列，即使是在大型基因组中，16bp 的识别序列仍是一个好的选择，我们也认为这种长度的识别序列足以满足整合需要。而这一观点在人类和小鼠细胞系中也被充分证实，相关实验显示 phiC31 整合酶能够介导相似的区段也就是伪 att 位点发生整合[4]（图 2）。

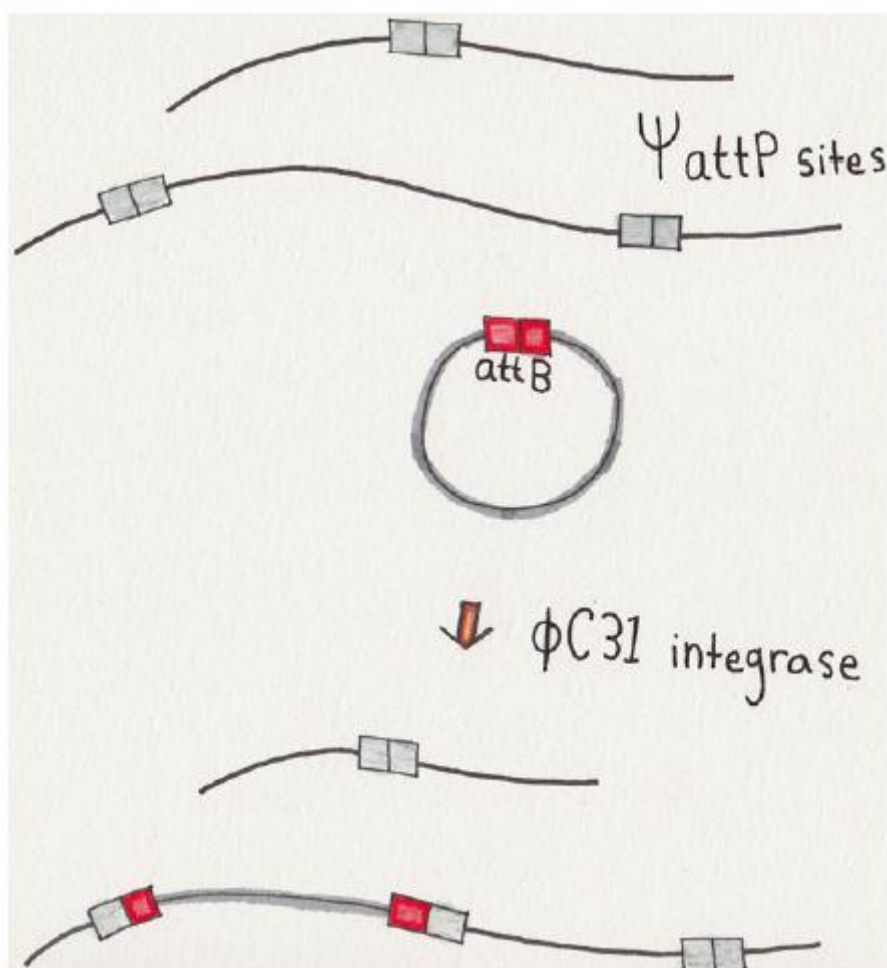


图2.伪位点插入反应。有一些丝氨酸噬菌体整合酶，尤其是phiC31整合酶，能够在att类似序列或具有部分一致的发生整合反应，这类序列称为伪att位点。在一段基因组中通常存在多个伪att位点，无论是在DNA序列中还是结构基因组中（如示意图中所示的三个伪att位点）。当phiC31整合酶和attB伪位点同时存在时，能够将质粒中的序列整合到伪位点中，整合通常发生在姊妹染色单体的一侧。在伪位点进行整合的精确度低于真实位点，通常在整合部位会出现小片段的缺失。

由 phiC31 整合酶介导的伪位点的半精确重组反应首先发现于哺乳动物细胞。虽然相较于同时期的随机 DNA 插入、逆转录病毒整合或转座子整合，噬菌体整合具有更高的特异性，但是仍然具有大概率的随机重组的发生。PhiC31 整合酶介导的伪位点整合对于基因整合产生了直接性的干扰，包括体内基因治疗和转基因生物体的构建。

基因组通常具有多个潜在的伪 att 位点,这些位点可能存在于 DNA 序列中也可能存在于结构基因组中;那些开放的、转录激活的伪 att 基因位点更有可能发生重组[5]。phiC31 介导的伪 att 位点的重组缺乏预测性,无法预测错配重组发生的位置,因而也就限制了其应用。而且,在伪 att 位点的整合通常是不精确的,在这一位点的整合经常具有多个碱基的缺失并有时会伴发大片段基因组的重组[5]。而且,由伪位点接到的不可预知的位点整合相较于同时期的其它系统的点特异性通常高多个数量级。

利用 phiC31 整合酶构建转基因生物体

噬菌体整合酶的一个常见应用是构建转基因的生物体。其中真正的 att 位点和伪 att 位点都具有应用。例如在两栖类非洲爪蟾的基因整合研究中就有伪 att 位点的应用[6]。在黑腹果蝇中, phiC31 attP 位点最早通过随机性的 P 部件整合到基因组中。随后利用 phiC31 整合酶编码 mRNA 共转导的方式,利用具有 attB 位点的质粒对 att 位点进行高效地靶向、整合[7]。这一系列的技术目前在更为广泛的领域中被用于产生转基因物种,包括鱼、鸟、两栖类动物、哺乳动物、昆虫和植物[8]。

伪位点整合的基因治疗研究

利用 phiC31 将 DNA 整合到未经修饰的哺乳动物基因组中为数不多的几个整合位点为基因治疗提供了一种新的机会。利用 phiC31 整合酶进行基因治疗的第一项研究是利用尾静脉高压注射的方式,将具有整合酶编码的质粒和人类 IX 基因因子简单高效地导入到肝脏中[9]。仅在一次注射之后 IX 因子的治疗效果就显现出来。疗法质粒在肝细胞中的点特异性整合功能已经被证实,而且整合过程具有很强的特异性,大多数整合行为只发生在基因组的一个热点区域。而且,整合行为具有稳定性,因子 IX 的产生能够持续很长时间[10]。这些特征显示利用 phiC31 整合酶进行血友病的治疗具有很好的临床转化前景。在血友病 A 和 B 模型小鼠中的研究也已经发表,人类 VIII 和 IX 因子在这些模型中都具有长时程的表达[11,12]。不幸的是,到目前为止将水溶性 DNA 导入大型动物肝脏的研究还无法达到疾病治疗和临床转化的水平。

很多其他的利用 phiC31 整合酶进行的动物水平的基因治疗研究也在尝试着将治疗基因导入到不同的组织和物种中(回顾[8,13,14])。例如,利用电转法将 DNA 导入到大鼠视网膜中进行长时程的标记基因的表达[15]。利用电转法将具有 DYSTROPHIN 基因的质粒导入到杜氏肌营养不良的模型小鼠中[16]。由于在染色体中长时间点特异性整合的概念在啮齿类动物中得到了验证,其临床转化有待于更高效的基因转导方式的开发及该疗法在更为大型的动物中相关实验的开展。

利用 phiC31 整合酶进行哺乳动物细胞重编程

有一种可以绕开低效的质粒 DNA 体内转导的方式是在首先体外进行 DNA 转导,对于大多数类型的细胞存在多种高效的体外转染方式,随后将基因整合后的细胞移植到体内。如果细胞是永生化的,可以进行离体单细胞克隆培养并对于整合位点进行确证以保证安全。

能够进行克隆培养的细胞包括多功能干细胞,例如胚胎干细胞(ESC)。在 2006 年,人工诱导的多功能干细胞(iPSC)被发现,即通过四种转录因子基因的插入能够诱导体细胞产生类似于 ESC 细胞的特征[17,18]。这些 iPSC 在再生医学中具有广泛的开发潜力,因为它们不具有免疫原性,能够扩增到很大的数量且易于进行基因编辑,并且免于法律或伦理的限制。对于遗传类疾病来说,病人来源的 iPSC 能够在体外进行编辑,例如对疾病相关基因的整合后通过适当的治疗方法在体外诱导分化,随后将其移植到适当的组织和器官中。

因此重编程技术提供了多种潜在的治疗策略。然而,早期 iPSC 的产生是通过逆转录病

毒将四种转录因子引入到基因组中。这一策略会导致基因组中多个随机性整合事件的发生，提升随机性突变插入的风险并导致肿瘤的产生[17,18]。为了克服这一缺陷，我们设计了一种策略利用 phiC31 整合酶一次性地介导四种转录因子插入到小鼠基因组中的一个单独的安全性位点中[19]。

在这个方法中，首先构建具有四种 Yamanaka 重编程转录因子的质粒（p4FLR，11.9kb），并同时引入重组识别位点。在 cMyc、Klf4、Oct4 和 Sox2 基因 cDNA 序列旁引入 2A 序列，从 CAG 启动子开始促进其多顺反子转录的发生。这一质粒同时也存在 phiC31 整合酶的 attB 位点，能够介导基因整合到伪 att 位点中（图 2）。在重编程功能域两侧存在两个 loxP 位点，使得其在编程后可通过瞬转 Cre 解离酶表达质粒被删除[19]。这一步骤非常重要，能够使 iPSC 不易于形成肿瘤而且更易于诱导分化。

由 phiC31 介导基因编辑方法成功的在小鼠胚胎纤维母细胞和成人间叶干细胞中得到了相较于逆转录病毒更为高效地应用，并且杜绝了由病毒随机性整合引起的危害和风险。利用连接法介导的 PCR 对人源化 iPSC 克隆的整合位点进行分析后发现，接近有三分之一的 iPSC 具有一个单独的整合位点 p4FLR。多种不同的鼠源性伪 attP 位点被用于进行 iPSC 的分析。在 14 个位点中有六个发生了整合，而在其中有两个被认为是安全位点，远离肿瘤发生基因和其它的损害。具有 Cre 游离酶表达的质粒转染到两种不同的 iPSC 细胞群落中，结果发现重编程功能域被精确敲除。在 Cre 敲除前后鉴定 iPSC 的多功能性，包括 iPSC 产生畸胎瘤和嵌合型小鼠的能力[19]。

治疗基因在前 attP 整合位点的点特异性整合

在早期的重编程研究中我们试图通过建立一种更为强大的重编程功能域获得更高编辑比例的 iPSC 克隆。我们也开展了有关于在 iPSC 细胞中将一段治疗性片段整合到基因组中的研究[20]。新的重编程质粒，pCOBLW，具有更易进行重编程的序列，随后配置有以 Oct 开头和以 Myc 结尾的基因片段（OSKM：图 3）。我们也引入了 WPRE 片段增强重编程功能域的转录效果。pCOBLW 能够更加高效更加特异性的对细胞进行重编程，有 93% 的被编程 iPSC 细胞都只具有唯一的插入片段[20]。

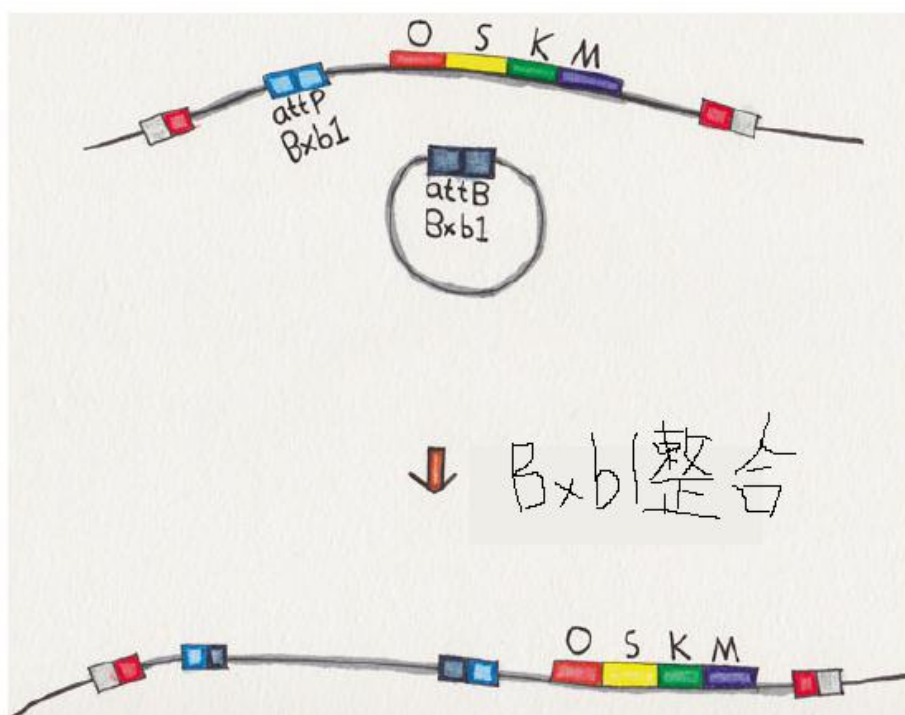


图3.利用phiC31对Bxb1位点进行整合。我们开发了一种具有Otc-Sox-Klf-Myc (OSKM) 重编程功能域和Bxb1 attP位点的质粒。PhiC31整合酶被用于在伪attP位点进行基因整合，诱导iPSC的产生。通过对DNA整合位点安全性和生物信息的检测，能够将具有Bxb1 attB位点的治疗基因骨架导入到iPSC细胞的安全整合位点中。Cre游离酶能够移除随机插入序列，通过对质粒中loxP位点的整合删除重编程基因和质粒骨架序列。

随着 loxP 位点在精确重编程基因组和质粒中的应用，我们目前已经利用 Bxb1 整合酶对 attP 位点进行了多种具有治疗靶向价值的重编程（图 3）。Bxb1 是一种和 phiC31 相关的丝氨酸整合酶，但是它的 att 位点和 phiC31 完全不同，也不存在重叠或交叉。Bxb1 能够在哺乳动物细胞中其独有的 att 位点被激活，但是目前还没有其在相似伪位点中具有整合活性的报道[21]。在整合序列一侧包含有绝缘体序列，能够降低治疗基因对自身及对临近基因组产生非预期性的影响。

为了实现这一重编程策略，我们利用质粒将 pCOBLW 核酸序列和一段 phiC31 整合酶序列导入到了 mdx 小鼠的离体纤维母细胞中。通过载体上的 phiC31 att 位点将载体序列导入到伪 att 位点。并挑选出一株具有安全位点整合的 iPSC 细胞系并导入治疗基因。在这一案例中，我们利用全长小鼠肌营养不良 cDNA 对杜氏肌营养不良模式小鼠进行治疗。我们利用质粒将具肌营养不良编码基因片段、非启动型嘌呤霉素抗性基因和 Bxb1 attB 位点和 Bxb1 整合酶编码序列通过核转染导入 iPSC 细胞，并随后对随机整合序列和质粒序列进行删除。之后，对于获得细胞进行概念性的分化诱导和移植，作为肌营养不良的一种潜在性治疗方式。

DICE 系统：利用 phiC31 和 Bxb1 诱导同源重组修复置换基因功能域

phiC31 介导的伪 att 位点的基因整合为基因组整合提供了一条非常简便的编辑思路，但是进行安全性基因编辑的目标克隆的筛选仍然十分费力。在多功能干细胞中，例如 ESC 和 iPSC 以及无限增殖细胞系中，还存在另一种借助于同源重组修复原理允许客户进行精确整合控制的策略。如果同源重组被用于 attP 位点进行基因整合，那么外源性基因功能域就能实现精确的插入。我们开发了一种名为双整合酶功能域置换或 DICE 的策略，能够实现 attP 位点功能域的精确整合[22]。

在 DICE 方法中，在 phiC31 整合酶识别的 attP 位点和 Bxb1 整合酶编码序列的两侧引入同源重组序列。在我们的研究中，我们利用了人类 22 号染色体上的一段重复排列的、安全的、具有转录活性的 H11 片段作为基因导入片段，并在 attP 位点之间导入 GFP 和新霉素抗性基因作为克隆筛选标记基因（图 4）。我们利用 TALENs 对 H11 进行靶向，诱导同源重组修复的产生。这一策略能够降低低效率同源重组产生的致病性，因而在 iPSC 细胞中格外具有价值。

一旦有一株细胞系完成了上述系列的基因编辑构建，可以随后利用 phiC31 和 Bxb1 整合酶实现高效率的点特异性的基因功能域置换。为了实现功能域置换的目的，编辑后细胞需要外源性导入在目的片段一侧具有 phiC31 和 Bxb1 整合酶 attB 识别位点及两种整合酶的编码序列的质粒。在功能域置换的过程中，新霉素和 GFP 基因将会被移除并在原位点导入目的基因（图 4）。目的基因可以是治疗基因和/或报道、筛选、移植示踪基因等。

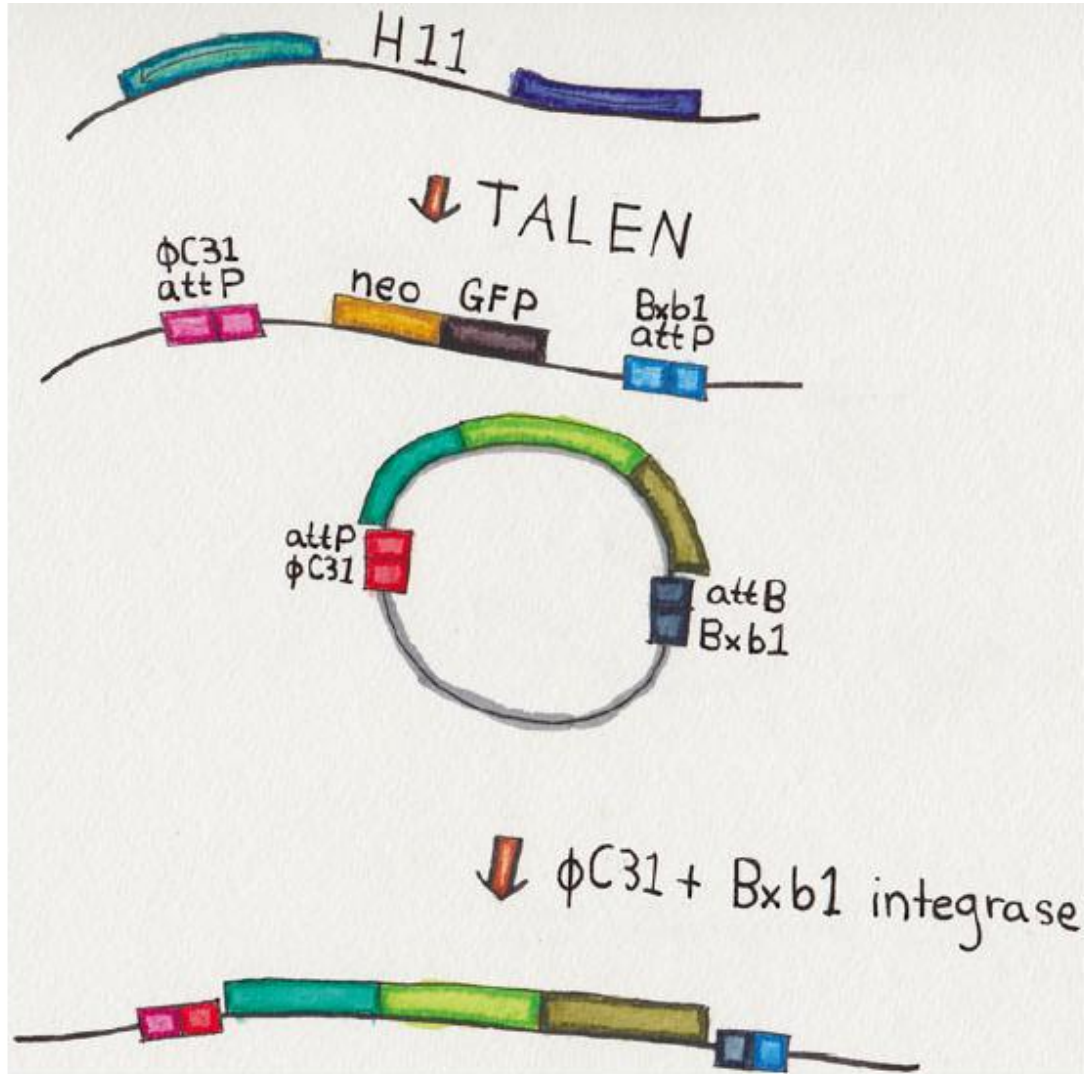


图4.DICE，双整合酶功能域置换。为了实现DICE基因整合，具有phiC31和Bxb1整合attP位点的基因片段通过同源重组插入目的序列。我们在人源性ESC和iPSC细胞22号染色体的H11位点——一个具有正常转录水平的安全位点导入这一序列。利用TALENs在目标位点造成双链损伤诱导高效率同源重组修复的产生。新霉素抗性或绿色荧光蛋白标记的报道基因或筛选基因被用于进行修复位点的鉴定。一旦该段基因完成了在基因组中的正确整合，下游具有phiC31和Bxb1 attB位点的目的基因就能够很容易的整合，例如治疗性基因、标记基因或转录因子，后面具有phiC31和Bxb1整合酶编码基因。点特异性的基因功能域的置换的结果是实现attB位点之间遗传信息的替换。

在将多种不同基因精确插入同一位点构建平行细胞系时，DICE系统尤其具有价值。在功能域替换的过程中，外源性基因的编码序列、插入方向和插入位置完全可控，编辑结果可预测。例如，这一方法能够允许我们迅速构建具有三种神经转录因子表达的人源性ESC和iPSC细胞，从而评估其在多巴胺能神经元分化中的作用[22]。

利用 phiC31 切除酶进行整合反应的逆转

利用 phiC31 整合酶对原有的 attB 和 attP 位点进行整合反应后, 结果产生的 attL 和 attR (图 1) 位点和原有的 attB 和 attP 位点不同, 并能再次进行整合[1,23]。噬菌体整合系统在一般情况下也包含有一种叫做切除酶或整合方向性因子 (RDF) 的小分子蛋白, 能够和整合酶进行连接并改变其特异性, 使其能对 attL 和 attR 序列进行整合从而逆转原有的整合反应。

phiC31 整合酶的 RDF 最近被发现[24], 该因子在哺乳动物细胞中具有有一定概率能够实现整合反应的逆转。我们构建了质粒用于验证这一假说并发现 phiC31 RDF 能够与 phiC31 联合有效逆转整合反应[25]。因此, 在未来的噬菌体整合酶研究中 phiC31 RDF 很可能成为一种非常有用的工具。

致谢 M.P.C.感谢 Victoria Ellis 提供的数据以及加利福尼亚研究所再生医学中心的支持。

参考文献

1. Thorpe HM, Smith MCM. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci U S A*. 1998;95:5505 – 10.
2. Groth AC, Olivares EC, Thyagarajan B, Calos MP. A phage integrase directs efficient sitespecific integration in human cells. *Proc Natl Acad Sci U S A*. 2000;97:5995 – 6000.
3. Rutherford K, Van Duyne G. The ins and outs of serine integrase site-specific recombination. *Curr Opin Struct Biol*. 2014;24:125 – 31.
4. Thyagarajan B, Olivares EC, Hollis RP, Ginsburg DS, Calos MP. Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Mol Cell Biol*. 2001;21:3926 – 34.
5. Chalberg TC, et al. Integration specificity of phage phiC31 integrase in the human genome. *J Mol Biol*. 2006;357:28 – 48.
6. Allen BG, Weeks DL. Transgenic *Xenopus laevis* embryos can be generated using phiC31 integrase. *Nat Methods*. 2005;2:975 – 9.
7. Groth AC, Fish M, Nusse R, Calos MP. Creation of transgenic *Drosophila* by using the sitespecific integrase from phage phiC31. *Genetics*. 2004;166:1775 – 82.
8. Geisinger J, Calos MP. Site-specific recombination using phiC31 integrase. In: Renault S, Duchateau P, editors. *Site-directed insertion of transgenes*. Dordrecht: Springer Science; 2013.
9. Olivares EC, et al. Site-specific genomic integration produces therapeutic factor IX levels in mice. *Nat Biotechnol*. 2002;20:1124 – 8.
10. Chavez C, et al. Kinetics and longevity of phiC31 integrase in mouse liver and cultured cells. *Hum Gene Ther*. 2010;21:1287 – 97.
11. Keravala A, et al. Long-term phenotypic correction in factor IX knockout mice by using phiC31 integrase-mediated gene therapy. *Gene Ther*. 2011;18:842 – 8.
12. Chavez C, et al. Long-term expression of human coagulation factor VIII in a tolerant mouse model using the phiC31 integrase system. *Hum Gene Ther*. 2012;23:390 – 8.
13. Chavez C, Calos M. Therapeutic applications of the phiC31 integrase system. *Curr Gene Ther*. 2011;11:375 – 81.
14. Karow M, Calos M. The therapeutic potential of phiC31 integrase as a gene therapy system. *Expert Opin Biol Ther*. 2011;11:1287 – 96.
15. Chalberg TC, Genise HL, Vollrath D, Calos MP. PhiC31 integrase confers genomic integration

- and long-term transgene expression in rat retina. *Invest Ophthalmol Vis Sci.* 2005;46:2140 – 6.
16. Bertoni C, et al. Enhancement of plasmid-mediated gene therapy for muscular dystrophy by directed plasmid integration. *Proc Natl Acad Sci U S A.* 2006;103:419 – 24.
 17. Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell.* 2007;131:861 – 72.
 18. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 2006;126:663 – 76.
 19. Karow M, et al. Site-specific recombinase strategy to create induced pluripotent stem cells efficiently with plasmid DNA. *Stem Cells.* 2011;29:1696 – 704.
 20. Zhao C, et al. Recombinase-mediated reprogramming and dystrophin gene addition in mdx mouse induced pluripotent stem cells. *PLoS One.* 2014;9(4), e96279.
 21. Keravala A, et al. A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. *Mol Genet Genomics.* 2006;276:135 – 46.
 22. Zhu F, et al. DICE, an efficient system for iterative genomic editing in human pluripotent stem cells. *Nucleic Acids Res.* 2014;42(5), e34.
 23. Thorpe HM, Wilson SE, Smith MCM. Control of directionality in the site-specific recombination system of the *Streptomyces* phage phiC31. *Mol Microbiol.* 2000;38:232 – 41.
 24. Khaleel T, Younger E, McEwan A, Varghese A, Smith M. A phage protein that binds phiC31 integrase to switch its directionality. *Mol Microbiol.* 2011;80:1450 – 63.
 25. Farruggio A, Chavez C, Mikell C, Calos M. Efficient reversal of phiC31 integrase recombination in mammalian cells. *Biotechnol J.* 2012;7(11):1332 – 6.

采用三链寡核苷酸和肽核酸进行精确基因组修饰

Raman Bahal , Anisha Gupta , and Peter M. Glazer

摘要 很多的遗传疾病是由于单碱基突变造成的缺陷性蛋白表达所产生的。除了基因替代性治疗外，在基因组 DNA 的多个疾病相关突变点中已经实现了点特异性功能性修复和表达。在这一方面，三链技术通过同源重组修复（HR）或突变诱导相关机制修复突变位点从而改变基因表达。在本章我们将会探讨寡聚核苷酸叠加三链技术（TFOs）和肽核酸（PNAs）在点特异性基因编辑中的发展。

关键词 TFOs • PNA • 重组 • 修复 • 突变产生 • 基因编辑

R. Bahal , Ph.D. • A. Gupta , Ph.D.

Department of Therapeutic Radiology , Yale University School of Medicine ,
New Haven , CT , USA

P. M. Glazer , M.D., Ph.D. (*)

Department of Therapeutic Radiology , Yale University , New Haven , CT , USA

Department of Genetics , Yale University , New Haven , CT , USA

e-mail: peter.glazer@yale.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_6

简介

很多遗传疾病都是由单碱基突变引起的缺陷蛋白表达产生的。除了碱基替代疗法之外，对基因组 DNA 的特异性位点进行编辑能够实现多个基因功能的操纵和表达的改变，从而对多种遗传缺陷产生影响。在这一方面，三链技术通过同源重组（HR）或靶向突变产生机制能够实现突变基因的修复和表达的改变。在本章中我们将要讨论寡聚核苷酸叠加三链技术（TFOs）和肽核酸（PNAs）在点特异性基因编辑领域中所取得的进步。

寡核苷酸叠加三链聚合技术

除了沃森-克里克配对之外，被人所熟知的 Hoogsteen 配对连接引导人们设计另一种模式的分子识别 DNA，也就是所谓的 TFOs。三螺旋概念最早由 Pauling 和 Corey 在 1953 年提出，而实际的三螺旋构象则由 Felsenfeld 于 1957 年报道[1]。通过 RNA 衍射研究发现，多聚尿嘧啶（U）和多聚腺嘌呤（A）在镁离子存在的情况下能够按照 2: 1 的比例进行整合[1]。高密度的同型嘌呤序列能够和特异性的片段整合形成 TFOs，使其可以靶向于独特的基因或介导基因功能发生改变。

连接编码

多聚嘧啶 TFOs 能够与多聚嘌呤 DNA 相连接，通过 Hoogsteen 氢键‘嘧啶基序’T 能和 AT 相连，C 能和 GC 相连（产生 T—A—T 和 C—G+—C 三链聚合体）（图 1），引起平行方向多聚化位点的重叠链的产生[3,4]。‘嘧啶基序’的方向由 Peter Dervan 及其同事发现，G 和 GC、A 和 AT（形成 A—A—T 和 G—G—C 三聚碱基）通过不可逆的氢键链接在 B-DNA 目标基因组中形成反向平行的多聚化结构[5-7]。通过 X 射线衍射[8]、核磁共振[9]和化学探针[10]研究发现 TFO 能够和基因组 DNA 结合引起螺旋发生扭曲[11]，从而引起细胞系统中突变的产生或同源重组修复的加强。

通过化学修饰提升 TFO 的连接亲和力和稳定性

TFOs 和双链 DNA 的连接受多种不同因子的影响，包括粒子条件、序列组成和靶向位点的易编辑性[12]。当多聚 TFOs 进行平行连接时，必须将胞嘧啶的 N3 进行质子化以便于和鸟嘌呤的 N7 产生最优化的 Hoogsteen 连接[13]。因此，TFOs 的嘧啶 pH 依赖性限制了其在细胞内的应用。另一方面，TFOs 多聚嘌呤与目标 DNA 的连接不依赖于 pH。然而，G 富集嘌呤 TFOs 由于其鸟嘌呤富集段在生理性高 K⁺条件下易于形成 G-四聚链[14-17]。细胞内高 K⁺条件能够促使 TFOs 形成稳定化的二级结构——G 四聚链。因此为了形成稳定化的三链聚合结构，因而需要持续性地引入多聚嘌呤[18]。嘧啶类的瞬间切断或单碱基的错配能够极大地破坏 TFO 连接的不稳定性。由于 TFOs 在同型嘧啶或同型嘌呤片段的使用上受到了限制，因此无法对复杂的 ds-B-DNA 结构的人源性基因进行编辑。而该技术的另一主要缺陷是三链复合体的多阴离子状态和细胞环境下可被编辑的位点状态对于 TFO 的设计仍是挑战。

因此，对于 TFOs 的化学修饰必须能够增强三链结构的稳定性并保护其不被降解。很多努力都试图对 TFO 的设计进行优化。在涉及到多聚嘧啶的 TFOs 设计时，会将普通的胞嘧啶置换成 5-甲基胞嘧啶或伪异胞嘧啶从而克服其 pH 依赖性[20,21]。其他的胞嘧啶替换策略还包括 8-氧腺嘌呤[22]，7,8-二氢-8-氧腺嘌呤[23]，6-氧胞嘧啶[24]，8-氧-2'-去氧腺嘌呤[25,26]和 8-胺鸟嘌呤[27]，能够提高结合亲和力。类似的对于胸苷酸的修饰包括 2'-去氧尿苷酸（dU），5-炔丙基胺-和 2'-胺乙氧基，5-炔丙基胺-dU 二取代衍生物，能够提高三链聚合体的稳定性[28-32]。对片段进行修饰的策略包括在 G 富集的 TFOs 中引入 2'-去氧-6-硫鸟嘌呤核苷酸或 7-去硫唑-2'-去氧黄嘌呤核苷抑制 G-四聚体的形成[15-17]。也有很多的修饰发生在糖基，例

如 2'-O-核苷酸[33,34]和 2'-O-胺乙基核糖[35-37]被发现具有核酸酶抗性并能促进三聚体稳定性。在骨架和碱基上不同的化学修饰能够提高 TFOs 的连接亲和性并拓宽生物应用的范围。用吗啉寡核苷酸替换磷酸糖类骨架，或其它的阳离子性氨基磷酸脂例如 N, N-二甲基-2-咪唑啉酮（DEED）或 N, N-二乙基-胺炔丙基胺，能够提高 TFO 在体外的亲和力[38-41]。而且，在 PNAs 中通过将骨架置换成非带电体、手性肽单位，极大推动了其在三链聚合领域的发展[42-44]。PNA 相关的三链聚合试剂将在这一章的其他部分讨论。

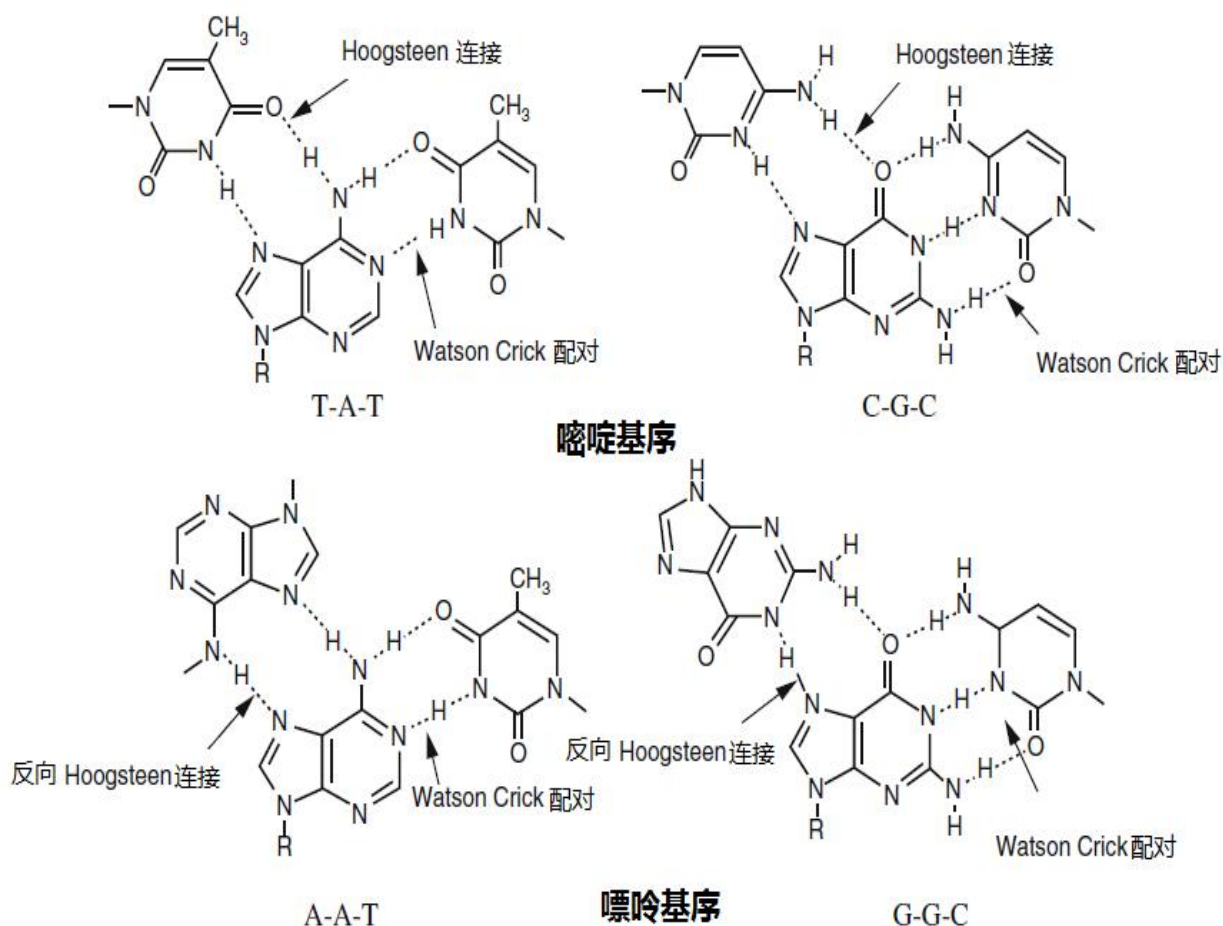


图1.三链聚合基序。(上)通过Hoogsteen氢键连接的三链平行DNA嘧啶聚合结构。(下)通过反式Hoogsteen氢键相连的三链反向平行嘌呤基序。如图为经典的三聚化结构。

三链介导的基因编辑

通过三链聚合体诱导目的基因突变

通过诱导基因组的特定位点发生突变，三链技术能够被用于诱导可遗传的基因功能和表达的改变。TFOs 直接与目的基因相连诱发 DNA 修复并在修复位点引入突变。共轭型补骨脂素（psf），一种 TFO 的光反应性突变诱导试剂，能够诱导 DNA 发生损伤[45-48]。在 UVA 辐射作用下，补骨脂素能够插入 DNA 与双链上的胸腺嘧啶发生共价连接[49]。在哺乳动物细胞中会发生 T: A 到 A: T 的置换[46,50-52]。

在哺乳动物细胞中采用基于质粒的实验检测突变的产生

为了分析 TFO 诱导产生的突变，构建了以 supF 基因为基础的报告系统。克隆到 SV40 载体上的 supF 基因能够编码具有 TFO 连接位点的琥珀抑制子酪氨酸 tRNA。利用 pso-TFO 和 UV A 辐射进行处理后，这一载体被整合到 COS-7 细胞中。突变的频率可以通过提取质粒和细菌转化进行量化。结果显示 6%的质粒发生了突变即有超过 55%的 T:A 转化为 A:T[45]。而且，有研究显示在诱变剂非依赖的条件下，单独的 TFO 也能够诱导突变的产生[53]。

TFO 在哺乳动物染色体中诱导产生的突变

TFPs 被发现能够在培养细胞染色体的靶向位点诱导基因突变的产生。Majumdar 等人报道了利用 Pso-TFOs 诱导 CHO 细胞内源性次黄嘌呤磷酸核糖转移酶 (hprt) 基因突变发生的事例[54]。他们发现在三链靶向区域有 80%的突变诱导率。Vasquez 等人也构建了一种转基因鼠，在 supF 突变报道质粒中包含有具有 30bp 三链靶向结构的多拷贝 λ 载体 (图 2) [55]。在利用 pso-TFOs 对细胞进行处理后，诱导点突变发生率可提高十倍。无论是在 UV A 辐照情况下还是非辐照情况下均观测到了突变的产生，提示 TFO 单独就具有诱导突变发生的潜力。

诱导同源重组的 TFOs

基于 HR 的基因编辑具有广泛性的应用。但是 HR 在哺乳动物中发生的低频率及在非靶向区域的随机插入限制了其应用 (回顾参考文献[56])。为了提升 HR 的频率，开发了多种诱导 DNA 损伤的，开发了多种方法，包括在目的基因位点进行 UV 辐照，利用内切酶对双链损伤或交联的经烷化处理或补骨脂素处理的 DNA 双链进行切割[57-59]。TFO 引起的 DNA 损伤有的是通过三链聚合体的形成或是通过突变诱导诱变剂提高基因重组修饰的频率。有关于利用 TFOs 通过 HR 进行不同的基因编辑实验的总结见表 1。

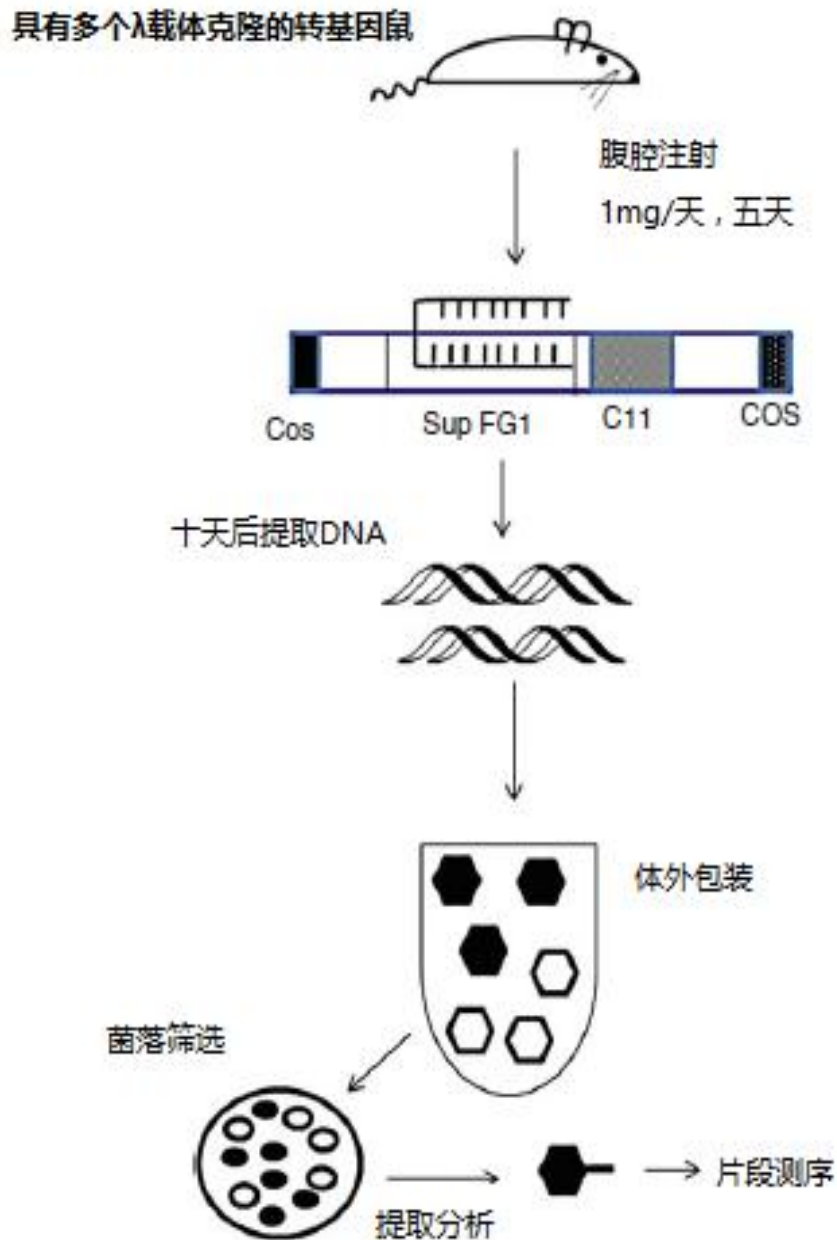


图2.在小鼠中利用TFO进行染色体突变检测的实验方法。转基因小鼠染色体中插入多个A supFG 1载体克隆，包括有30bp的三链复合体链接。腹腔注射TFOs。处理后收集组织，提取DNA并进行分析。如果没有突变产生，在IPTG和X-Gal存在的条件下显示出蓝色。如果产生突变，菌落呈现白色。

分子内重组

为了检测同源重组效率的提高程度，构建了多种不同的报告系统。我们实验室利用 Pso-TFOs 同时诱导质粒上报告基因和染色体位点上的突变产生[60]。为了利用 TFO 直接对交联聚合体进行靶向重组，Faruqi 等人设计了具有 supF 两种不同点突变和一个 TFO 连接位点的突变基因的 SV40 载体质粒[61]。Later, Luo 等人在小鼠染色体中进行了相似结构的构建[62]。由 TFO 诱导的重组可以通过 β -半乳糖苷酶筛选实验的菌株克隆实验进行检测。在这一实验中利用游离型和补骨脂素共轭结合型的 TFOs。实验结果发现由补骨脂素共轭结合的 TFO 能够诱导更高频率的链内交联的产生。而且，有关于染色体内重组的机制已经进行过研究。三链聚合体诱导的重组效率的提升主要依赖于核酸切除修复（NER）途径[61]。在着色性干皮病组 A（XPA）修复因子缺陷的细胞系中并没有显著性的重组修复的发生[63]。由补骨脂素共轭 TFO 诱导的基因重组也是部分依赖于 XPA，并也于其他多条修复通路有关。

表 1. 在哺乳动物细胞中由 TFO 介导的 HR

TFO	重组类型	报告基因	重组频率	细胞系
AG-30	分子内重组	supF（细菌抑制性 tRNA）	0.37	COS-7
Pso-AG30	分子内重组	SupF	0.58	COS-7
AG30	染色体内重组	TK（胸苷激酶）	1.2	LTK-
AG30+51mer 外源 DNA	分子间重组	Flic（Firefly 荧光素酶基因）	0.05	CHO 衍生
22 mer 嘌呤富集 TFO	分子间重组	SupF	~0.05	Jurkat T 类淋巴细胞
补骨脂素共轭 19mer 嘌呤 TFOs	分子内整合	SupF	0.002	Hela

分子内整合

Chan 等人利用了一种基于 TFOs 的方法将连接的外源性 DNA 片段同源重组到靶向位点（目的在于对配对碱基进行修复）并实现了分子间的基因重组[64]。通过这一策略，三链聚合体的形成能够诱导靶向位点的敏感性重组和修复[65]。在外源性 DNA 上还连有用于进行重组的 TFO 功能域。随后研究证明外源 DNA 和 TFOs 之间的连接不是必须的[66-69]。

肽核酸（PNAs）

以 TFOs 为基础的 DNA 编辑的主要问题是酶学降解。在过去，为了提高酶学稳定性，发展了多种核酸衍生物。有一类具有潜力的核酸衍生物是 PNA。在结构上，PNA 和 DNA 具有同源性，只是将 DNA 的磷酸二聚骨架替换成了不带点的胺乙基甘氨酸骨架[42]。由于 PNAs 电中性的手性结构，相对更难被降解（图 3）。PNA 能够和 DNA/RNA 目标序列进行互补配对，并具有更高的稳定性（PNA/DNA 或 PNA/RNA）。PNAs 也能插入 Watson-Crick 配对的 DNA 双链中通过 Hoogsteen 连接形成 PNA-DNA 三链聚合体[70]。互补的两条 PNA 链能通过复杂的连接聚合作用形成夹钳结构，并于目标双链 DNA 形成更为稳定的复合体。

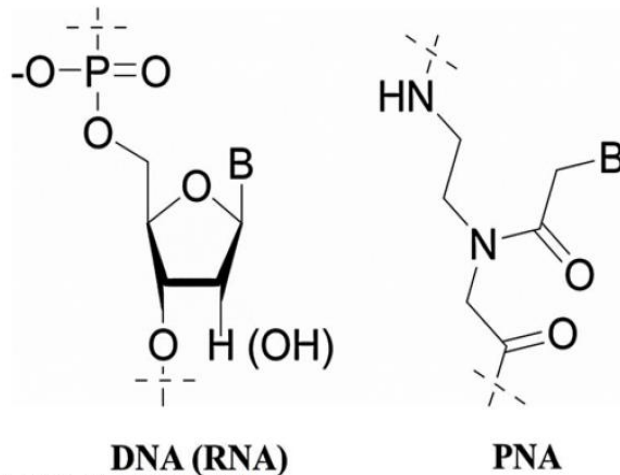


图3. DNA (RNA) 和PNA的化学结构。

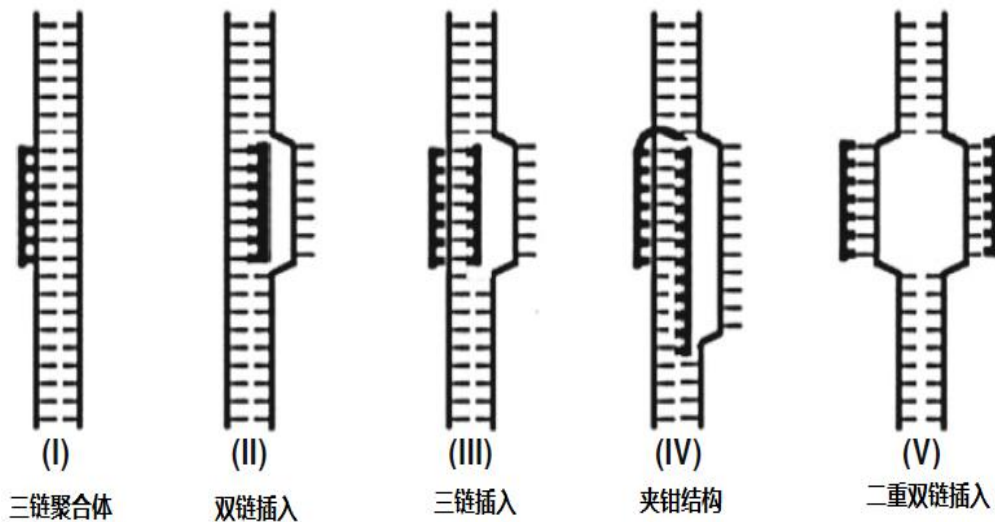


图4. PNA和双链DNA不同的连接模式。

最近有研究发现 PNA 能够插入特定序列的双链 DNA 中。PNA 与双链 DNA 结合的机制主要依赖于序列的互补配对。至少有物种不同的连接模式（图 4）。对于胞嘧啶富集 PNA，连接主要通过 Hoogsteen 配对连接，比例为 1: 1。同型嘌呤 PNA 也能和双链 DNA 按照 1: 1 的比例结合，但是 Watson-Crick 配对的结果能够导致同源 DNA 序列原有碱基的替换，产生 D 环构象[71]。从另一方面来讲，同型嘧啶 PNA 也能够按照 2: 1（PNA: DNA）的比例和双链 DNA 中的一条按照 Watson-Crick 配对结合，另一条链按照 Hoogsteen 连接配对，形成 PNA₂-DNA 复合体和进行局部替换的 D 环结构。同型嘧啶/同型嘌呤的连接模式促进了“尾端夹钳”发卡结构的发展，该结构能够识别部分互补配对的双链 DNA 结构[44]。Peter Nielsen 及其同事发现了第五种依赖于伪互补（pcPNAs）PNAs 刺激 DNA 双链交联的连接模式。这一策略也被称为二重双链插入策略[70]。这一方式的结合选择性更为复杂，而且引入两种特异性 pcPNA 进行结合的方式更为复杂，存在两条链相互作用的可能性，因而其应用也存在限制。

PNAs 作为转录调控、蛋白连接抑制子、DNA 杂交阻遏子具有巨大的发展潜力。然而如何将 PNAs 导入细胞是其进行治疗应用的主要障碍。

利用 PNA 进行修复和重组

PNA 由于和互补序列具有高亲和力，能够诱导突变产生及 HR 修复。在过去，Faruqi 等人利用一种单链二聚化 PNA 对小鼠纤维母细胞中代偿性 λ 噬菌体穿梭质粒整合染色体的 supFG1 突变报告基因进行靶向整合[72]。PNA 在 supFG1 基因形成 8bp 或 10bp 的二聚体和三聚体螺旋夹钳结构并可诱导 0.1% 频率的突变产生，比原有背景高 10 倍[72]。

Rogers 等人为证明二聚化 bis-PNA 能够促进基因重组作了详尽的研究[73]。一般来讲，bis-PNAs 既能发生单链插入也能形成三链聚合体，并能与靶向双链 DNA 形成夹钳结构。在这一研究中，他们发现 bis-PNA 与 40-核苷酸长度的外源 DNA 共轭连接后能够诱导直接性的基因整合。PNA-外源性 DNA 共轭连接可通过马来酰亚胺为基础的化学反应形成。PNA-DNA 共轭体能够介导人类离体细胞 supFG1 报告基因的体外特异性整合，并导致该背景下至少提高六十倍的突变修复[73]。类似的点特异性重组的诱导也能通过非结合的 bis-PNA 和外源性 DNA 完成。Bis-PNA 及 bis-PNA-外源 DNA 共轭体也被发现能够诱导目标质粒的高频率 DNA 修复。有研究发现无论是 PNA 诱导的重组还是基因修复都依赖于核酸删除修复因子，XPA（着色性干皮病代偿性 A 组蛋白）[73]。PNA 通过与 DNA 形成二聚化的夹钳结构能够导致原有螺旋结构的变形，进而大概率的诱导 DNA 修复的产生并因此使目的 DNA 易于进行重组。

Wang 等人设计了一系列的二聚化 PNAs 在大肠杆菌 supFG1 基因形成 10bp 的同型嘌呤/同型嘧啶二聚化夹钳结构[74]。为了研究在整合机制的概率，在小鼠纤维母细胞 15 号染色体上整合了具有 supFG1 基因的 λ 噬菌体穿梭质粒克隆。SupFG1 基因的突变形成能够通过在大肠杆菌 ClacZ125 (am) 细菌的转化后计数进行检测。有研究发现，二聚化的 PNAs 能够诱导小鼠细胞染色体中 supFG1 基因发生 0.1% 频率的基因突变（高于原有背景十倍）[74]。序列分析证实大多数突变发生于 PNA 结合位点。进一步的研究涉及了 PNAs 进行点特异性突变诱导所应用的诱变剂，例如苯甲酮、葱醌或补骨脂素能够被用于体外 DNA 编辑实验。

在其他的研究中则主要应用了二聚化的 bis-PNAs-补骨脂素共轭体。PNA-补骨脂素联接体能够与 supFLSG3 报告基因聚合形成三链结构并可由补骨脂素介导发生点特异性的光加成反应[75]。光加成反应在体外实验中被证实，而目标位点的突变形成主要利用 SV 40 为基础的游离穿梭质粒实验。PNAs 与补骨脂素共轭体的光加成反应能够诱导突变发生的概率为 0.46%（比背景高 6.5 倍）。对于细胞内的游离穿梭质粒，补骨脂素-PNA 的突变诱导效率为 0.13%（高于背景 3.5 倍）。在该研究中发现大多数的突变诱导类型是 PNA 连接位点或光加成反应整合位点的基因删除或单碱基替换[75]。这一系列研究为 PNAs 成为一种基因靶向应用的研究工具提供了基础。

除此之外，还探讨了 PNA 和双链 DNA 之间不同的连接模式。除了 bis PNAs 之外，pcPNA 也被用于细胞内多种片段的靶向。有充分研究显示，由于空间阻遏作用的存在，pcPNA 无法形成 pcPNA-pcPNA 二聚体，但是能够和互补的 DNA 序列通过 Watson-Crick 配对插入双链二聚体。有研究显示与补骨脂素共轭连接的 pcPNA 能够介导点特异性的光加成反应并对哺乳动物细胞染色体 DNA 中及游离质粒 DNA 进行编辑，且不会产生脱靶效应[76]。补骨脂素-pcDNA 在连接位点所诱导的突变主要为单碱基替代突变或基因删除。PcPNAs 单独不能诱导突变产生而且相同片段中 PNA 的互补序列也无法产生突变。

PNAs 能够和外源性 DNAs 联合形成三链结构通过同源重组机制诱导基因纠正反应，这一手段可在多种疾病模型上进行应用，包括 β -地中海贫血病[69]。通常来讲， β -球蛋白基因突变能够导致异常转录并减弱 β -球蛋白的功能并最终引起 β -地中海贫血病的发生。有研究发现当 bis-PNA 和外源整合性突变纠正 DNA 片段进行共转染时，能够促进地中海贫血病常见相关突变—— β -球蛋白二号内含子单碱基修复。绿色荧光蛋白- β -球蛋白融合基因被用于进行修复基因的检测。有研究显示利用亲溶酶体试剂、氯喹能够提高 PNAs/DNA 的整合频率

[69]。

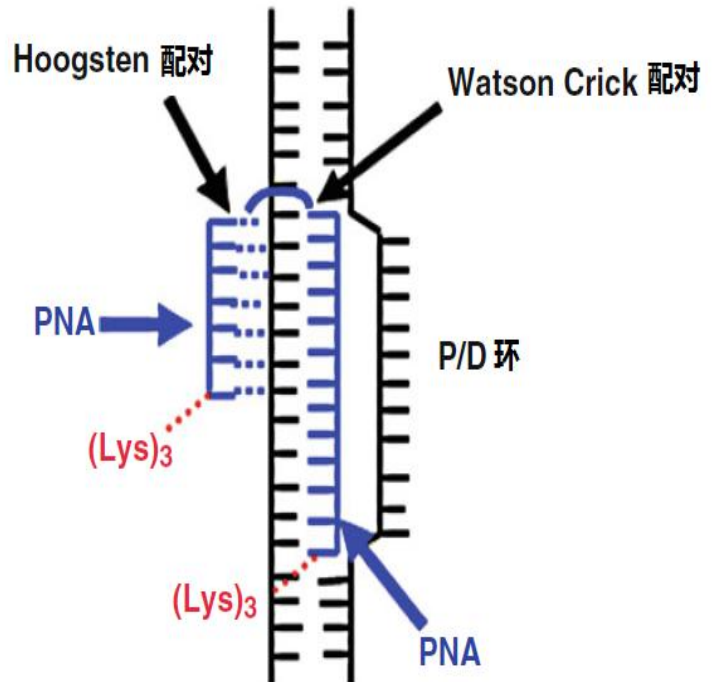
Later on、Lonkar 及其同事也证实 pcPNAs 与外源性 DNA 片段共转染时，能够促进 β -球蛋白基因二号内含子初始段单碱基对突变[68]。基因编辑能够通过荧光素酶蛋白- β 球蛋白融合基因的转录产物功能测定进行检测。除此之外，pcPNAs 能够刺激人源性纤维母细胞不依赖于核酸切除修复因子——XPA 的重组。这些结果显示 pcPNA 能够作为哺乳动物细胞非限制性片段点特异性突变诱导的工具。

除了 β -球蛋白基因，基因编辑技术也可被用于 CCR5 基因，CCR5 是一种趋化因子受体，在 HIV-1 入侵人类细胞是具有重要功能，而该基因中如带有突变可对 HIV-1 的感染产生抗性。Schleifman 等人发现利用 bisPNAs/ 外源 DNA 靶向 CCR5 基因诱导类似自然状态下的 CCR5-delta32 终止子突变的频率可达到 2.46%[67]。在系列实验中，CCR5 基因的编辑可发生在 DNA、RNA 和蛋白水平[67]。研究显示终止子的导入能够产生 HIV-1 抗性。这一工作显示 PNA 诱导的基因编辑能够用于 HIV-1 感染个体的 CCR5 敲除疗法。

而且 Rogers 等人还发现通过肽核酸（PNA）和转运肽、触角足（Antp）共轭形成三聚体能够成功实现造血祖细胞体内染色体的基因编辑[77]。而且，在造血祖细胞进行基因编辑过后仍然具有分化的能力。这一策略排除了利用 PNA/外源性 DNA 进行转染的步骤。

McNeer 等人提出了利用多聚（乳酸聚羟基乙酸）（PLGA）纳米粒子进行 PNA/外源性 DNA 分子转染进行基因编辑的方法[78]。PLGA 是一种被 FDA 批准的生物相容性多聚物，临床上被用于多种疾病的药物递送，包括前列腺癌的治疗（Lupron 和 Trelstar）。之前的研究证实 PLGA 纳米颗粒能够被用于细胞内核苷酸多聚物和寡核苷酸的递送，包括质粒 DNA 和进行基因沉默研究的 siRNAs[79]。PLGA 纳米颗粒能够利用复乳溶剂蒸发技术获得。PLGA 纳米颗粒能够包裹 tcPNA（图 5）、DNA（利用反荷离子亚精胺进行电中性处理）或同时包裹 tcPNA 和 DNA（利用赖氨酸同时共轭 PNA 的 C 端和 N 端作为 DNA 的反荷离子）。而且，可以利用包裹有 0.25-2mg PLGA/mL 浓度的核苷酸的纳米粒子处理人源性 CD34+ HSCs。结果显示利用纳米颗粒处理过的细胞具有更强的恢复力并可用于核转染实验。纳米颗粒处理也能够导致更高比例的重组发生，能够将效率提高 60 倍以上。

图5.赖氨酸共轭尾端夹钳PNA与双链DNA结合。



利用纳米颗粒将 PNA/外源 DNA 导入体内人源性细胞进行基因编辑的方式在造血祖细胞移植免疫缺陷小鼠中取得了成功。通过静脉注射具有 PNAs/DNAs 的粒子能够对小鼠中人源细胞的 CCR5 基因进行编辑。深入进行序列分析的结果发现在脾脏和骨髓内成功进行 CCR5 基因编辑的频率在 0.1-0.5% 之间。与此同时，在同源性 CCR2 基因处发生脱靶的概率低两个数量级。利用纳米技术进行 PNA 点特异性突变具有很多优势。首先，这一策略为 HSCs 体内直接性基因编辑的复杂系统提供了发展的框架。其次，这一工作也为以人类造血祖细胞为靶标的药物递呈提供了一种通用性方法。

tcPNA 的靶向通常需要目标区域具有富集的同型嘌呤或同型嘧啶区域。

在其他的策略研究中我们也发现新一代的化学修饰 γ PNAs 能够用于靶向多种 DNA 基因组片段并能诱导基因修饰的产生[80]。在 γ PNAs 中，手性中心位于 PNA 骨架的 γ 位点(图 6) [81-83]。在 γ 位手性中心的产生能够增强 PNAs 和与其互补的 DNA 或 RNA 之间的亲和力 [84,85]。

为了证实这一观点，在具有 β -球蛋白/EGFP 融合基因的转基因鼠模型的人源性 β -球蛋白 2 号内含子插入了 GFP 编码序列[86,87]。在内含子区域中 IVS2-654 (C->T) 突变是东南亚地区遗传性地中海贫血病的常见致病原因。这一突变能够导致在内含子形成一个隐蔽性剪切位点并阻止基因表达。利用纳米颗粒递送系统我们成功地将设计好的 IVS2-654 附近区域靶向的单链 γ PNAs 与外源性 DNA 导入基因组 DNA，诱导小鼠离体骨髓细胞中 0.8% 的点特异性基因编辑和 IV 体内注射中 0.1% 的体内诱导突变，并未检测到毒性的产生。我们已经成功为 γ PNAs 通过 Watson-Crick 识别在非限制性位点实现基因编辑迈出了重要性的一步。

总结

总之，TFOs 和 PNAs 在多种策略的基因编辑中都具有重要的作用。无论是在体内实验中还是在体外实验中都取得了非常激动人心的结果。通过利用化学/生物学的反应过程，研究人员提出了多种提高基因编辑频率的策略。虽然相较于 TFOs，PNAs 具有更加优越的化学特征，但是还需要进行还多工作从而获得更高的基因编辑效率，尤其是和临床具有相关性的工作。

参考文献

1. Felsenfeld G, Davies DR, Rich A. Formation of a three-stranded polynucleotide molecule. *J Am Chem Soc.* 1957;79:2023 - 4.
2. Francois JC, Saison-Behmoaras T, Thuong NT, Helene C. Inhibition of restriction endonuclease cleavage via triple helix formation by homopyrimidine oligonucleotides. *Biochemistry.* 1989;28(25):9617 - 9.
3. Hoogsteen K. The structure of crystals containing a hydrogen-bonded complex of 1- methylthymine and 9-methyladenine. *Acta Crystallogr.* 1959;12:822 - 3.
4. Moser HE, Dervan PB. Sequence-specific cleavage of double helical DNA by triple helix formation.
5. Letai AG, Palladino MA, Fromm E, Rizzo V, Fresco JR. Specificity in formation of triplestranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns. *Biochemistry.* 1988;27(26):9108 - 12.
6. Beal PA, Dervan PB. Second structural motif for recognition of DNA by oligonucleotide directed triple-helix formation. *Science.* 1991;251(4999):1360 - 3.
7. Durland RH, Kessler DJ, Gunnell S, Duvic M, Hogan ME, Pettitt BM. Binding of triple helix forming oligonucleotides to sites in gene promoters. *Biochemistry.* 1991;30(38):9246 - 55.
8. Arnott S, Selsing E. Structures for the polynucleotide complexes poly dA.poly dT and poly

- dT.poly dA.poly dT. *J Mol Biol.* 1974;88(2):509 – 21.
9. Radhakrishnan I. De, I. S. C.; Patel, D. J., Nuclear magnetic resonance structural studies of intramolecular purine • purine • pyrimidine DNA triplexes in solution. Base triple pairing alignments and strand direction. *J Mol Biol.* 1991;221(4):1403 – 18.
 10. Francois JC, Saison-Behmoaras T, Helene C. Sequence-specific recognition of the major groove of DNA by oligodeoxynucleotides via triple helix formation. Footprinting studies. *Nucleic Acids Res.* 1988;16(24):11431 – 40.
 11. Hartman DA, Kuo SR, Broker TR, Chow LT, Wells RD. Intermolecular triplex formation distorts the DNA duplex in the regulatory region of human papillomavirus type-11. *J Biol Chem.* 1992;267(8):5488 – 94.
 12. Pilch DS, Brousseau R, Shafer RH. Thermodynamics of triple helix formation: spectrophotometric studies on the d(A)₁₀ • 2d(T)₁₀ and d(C + 3T4C + 3) • d(G3A4G3) • d(C3T4C3) triple helices. *Nucleic Acids Res.* 1990;18(19):5743 – 50.
 13. Lee JS, Johnson DA, Morgan AR. Complexes formed by (pyrimidine)_n. (purine)_n DNAs on lowering the pH are three-stranded. *Nucleic Acids Res.* 1979;6(9):3073 – 91.
 14. Cheng AJ, Van DMW. Monovalent cation effects on intermolecular purine-purine-pyrimidine triple-helix formation. *Nucleic Acids Res.* 1993;21(24):5630 – 5.
 15. Milligan JF, Krawczyk SH, Wadwani S, Matteucci MD. An anti-parallel triple helix motif with oligodeoxynucleotides containing 2' -deoxyguanosine and 7-deaza-2' -deoxyxanthosine. *Nucleic Acids Res.* 1993;21(2):327 – 33.
 16. Olivas WM, Maher III LJ. Overcoming potassium-mediated triplex inhibition. *Nucleic Acids Res.* 1995;23(11):1936 – 41.
 17. Rao TS, Durland RH, Seth DM, Myrick MA, Bodepudi V, Revankar GR. Incorporation of 2' -deoxy-6-thioguanosine into G-rich oligodeoxyribonucleotides inhibits G-tetrad formation and facilitates triplex formation. *Biochemistry.* 1995;34(3):765 – 72.
 18. Cheng A-J, Van DMW. Oligodeoxyribonucleotide length and sequence effects on intermolecular purine-purine-pyrimidine triple-helix formation. *Nucleic Acids Res.* 1994;22(22):4742 – 7.
 19. Brown PM, Fox KR. Nucleosome core particles inhibit DNA triple helix formation. *Biochem J.* 1996;319(2):607 – 11.
 20. Lee JS, Woodsworth ML, Latimer LJP, Morgan AR. Poly(pyrimidine) • poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. *Nucleic Acids Res.* 1984;12(16):6603 – 14.
 21. Singleton SF, Dervan PB. Influence of pH on the equilibrium association constants for oligodeoxyribonucleotide- directed triple helix formation at single DNA sites. *Biochemistry.* 1992;31(45):10995 – 1003.
 22. Miller PS, Bhan P, Cushman CD, Trapane TL. Recognition of a guanine-cytosine base pair by 8-oxoadenine. *Biochemistry.* 1992;31(29):6788 – 93.
 23. Jetter MC, Hobbs FW. 7,8-Dihydro-8-oxoadenine as a replacement for cytosine in the third strand of triple helices. Triplex formation without hypochromicity. *Biochemistry.* 1993;32(13):3249 – 54.
 24. Berressem R, Engels JW. 6-Oxocytidine--a novel protonated C-base analog for stable triple helix formation. *Nucleic Acids Res.* 1995;23(17):3465 – 72.
 25. Krawczyk SH, Milligan JF, Wadwani S, Moulds C, Froehler BC, Matteucci analog exhibiting

- pH-independent binding within the physiological range. *Proc Natl Acad Sci USA*. 1992;89(9):3761 – 4.
26. Ishibashi T, Yamakawa H, Wang Q, Tsukahara S, Takai K, Maruyama T, Takaku H. Triple helix formation with oligodeoxyribonucleotides containing 8-oxo-2' -deoxyadenosine and 2' -modified nucleoside derivatives. *Nucleic Acids Symp Ser*. 1995;34:127 – 8 (Twentysecond Symposium on Nucleic Acids Chemistry, 1995).
27. Soliva R, Garcia RG, Bias JR, Eritjal R, Asensio JL, Gonzalez C, Luque FJ, Orozco M. DNATriplex stabilizing properties of 8-aminoguanine. *Nucleic Acids Res*. 2000;28(22):4531 – 9.
28. Froehler BC, Wadwani S, Terhorst TJ, Gerrard SR. Oligodeoxyribonucleotides containing C-5 propyne analogs of 2' -deoxyuridine and 2' -deoxycytidine. *Tetrahedron Lett*. 1992;33(37): 5307 – 10.
29. Bijapur J, Keppler MD, Bergqvist S, Brown T, Fox KR. 5-(1-propargylamino)-2' -deoxyuridine (UP): a novel thymidine analogue for generating DNA triplexes with increased stability. *Nucleic Acids Res*. 1999;27(8):1802 – 9.
30. Sollogoub M, Darby RAJ, Cuenoud B, Brown T, Fox KR. Stable DNA triple helix formation using oligonucleotides containing 2' -aminoethoxy,5-propargylamino-U. *Biochemistry*. 2002;41(23):7224 – 31.
31. Osborne SD, Powers VEC, Rusling DA, Lack O, Fox KR, Brown T. Selectivity and affinity of triplex-forming oligonucleotides containing 2' -aminoethoxy-5-(3-aminoprop-1-ynyl)uridine for recognizing AT base pairs in duplex DNA. *Nucleic Acids Res*. 2004;32(15):4439 – 47.
32. Rusling DA, Rachwal PA, Brown T, Fox KR. The stability of triplex DNA is affected by the stability of the underlying duplex. *Biophys Chem*. 2009;145(2-3):105 – 10.
33. Inoue H, Hayase Y, Imura A, Iwai S, Miura K, Ohtsuka E. Synthesis and hybridization studies on two complementary nona(2' -O-methyl)ribonucleotides. *Nucleic Acids Res*. 1987;15(15): 6131 – 48.
34. Shimizu M, Konishi A, Shimada Y, Inoue H, Ohtsuka E. Oligo(2' -O-methyl)ribonucleotides. Effective probes for duplex DNA. *FEBS Lett*. 1992;302(2):155 – 8.
35. Blommers MJJ, Natt F, Jahnke W, Cuenoud B. Dual recognition of double-stranded DNA by 2' -aminoethoxy-modified oligonucleotides: the solution structure of an intramolecular triplex obtained by NMR spectroscopy. *Biochemistry*. 1998;37(51):17714 – 25.
36. Cuenoud B, Casset F, Huesken D, Natt F, Wolf RM, Altmann K-H, Martin P, Moser HE. Dual recognition of double-stranded DNA by 2' -aminoethoxy-modified oligonucleotides. *Angew Chem Int Ed*. 1998;37(9):1288 – 91.
37. Puri N, Majumdar A, Cuenoud B, Natt F, Martin P, Boyd A, Miller PS, Seidman MM. Targeted gene knockout by 2' -O-aminoethyl modified triplex forming oligonucleotides. *J Biol Chem*. 2001;276(31):28991 – 8.
38. Xodo L, Alunni-Fabbroni M, Manzini G, Quadrioglio F. Pyrimidine phosphorothioate oligonucleotides form triple-stranded helices and promote transcription inhibition. *Nucleic Acids Res*. 1994;22(16):3322 – 30.
39. Escude C, Giovannangeli C, Sun J-S, Lloyd DH, Chen J-K, Gryaznov SM, Garestier T, Helene C. Stable triple helices formed by oligonucleotide N3' → P5' phosphoramidates inhibit transcription elongation. *Proc Natl Acad Sci U S A*. 1996;93(9):4365 – 9.
40. Lacroix L, Arimondo PB, Takasugi M, Helene C, Mergny J-L. Pyrimidine morpholino oligonucleotides form a stable triple helix in the absence of magnesium ions. *Biochem Biophys*

Res Commun. 2000;270(2):363 – 9.

41. Michel T, Debart F, Vasseur J-J, Geinguenaud F, Taillandier E. FTIR and UV spectroscopy studies of triplex formation between α -oligonucleotides with non-ionic phosphoramidate linkages and DNA targets. *J Biomol Struct Dyn*. 2003;21(3):435 – 45.

42. Nielsen PE, Egholm M, Berg RH, Buchardt O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*. 1991;254(5037):1497 – 500.

43. Demidov VV, Potaman VN, Frank-Kamenetskii MD, Egholm M, Buchard O, Sonnichsen SH, Nielsen PE. Stability of peptide nucleic acids in human serum and cellular extracts. *Biochem Pharmacol*. 1994;48(6):1310 – 3.

44. Nielsen PE. Targeted gene repair facilitated by peptide nucleic acids (PNA). *Chembiochem*. 2010;11(15):2073 – 6.

45. Havre PA, Glazer PM. Targeted mutagenesis of simian virus 40 DNA mediated by a triple helix-forming oligonucleotide. *J Virol*. 1993;67(12):7324 – 31.

46. Havre PA, Gunther EJ, Gasparro FP, Glazer PM. Targeted mutagenesis of DNA using triple helix-forming oligonucleotides linked to psoralen. *Proc Natl Acad Sci U S A*. 1993;90(16):7879 – 83.

47. Gasparro FP, Havre PA, Olack GA, Gunther EJ, Glazer PM. Site-specific targeting of psoralen photoadducts with a triple helix-forming oligonucleotide: characterization of psoralen monoadduct and crosslink formation. *Nucleic Acids Res*. 1994;22(14):2845 – 52.

48. Sandor Z, Bredberg A. Repair of triple helix directed psoralen adducts in human cells. *Nucleic Acids Res*. 1994;22(11):2051 – 6.

49. Takasugi M, Guendouz A, Chassignol M, Decout JL, Lhomme J, Thuong NT, Helene C. Sequence-specific photo-induced cross-linking of the two strands of double-helical DNA by a psoralen covalently linked to a triple helix-forming oligonucleotide. *Proc Natl Acad Sci U S A*. 1991;88(13):5602 – 6.

50. Sancar A, Tang M-S. Nucleotide excision repair. *Photochem Photobiol*. 1993;57(5):905 – 21.

51. Wood RD, Aboussekhra A, Biggerstaff M, Jones CJ, O' Donovan A, Shivji MK, Szymkowski DE. Nucleotide excision repair of DNA by mammalian cell extracts and purified proteins. *Cold Spring Harb Symp Quant Biol*. 1993;58:625 – 32.

52. Wang G, Levy DD, Seidman MM, Glazer PM. Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation. *Mol Cell Biol*. 1995;15(3):1759 – 68.

53. Seidman MM, Glazer PM. The potential for gene repair via triple helix formation. *J Clin Invest*. 2003;112(4):487 – 94.

54. Majumdar A, Khorlin A, Dyatkina N, Lin FLM, Powell J, Liu J, Fei Z, Khripine Y, Watanabe KA, George J, Glazer PM, Seidman MM. Targeted gene knockout mediated by triple helix forming oligonucleotides. *Nat Genet*. 1998;20(2):212 – 4.

55. Vasquez KM, Wang G, Havre PA, Glazer PM. Chromosomal mutations induced by triplexforming oligonucleotides in mammalian cells. *Nucleic Acids Res*. 1999;27(4):1176 – 81.

56. Vasquez KM, Marburger K, Intody Z, Wilson JH. Manipulating the mammalian genome by homologous recombination. *Proc Natl Acad Sci U S A*. 2001;98(15):8403 – 10.

57. Vos JMH, Hanawalt PC. DNA interstrand cross-links promote chromosomal integration of a selected gene in human cells. *Mol Cell Biol*. 1989;9(7):2897 – 905.

58. Bhattacharyya NP, Maher VM, McCormick JJ. Intrachromosomal homologous

recombination in human cells which differ in nucleotide excision-repair capacity. *Mutat Res.* 1990;234(1):31 – 41.

59. Tsujimura T, Maher VM, Godwin AR, Liskay RM, McCormick JJ. Frequency of intrachromosomal homologous recombination induced by UV radiation in normally repairing and excision repair-deficient human cells. *Proc Natl Acad Sci U S A.* 1990;87(4):1566 – 70.

60. Faruqi AF, Seidman MM, Segal DJ, Carroll D, Glazer PM. Recombination induced by triplehelix-targeted DNA damage in mammalian cells. *Mol Cell Biol.* 1996;16(12):6820 – 8.

61. Faruqi AF, Datta HJ, Carroll D, Seidman MM, Glazer PM. Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway. *Mol Cell Biol.* 2000;20(3):990 – 1000.

62. Luo Z, Macris MA, Faruqi AF, Glazer PM. High-frequency intrachromosomal gene conversion induced by triplex-forming oligonucleotides microinjected into mouse cells. *Proc Natl Acad Sci USA.* 2000;97(16):9003 – 8.

63. Vasquez KM, Christensen J, Li L, Finch RA, Glazer PM. Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. *Proc Natl Acad Sci USA.* 2002;99(9):5848 – 53.

64. Knauert MP, Kalish JM, Hegan DC, Glazer PM. Triplex-stimulated intermolecular recombination at a single-copy genomic target. *Mol Ther.* 2006;14(3):392 – 400.

65. Chan PP, Lin M, Faruqi AF, Powell J, Seidman MM, Glazer PM. Targeted correction of an episomal gene in mammalian cells by a short DNA fragment tethered to a triplex-forming oligonucleotide. *J Biol Chem.* 1999;274(17):11541 – 8.

66. Majumdar A, Muniandy PA, Liu J, Liu J-L, Liu S-T, Cuenoud B, Seidman MM. Targeted gene knock in and sequence modulation mediated by a psoralen-linked triplex-forming oligonucleotide. *J Biol Chem.* 2008;283(17):11244 – 52.

67. Schleifman EB, Bindra R, Leif J, del Campo J, Rogers FA, Uchil P, Kutsch O, Shultz LD, Kumar P, Greiner DL, Glazer PM. Targeted disruption of the *ccr5* gene in human hematopoietic stem cells stimulated by peptide nucleic acids. *Chem Biol.* 2011;18(9):1189 – 98.

68. Lonkar P, Kim K-H, Kuan JY, Chin JY, Rogers FA, Knauert MP, Kole R, Nielsen PE, Glazer PM. Targeted correction of a thalassemia-associated β -globin mutation induced by pseudocomplementary peptide nucleic acids. *Nucleic Acids Res.* 2009;37(11):3635 – 44.

69. Chin JY, Kuan JY, Lonkar PS, Krause DS, Seidman MM, Peterson KR, Nielsen PE, Kole R, Glazer PM. Correction of a splice-site mutation in the β -globin gene stimulated by triplexforming peptide nucleic acids. *Proc Natl Acad Sci U S A.* 2008;105(36):13514 – 9. S13514/1-S13514/7.

70. Egholm M, Christensen L, Dueholm KL, Buchardt O, Coull J, Nielsen PE. Efficient pH independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA. *Nucleic Acids Res.* 1995;23(2):217 – 22.

71. Nielsen PE, Egholm M, Buchardt O. Evidence for (PNA)₂/DNA triplex structure upon binding of PNA to dsDNA by strand displacement. *J Mol Recognit.* 1994;7(3):165 – 70.

72. Faruqi AF, Egholm M, Glazer PM. Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells. *Proc Natl Acad Sci U S A.* 1998;95(4):1398 – 403.

73. Rogers FA, Vasquez KM, Egholm M, Glazer PM. Site-directed recombination via bifunctional PNA-DNA conjugates. *Proc Natl Acad Sci U S A.* 2002;99(26):16695 – 700.

74. Wang G, Xu X, Pace B, Dean DA, Glazer PM, Chan P, Goodman SR, Shokolenko I. Peptide

- nucleic acid (PNA) binding-mediated induction of human γ -globin gene expression. *Nucleic Acids Res.* 1999;27(13):2806 – 13.
75. Kim K-H, Nielsen PE, Glazer PM. Site-Specific Gene Modification by PNAs Conjugated to Psoralen. *Biochemistry.* 2006;45(1):314 – 23.
76. Kim K-H, Nielsen PE, Glazer PM. Site-directed gene mutation at mixed sequence targets by psoralen-conjugated pseudo-complementary peptide nucleic acids. *Nucleic Acids Res.* 2007;35(22):7604 – 13.
77. Rogers FA, Lin SS, Hegan DC, Krause DS, Glazer PM. Targeted Gene Modification of Hematopoietic Progenitor Cells in Mice Following Systemic Administration of a PNA-peptide Conjugate. *Mol Ther.* 2012;20(1):109 – 18.
78. McNeer NA, Chin JY, Schleifman EB, Fields RJ, Glazer PM, Saltzman WM. Nanoparticles Deliver Triplex-forming PNAs for Site-specific Genomic Recombination in CD34+ Human Hematopoietic Progenitors. *Mol Ther.* 2011;19(1):172 – 80.
79. Blum JS, Saltzman WM. High loading efficiency and tunable release of plasmid DNA encapsulated in submicron particles fabricated from PLGA conjugated with poly-L-lysine. *J Control Release.* 2008;129(1):66 – 72.
80. Bahal R, Quijano E, McNeer NA, Liu Y, Bhunia DC, Lopez-Giraldez F, Fields RJ, Saltzman WM, Ly DH, Glazer PM. Single-Stranded γ PNAs for In Vivo Site-Specific Genome Editing via Watson-Crick Recognition. *Curr Gene Ther.* 2014;14(5):331 – 42.
81. Rapireddy S, Bahal R, Ly DH. Strand Invasion of Mixed-Sequence, Double-Helical B-DNA by γ -Peptide Nucleic Acids Containing G-Clamp Nucleobases under Physiological Conditions. *Biochemistry.* 2011;50(19):3913 – 8.
82. Dragulescu-Andrasi A, Rapireddy S, Frezza BM, Gayathri C, Gil RR, Ly DH. A Simple γ -Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure. *J Am Chem Soc.* 2006;128(31):10258 – 67.
83. He G, Rapireddy S, Bahal R, Sahu B, Ly DH. Strand Invasion of Extended, Mixed-Sequence B-DNA by γ PNAs. *J Am Chem Soc.* 2009;131(34):12088 – 90.
84. Sahu B, Sacui I, Rapireddy S, Zanotti KJ, Bahal R, Armitage BA, Ly DH. Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing γ -peptide nucleic acids with superior hybridization properties and water solubility. *J Org Chem.* 2011;76(14):5614 – 27.
85. Bahal R, Sahu B, Rapireddy S, Lee C-M, Ly DH. Sequence-unrestricted, Watson-Crick recognition of double helical B-DNA by (R)-MiniPEG- γ PNAs. *Chembiochem.* 2012;13(1):56 – 60
86. Svasti S, Suwanmanee T, Fucharoen S, Moulton HM, Nelson MH, Maeda N, Smithies O, Kole R. RNA repair restores hemoglobin expression in IVS2-654 thalassemic mice. *Proc Natl Acad Sci U S A.* 2009;106(4):1205 – 10.
87. Sazani P, Gemignani F, Kang S-H, Maier MA, Manoharan M, Persmark M, Bortner D, Kole R. Systemically delivered antisense oligomers upregulate gene expression in mouse tissues. *Nat Biotechnol.* 2002;20(12):1228 – 33.

通过适配体介导的基因靶向（AGT）

Patrick Ruff and Francesca Storici

摘要 DNA 适配体是一段具有独特二级结构和目标序列结合能力的 DNA 片段。适配体技术最近才被用于基因修复。利用适配体进行基因编辑的效率只要来自于其通用性，因为适配体能够被用于和已知的基因编辑系统进行连接。在这里我们主要介绍 DNA 适配体是如何增加外源性 DNA 的获得效率，并是如何将外源性 DNA 分子整合到已知的染色体位点中。虽然相对于其它的基因编辑系统，这一技术仍然处于早期，但是适配体介导的基因编辑（AGT）为基因编辑提供了新的领域和方向。

关键词 适配体介导的基因编辑（AGT）• 适配体 • I-SceI • 基因修复 • 基因编辑 • DNA 寡核苷酸 • 外源性 DNA • 酵母 • 人类细胞

P. Ruff, Ph.D.

Department of Microbiology and Immunology, Columbia University Medical Center,
New York, NY, USA

F. Storici, Ph.D. (*)

School of Biology, Georgia Institute of Technology,
310 Ferst Drive, Atlanta, GA 30332, USA

e-mail: storici@gatech.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_7

简介

DNA 双链损伤 (DSBs) 是 DNA 损伤的一种形式, 如果修复发生问题, 会造成突变、染色体重组甚至是细胞死亡。细胞必须对这种损伤进行修复, 即使是这种修复有可能造成损伤位点的突变。利用细胞修复的这一机制, 具有点特异性的归巢内切酶能够被用于作为诱导基因突变产生的工具。这一过程需要能够产生 DSB 的内切酶和一个可以作为 DSB 修复模板的外源性 DNA 片段。在被编辑位点诱导 DSB 的产生能够将基因编辑效率提高数个数量级, 在细菌[1]、酵母[2]、植物[3]、果蝇[4]、小鼠[5]、人类胚胎干细胞[6]和其他类型细胞中被证明具有很高的效率。由于在自然状态下的归巢内切酶只能靶向于特定的识别区域, 因此对于研究者来说如何找到方法在其它目标位点造成 DSBs 就十分重要。很多研究策略利用自然界已有的核酸内切酶, 如 Cre[7]和 Flp[8], 并与具有 DNA 识别功能的蛋白融合, 如 FokI (ZFNs[9,10]和随后的 TALENs[11,12]) 以及最近发现的 RNA 介导的短促回文序列簇 (CRISPR) /Cas9 系统[13]。产生点特异性 DSBs 的优势十分明显, 而相关信息可以查阅本书的其它章节来理解这些基因编辑工具的工作原理及其潜在的脱靶效应。

虽然点特异性 DSBs 具有提高编辑效率的优势, 但是对于其它的基因靶向的必要构建也需要获得重视: 如进行必要性整合的外源性 DNA 片段。为了解决外源性 DNA 获取的问题, 我们开发了一种全新的基因靶向方法, 也就是适配体接到的基因编辑——AGT, 利用该技术我们能够将归巢内切酶 I-SceI 和外源性 DNA 通过适配体进行融合, 对 I-SceI 切除位点旁边的外源 DNA 的同源部分进行靶向[14]。DNA 适配体是一段 DNA 序列, 具有独特的二级结构, 能够和相应的目标片段高亲和力连接[15]。通过引导外源性 DNA 整合到基因编辑位点附近, AGT 能够提高基因的编辑效率。通过将外源性 DNA 和 I-SceI 相连, 能够对 18-bp 长度的片段进行识别并产生 DSB[16], 我们成功在酵母和人类细胞中实现了外源性 DNA 在 I-SceI 造成的 DSB 位点的基因整合, 促进了由外源性 DNA 介导的基因修复的进行[14]。与 I-SceI 蛋白进行连接的 DNA 适配体能够通过配体指数富集电泳 (CE-SELEX), 也称为“非 SELEX” [17]进行选择。通过利用具有 I-SceI 适配体片段和进行 I-SceI DSB 靶向基因修复的同源片段的寡聚去氧核苷酸 DNA (oligos), 我们能够将酵母中的基因整合效率相较于非整合条件提高 32 倍, 将人类细胞中的基因整合效率提高 16 倍[14]。我们的策略为提高基因靶向效率提供了新的方法, 也是第一个成功利用适配体进行基因修复的研究。AGT 工作的一般性模型见图 1。

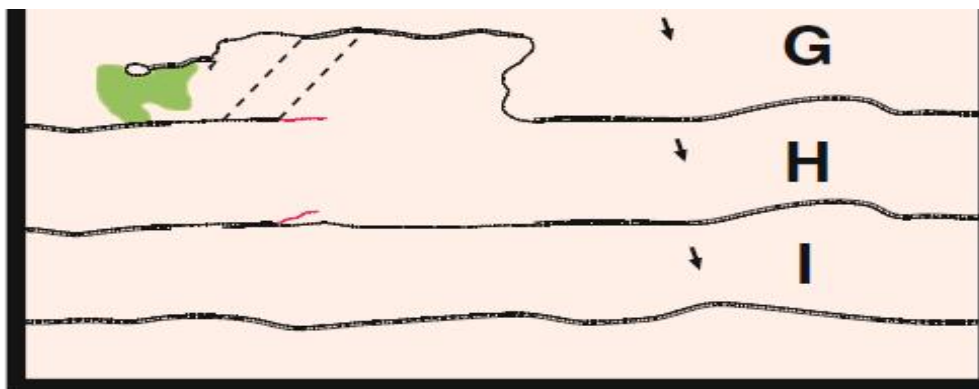
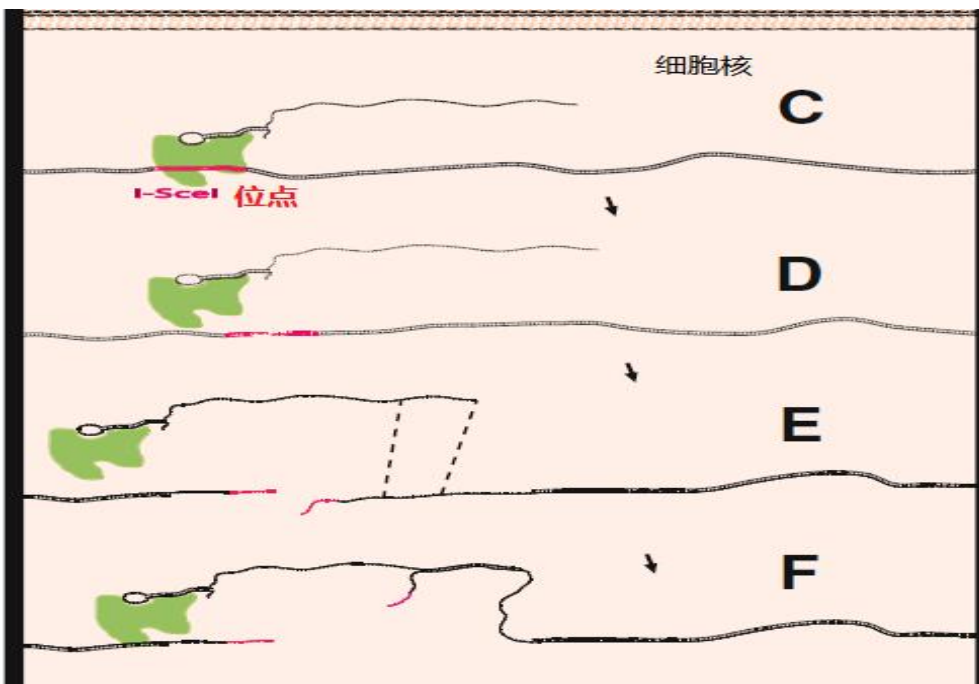
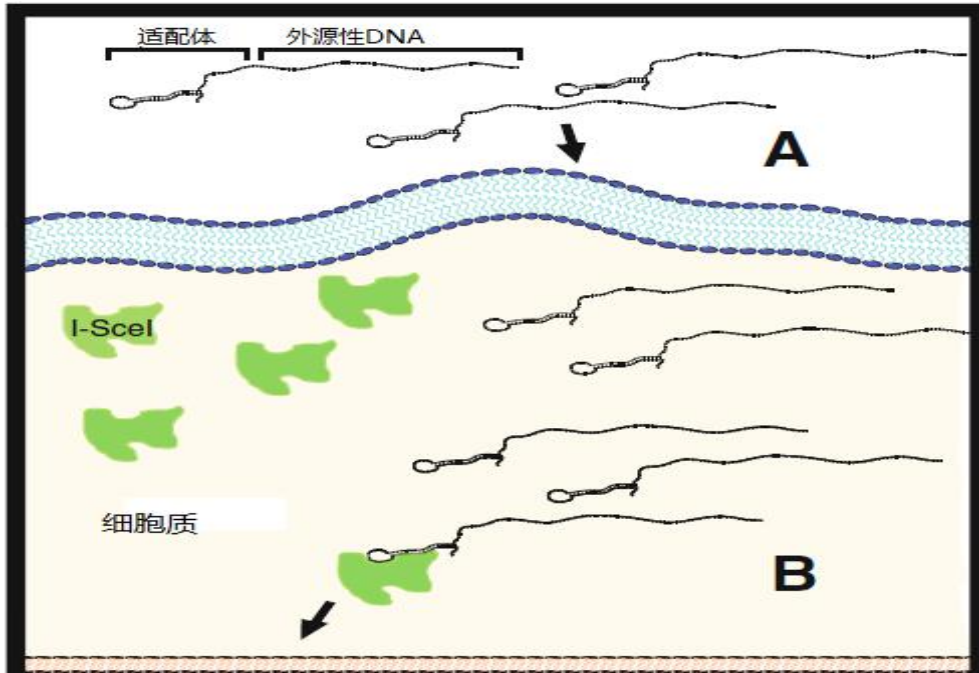


图1. 适配体介导的基因重组模型。 (a) 具有生物靶向性的寡聚功能性片段的5'端连接有A7适配体和对细胞目的缺陷基因进行功能性修复的同源片段。I-SceI核酸内切酶或是由酵母细胞染色体产生,或是由功能表达性质粒产生(人类细胞)。(b) A7适配体随后与I-SceI蛋白相连,或在细胞质(如图)或细胞核中结合。(c) I-SceI引导两种功能的寡聚片段到达I-SceI的切割位点,并(d)在I-SceI识别位点造成双链损伤。(e)切除DSB的5'端产生3'尾端结构。寡聚双功能片段的3'部分与损伤的互补片段发生退火互补连接,并将非同源片段切除。(g)在模板结合处继续DNA合成。(h)在寡聚链解螺旋后,在延伸链的3'端和DSB的另一3'端进行第二步退火反应。(i)进一步连接,DNA空隙填充及随后目标位点的完整性修复。图片来源于Ruff等人,2014[14]

AGT 起源

在 I-SceI-DSB 酵母系统中, Storici 等人[18]发现即使是很小同源片段的 DNA 寡聚片段能够有效修复 DSB。这些寡聚片段所诱导产生的修复被认为是经过两步依赖于 Rad52 的退火反应,这一基因整合过程被认为和 Rad51 整合酶无关[18]。实际上, RAD51 的删除能够提高寡聚片段的整合效率数倍,因为在 G2 期,它通过未切割的姊妹染色体抑制了 DSB 修复而且并不同时性诱导两条染色单体上整合的发生[18]。因此,未发生切割的姊妹染色体是寡聚体的重要竞争者,因为它距离切割位点更近。相似的情况发生在人类细胞 SMC1 的敲除,基因靶向效率的提高也被认为姊妹染色体之间 HR 效率的下降[19]。然而在这种情况下,姊妹染色体之间发生 HR 修复仍然具有可能,但是在缺少 hSMC1 (部分黏联蛋白复合物)的情况下,姊妹染色体之间的修复不会发生。

有研究发现在酵母中通过提高靶向拷贝基因修复效率或改变 DSB 位点和外源片段之间距离(从 20kb 到 200bp),外源片段进行 HR 修复的频率有所提高[21]。最近, Renkawitz 等人[22]利用酵母中的时间分辨染色体免疫沉淀反应,发现 HR 更容易发生在修复模板距离 DSB 位点较近的情况下。与此类似,在小鼠中,DSB 发生后,基因易位主要发生在损伤位点附近(也可发生在损伤远端,但是效率更低)[23]。在 VDJ 重组情况下,RAG 重组酶能够拉近 T 细胞受体 $\alpha \delta$ (Tcra/d) 和 Igh 位点之间的距离,并有助于这些位点之间的重组[24]。

基于高效的 HR 外源片段的开发,我们提出了一种全新的方法,利用适配体也就是 AGT 将外源 DNA 引入 DSB 位点附近[14]。下一部分将对适配体进行一个简单的回顾。

适配体

核酸适配体是一种短片段的单链 DNA 或 RNA 寡聚体,其二级结构与配体(蛋白、小分子或活细胞)之间具有高亲和力。大多数 DNA 和 RNA 都具有二级结构,然而很少有片段具有和特异性靶标连接的亲和力。然而适配体不但具有亲和力,而且具有高度的特异性,例如有一种适配体只能和茶碱相连而无法和咖啡因相连,而两者之间的差异只有一个甲基[25]。虽然有时候适配体会被认为是一种人工合成的抗体,但是相对于抗体适配体具有很多优势,包括生产的简便性和低成本,且对于动物影响很小。适配体相对于抗体免疫原性弱,很少被直接用于人类的疾病治疗[26]。

适配体的产生经过了严格的筛选,适配体来源于随机的 DNA 或 RNA 池,其中只有很少一部分序列(几乎不存在)具有和靶向起始序列结合的活性(通常是 10^{14} 或更高)。经典的随机文库经常和复杂的引物端相连,其结构主要为 5'-引物段 1-N₂₀₋₆₀-引物段 2(反向互补序列)-3',其中 N 是随机性碱基[28]。引物被用于进行文库序列的 PCR 扩增。适配体的体外筛选过程由 Szostak[29]和 Gold[30]课题组于 1990 年开发,而这一过程也随着配体的系统性升级发展为指数富集法(SELEX)。SELEX 实验首先将 DNA/RNA 序列的随机性文库与目标序

列进行孵育, 随后将未结合的片段除去, 之后将结合片段洗脱下来并对文库序列进行扩增(见图 2 所示简图)。几年之后, 发展出多种 SELEX 方法。有一种 SELEX 利用毛细管电泳 (CE) 使得 SELEX 分选的效率更高, 不但节省了时间也阻止了适配体和配体之间的结合 (配体在缓冲液中处于游离状态)。CE-SELEX 利用少于五次的最少循环次数的分选[31], 能够分选出高亲和力的适配体, 而经典的 SELEX 则需要不少于十次的循环分选。由 CE-SELEX 分选产生的适配体能够在 nM 甚至是 pM 水平的解离常数。而 CE-SELEX 技术带给我们的简便性和高效率能够为我们更加简便的选择与 I-SceI 核酸内切酶相连的适配体提供帮助。

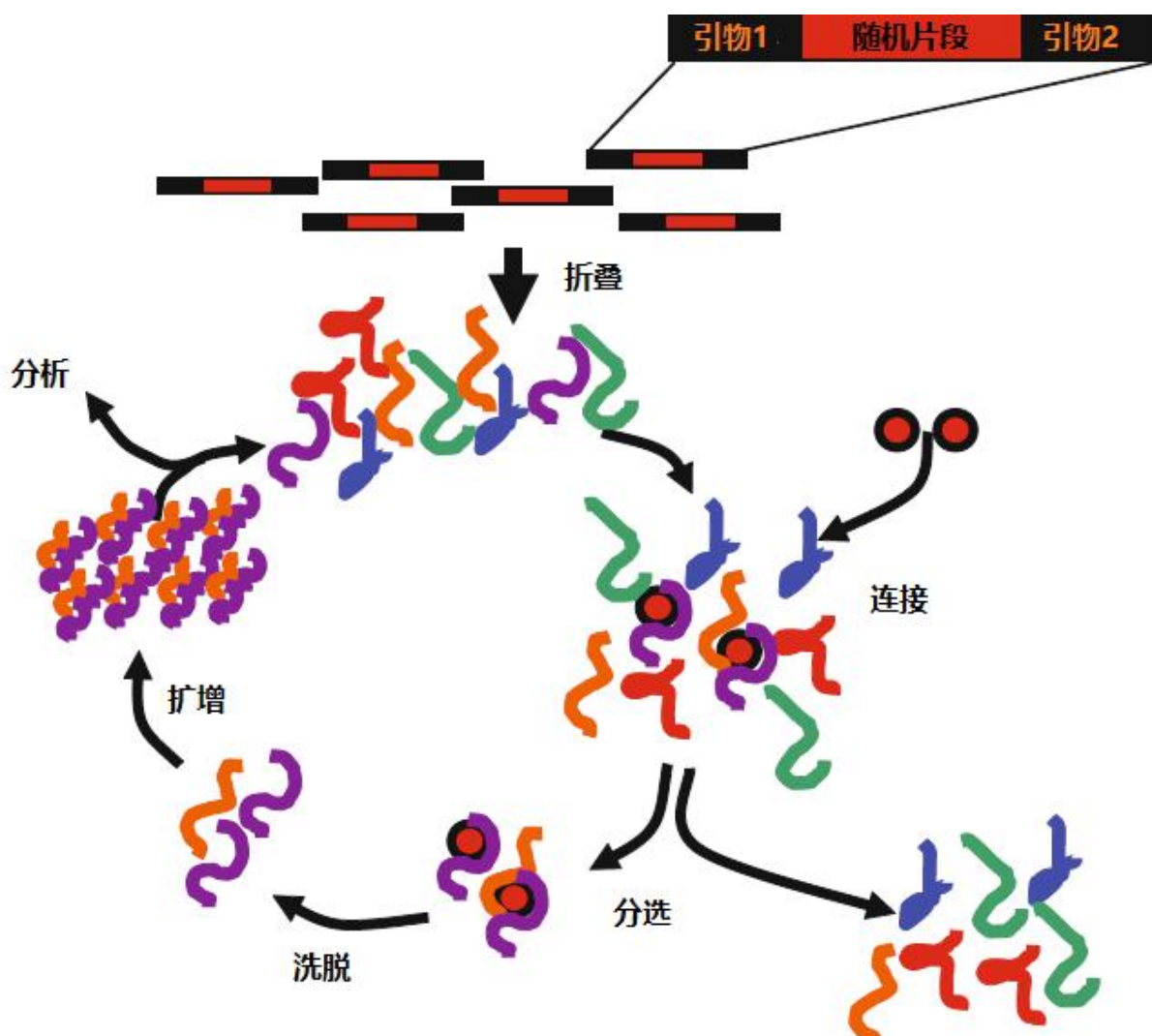


图2.配体系统性分选 EX潜在富集 (SELEX) 过程。目标配体, 以红色展示, 孵育的随机性 ssDNA 分子双链文库在确定性随机片段两侧具有已知的引物序列, 能够被用于 PCR 扩增。序列形成的二级结构, 用紫色和橘色的 DNA 片段与配体连接的形式展示。这些片段通过连接键的强弱进行分选并洗脱。在对这些连接片段进行扩增后, 重复进行这一过程。

点特异性核酸内切酶 I-SceI

DNA 连接蛋白几乎存在于所有形式的生命中，但是在其中只有很少一部分蛋白具有对双链 DNA 的特异性位点进行识别并切割的作用活性。能够造成 DNA 双链损伤的核酸内切酶成为归巢内切酶，而且这些酶通常对于长片段的识别片段（12-40bp）具有更高的特异性[33]。归巢内切酶的研究自 1970s 年代开始，最先被研究的归巢内切酶最初被称为“Omega”，随后被命名为 I-SceI[34]。在自然状态下，I-SceI 内切酶能识别酵母线粒体内含子中的 5'TAG GGA TAA CAG GGT AAT 3'并在该片段中产生 DNA DSB，并通过内含子整合和基因转换介导等位基因之间发生整合和覆盖[35]。由于这一发现，I-SceI 被用于用于从细菌到人类细胞等所有模型系统诱导基因产生 DSB 损伤和修复。由于 I-SceI 的广泛性应用，我们选择 I-SceI 作为 AGT 连接的点特异性核酸酶。

作为外源性 DNA 的单链寡聚 DNA

合成的单链寡聚体 DNA 是短片段的 DNA，通常具有 90nt 或更短。寡核苷酸通过同源重组将突变导入到目标序列中。寡核苷酸能够被化学合成且成本低廉，相较于长片段具有相似的效率，包括外源性质粒或 PCR 产物[36]。通过寡聚核苷酸进行同源重组基因修复的片段甚至可以小于 30 个核苷酸[14,37]。在很多细胞系中寡核苷酸都被成功的运用进行高效的基因修复[38,39]。对于 AGT，我们利用两部分构成的寡核苷酸，一部分与 I-SceI 相连（适配体区域），另一部分为 DSB 进行同源修复的模板并对目标序列进行选择性的编辑[14]。

AGT 的特征

在 AGT 系统设计完成且经过了 I-SceI 的适配体筛选后，有一些片段被发现能够在体外与 I-SceI 连接。在其中，有两种片段结合能力最强，称为适配体 A3 和适配体 A7，最早于啤酒酵母的系统检测中得到。第一个进行该技术靶向的目标序列是酵母中的 trp5 基因，实验在该基因中整合入一段功能性片段并测量 Trp⁺的转录数来表征编辑效率。Trp5 位点功能域整合的简图见图 3A。进行 AGT 设计的第一个步骤是决定适配体对于外源性 DNA 寡核苷酸的位置导向。早期研究发现寡核苷酸片段的 5'端相较于 3'端具有更高的效率，其原因很可能是下游 3'端的外来同源序列能够促进基因整合的发生。

进行 AGT 设计的第二步是对和适配体区域连接的引物区域进行检测。如上所述，在适配体筛选中，每一个随机性 DNA 序列都连接有两条复杂的引物区域用于进行 PCR 扩增（图 2），而这些引物本身对于适配体与靶标的连接没有任何用处[40-42]。基于这一点，我们移除了分选后 I-SceI 适配体中的引物区域并同时提高了靶向 DNA 的同源长度。而且，在 trp5 位点中，适配体 A7 和适配体 A4 和外源性片段连接成为 A7.TRP5.54 和 A4.TRP5.54，在 trp5 的 5'端进行 54 个同源碱基的整合。而且，还进行了同样长度的适配体对照试验，对照组不与 I-SceI 相连，称为 C.TRP5.54。适配体 A4 连接力较弱，和对照组之间没有显著性差别，然而 A7 的连接力较强，相对于对照组编辑效率稳定提高七倍[14]。在 I-SceI 进行表达但是不存在 I-SceI 识别位点的情况下或存在 I-SceI 识别位点但是不存在 I-SceI 表达的情况下，或是在 GAL1-10 引起的葡萄糖阻遏的作用下，A7.TRP5.54 和 C.TRP5.54 之间均不存在基因编辑效率的显著性差别[14]。

在明确了 A7 适配体能够提高基因靶向效率之后，有假设认为当同源片段更短或更长时能够得到更为理想的整合效率，因为更长的同源序列有可能对 I-SceI 和适配体之间的连接产生影响。又或者由更短序列的外源性 DNA 介导的适配体连接整合更为精确。为了检测 I-SceI 适配体是否与更短序列的外源性 DNA 结合后可以产生更高的整合效率，将寡核苷酸的同源性长度减少到 40 个碱基合成了新的 A7.TRP5.40 和 C.TRP5.40。新合成的两种适配体的基因编

辑效率均下降，而 A7 适配体介导的基因编辑效率相较于对照组提高了 9.2 倍（图 3b）。有研究称结合于 5'端的适配体分子，在寡核苷酸不纯净的条件下可能并不完整。确实，不纯净的 100-mer 的寡核苷酸链可能只存在约 60%的完整片段，而其余 40%的片段 5'端可能被截短 [43]。这也意味着 5'端截短后的 A7.TRP5.40 寡核苷酸有可能在基因靶向位点具有更高效率的整合。为验证这一说法，采用了聚丙烯酰胺凝胶电泳（PAGE）纯化过的寡核苷酸，检测了 54 个碱基的同源片段和 40 个碱基的同源片段的寡核苷酸的长度。对于两种 PAGE-纯化后的寡核苷酸片段来说，基因整合频率稳定性提高（纯化后的 A7.TRP5.54 相较于 C.TRP5.54 编辑效率提升 27 倍，A7.TRP5.40 相较于 C.TRP5.40 编辑效率提升 30 倍）[14]。这一结果显示适配体对于基因的靶向具有重要的作用；同源片段的长度也有一定的影响，而对短片段的寡核苷酸片段适配体的作用更为重要。

AGT 不只是在酵母染色体的 *trp5* 基因中进行了应用，也在酵母的 *ade2* 和 *leu2* 基因位点和人源胚肾细胞中（HEK-293）质粒上的 DsRed2（一种红色荧光蛋白）基因位点进行了应用。在每次应用中，A7 适配体连接的 54 碱基同源寡核苷酸序列的基因编辑效率都得到了显著提高（图 3B，图 4A 和 [14]）。在这些实验中寡核苷酸片段的序列也不尽相同。对于 DsRed2 位点，A.Red.40（A7 适配体连接的 40 碱基同源片段）具有 6.2 倍效率的提高，而 A.Red.30 寡核苷酸具有 16 倍的提高（图 4A）。在不同的编辑位点 AGT 的不同可能主要是由于寡核苷酸二级结构的不同造成的。为了进行 I-SceI 连接，适配体区域必须具有特定的二级结构而且使长片段 ssDNA 同源区段有可能和这一结构发生相互作用。有关于进一步适配体二级结构的研究，请参见 Ruff 等人的研究 [14]。

为了进一步研究在 I-SceI 存在的情况下能够刺激特定基因片段进行整合的 I-SceI 适配体 AGT，设计利用不同于 I-SceI 的另一种核酸酶的对照试验（实验利用酵母进行）。在酵母中，归巢内切酶 HO 被用来代替 I-SceI 在 *trp5* 位点产生 DSB（图 3A）。HO 相较于 I-SceI 产生 DSB 的效率更高，特异性也更好 [44]。然而，由适配体连接的外源寡核苷酸 A7.TRP5.40 介导基因修复只有在 I-SceI 诱导的条件下发生的频率更高（图 3B）。与之相似，在人类细胞中具有 DsRed2 基因的质粒在体外用 I-SceI 消化过夜。由于 HEK-293 细胞中没有 I-SceI 的表达，A.Red.40 和 A.Red.30 寡核苷酸之间没有显著性差异，A.Red.54 寡核苷酸效率只有 1.8 倍的提高（图 4B）。这些结果显示 AGT 能够提高外源性 DNA 在特定目的基因位点的整合能力。

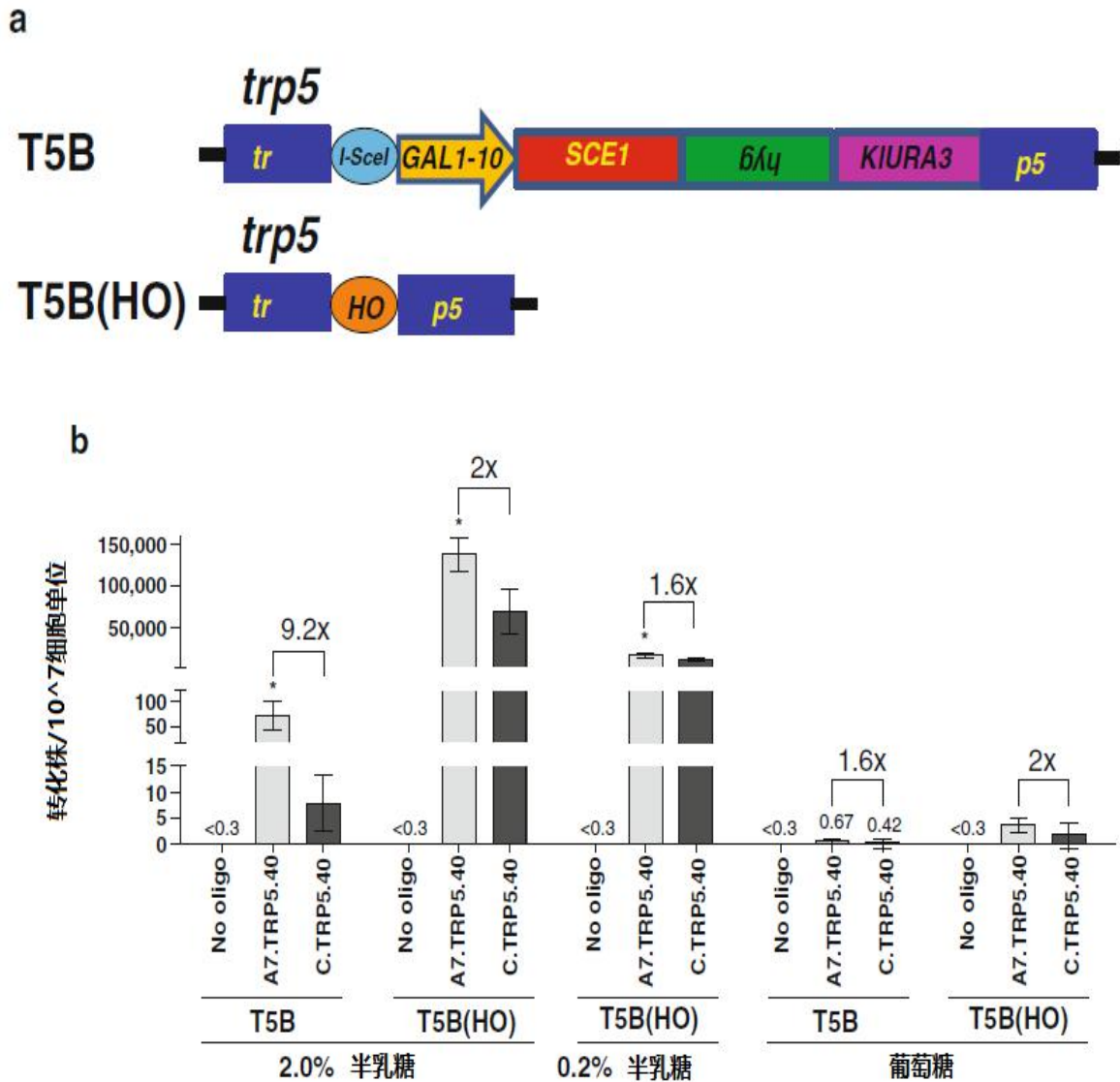


图3, AGT刺激酵母中的基因靶向。(a) 酵母靶向位点简图。FRO-155/156链, 如T5B所示, 将I-SceI损伤位点(蓝色椭圆)、处于GAL1-10启动子下游的I-SceI功能域SCE1、潮霉素抗性基因hyg以及克隆选择性的K.I. URA3基因插入到已知的TRP5基因中。链HK-225/226, 如图T5B(HO)所示, 将HO损伤位点(记有HO字样的橘色椭圆)插入到TRP5基因中。(b) 利用具有适配体的寡核苷酸诱导的酵母中基因修复频率用亮灰色进行标识, 而无适配体连接的对照组用暗灰色标识(X轴), 转化细胞数按每10⁷个细胞进行计算(Y轴), 以无寡聚链组作为对照基准值。基因重组发生在I-SceI链(T5B)或HO链(T5B(HO))在半乳糖存在的条件下以及葡萄糖抑制条件下的trp5基因位点。平均值的矫正值、置信区间在95%之内。星号表示适配体连接组的寡核苷酸相对于无适配体连接组具有显著性差异(*代表p<0.05), 基因修复的频率倍数如图。本图的部分数据来自于Ruff等人的研究[14]。

利用 AGT 进行基因编辑

如果存在一种 DNA 序列能够同时具有识别结合靶向序列及进行目标序列的双重修复功能，那么将在基因修复过程中取得优势。AGT 利用的 DNA 适配体，自 1990s 早期以来几乎还是新生事物[29,30]。适配体可以用于生物传感器[45,46]和临床治疗[47]，但还几乎未涉足基因靶向领域。我们的适配体能够连接 I-SceI，能够对特定的基因位点进行靶向。这目前还是一个全新的基因靶向策略，还需要更多的实例进行验证。

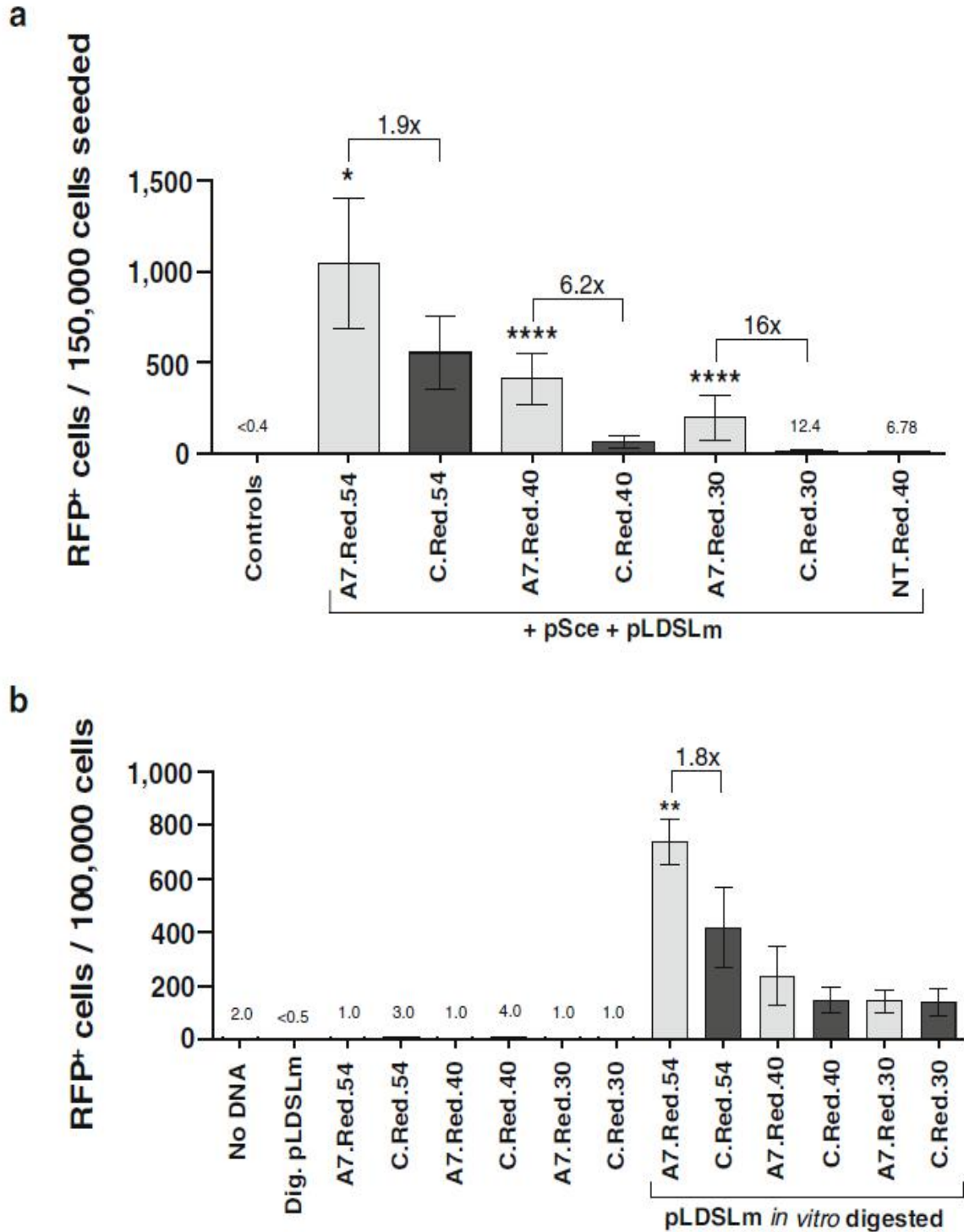


图4. 在人类细胞中AGT刺激基因靶向。利用I-SceI表达载体 (pSce) 顺转HEK-293细胞, 载体 (pLDSLm) 具有双终止子导入失活的DsRed2基因和I-SceI识别位点和寡核苷酸序列。阴性对照为HEK-293细胞本身 (没有DNA, 只加入转染试剂), 和只加入pSce, 只加入pLDSLm, 以及只加入寡核苷酸片段。 (a) 流式细胞术统计不同长度寡核苷酸链诱导的细胞背景中RFP+的细胞数 (见[14])。X轴中不同样本所引起的15, 0000细胞背景中RFP+细胞数目的差异见Y轴。阴性对照未发现任何RFP+细胞。 (b) 在体外对pLDSLm载体转染的细胞进行流式分析。X轴所示的不同的样本引起的每10, 0000单位细胞中RFP+细胞数的差异显示在Y轴中。阴性对照包括独立细胞组 (无DNA)、寡核苷酸独立组。消化后载体和寡核苷酸片段的联合转染是不尽人意的, 平均值趋势一致, 置信区间小于95%。星号表示具有适配体的寡核苷酸组和没有适配体连接的非对接组具有显著性差异 (*代表 $p < 0.05$, **代表 $p < 0.01$, ***代表 $p < 0.001$)。差异倍数如图所示, 数据来源于Ruff等人的研究, 2014[14]。

AGT 的未来发展

这里所讨论的 AGT 系统为利用适配体进行基因靶向提供了新的原理思路, 基于早期的工作, 这一技术未来还需要进行进一步的阐释、证实和系统提升。例如, 不同的单链同源序列有可能会破坏适配体的二级结构并抑制其与靶位的连接, 并降低 AGT 的效率。而且, 适配体的亲和力还有继续优化的空间。除此之外, 目前的 AGT 主要是用于 I-SceI 核酸酶的连接。

有很多潜在的方法可以提升 AGT 系统。例如, 减少适配体区域和同源区段之间可能发生的退火反应, 外源性区段可以利用双链代替单链。碱基之间的连接序列可以是一段既非适配体片段也非同源区段的 DNA 序列, 或者可以是硫化物连接[48]或者碳化连接[49]。而且, 适配体本身也可进行改进增强其与靶点之间的亲和力。A7 适配体和 I-SceI 之间的亲和力大约在 $3.16 \mu\text{M}$ 左右, 相较于 A4 适配体的亲和力大约提高 15 倍 ($52.49 \mu\text{M}$)。虽然 I-SceI 的适配体 A7 是从寡核苷酸混合池中筛选得到的, 但是仍然存在可能有亲和力更高的适配体并未被筛选出来。如果想要提升现有适配体的亲和力, 还可以对 A7 适配体序列进行化学修饰。从 DNA 靶向分子和寡核苷酸稳定性上来看, 无论是哪个因素制约了基因修复的效率, 都可以通过对寡核苷酸链进行化学修饰而得到解决。对于寡核苷酸的修饰, 可以利用硫代磷酸键提升外源 DNA 的稳定性从而提高基因整合的效率。一种用于 I-SceI 酶特异性的适配体 AGT, 原则上来讲也可能适用于 FokI (ZFNs 和 TALENs) 或 Cas9 (CRISPRs) 的基因靶向。AGT 提高基因靶向的能力也有可能和核酸酶无关, 但是从理论上来说其它的 DNA 连接酶有可能影响基因整合。我们期待着筛选结合力强的 DNA 连接因子的适配体能够成为外源 DNA 整合一种高效的策略并可由此提高基因编辑的效率。

致谢 我们要感谢来自乔治亚州肿瘤联合基金 R9028、美国国立卫生基金 R21EB9228、和乔治亚州技术资金研究和教育款项 GT-FIRE-1021763 的支持。

参考文献

1. Nussbaum A, Shalit M, Cohen A. Restriction-stimulated homologous recombination of plasmids by the RecE pathway of Escherichia coli. *Genetics*. 1992;130(1):37 - 49.
2. Storici F, Durham CL, Gordenin DA, Resnick MA. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. *Proc Natl Acad Sci U S A*. 2003;100(25):14994 - 9.
3. Puchta H, Dujon B, Hohn B. Homologous recombination in plant cells is enhanced by in vivo

- induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res.* 1993;21(22):5034 – 40.
4. Banga SS, Boyd JB. Oligonucleotide-directed site-specific mutagenesis in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A.* 1992;89(5):1735 – 9.
 5. Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A.* 1994;91(13):6064 – 8.
 6. Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res.* 1995;23(24):5012 – 9.
 7. Santoro SW, Schultz PG. Directed evolution of the site specificity of Cre recombinase. *Proc Natl Acad Sci U S A.* 2002;99(7):4185 – 90.
 8. Voziyanov Y, Konieczka JH, Stewart AF, Jayaram M. Stepwise manipulation of DNA specificity in Flp recombinase: progressively adapting Flp to individual and combinatorial mutations in its target site. *J Mol Biol.* 2003;326(1):65 – 76.
 9. Palant NJ, Dudzinski D. Zinc finger nucleases: looking toward translation. *Gene Ther.* 2013;20(2):121 – 7.
 10. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A.* 1996;93(3):1156 – 60.
 11. Carlson DF, Fahrenkrug SC, Hackett PB. Targeting DNA With Fingers and TALENs. *Mol Ther Nucleic Acids.* 2012;1, e3.
 12. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science.* 2009;326(5959):1501.
 13. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339(6121):819 – 23.
 14. Ruff P, Koh KD, Keskin H, Pai RB, Storici F. Aptamer-guided gene targeting in yeast and human cells. *Nucleic Acids Res.* 2014;42(7), e61.
 15. Ni X, Castanares M, Mukherjee A, Lupold SE. Nucleic acid aptamers: clinical applications and promising new horizons. *Curr Med Chem.* 2011;18(27):4206 – 14.
 16. Niu Y, Tenney K, Li H, Gimble FS. Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. *J Mol Biol.* 2008;382(1):188 – 202.
 17. Berezovski M, Musheev M, Drabovich A, Krylov SN. Non-SELEX selection of aptamers. *J Am Chem Soc.* 2006;128(5):1410 – 1.
 18. Storici F, Snipe JR, Chan GK, Gordenin DA, Resnick MA. Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing. *Mol Cell Biol.* 2006;26(20):7645 – 57.
 19. Potts PR, Porteus MH, Yu H. Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. *EMBO J.* 2006;25(14):3377 – 88.
 20. Wilson JH, Leung WY, Bosco G, Dieu D, Haber JE. The frequency of gene targeting in yeast depends on the number of target copies. *Proc Natl Acad Sci U S A.* 1994;91(1):177 – 81.
 21. Coic E, Martin J, Ryu T, Tay SY, Kondev J, Haber JE. Dynamics of homology searching during gene conversion in *Saccharomyces cerevisiae* revealed by donor competition. *Genetics.* 2011;189(4):1225 – 33.
 22. Renkawitz J, Lademann CA, Kalocsay M, Jentsch S. Monitoring homology search during

- DNA double-strand break repair in vivo. *Mol Cell*. 2013;50(2):261 – 72.
23. Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T. Spatial dynamics of chromosome translocations in living cells. *Science*. 2013;341(6146):660 – 4.
24. Rocha PP, Chaumeil J, Skok JA. Molecular biology. Finding the right partner in a 3D genome. *Science*. 2013;342(6164):1333 – 4.
25. Jenison RD, Gill SC, Pardi A, Polisky B. High-resolution molecular discrimination by RNA. *Science*. 1994;263(5152):1425 – 9.
26. Singerman LJ, Masonson H, Patel M, Adamis AP, Buggage R, Cunningham E, et al. Pegaptanib sodium for neovascular age-related macular degeneration: third-year safety results of the VEGF Inhibition Study in Ocular Neovascularisation (VISION) trial. *Br J Ophthalmol*. 2008;92(12):1606 – 11.
27. Hall B, Micheletti JM, Satya P, Ogle K, Pollard J, Ellington AD. Design, synthesis, and amplification of DNA pools for in vitro selection. *Curr Protoc Mol Biol*. 2009;Chapter 24:Unit 24.2.
28. Sabeti PC, Unrau PJ, Bartel DP. Accessing rare activities from random RNA sequences: the importance of the length of molecules in the starting pool. *Chem Biol*. 1997;4(10):767 – 74.
29. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. *Nature*. 1990;346(6287):818 – 22.
30. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. 1990;249(4968):505 – 10.
31. Berezovski M, Drabovich A, Krylova SM, Musheev M, Okhonin V, Petrov A, et al. Nonequilibrium capillary electrophoresis of equilibrium mixtures: a universal tool for development of aptamers. *J Am Chem Soc*. 2005;127(9):3165 – 71.
32. Mosing RK, Mendonsa SD, Bowser MT. Capillary electrophoresis-SELEX selection of aptamers with affinity for HIV-1 reverse transcriptase. *Anal Chem*. 2005;77(19):6107 – 12.
33. Belfort M, Roberts RJ. Homing endonucleases: keeping the house in order. *Nucleic Acids Res*. 1997;25(17):3379 – 88.
34. Stoddard BL. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure*. 2011;19(1):7 – 15.
35. Colleaux L, D'Auriol L, Galibert F, Dujon B. Recognition and cleavage site of the intron-encoded omega transposase. *Proc Natl Acad Sci U S A*. 1988;85(16):6022 – 6.
36. Wang Z, Zhou ZJ, Liu DP, Huang JD. Double-stranded break can be repaired by single-stranded oligonucleotides via the ATM/ATR pathway in mammalian cells. *Oligonucleotides*. 2008;18(1):21 – 32.
37. Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods*. 2011;8(9):753 – 5.
38. Olsen PA, Randol M, Krauss S. Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides. *Gene Ther*. 2005;12(6):546 – 51.
39. Papaioannou I, Simons JP, Owen JS. Oligonucleotide-directed gene-editing technology: mechanisms and future prospects. *Expert Opin Biol Ther*. 2012;12(3):329 – 42.
40. Cowperthwaite MC, Ellington AD. Bioinformatic analysis of the contribution of primer sequences to aptamer structures. *J Mol Evol*. 2008;67(1):95 – 102.
41. Pan W, Clawson GA. The shorter the better: reducing fixed primer regions of oligonucleotide libraries for aptamer selection. *Molecules*. 2009;14(4):1353 – 69.

42. Legiewicz M, Lozupone C, Knight R, Yarus M. Size, constant sequences, and optimal selection. *RNA*. 2005;11(11):1701 – 9.
43. Stafford P, Brun M. Three methods for optimization of cross-laboratory and cross-platform microarray expression data. *Nucleic Acids Res*. 2007;35(10), e72.
44. Plessis A, Perrin A, Haber JE, Dujon B. Site-specific recombination determined by I-SceI, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics*. 1992;130(3):451 – 60.
45. Cheng AK, Ge B, Yu HZ. Aptamer-based biosensors for label-free voltammetric detection of lysozyme. *Anal Chem*. 2007;79(14):5158 – 64.
46. Stojanovic MN, Kolpashchikov DM. Modular aptameric sensors. *J Am Chem Soc*. 2004;126(30):9266 – 70.
47. Barbas AS, Mi J, Clary BM, White RR. Aptamer applications for targeted cancer therapy. *Future Oncol*. 2010;6(7):1117 – 26.
48. Chu TC, Twu KY, Ellington AD, Levy M. Aptamer mediated siRNA delivery. *Nucleic Acids Res*. 2006;34(10), e73.
49. Zhou J, Neff CP, Swiderski P, Li H, Smith DD, Aboellail T, et al. Functional in vivo delivery of multiplexed anti-HIV-1 siRNAs via a chemically synthesized aptamer with a sticky bridge. *Mol Ther*. 2013;21(1):192 – 200.

通过双链断裂修复刺激 AAV 基因的编辑

Angela M. Mitchell , Rachel Moser , Richard Jude Samulski ,and Matthew Louis Hirsch

摘要 最近在哺乳动物基因组编辑技术的进步已经实现了在染色体水平上进行疾病相关基因的精确操作。事实上，锌指核酸酶（ZFNs）能够在体外诱导患者 T 细胞 HIV CCR5 受体的翘楚并在临床上促进细胞融合的发生。在这些应用中，腺病毒递呈载体中的一类腺相关病毒（AAV）基因递呈已经进行过了 100 次以上的临床基因治疗试验且并未发现任何载体相关毒性，目前的应用已经变得非常普遍。这一章回顾了 AAV 在临床基因治疗中的发展，AAV 基因的早期实验结果，目前诱导 DNA 双链损伤点特异性修复的应用。以及在 AAV 诱导的载体整合和点特异性核酸酶编辑中仍存在的障碍。

关键词 AAV • 基因治疗 • 锌指核酸酶 • Cas9/CRISPR • TALEN • DNA 修复 • 同源重组 • 基因编辑 • 基因靶向

A. M. Mitchell , Ph.D.

Gene Therapy Center , University of North Carolina at Chapel Hill , Chapel Hill , NC , USA

e-mail: angelamm@princeton.edu

R. Moser , B.A.

Department of Ophthalmology , University of North Carolina at Chapel Hill ,

Chapel Hill , NC , USA

e-mail: rmoser@med.unc.edu

M. L. Hirsch , Ph.D. (*)

Gene Therapy Center , University of North Carolina at Chapel Hill , Chapel Hill , NC , USA

Department of Ophthalmology , University of North Carolina at Chapel Hill ,

Chapel Hill , NC , USA

e-mail: mhirsch@email.unc.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental

Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_8

腺相关病毒的载体化

腺相关病毒（AAV）发现于 1965 年，是一种发现于腺体中的污染病毒并被认定为一类新品种的病毒，能够依赖于腺病毒完成自己的复制周期[1]。AAV 是细小病毒家族中的成员之一，具有一条附着在衣壳蛋白上的单链 DNA，其也可被进一步被分类为依赖病毒属，因为其无法在辅助病毒缺失的情况下独立完成复制[2]。AAV 感染非常常见，很多人在非常年轻时就具有 AAV 血清阳性[3]，然而，AAV 并不具有致病性。

野生的 AAV 基因组有 4.7kb，至少具有两种编码基因，Rep 和 Cap，并连接有末端反向重复序列（ITRs），是病毒组件中唯一需要顺式折叠的基因组[4]。Rep 基因编码四种没有空间结构的 Rep 蛋白，这些蛋白对于 AAV 基因组的复制和包装非常重要[5,6]，由于 Cap 基因能够编码的 VP1、VP2 和 VP3 蛋白能够组成十二面体衣壳，包装活性蛋白（AAP）对于十二面体的组装和基因组的包装非常重要[7]。有多种的 AAV 血清型被发现，其中有 49%到 99% 的病毒能够通过它们的衣壳氨基酸序列进行鉴别[8]而且它们的一些生物性能也不尽相同，例如生物趋向性和转导动力学[9,10]。对于基因治疗用途的载体，病毒基因被移除并置换成转基因功能域[4]。重组性 AAV（rAAV）载体的产物和特征在之前的综述中已经进行了很详细的描述[11]。

AAV 通过与主要的多糖受体连接起始转录途径，例如硫酸乙酰肝素，或者与第二受体作用，如生长因子受体或整合素。AAV 能够结合的受体类型取决于血清型和被转染的细胞类型（更详细的回溯见[12]）。AAV 随后通过受体介导的内吞进入细胞[13]并通过溶酶体途径到达细胞核周围微管组织中心[14]。在核小体形成前和形成过程中病毒颗粒都停留在这一区域以逃避磷脂酶对 VP1 N 末端区域的降解[14-16]。在被核内体释放出来后，完整的 AAV 颗粒会进入解聚化的细胞核中[17]。由于 AAV 是单链病毒，第二链的合成必须发生在基因组转录激活之前，而 AAV 是依赖于细胞的聚合酶完成第二链的合成的[18]。在辅助病毒缺失的情况下，野生型基因组的转录受到 Rep 蛋白的抑制，能够允许病毒进行潜伏直到辅助病毒能够促进 AAV 的生命周期的完成[19]。对于 rAAV，能够利用持续性激活的启动子在第二链合成后表达外源性导入的基因功能域。

AAV 基因组的游离基因能够形成游离单体，而串联体能够长时间的存在，虽然他们无法复制[20-22]。除此之外，AAV2 基因组能够点特异性地整合到 19 染色体 AAVS1 位点，该过程具有 Rep 依赖性[23,24]。然而，在 Rep 缺失的情况下，只有低水平的（小于 0.5%）的不规则集成的发生[25-27]。这一整合经常能够影响 ITR 和宿主 DNA 之间的关系并能够导致宿主 DNA 小片段的删除及载体基因组中片段的插入[28]。虽然这一整合作用被发现能够发生于全基因组，目前的研究发现基因、转录起始位点、核糖体 DNA 序列、CpG 岛和回文序列是 rAAV 整合位点中的热点[27,28]。除了整合的低水平之外，很多研究并未发现 rAAV 具有致癌作用[25,26]，然而也有研究发现转基因和特异性的启动子具有低水平的致癌性[25-27]。

通过对 rAAV 转化途径的深入了解，也发明了多种策略利用 rAAV 的生物特性产生更高效率的载体，其中效率最高的是自我互补的 rAAV[29,30]。自我互补的 rAAV 的发展起源于对 rAAV 转化的研究，研究发现 rAAV 的转化水平在腺病毒进行共转染时有所提高，即使是排除了腺病毒作为辅助病毒的功能（例如转录激活），rAAV 转化也有提高[18]。腺病毒共转染造成的转化效率的提高一部分是由于 E4Orf6 蛋白造成的第二链 DNA 合成增加引起的[18]，也说明第二链的合成是 rAAV 转化的限速步骤，如果第二链合成的速率能够提高，则转化效率也会提高。

为了避免第二链合成速率的限制，载体的尺寸设计少于 rAAV 包装能力的一半，使得单链复制形成的二聚体转基因功能域能够被折叠[29]。当从病毒衣壳被释放出来时，这些互补的二聚体能够进行链内互补配对激活转录，从而对单链 DNA 进行复制。当在一个 ITR 的 Rep

基因处造成损伤（Rep 缺刻的 D 部分）引起该基因的删除能够使得这一过程更具有效率，并由此产生了自我互补二聚体[31]。自我互补载体能够引起体内和体外的更为快速的转化动力学和更高水平的转基因表达[31]在基因编辑应用中非常有用[32]，在临床基因治疗中具有良好的前景。

作为临床基因治疗载体的 AAV

由于具有多种优势特征，AAV 已经被发展用于临床基因治疗的载体（rAAV）。AAV 及 AAV 辅助病毒目前还未发现相关致病性方面的表征数据，提示 AAV 来源载体应该具有良好的安全性。而且，相对于其他病毒来说，AAV 的免疫原性弱，病毒接种引发的免疫相关症状和免疫反应不明显[33]。除此之外，rAAV 中所有的病毒基因都被移除，降低的病毒基因产物有可能引起的免疫应答反应。AAV 能够同时感染处于分裂期和非分裂期的细胞且能够引起转导基因的长时程表达[3,34,35]。而且，与 AAV2 ITRs 相连的转基因功能域能够被多种血清亚型的 AAV 病毒衣壳包装，该过程被称为衣壳转导[36,37]，在这个过程中载体产物可以具有高度异质性的生物趋向性和完全不同的免疫特征。多种不同血清亚型的 rAAV 的转化能力已经得到了详细的研究，使得在不同应用中对于 rAAV 血清亚型的选择更加合理[9,10,38]。基于这些特征，rAAV 已经被用于超过 100 项临床实验（<http://www.abedia.com/wiley>），目前已有多项实验能够证明 rAAV 介导的基因编辑具有安全性[39]。这些实验所涉及的疾病主要包括单基因突变疾病，例如血友病[40,41]、中枢神经系统疾病[42]和心脏疾病[43]，也体现出 AAV 基因治疗具有广泛的应用前景。

最近，有很多研究证实了 rAAV 作为一种临床载体在提高治疗成功率方面具有潜力。事实上，第一款基因治疗产品首先在欧洲获得了应用审批，Glybera[®]利用 rAAV 转导脂蛋白脂肪酶（LPL）基因治疗脂蛋白脂肪酶缺陷（LPLD）[44]。LPLD 是一种由 LPL 的功能性突变引发的 LPL 本身或其他甘油三酯清除相关基因的功能缺陷的罕见疾病，该疾病会导致血清中高水平的甘油三酯并诱发急性或慢性的胰腺炎[44]。Glybera[®]具有肌肉趋向性的 AAV1 衣壳，携带一个具有自然功能突变的 LPL（LPL^{S477X}），通过肌肉注射的方式进行给药[44]。虽然 LPLD 的患者人群很小，Glybera[®]的在安全性评价和药效学评价的临床试验中并未发现安全问题。而且，有研究发现这种疗法能够降低胰腺炎的发病风险和严重程度，并可避免由胰腺炎导致的住院治疗，扩大了这种疗法的临床疗效[44-47]。基于这些临床结果，Glybera[®]于 2012 年在欧洲被批准实现商业化。

另一项已成功获得批准的由 rAAV 介导的基因疗法是视网膜基因治疗。通常来讲，眼科疾病由于其目标器官眼睛具有尺寸较小、免疫环境独特[49]、易于载体给药、技术评价体系成熟（X 射线断层摄影术和视网膜电图）[50]等特点，因而是基因治疗非常理想的靶标[48]。而且，有多种血清亚型的 rAAV 易于导入不同的视网膜细胞，使得这一导入系统能够满足特殊的治疗需求[51]。

在视网膜疾病治疗中，rAAV 介导的 Lener's 先天性黑内障的（LCA）基因治疗在临床试验中成功的提高了患者视力，受到了极大的关注。LCA 能够导致早期的视网膜退化，主要发生在儿童时期造成儿童盲[51]。虽然 LCA 可能由多种基因突变导致，但是临床上运用的 rAAV 疗法主要是修复引起 LCA 的 RPE65 基因功能突变[52-57]。RPE65 能够将全反式视网膜中的光感受器通路转化为 11-反式-视网膜，使得光感受器发生重复[58]。LCA 的基因治疗利用了 rAAV2 载体进行了 RPE65 基因的功能克隆转导。目前已经利用该疗法治疗了 30 名患者，这些患者在视力、眼球震颤、暗视野测量和暗迁移等方面都得到了持续性的改善[52-57]。

基于这些实验所取得的成功，有不少于八项的利用 rAAV 治疗视网膜疾病的临床试验已经启动，包括一项治疗 LCA 的 III 期临床试验（<http://www.abedia.com/wiley>）。这些实验结果，包括在 LPLD 中所取得的结果，证明了由 rAAV 介导的基因治疗具有很好的应用前景。基

于这一前景，AAV 的研究目前不只局限于基因导入策略，也被用于基于双链损伤修复的疾病突变纠正策略。

AAV 载体和基因编辑

在哺乳动物细胞中，人源基因组的精确突变诱导受限于同源重组的低效率，该效率只接近于每百万细胞一例。到本世纪初期，有研究发现 AAV 载体基因组在感染人类分裂期细胞后，相较于质粒修复策略 HR 诱导率提高了 1000 倍[59,60]。AAV 基因编辑的特征显示特定的载体修饰能够进一步提高 AAV 介导 HR 的效率，包括：（1）提高 AAV 载体基因组和靶向位点之间的同源重组效率，（2）将非同源末端片段或修饰后片段引入载体基因组的中心序列以及（3）HR 诱导频率和时间之间具有直接性的关系[61]。而且，自然状态下发生的突变影响 AAV 介导的 HR。有研究显示由 AAC 载体通过 HR 介导的基因删除相较于靶向基因插入具有更高的效率（14 倍）[60]。然而，在这一点上的研究也存在分歧，有研究称自然状态下的 HR 对于不同的目标突变（基因敲除 vs 插入 vs 点突变）的诱导效率并不总是具有差异[60,61]。另一个能够影响 AAV 基因编辑效率的是细胞周期动力学。通常来讲，AAV 感染在细胞进入 S 期时的效率更高[62]。例如，在早期研究中发现细胞分裂对于 AAV 基因靶向是必要的[63]而且在骨骼肌纤维中也没有发现基因靶向的发生。也有研究发现细胞发生转染的时期处于 G1/S 期附近的细胞越多，AAV 介导的基因靶向效率也会提高[63]。利用遗传毒性压力试剂能够提高 AAV 载体的感染效率，然而这种方法并不能刺激 AAV 基因的编辑，结果提示染色体损伤和 DNA 修复环境对于 AAV 基因组介导 HR 发生的能力没有影响[61,64]。目前还无法得知为何 AAV 基因组能够提高 HR 效率，有推测可能是由于 AAV 基因组自然状态下的单链结构能够在感染细胞的基础上诱导 DNA 损伤修复的进行，或者是由于 AAV 基因组在不同的细胞核区域包括核仁中的定位[17,64-67]。

AAV 除了能够提高基因编辑的效率之外，近些年还有研究发现称 AAV 的介导效率在不同的种属中差异十分巨大，早期研究称效率范围从无法检测到之 1/1000 之间，而其中对 1% 的细胞缺乏选择性[59,63]。实际上，最近有报告显示在试验方案缺乏例如提高编辑细胞的适应性等筛选性原则[68]的基础上，AAV 基因靶向编辑效率还是达不到大多数应用的需要。因此，在研究中有趋势将 AAV 与点特异性核酸酶进行联合应用，从而提高目标位点的编辑效率[69,70]。

通过点特异性的 DNA 损伤刺激 AAV 基因编辑

由于简便性的 AAV 感染能够高效率的诱导靶向同源重组的产生，有两个课题组在 2003 年报道了利用点特异性核酸酶刺激 AAV 基因编辑的方法[69,79]。为了达到这一目的，具有 I-SceI 识别位点的缺陷报告基因被整合到人类基因组中，并将其作为 AAV 介导基因编辑的目标位点。在这些试验中进行同源重组的 AAV 和 I-SceI 基因分别独立进行导入。与之前的转染结果一致的是，在没有筛选试剂的条件下报告基因修复的比例提高了 100 倍。通过 Southern-blot 对于点特异性的 I-SceI 介导的基因损伤过程进行量化发现在检测条件下由 I-SceI 造成的损伤中有五分之一的损伤是由 rAAV 介导修复的[69]。除了在 I-SceI 潜在的非目标错配位点外，AAV 载体整合效率并没有因为核酸酶而提高[69,70]。这些细胞培养数据显示，黄金比例的核酸酶能够刺激 AAV 介导特定位点 DSBs 的 HR 效率的提高，可用于特定遗传疾病的治疗。进一步来讲，利用自我互补 AAV 载体进行 I-SceI 的导入能够显著提高基因修复效率近十倍，也说明了高效载体导入的必要性[32]。

随着 AAV 介导 DSB 损伤部位 HR 实验中振奋人心的报道，利用核酸内切酶平台进行稳定的程序性的 DNA 靶向是必要的。与此同时，ZFNs 也被广泛性的应用，最初报道的利用 AAV-ZFN 进行 HR 诱导的实验是在血友病 B 的小鼠模型上进行的[71]。而且，血友病 B 也非常适合 AAV

进行 HR 诱导，因为在病人过表达的因子 9 (F9) 基因中具有一个特殊的点突变，利用 IV 型给药系统，使得 AAV 感染肝脏的效果非常不错，在一些案例中干细胞还可以进行复制，只需要低水平的具有功能的 F9 就能显著缓解流血症状。而且，在肝 zhong1 直接导入 AAV-F9 基因的策略目前在临床上也已经显示出不错的前景[40]。为了扩大这种方法在多种 F9 突变患者中的应用，有一种策略试图在 F9cDNA 精确插入 2-8 外显子从而修复 F9 的功能性表达[71]。除了载体所需的必要组成外，还有额外的试剂被包装进入 AAV8 衣壳：(1) ZFN1，(2) ZFN2，和 (3) 具有 F9 外显子 2-8 编码序列及 F9 内源性同源片段连接的修复底物。因此，每种细胞都需要进行三种不同载体的转录使得 ZFN 介导的修复过程顺利发生。进一步的缺陷还包括肝细胞介导 HR 发生的固有能力和通常情况下由于肝细胞复制在非诱导条件下受限。除了以上因素外，还有研究发现 AAV8-ZFNs 共转染能够诱导目标序列的切割并在 AAV 修复底物存在的情况下刺激 HR 的发生[71]。在三种质粒转染完毕之后，血友病模型小鼠的循环系统中 F9 的水平足以克服凝块缺陷。重要的是，WT 型及 F9 缺陷型的小鼠在经核酸酶和 AAV 修复载体处理后都未发现治疗相关毒性并能长时间耐受[71]。

在 (2012) 随后的几年中，细胞培养的相关研究拓展了多种 AAV 的应用，尤其是核酸酶介导的基因编辑[72, 73]。特别是 AAV-ZFN 介导的目标基因的破坏和，尤其是前病毒的删除处理的细胞分别占 30%和 12%[73]。于此相反，同源重组修复的发生率很低 (1-5%)，这对 NHEJ 效率的影响会高于 HR，或许需要另外的 AAV 修复载体的转染[72, 73]。为了解决后者的缺陷，单独的启动子平台依赖于 T2A 核糖体跳跃序列和两个 ZFNs 相连形成三个部分 (ZFN1、ZFN2 和修复片段) 的编码链包装在同一个 AAV 病毒颗粒中 (<5kb) [74]。重要的是，单独的 AAV-ZFN1-ZFN2 修复载体形式能够确保所有的被感染细胞都具有必要的进行目的修饰的因子。与观察结果一致的是衣壳和基因组增强元件能够提高转化效率，在不使用药物制剂的情况下也能提高 AAV 基因编辑的效率[32, 72, 74-76]。

在最近这年中，开发用于核酸酶诱导基因编辑的 AAV 载体在这些蛋白设计上显示出了优势。例如，在血友病 B 的 F9 修复的相同实验组中，相较于在成年鼠中采用相似的 AAV 载体策略，利用相应的异二聚体 ZFNs 能够降低突变的脱靶诱导率[76]。有一篇值得引起重视的报道称在非分裂的细胞中也存在 HR 修复，但是缺乏严谨的证实[76]。虽然，能够进行人工设计的核酸酶领域已经拓宽到了 TALENs、归巢内切酶和 Cas9/CRISPR 系统，由 AAV 介导的这些新技术的应用研究还比较匮乏。这可能是由于 TALEN 和 Cas9/介导链尺寸较大，无法利用有限大小 AAV 衣壳进行包装 (<5kb)。然而，Cas9、TALEN 的大小也存在种属性差异，我们也可以利用小尺寸的工具进行独立的 AAV 病毒颗粒的包装并对相对较难编辑的细胞进行高效的导入 (Cong 等人。2014 ASGCT 摘要) [77, 78]。

通过 DSB 修复进行 AAV 基因编辑的挑战

通常来讲，AAV 载体的临床接受度最高，是哺乳动物细胞中进行基因导入的一种高效的方法。然而，使这种基因编辑方式变得流行的原因同样在 AAV 载体介导的基因编辑方面存在缺陷。例如，在 LCA 眼试验中，在接受了单次载体给药后，基因表达和表型纠正要好几年后才会被观测到。在具有突变诱导能力的核酸酶存在的情况下，不需要 AAV 具有显著的长时程游离表型的倾向性，在应用这一转导系统时，核酸内切酶产物的时序调节是非常重要的。现行的 AAV 的另一显著性缺点是，尤其是在间接体内疗法应用中，其无法对特定类型的祖细胞进行高效的转导。例如，虽然近 30 年对 AAV 载体进行了多次优化，只有少量研究成功实现了对造血祖细胞的感染，而且我们已经注意到 AAV ITR 在人类胚胎干细胞中具有一定的毒性[79-82]。因此，基于 AAV 载体感染的细胞应答，目前仍有部分的细胞类型和处于某些特殊状态的多功能干细胞很难用 AAV 介导点特异性 DNA 损伤位点的编辑。在载体基因组水平，虽然有很多策略实现了大尺寸或超大尺寸的基因的 AAV 感染，相较于单独病毒颗粒的

AAV 转染，它们的效率在临床应用中显著性的下降[83-88]。因此，对转基因调控设计总结如下：扩大病毒衣壳的容积，在临床重要的细胞类型中对 AAV 载体的生物特征加深理解，更小尺寸的核酸酶有可能掀起 AAV 载体通过 DSB 修复进行基因编辑的下一轮应用浪潮。

参考文献

1. Atchison RW, Casto BC, Hammon WM. Adenovirus-associated defective virus particles. *Science*. 1965;149(3685):754 - 6.
2. Fields BN, Knipe DM, Howley PM. *Ovid technologies Inc. Fields' virology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2007. Available from: http://resources.lib.unc.edu/external_db/external_database_auth.html?A=P%7CF=N%7CID=162631%7CREL=HSL%7CSO=HSL%7CURL=http://libproxy.lib.unc.edu/login?url=http://gateway.ovid.com/ovidweb.cgi?T=JS&NEWS=N&PAGE=booktext&D=books&AN=00139921/5th_Edition/ .
3. Berns KI, Pinkerton TC, Thomas GF, Hoggan MD. Detection of adeno-associated virus (AAV)-specific nucleotide sequences in DNA isolated from latently infected Detroit 6 cells. *Virology*. 1975;68(2):556 - 60.
4. Xiao X, Xiao W, Li J, Samulski RJ. A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. *J Virol*. 1997;71(2):941 - 8.
5. King JA, Dubielzig R, Grimm D, Kleinschmidt JA. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J*. 2001;20(12):3282 - 91.
6. Kyostio SR, Owens RA, Weitzman MD, Antoni BA, Chejanovsky N, Carter BJ. Analysis of adeno-associated virus (AAV) wild-type and mutant Rep proteins for their abilities to negatively regulate AAV p5 and p19 mRNA levels. *J Virol*. 1994;68(5):2947 - 57.
7. Sonntag F, Schmidt K, Kleinschmidt JrA. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci*. 2010;107(22):10220 - 5.
8. Gao G, Vandenberghe LH, Alvira MR, Lu Y, Calcedo R, Zhou X, et al. Clades of adenoassociated viruses are widely disseminated in human tissues. *J Virol*. 2004;78(12):6381 - 8.
9. Ellis B, Hirsch M, Barker J, Connelly J, Steininger R, Porteus M. A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adenoassociated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virology*. 2013;10(1):74. doi: 10.1186/1743-422X-10-74 .
10. Zincarelli C, Soltys S, Rengo G, Rabinowitz JE. Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol Ther*. 2008;16(6):1073 - 80.
11. Grieger JC, Choi VW, Samulski RJ. Production and characterization of adeno-associated viral vectors. *Nat Protoc*. 2006;1(3):1412 - 28. PubMed Epub 2007/04/05. eng.
12. Asokan A, Schaffer DV, Samulski RJ. The AAV vector toolkit: poised at the clinical crossroads. *Mol Ther*. 2012;20(4):699 - 708. PubMed Epub 2012/01/26. eng.
13. Bartlett JS, Wilcher R, Samulski RJ. Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J Virol*. 2000;74(6):2777 - 85.
14. Xiao PJ, Samulski RJ. Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *J Virol*. 2012;86(19):10462 - 73. PubMed Epub 2012/07/20. eng.
15. Popa-Wagner R, Porwal M, Kann M, Reuss M, Weimer M, Florin L, et al. Impact of VP1-

- specific protein sequence motifs on adeno-associated virus type 2 intracellular trafficking and nuclear entry. *J Virol.* 2012;86(17):9163 – 74.
16. Rabinowitz JE, Samulski RJ. Building a better vector: the manipulation of AAV virions. *Virology.* 2000;278(2):301 – 8.
17. Johnson JS, Samulski RJ. Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J Virol.* 2009;83(6):2632 – 44. PubMed Pubmed Central PMCID: 2648275, Epub 2008/12/26. eng.
18. Ferrari FK, Samulski T, Shenk T, Samulski RJ. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J Virol.* 1996;70(5):3227 – 34.
19. Beaton A, Palumbo P, Berns KI. Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. *J Virol.* 1989;63(10):4450 – 4. PubMed Epub 1989/10/01. eng.
20. Miao CH, Snyder RO, Schowalter DB, Patijn GA, Donahue B, Winther B, et al. The kinetics of rAAV integration in the liver. *Nat Genet.* 1998;19(1):13 – 5. PubMed Epub 1998/05/20. eng.
21. Cheung AK, Hoggan MD, Hauswirth WW, Berns KI. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. *J Virol.* 1980;33(2):739 – 48.
22. Hirsch ML, Storici F, Li C, Choi VW, Samulski RJ. AAV recombineering with single strand oligonucleotides. *PLoS One.* 2009;4(11), e7705. PubMed Pubmed Central PMCID: 2765622, Epub 2009/11/06. eng.
23. Kotin RM, Siniscalco M, Samulski RJ, Zhu XD, Hunter L, Laughlin CA, et al. Site-specific integration by adeno-associated virus. *Proc Natl Acad Sci U S A.* 1990;87(6):2211 – 5.
24. Samulski RJ, Zhu X, Xiao X, Brook JD, Housman DE, Epstein N, et al. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J.* 1991;10(12):3941 – 50. PubMed Epub 1991/12/01. eng.
25. Deyle DR, Russell DW. Adeno-associated virus vector integration. *Curr Opin Mol Ther.* 2009;11(4):442 – 7. PubMed Epub 2009/08/04. eng.
26. McCarty DM, Young Jr SM, Samulski RJ. Integration of adeno-associated virus (AAV) and recombinant AAV vectors. *Annu Rev Genet.* 2004;38:819 – 45. PubMed Epub 2004/12/01. eng.
27. Rosas LE, Grieves JL, Zaraspe K, La Perle KMD, Fu H, McCarty DM. Patterns of scAAV vector insertion associated with oncogenic events in a mouse model for genotoxicity. *Mol Ther.* 2012;20(11):2098 – 110.
28. Miller DG, Trobridge GD, Petek LM, Jacobs MA, Kaul R, Russell DW. Large-scale analysis of adeno-associated virus vector integration sites in normal human cells. *J Virol.* 2005;79(17):11434 – 42.
29. McCarty DM, Monahan PE, Samulski RJ. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.* 2001;8(16):1248 – 54. PubMed Epub 2001/08/18. eng.
30. Wang Z, Ma HI, Li J, Sun L, Zhang J, Xiao X. Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors in vitro and in vivo. *Gene Ther.* 2003;10(26):2105 – 11. PubMed Epub 2003/11/20. eng.
31. McCarty DM, Fu H, Monahan PE, Toulson CE, Naik P, Samulski RJ. Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the

- ratelimiting step to transduction in vivo. *Gene Ther.* 2003;10(26):2112 – 8. PubMed Epub 2003/11/20. eng.
32. Hirsch ML, Green L, Porteus MH, Samulski RJ. Self-complementary AAV mediates gene targeting and enhances endonuclease delivery for double-strand break repair. *Gene Ther.* 2010;17(9):1175 – 80. PubMed Pubmed Central PMCID: 3152950, Epub 2010/05/14. eng.
33. Mueller C, Flotte TR. Clinical gene therapy using recombinant adeno-associated virus vectors. *Gene Ther.* 2008;15(11):858 – 63.
34. Podsakoff G, Wong KK, Chatterjee S. Efficient gene transfer into nondividing cells by adenoassociated virus-based vectors. *J Virol.* 1994;68(9):5656 – 66.
35. Xiao X, Li J, Samulski RJ. Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. *J Virol.* 1996;70(11):8098 – 108.
36. Rabinowitz JE, Bowles DE, Faust SM, Ledford JG, Cunningham SE, Samulski RJ. Crossdressing the virion: the transcapsidation of adeno-associated virus serotypes functionally defines subgroups. *J Virol.* 2004;78(9):4421 – 32. PubMed Epub 2004/04/14. eng.
37. Rabinowitz JE, Rolling F, Li C, Conrath H, Xiao W, Xiao X, et al. Cross-packaging of a single adeno-associated virus (AAV) type 2 vector genome into multiple AAV serotypes enables transduction with broad specificity. *J Virol.* 2002;76(2):791 – 801. PubMed Epub 2001/12/26. eng.
38. Ellis BL, Hirsch ML, Barker JC, Connelly JP, Steininger 3rd RJ, Porteus MH. A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virol J.* 2013;10:74. PubMed Pubmed Central PMCID: 3607841.
39. Mitchell AM, Nicolson SC, Warischalk JK, Samulski RJ. AAV ' s anatomy: roadmap for optimizing vectors for translational success. *Curr Gene Ther.* 2010;10(5):319 – 40. PubMed Epub 2010/08/18. eng.
40. Nathwani AC, Tuddenham EG, Rangarajan S, Rosales C, McIntosh J, Linch DC, et al. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med.* 2011;365(25):2357 – 65. PubMed Pubmed Central PMCID: 3265081, Epub 2011/12/14. eng.
41. Manno CS, Pierce GF, Arruda VR, Glader B, Ragni M, Rasko JJ, et al. Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med.* 2006;12(3):342 – 7. PubMed Epub 2006/02/14. eng.
42. Gray SJ. Gene therapy and neurodevelopmental disorders. *Neuropharmacology.* 2013;68:136 – 42. PubMed Epub 2012/07/04. eng
43. Zouein FA, Booz GW. AAV-mediated gene therapy for heart failure: enhancing contractility and calcium handling. *F1000Prime Rep.* 2013;5:27. PubMed Epub 2013/08/24. eng.
44. Gaudet D, Methot J, Kastelein J. Gene therapy for lipoprotein lipase deficiency. *Curr Opin Lipidol.* 2012;23(4):310 – 20. PubMed Epub 2012/06/14. eng.
45. Carpentier AC, Frisch F, Labbe SM, Gagnon R, de Wal J, Greentree S, et al. Effect of alipogene tiparvovec (AAV1-LPL(S447X)) on postprandial chylomicron metabolism in lipoproteinlipase-deficient patients. *J Clin Endocrinol Metab.* 2012;97(5):1635 – 44. PubMed Epub 2012/03/23. eng.
46. Gaudet D, Methot J, Dery S, Brisson D, Essiembre C, Tremblay G, et al. Efficiency and longterm safety of alipogene tiparvovec (AAV1-LPLS447X) gene therapy for lipoprotein lipase deficiency: an open-label trial. *Gene Ther.* 2013;20(4):361 – 9. PubMed Epub 2012/06/22. eng.

47. Stroes ES, Nierman MC, Meulenberg JJ, Franssen R, Twisk J, Henny CP, et al. Intramuscular administration of AAV1-lipoprotein lipase S447X lowers triglycerides in lipoprotein lipase-deficient patients. *Arterioscler Thromb Vasc Biol.* 2008;28(12):2303 – 4. PubMed Epub 2008/09/20. eng.
48. Ali RR, Reichel MB, Thrasher AJ, Levinsky RJ, Kinnon C, Kanuga N, et al. Gene transfer into the mouse retina mediated by an adeno-associated viral vector. *Hum Mol Genet.* 1996;5(5):591 – 4. PubMed Epub 1996/05/01. eng.
49. Barker SE, Broderick CA, Robbie SJ, Duran Y, Natkunarajah M, Buch P, et al. Subretinal delivery of adeno-associated virus serotype 2 results in minimal immune responses that allow repeat vector administration in immunocompetent mice. *J Gene Med.* 2009;11(6):486 – 97.
50. Jacobson SG, Cideciyan AV, Ratnakaram R, Heon E, Schwartz SB, Roman AJ, et al. Gene therapy for leber congenital amaurosis caused by RPE65 mutations: safety and efficacy in 15 children and adults followed up to 3 years. *Arch Ophthalmol.* 2012;130(1):9 – 24. PubMed Epub 2011/09/14. eng.
51. McClements ME, MacLaren RE. Gene therapy for retinal disease. *Transl Res.* 2013;161(4):241 – 54. PubMed Epub 2013/01/12. eng.
52. Testa F, Maguire AM, Rossi S, Pierce EA, Melillo P, Marshall K, et al. Three-year follow-up after unilateral subretinal delivery of adeno-associated virus in patients with Leber congenital Amaurosis type 2. *Ophthalmology.* 2013;120(6):1283 – 91. PubMed Pubmed Central PMCID: 3674112.
53. Bainbridge JW, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, et al. Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med.* 2008;358(21):2231 – 9. PubMed Epub 2008/04/29. eng.
54. Hauswirth WW, Aleman TS, Kaushal S, Cideciyan AV, Schwartz SB, Wang L, et al. Treatment of leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. *Hum Gene Ther.* 2008;19(10):979 – 90. PubMed Epub 2008/09/09. eng.
55. Maguire AM, High KA, Auricchio A, Wright JF, Pierce EA, Testa F, et al. Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial. *Lancet.* 2009;374(9701):1597 – 605. PubMed Epub 2009/10/27. eng.
56. Maguire AM, Simonelli F, Pierce EA, Pugh Jr EN, Mingozzi F, Bennicelli J, et al. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med.* 2008;358(21):2240 – 8. PubMed Epub 2008/04/29. eng.
57. Bennett J, Ashtari M, Wellman J, Marshall KA, Cyckowski LL, Chung DC, et al. AAV2 gene therapy readministration in three adults with congenital blindness. *Sci Transl Med.* 2012;4(120):120ra15.
58. Cai X, Conley SM, Naash MI. RPE65: role in the visual cycle, human retinal disease, and gene therapy. *Ophthalmic Genet.* 2009;30(2):57 – 62. PubMed Epub 2009/04/18. eng.
59. Russell DW, Hirata RK. Human gene targeting by viral vectors. *Nat Genet.* 1998;18(4):325 – 30. PubMed Pubmed Central PMCID: 3010411, Epub 1998/04/16. eng.
60. Russell DW, Hirata RK. Human gene targeting favors insertions over deletions. *Hum Gene Ther.* 2008;19(9):907 – 14. PubMed Pubmed Central PMCID: 2940567.
61. Hirata RK, Russell DW. Design and packaging of adeno-associated virus gene targeting vectors. *J Virol.* 2000;74(10):4612 – 20. PubMed Pubmed Central PMCID: 111981, Epub

2000/04/25. eng.

62. Russell DW, Miller AD, Alexander IE. Adeno-associated virus vectors preferentially transduce cells in S phase. *Proc Natl Acad Sci U S A.* 1994;91(19):8915 – 9. PubMed Pubmed Central PMCID: 44717.

63. Trobridge G, Hirata RK, Russell DW. Gene targeting by adeno-associated virus vectors is cellcycle dependent. *Hum Gene Ther.* 2005;16(4):522 – 6. PubMed.

64. Liu X, Yan Z, Luo M, Zak R, Li Z, Driskell RR, et al. Targeted correction of single-base-pair mutations with adeno-associated virus vectors under nonselective conditions. *J Virol.* 2004;78(8):4165 – 75. PubMed Pubmed Central PMCID: 374254.

65. Cervelli T, Palacios JA, Zentilin L, Mano M, Schwartz RA, Weitzman MD, et al. Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. *J Cell Sci.* 2008;121(Pt 3):349 – 57. PubMed.

66. Inagaki K, Ma C, Storm TA, Kay MA, Nakai H. The role of DNA-PKcs and artemis in opening viral DNA hairpin termini in various tissues in mice. *J Virol.* 2007;81(20):11304 – 21. PubMed Pubmed Central PMCID: 2045570, Epub 2007/08/10. eng.

67. Schwartz RA, Carson CT, Schuberth C, Weitzman MD. Adeno-associated virus replication induces a DNA damage response coordinated by DNA-dependent protein kinase. *J Virol.* 2009;83(12):6269 – 78. PubMed Pubmed Central PMCID: 2687378.

68. Paulk NK, Loza LM, Finegold MJ, Grompe M. AAV-mediated gene targeting is significantly enhanced by transient inhibition of nonhomologous end joining or the proteasome in vivo. *Hum Gene Ther.* 2012;23(6):658 – 65. PubMed Pubmed Central PMCID: 3392621.

69. Miller DG, Petek LM, Russell DW. Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks. *Mol Cell Biol.* 2003;23(10):3550 – 7. PubMed Pubmed Central PMCID: 164770.

70. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science.* 2003;300(5620):763. PubMed Epub 2003/05/06. eng.

71. Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature.* 2011;475(7355):217 – 21. PubMed Pubmed Central PMCID: 3152293.

72. Asuri P, Bartel MA, Vazin T, Jang JH, Wong TB, Schaffer DV. Directed evolution of adenoassociated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol Ther.* 2012;20(2):329 – 38. PubMed Pubmed Central PMCID: 3277219.

73. Handel EM, Gellhaus K, Khan K, Bednarski C, Cornu TI, Muller-Lerch F, et al. Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adenoassociated viral vectors. *Hum Gene Ther.* 2012;23(3):321 – 9. PubMed Pubmed Central PMCID: 3300077.

74. Ellis BL, Hirsch ML, Porter SN, Samulski RJ, Porteus MH. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. *Gene Ther.* 2013;20(1):35 – 42. PubMed Epub 2012/01/20. Eng.

75. Rahman SH, Bobis-Wozowicz S, Chatterjee D, Gellhaus K, Pars K, Heilbronn R, et al. The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. *Hum Gene Ther.* 2013;24(1):67 – 77. PubMed Pubmed Central PMCID: 3555098.

76. Anguela XM, Sharma R, Doyon Y, Miller JC, Li H, Haurigot V, et al. Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood*. 2013;122(19):3283 – 7. PubMed Pubmed Central PMCID: 3821724.
77. Beurdeley M, Bietz F, Li J, Thomas S, Stoddard T, Juillerat A, et al. Compact designer TALENs for efficient genome engineering. *Nat Commun*. 2013;4:1762. PubMed Pubmed Central PMCID: 3644105.
78. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods*. 2013;10(11):1116 – 21. PubMed Pubmed Central PMCID: 3844869.
79. Hirsch ML, Fagan BM, Dumitru R, Bower JJ, Yadav S, Porteus MH, et al. Viral single-strand DNA induces p53-dependent apoptosis in human embryonic stem cells. *PLoS One*. 2011;6(11), e27520. PubMed Pubmed Central PMCID: 3219675, Epub 2011/11/25. eng.
80. Song L, Li X, Jayandharan GR, Wang Y, Aslanidi GV, Ling C, et al. High-efficiency transduction of primary human hematopoietic stem cells and erythroid lineage-restricted expression by optimized AAV6 serotype vectors in vitro and in a murine xenograft model in vivo. *PLoS One*. 2013;8(3), e58757. PubMed Pubmed Central PMCID: 3597592.
81. Maina N, Han Z, Li X, Hu Z, Zhong L, Bischof D, et al. Recombinant self-complementary adeno-associated virus serotype vector-mediated hematopoietic stem cell transduction and lineage-restricted, long-term transgene expression in a murine serial bone marrow transplantation model. *Hum Gene Ther*. 2008;19(4):376 – 83. PubMed.
82. Srivastava A. Hematopoietic stem cell transduction by recombinant adeno-associated virus vectors: problems and solutions. *Hum Gene Ther*. 2005;16(7):792 – 8. PubMed.
83. Duan D, Yue Y, Engelhardt JF. Expanding AAV packaging capacity with trans-splicing or overlapping vectors: a quantitative comparison. *Mol Ther*. 2001;4(4):383 – 91. PubMed Epub 2001/10/11. eng.
84. Hirsch ML, Agbandje-McKenna M, Samulski RJ. Little vector, big gene transduction: fragmented genome reassembly of adeno-associated virus. *Mol Ther*. 2010;18(1):6 – 8. PubMed Pubmed Central PMCID: 2839225, Epub 2010/01/06. eng.
85. Hirsch ML, Li C, Bellon I, Yin C, Chavala S, Pryadkina M, et al. Oversized AAV transduction is mediated via a DNA-PKcs independent, Rad51C-dependent repair pathway. *Mol Ther*. 2013;13. PubMed.
86. Lai Y, Yue Y, Duan D. Evidence for the failure of adeno-associated virus serotype 5 to package a viral genome > or = 8.2 kb. *Mol Ther*. 2010;18(1):75 – 9. PubMed Pubmed Central PMCID:2839223, Epub 2009/11/12. eng.
87. Wu Z, Yang H, Colosi P. Effect of genome size on AAV vector packaging. *Mol Ther*. 2010;18(1):80 – 6. PubMed Pubmed Central PMCID: 2839202, Epub 2009/11/12. eng.
88. Halbert CL, Allen JM, Miller AD. Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nat Biotechnol*. 2002;20(7): 697 – 701. PubMed Epub 2002/06/29. eng.

核苷酸酶编辑和三核苷酸重复序列疾病

John H. Wilson , Christopher Moye , and David Mittelman

摘要 串联重复序列是一组连续重复性的核苷酸片段,可以发生重复片段单位的增加和缺失。串联性序列能够调节基因功能并在包括人类在内的多个物种中保持遗传特征。发生的重复序列的突变,大多是三核苷酸重复突变,能够引起人类神经功能的损伤和骨骼肌异常。尤其是能够引起多种神经异常的,在正常情况下少于 30 个单位重复的 CAG 序列的扩增,在疾病产生时这一重复单位可扩增至上百个,有时甚至是上千个。由于这一重复单位能够引起多种异常,因而是治疗的理想靶标,但是在实际靶向时仍存在很多障碍。在这里我们介绍了我们和其他实验室有关利用核苷酸编辑破坏 CAG 重复单位的研究。我们讨论了这一疗法的现行挑战并关注了下一代测序方法作为重复突变测序和核酸内切酶介导基因编辑效率的有力工具的可能。

关键词 三核苷酸重复疾病 · 微卫星重复 · 编辑核酸酶 · 下一代测序技术

J. H. Wilson , Ph.D. (*) • C. Moye , Ph.D.

Verna and Marrs McLean Biochemistry and Molecular Biology , Baylor College of Medicine ,
One Baylor Plaza , Houston , TX 77030 , USA
e-mail: jwilson@bcm.edu

D. Mittelman

Department of Biological Sciences , Virginia Tech , Blacksburg , VA , USA
Virginia Bioinformatics Institute , Blacksburg , VA , USA

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_9

缩略语

APRT	腺嘌呤磷酸核糖转移酶
CRISPR	短促回文整合性调控簇
DM1	强直性肌营养不良类型 1
DSBs	双链损伤
FRAXA X	染色体易损综合征
GFP	绿色荧光蛋白
HD	亨廷顿疾病
HPRT	次黄嘌呤鸟嘌呤脱氧核糖转移酶
PCR	聚合酶链式反应
PolyQ	多聚甘氨酸
SCA	脊髓小脑性共济失调
TALENs	转录激活样效应核酸酶
TC-NER	转录伴随性核苷酸删除修复
TNRs	三核苷酸重复
3'UTRs	3'非翻译序列
5'UTRs	5'非翻译序列
ZFNs	锌指核酸酶
ZFN 切口酶	锌指缺口酶

介绍

人类基因组中有一种普遍存在的简单的重复序列，称为微卫星。总的来讲，这些重复序列占人类基因组的约 3%，在基因组中占比 3%[1]。微卫星重复序列是基因组中最常发生突变的部分，它们能够扩增出更多的重复单位（放大）或者发生丢失（收缩），发生比率在 0.001-1% 之间[2-4]。这些序列之间的差异性注定其无法作为编码结构域，在人类中这类基因具有 17% 的比例[3]。诚然，有一组微卫星重复序列——三核苷酸重复（TNRs），在外显子区域中出现的比例很高，尤其是位于转录因子和其他调控蛋白上的一切区段，提示 TNRs 可能会带来进化上的好处[1,5-8]。

然而这种在进化中所具有的潜在性的福利，也让人类付出了巨大的代价。在 25 条染色体中，TNP 的扩增能够导致遗传性疾病，例如广为人知的 X 染色体易损性疾病、遗传性共济失调、强直性肌营养不良和亨廷顿疾病[9-13]。在染色体易损性疾病和遗传性共济失调中，重复单位的扩增能够干扰基因的表达，然而在强直性肌营养不良和亨廷顿疾病中，重复单位的扩增能够产生具有毒性的产物。对于缺陷型细胞和疾病表型之间的研究为药物研发铺垫了具有高效率的道路，该方法在模型系统中已经取得了很多有意义的成果，然而直接用于病人治疗的还几乎没有。

有一种直接的治疗方式是直接阻断驱动重复单位进行扩增的机制，在已知的疾病产生分子机制的基础上第一时间阻止或减少其发生[14]。自从 TNR 的不稳定性被发现之后这一方式变得十分简洁。在模式生物中的研究涉及到了 DNA 的很多领域，包括有助于 TNR 不稳定性的复制、重组、DNA 修复和转录[1,3,15,16]。事实上，很少有过程能在不影响 TNR 稳定性的基础上影响 DNA。通过这一理论研究发现小鼠的不同的组织中具有不同的 TNR 不稳定性机制[17-21]。在多种机制途径中找到合适于靶位点的方法就成为了一项重要的挑战。

第三种选择是开发可以缩小重复单位序列的试剂,一种可以被患者长时间接受的不会对细胞产生过大损伤的治疗方法。最初看来,这种方法也需要对于不稳定机制深入的知识了解:然而,最近在定制化 DNA 核酸酶方面的进展为 TNR 不稳定性生物途径提供了一种常规化的解决手段。在重复单位处引入双链损伤(DSB)能够刺激 TNR 在损伤位点发生收缩[22,23]。我们常用的编辑核酸酶工具包括锌指核酸酶(ZFNs)、转录激活样核酸酶(TALENs)、和短促回文调控簇(CRISPR)/Cas 核酸酶,以及其他的可以造成 TNRs 损伤的分子工具。而这一方法的挑战就在于在易于损伤处,点特异性的引入高频率的损伤并最小化随之而产生的伤害。

编辑化核酸酶能够基因操作的可变化工具,但是它们真的能够被转化用于人类的基因手术吗?为了解决这一挑战,我们对人类基因组中的重复性序列、TNRs 固有的不稳定性机制和 TNR 扩增引起疾病的过程进行了回顾。

微卫星重复序列和 TNR 不稳定性

重复性 DNA 序列或多或少的存在于几乎所有的基因组中,但是它们在人类基因组中出现的概率异常高,约占我们基因组的一半[1]。串联重复是这一类型片段的子范畴,这些重复单位在一条链上尾对尾的像是棚车一样排列。如果重复序列超过了 10 个核苷酸,这一重复性序列可被成为微卫星重复[3]。能够产生疾病的 TNRs 也属于微卫星的范畴。微卫星由其总长进行定义,其范围在几个单位到数千个单位之间,而这一序列的纯粹性则是由该固定序列所占有的总核苷酸的比例决定的。TNR 进行扩增或收缩的倾向性主要取决于重复片段的总长度、重复片段的纯度以及重复性单位的组成。

在 TNRs 处发生突变的比率相较于单拷贝序列高出 100 到 100 倍,最主要的形式是重复区域重复单位的数量的改变[2,3]。因此,TNR 突变具有可预测性:它们能够改变重复片段的总长、这也预示这一过程易于发生可逆,以及在进化适应性中不同的长度的重复片段可以被很容易检测到[7,8,24]。在大部分的重复性片段中,突变发生率受以下的因素影响:在长片段中的突变发生率高于短片段,纯粹性高的片段突变率高于低的[4]。然而在基因中,在不同的位点中相似的 TNRs 具有不同的突变率。虽然具体的原因尚不清楚,有可能是这些环境也是影响 TNR 稳定性机制的原因,这些原因包括距离复制位点远近、启动子、染色体结表现状态以及其他的可能性[1]。

在细菌、酵母、果蝇、哺乳动物细胞和小鼠中的研究发现与 DNA 相关的全部领域,包括 DNA 复制、重组、修复和转录,都会对微卫星的不稳定性产生影响[1,3,15,16]。通过 DNA 单链的暴露,这些过程能够允许重复片段内部发生位移整合,并能够引起 DNA 结构多种形式的变化。例如,在复制时滑动前导链相较于模板链的位置能够延长或者缩短重复片的长度(图 1A) [1]。与之类似,在酵母和人类细胞中,在 CAG 处发生的损伤能够通过单链退火途径进行重组从而促进 TNR 的不稳定性(图 1B) [22,25]。最后,也可能是最为特殊的是转录也能使 TNRs 变得不稳定[15,26-33]。在人类细胞中,转录通过刺激 R 环的形成介导 TNR 的不稳定性需要错配修复中 MutS β 识别部件的参与,并依赖于转录伴随核苷酸删除修复(TC-NER)中的互补途径(图 1C) [34,35]。

引起 TNR 不稳定的真实差异相较于图 1 中所示可能更为复杂,在模型组织和游离细胞系统中可能有更多的因族和过程能够影响 TNR 的不稳定性[36]。目前已知的影响 CAG TNRs 不稳定的因素有:复制聚合酶[37]、旗状特异性核酸内切酶[25]、复制因子[37-42]、跨损伤修复[43]、超螺旋和拓扑异构酶[44,45]、解旋酶[46-48]、错配蛋白[27, 49-53]、碱基删除修复组件[19,20,54]、核苷酸删除修复[21,33,35,55]、单链损伤修复[45]、双链损伤修复和同源重组修复[22,25,56-59]、转录[27,33,60]、聚敛转录[31,32]、R 环[34,61]、E3 泛素化连接和蛋白酶体[35]、CpG 甲基化[17,62]、组蛋白去乙酰化酶[63]、Hsp90 分子伴侣[8,58]和 DNA 损伤

检查点途径[31,42,64]。理解这些过程如何发挥作用是一个令人望而生畏的挑战，但是在以缩小重复单位为目的的疗法中这些问题又不得不对这些问题进行深入思考。

TNRs 和人类疾病

在人类基因组中的 25 条染色体上 TNRs 的扩增能够引起严重的神经系统疾病、肌肉疾病或者其余潜在性疾病[9-13]。疾病相关的重复单位扩增发现于基因的转录部分，包括外显子、内含子、5'UTRs 和 3'UTRs (图 2)。一些重复片段的扩增能够影响基因的功能，产生隐性的、功能性表型的缺失[9,36]。例如，在 X 染色体易损伤性疾病 (FRAXA) 中，CGG 在 5'UTR 中的扩增能够引起启动子的超甲基化和失活[65]。与之相类似，在遗传性共济失调 (FRDA) 中，内含子中 GAA 的扩增能够干扰转录延伸[66-68]。然而更为常见的是，TNR 等位基因的扩增与蛋白或 RNA 产物的毒性相关，能够引起主要的、获得性功能的表型的产生。

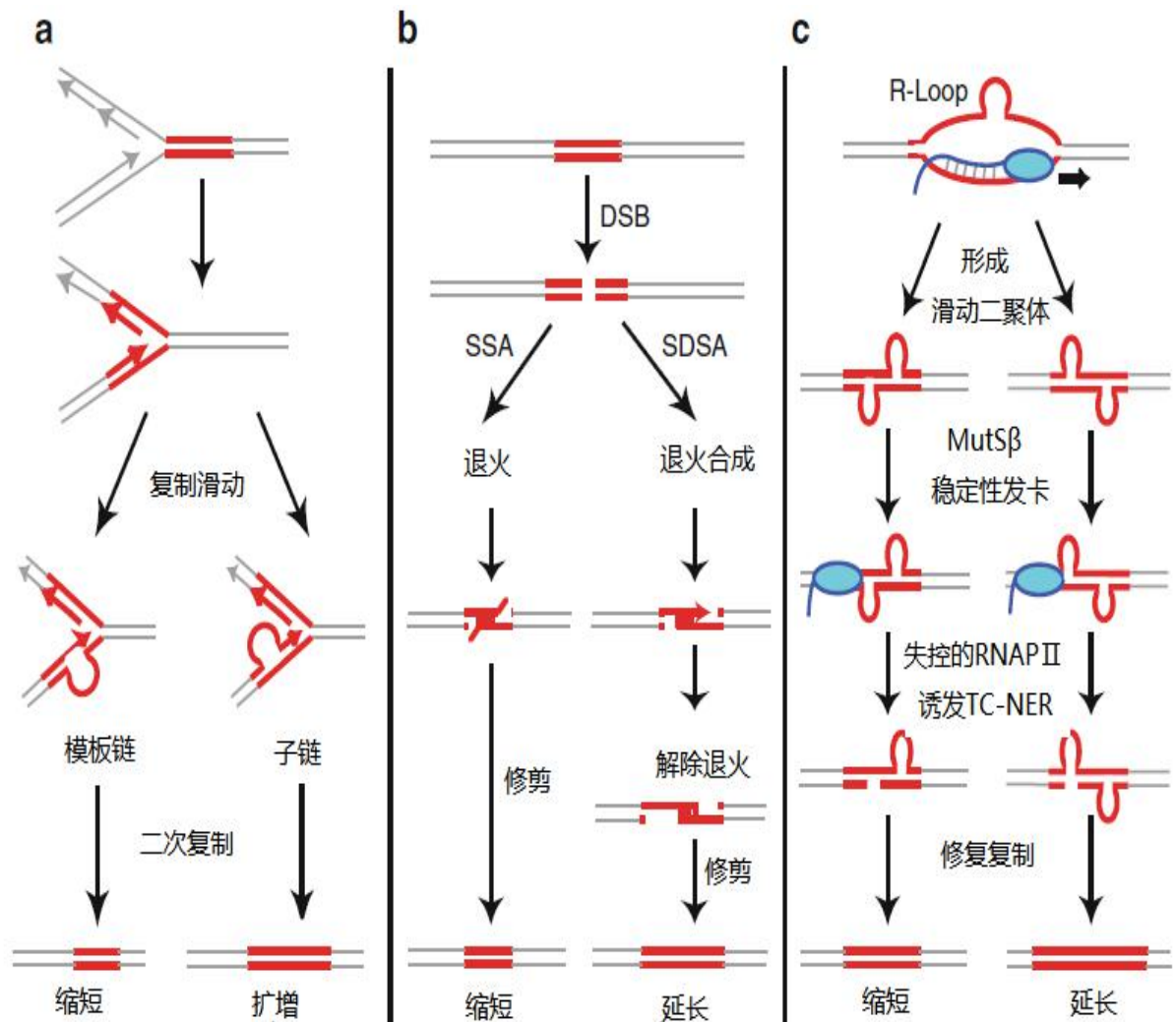


图1.引起微卫星重复序列不稳定的机制。(a)复制滑动。子链和模板链发生错配能够引起延展段重复序列成环。随后的复制将会产生片段的延长或缩短。(b)同源重组。DSBs引起的单链暴露可直接通过单链退火途径(SSA)进行修复,并引起序列的缩短。如果一条链通过DNA合成得到了延长,另一条链也会进行延长修复并通过合成依赖性退火(SDSA)途径产生延长性的重复序列[119]。(c)转录。R-loop的产生在RNA聚合酶之后,能够在非编码链形成发夹结构,当R-loop降解时会引起双链之间的相对滑动。MutS β 和CAG、CTG之间连接形成的发夹结构能够阻遏RNA聚合酶的作用,是TC-NER的起始信号[49,120,121]。修复和随后的复制能够引起片段的扩增和缩短,图1改编自Chatterjee, N., Santillan, B.A., and Wilson, J.H. (2013) *Microsatellite Repeats: Canaries in the Coalmine*. In *Stress-Induced Mutagenesis*, D. Mittelman, ed, Springer Publishing Co, New York, NY, pp. 119–150.图片改编已获得许可。

最常见的功能获得性表征是亨廷顿疾病(HD)和多种脊髓小脑共济失调疾病(SCAs),主要由于内含子中CAG重复序列所编码的多聚谷氨酰胺(polyQ)突变产生的“蛋白毒性”所引起的。在每一个案例中,被延长的polyQ序列都被认为改变了原有蛋白的特征,使得其具有毒性并诱发疾病的产生[9]。“RNA毒性”的现象最早发现于I型强直性肌营养不良(DM1)位点(在3'UTR部位的CTG)[69,70]。这些DM RNAs的扩增序列能够与不同的盲肌剪切因子相连,引起多种其它转录片段的一场剪切,从而为这些疾病的产生提供了基础[69]。有趣的是,位于X染色体上的易损位点CGG,当单位扩增数量达到55到200时,能够引起功能获得性疾病,易损伤X震动和共济失调综合征(FXTAS),并且由于RNA的过表达,也能够引起细胞毒性[71,72]。最近有研究发现,由CAG重复序列引起的汇聚性转录在一些疾病中能够通过提高“DNA毒性”的可能性引起细胞死亡[31,73],尤其是在人类基因组中的反义转录所达到的意想不到的高频率。

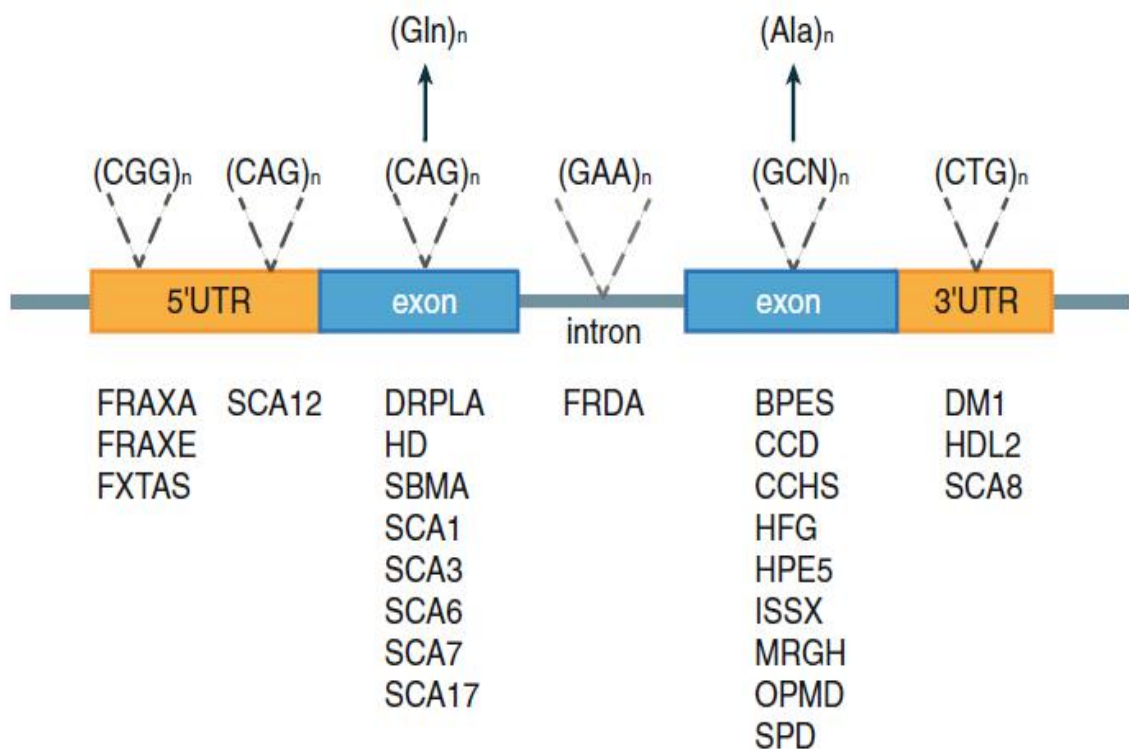


图2. 由TNRs引起的人类疾病。与特异性的TNRs相关的疾病标示于重复单位下。BPES脸部狭小、下垂和内眦赘皮矫正。CCD颅骨发育不良, CCHS先天性中枢换气不足综合症、DM1 I型强直性肌营养不良、DRPLA 齿状核红核苍白球丘脑下部核萎缩、FRAXA 易损性X染色体综合征、FRAXE FRAXE位点易损性X染色体精神发育迟滞、HD 亨廷顿疾病、HDL2 2类亨廷顿疾病、HFG 手足生殖器疾病、HPE5 前脑无裂畸形 5、ISSX X相关婴儿痉挛症、MRGH 生长激素缺陷引起的智力迟钝、OPMD 眼咽肌肉营养不良、SBMA 脊髓和延髓肌肉异常、1、3、6、7、8、12、17型SCA 脊髓小脑共济失调、SPD 并指/趾多症。图3改编自Chatterjee,N., Santilan, B.A., and Wilson, J.H. (2013) Microsatellite Repeats: Canaries in the Coalmine. In *Stress-Induced Mutagenesis*, D. Mittelman, ed, Springer Publishing Co, New York, NY, pp. 119-150。图片改编获得授权。

疾病相关的 TNRs 只包括 CNG 和 GAA 重复。与其它可能的三核苷酸重复不同, 这些 TNRs 能够很容易形成非 B 型 DNA 结构, 包括发夹结构和双链滑动(CAG 和 CGG)、G-四重叠(CGG)和三聚体(GAA) [76,77]。这些 TNRs 由于在复制、修复、重组和转录过程中易于形成瞬时的单链结构, 无法保持稳定的 B 型 DNA 结构因而具有不稳定性。形成复杂二级结构的倾向性在一些情况下可以提高 TNR 不稳定性近 100%[78,79]。但是在不同组织中相同片段的不稳定性存在显著性差异。

疾病相关 TNRs 在种系基因组中发生高频率的扩张或缩短在不同的亲本中经常具有不同的趋向性, 这和疾病的种类相关。具有 HD 或严重 SCAs (均有 CAG 重复序列) 的亲本在父系基因组中经常具有扩增的倾向。与此相反, 在 FRAXA (CGG 重复)、FRDA (GAA 重复) 或 DM1 (CTG 重复) 患者的片段扩增具有母系倾向[11]。片段扩增这一倾向性与早期氛围有关, 临床研究发现疾病症状在一代传给下一代时会变得更加严重。虽然疾病表型经常来源于 TNR 的扩增, 但是 TNR 的不稳定性并不总是会引起片段扩增; 例如, 在 FRDA 和 FRAXA 患者中, 在亲本种系中具有显著性的收缩的倾向[11,80]。这些亲本原发效应目前还无法理解。

TNR 不稳定性也可见于体细胞组织, 改效应和年龄有关, 尤其是大脑, 可以加速神经紊乱和死亡的进程[79,81,82]。由 CAG 和 CTG 造成的 TNR 疾病具有相似的不稳定模式[83]。在白血球细胞和心肌细胞中的基因组微卫星序列趋于稳定, 而在肝和肾中的基因组则表现为不稳定的中间态。与此类似, 在大脑中, 在纹状体中的重复序列特别不稳定, 而在大脑皮层和海马区的重复序列则比较不稳定, 而小脑中的重复序列则是较稳定的[83]。与此相反, 在 FRDA 患者中的 GAA 序列在血液和小脑中非常不稳定, 这也是将其与 CAG 重复序列区分开来的重要特征[84]。除此之外, GAA 重复序列的不稳定性在所有组织中都趋向于片段缩短; 然而, 在退化能引起疾病的背根神经节。片段延伸的频率更高[81,84]。在 FRAXA 患者中 (CGG 重复), 重复长片段的 CpG 位点经常产生甲基化并表现出较弱的不稳定性[80,85], 虽然比较罕见, 但是非甲基化的互补链非常不稳定[86,87]。

在不同的组织和疾病中 TNR 不稳定性的差异使其最令人困惑的特征之一[83]。有研究发现 TNR 的不稳定性所介导的多种机制和重复性的序列、基因片段的内容、组织类型和发育状态有关。例如在男性扩增细胞的基因组中, 通过复制滑动进行 TNRs 的扩增[16], 但是在高度分化的脑组织中, 由于 DNA 不进行复制, 类似的扩增过程与转录时发生的 DNA 修复有关 (图 1) [15,21,88]。在 CAG 重复疾病的鼠模型中, 遗传性实验强化了多机制的假说。错配修复的识别部件 (MutS β 复合体中的 Msh2 和 Msh3 组件), 能够影响男性和女性多个体细胞组织中基因组的不稳定性; 主要形式的 DNA 甲基转移酶——Dnmt1, 能够影响体细胞外的基因组的不稳定性[17]; DNA 连接酶 1 在母系基因组中改变重复片段的不稳定性, 但是在父系基因组中则无此影响[18]; 糖苷酶 Ogg1 能够有选择性的改变体细胞组织的不稳定性

[19,20]; 二核苷酸删除修复中的关键性组分 Xpa 则可影响神经组织中的不稳定性[21]。

TNR 在不同组织中所具有的不稳定差异和控制 TNR 不稳定性的蛋白网络的理解对 TNR 不稳定性调控是一项巨大的挑战。然而最主要的问题是,我们是否可以采用某种治疗策略绕过这些复杂的机制,直接攻击致病性的 TNR。而利用编辑后的核酸酶进行直接性的重复片段 DNA 损伤诱导很可能成为这样的理想策略。

编辑化的核酸酶和 TNR 缩短

在已有的编辑核酸酶——ZFNs、TALENs 和 CRISPR/Cas 核酸酶中,只有 ZFNs 和 TALENs 已经被用于进行细胞中 TNRs 的破坏。在最早的报告中, Mittelman 等人构建了两个具有 CAG 特异性的 ZFN: zFGCA 和 zFGCT,能够分别识别 CAG 和 CTG 片段(图 3A) [22]。对独特的锌指识别区域进行连接优化将两个组件排在相距七个核苷酸的距离上,虽然较之首选的六核苷酸距离较远,但是能够完成切割目的[89,90]。对于纯粹性的组分的切割显示出 zFGCA 和 zFGCT 复合体能够如预期一般对具有重复单位的质粒 DNA 进行高效切割;然而,在与具有活性的 FokI 裂解区域进行整合后,单独的组分也能凭借自身完成切割。这一结果对于 GCA 锌指和 GCT 片段之间具有的交叉反应性并不完全意外,而反之亦然[91,92],这一结果在体外系统中同样适用[93]。

为了评价哺乳动物细胞中 ZFN 介导的 CAG 重复单位的切割效率,选择了两种试验用来检测 CAG 重复序列缩短的情况。在两种实验中,一项是在鼠细胞中以 APRT 基因为基础,另一项是在人类细胞中以 HPRT 微基因为基础——当 CAG₉₅ 序列在内含子中产生时会导致基因失活,其作用类似于一个额外的外显子,能够引起异常的剪切并阻止基因的表达[27,94]。当重复序列缩短到 40 个单位以下时,能够使细胞进行足够量的表达。利用 zFGCA+zFGCT 能够提高细胞存活的频率 15 倍以上,总发生率约为 0.01%左右[22]。虽然单独利用 zFGCA 进行治疗没有效果,但是 zFGCT 本身就能达到完整核酸酶 70%的效率,在体外的结果也大致相同。

这些研究也检查了重复片段长度对 ZFNs 产生 APRT⁺或 HPRT⁺克隆的影响。当活细胞中重复片段缩短 30% (从 CAG₉₅ 到 CAG₆₁ 或 CAG₆₈) 时,效率降低了 2-3 倍[22]。长片段的这一倾向在其它的研究中也有所体现,ZFNs 对于 CAG₁₀₂ 的切割效率很高,但是无法造成同一细胞中 CAG₁₂ 重复序列的不稳定性[95]。显著性的长片段依赖显示出典型的 TNR 长片段疾病可能更适于该类靶向疗法。

之前研究中对 APRT⁺和 HPRT⁺活细胞克隆中 CAG 的分析发现克隆中缩短后的片段几乎全部少于 40 个重复单位[27,35,94]。令人意外的是,对于克隆中重复片段的分析发现只有 55% 的克隆具有所期望的缩短片段;其余克隆中的 CAG 片段或是删除或是插入外源片段,从而阻止其被排斥于 mRNA 之外。片段的删除和插入说明 ZFN 能够诱导 DSBs,但是同时提高了 TNRs 精确删除的潜在风险,因为 DSBs 也有可能造成重复片段之外的未知基因序列的改变。

在另一项研究的结果中发现能够靶向于 CAG 的 ZFNs 具有选择性缩短 CAG 重复序列但不引起 DSBs 的潜在途径。Liu 等人设计了两组件的和上文中类似的 ZFN，只是利用 zfAGC 代替了 zfGCA[95]。这一改变使得 ZFN 的两个组件位于最为合理的四核苷酸区间，从而能够进行效率更高的切割（图 3B）[89,90]。正如同所期待的那样，zfAGC 和 zfGCT 的复合体能够对具有 CAG₁₀₂ 的重复序列（转染时间在第 0 天和第 3 天，分析时间在第六天）的 PCR 产物进行高效的切割，能够消除几乎所有的 CAG₁₀₂ 等位基因[95]。

Liu 等人利用小通量 PCR 分析对 CAG₁₀₂ 重复序列进行检测。在原理上来讲，这一技术能够检测所有的突变——包括缩短和延伸、删除和插入。然而实际上，具有更少 CAG 单位的 PCR 产物更易于缩短，因此，缩短和删除发生的频率要高于延伸和插入。基于这一特性，ZFN 切割产物较之亲本 CAG₁₀₂ 大多更短，更易于发生片段缩短[95]。因此能够检测出 ZFN 处理后的细胞等位基因修复所产生的全部形式，包括缩短、延伸、删除和插入，这对于评估 ZFN 靶向性 DSBs 作为 TNR 疾病的一种临床治疗手段的可行性具有重要的价值。

然而最令人感到惊讶的结果是独立部件的获得[95]。与 Mittelman 等人所采用的独立部件不同，Liu 等人所采用的无论是 zfGCT 还是 zfAGC 都不能切割二聚化的 DNA。然而，它们能够切割由 CTG 和 CAG 发卡结构形成的重复片段（图 3C）。这一发卡结构切割活性最早被发现于具有 CAG₁₀₂ 重复片段的 PCR 扩增产物，该研究发现多种类似结构均对滑链错配所产生的多个发卡结构具有很高的亲和力[96]。与每个单独的组件共同孵育都会引起发卡结构的消失且对于双链 DNA 主要的结构不会产生明显的影响。这些现象在 CAG 或 CTG 重复链上更为明显，进一步研究结果提示 zfGCT 能够特异性切割 CTG 发卡结构而 zfAGC 能够特异性的切割 CAG 发卡结构。

这些结果在结构学角度解释更为明确。一般情况下，每个锌指结构域能够和 DNA 上三个相邻碱基的大沟结合；然而，然而，CTG 和 CAG 发卡结构正位于每个三碱基错配碱基的主要大沟上（分别为 T: T 和 A: A）。很明显，这些错配在最低限度下改变了同型二聚体 zfGCT 和同型二聚体 zfAGC 与 TNR 发卡结构之间的结合和切割能力[95]。

发卡特异性试剂为探索 TNR 不稳定性提供了可能性。Liu 等人发现当独立性部件 zfGCT 和 zfAGC 转染到细胞中时，能够缩短 CAG₁₀₂ 的长度，当将两种组件进行混合转染时缩短程度更高[95]。与 zfGCT 和 zfAGC 的混合结果比较结果不同，与亲本的 CAG₁₀₂ 重复区域相比，独立部件均能引起亲本重复片段的缩短。研究者认为细胞中发卡结构形成的原因是因为原有序列的相邻区域与原始结构发生错位。为了验证他们的假设，研究者发现血浆中不发生复制的 CAG₁₀₂ 片段对于独立性的组件不敏感，但能够被组件的混合物完全切割。Liu 等人并不认为重复片段的缩短是短片段产物产生的唯一原因，这只是众多可能性中的一种。假定发卡结构的切割只能够引起单链的损伤，那么单一组件的处理就能避免 DSBs，也可能是片段删除和插入的来源。从治疗的角度来讲，只进行发卡结构特异性切割的 ZFNs 相较于能够实现完全性切割的 ZFNs 具有两种优势。首先，它们产生的片段缩短具有独立性，而不像是 ZFN 诱导的 DSBs，在进行片段删除的同时还会发生具有毒害作用的片段扩增，这一作用有可能造成患者疾病表型的进一步恶化。其次，发卡结构特异性的 ZFNs 自然状态下更倾向于靶向相对与发病较为密切的较长的片段，因为长片段形成的发卡结构相对稳定[97,98]。发卡结构特异性的 ZFNs 的一个缺陷是需要 DNA 复制为其切割提供条件[95]，这一缺陷使得它们可能不适合不再进行分裂的成熟分化细胞的处理。然而在这样的细胞中，重复片段转录的过程中或许也会有发卡结构的产生[21,27,31,35]。毫无疑问，发卡结构特异性的 ZFNs 需要在体内进一步的检验其效果并确定其作用特征。

Moye 等人采用了一个不同的方法——开发了一种锌指缺口酶（ZFNickases），试图只对单链造成损伤从而避免产生 DSBs 引起相关潜在性的问题[99]。通常来讲，ZFNickases 的构建需要 FokI 失活其两个功能域中的一个组分的切割活性，从而保证只有一个功能域发挥了切

割活性。除了诱导产生这样的切割致死 (cd) 突变之外, 对 CAG 重复片段进行 ZFNickases 的设计还需要修饰消除独立组件二聚化产生的间隙[22]。通过在 FokI 二聚化区域诱导产生突变, Moyer 等人还合成可无法同源二聚化的 zfGCT 和 zfGCA 的变构体, 也就是所谓的“RR” (D483R) 和“DD” (R487D) (图 3D) [100]。

为了检测不同组分的活性, Moyer 等人采用了一种具有 CAG₅₄ 重复片段的 582bp 的超螺旋微环 (图 4a)。每个锌指部件都在体外进行转录和整合, 并与超螺旋 DNA 微环进行孵育。将独立的未修饰组分 (野生型) 的缺口部分连接形成混合连接物 (图 4b, 跑道 1 和跑道 2) [99]。连接产物的存在排除了 zfGCT 和 zfGCA 特异性的靶向于 CTG 或 CAG 发卡结构只产生缺口螺旋的可能, 因为损伤有可能被受限于单链。Moyer 和 Liu 等人有关于 zfGCT 的不同行为的发现有可能影响锌指组件构建过程中实际锌指功能域的设计。

与原有期待相符, 独立修饰的组件——zfGCA-RR、zfGCT-DD、zfGCA-DD 和 zfGCT-RR 没有对超螺旋微环进行切割 (图 4b, 跑道 4,5,12,和 13)。修饰后组件进行合适的组合 (zfGCA-RR 和 zfGCT-DD 的组合以及 zfGCA-DD 和 zfGCT-RR 之间的组合) 能够对混合连接和带有缺口的微环进行切割 (跑道 6 和 14)。修饰后的组件对于微环的切割效率是野生型组件的三分之一, 但是能够将微环的连接序列变得更短并移除大部分的 CAG 重复序列 (图 4b, 跑道 3) [99]。在其他的专一性异二聚体 ZFNs 中也发现了切割效率的降低[101]。

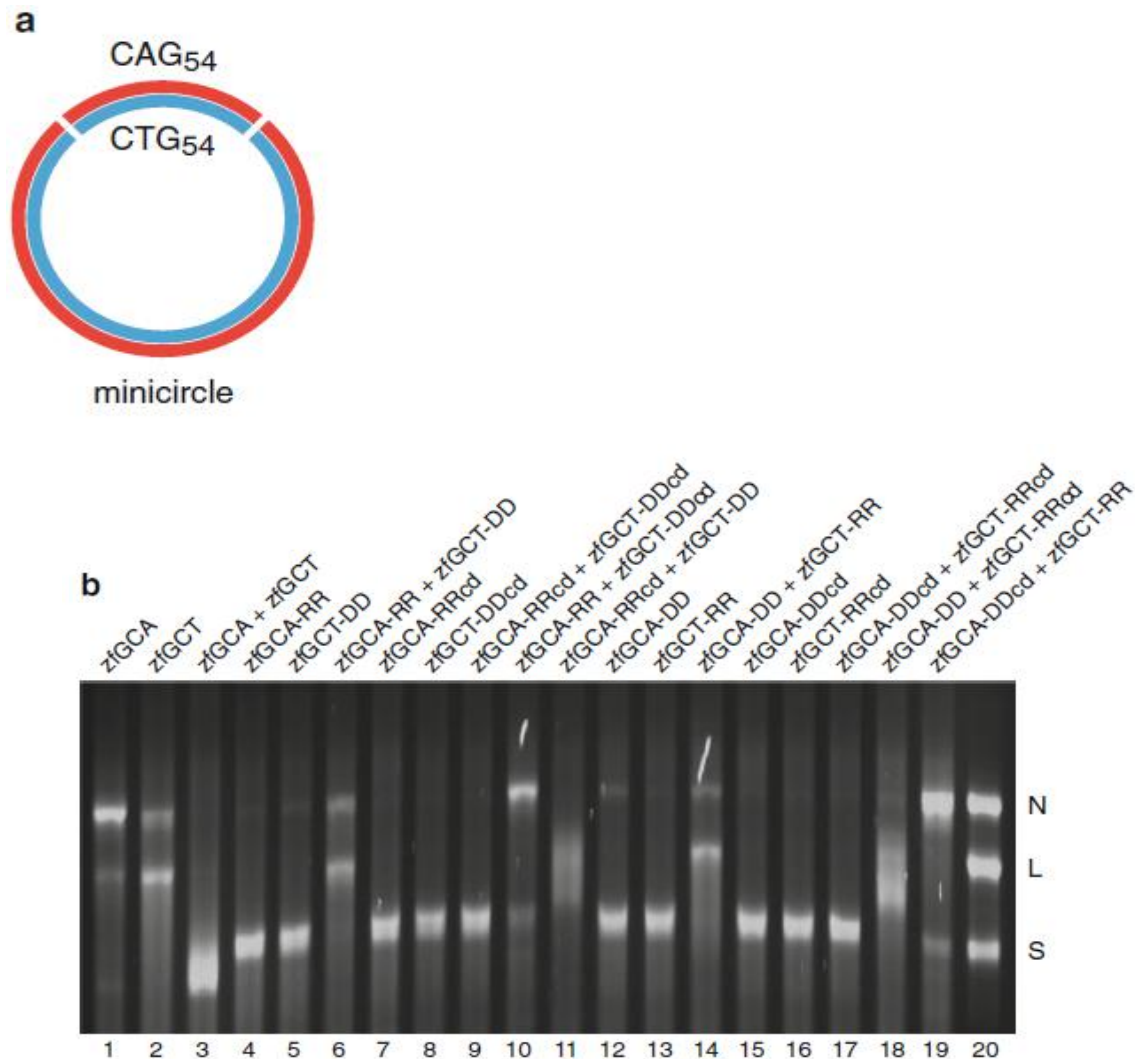


图4. ZFNs和ZFN缺口酶的体外实验活性评价。(a) 具有CAG54TNR的微环。(b) 不同ZFN组件的切割。在体外对于ZFN组件进行转录和翻译并于微环进行共通孵育60min, 产物通过琼脂糖凝胶电泳进行分选。缺口环(N), 线性环(L)和超螺旋(S)。线性链的迁移在超螺旋微环之后; 缺口链的迁移在线性链之后。

为了产生 ZFN 缺刻酶, 对应于专属异二聚体突变, 在 FokI 激活区域也产生了相应的突变 (D450A)。与预期相一致的是, 无论是独立切割致死组件还是切割致死组件的混合物都无法切割微环 (图 4b, 跑道 7,8,9,15,16 和 17)。与之相反, zfGCA-RR 和 zfGCT-DDcd 的组合 (跑道 10) 和 zfGCA-DDcd 和 zfGCT-RR 的组合 (跑道 19), 能够切割几乎所有的超螺旋形成缺口环。出人意料的是, 类似结构的 zfGCA-RRcd 和 zfGCT-DD 组合 (跑道 11) 与 zfGCA-DD 和 zfGCT-RRcd (跑道 18) 能够产生一条全新的条带, 被证明是一种非互补双链, 其中一条链相对于另一条链重复片段明显缩短[99]。这一结果与缺口产生的比例相一致, 相对于缺刻, 缺口的产生速度要提高 3-4 倍[99]。这一结果说明 RR 修饰能够降低 FokI 切割结构域的活性。

Moye 等人利用 GFP 试验[102]解决了 ZFN 缺刻在 CAG 重复序列中不稳定的关键性问题, 该试验与 Mittelman 等人所采用的 HPRT 和 APRT 试验相类似[22,99]。其中最主要的差别在于具有更短的 CAG 重复序列的细胞能够进行排序而非细胞分选。利用野生型 ZFN (zfGCA 加上 zfGCT) 能够提高细胞 GFP+强度 9.3 倍。利用活性更强的 ZFNickase (zfGCA-DD 加 zfGCT-RRcd 的组合)能够提高细胞 GFP+强度 3.5 倍, 然而利用活性较弱的 ZFNickase (zfGCA-DDcd 加 zfGCT 加 zfGCT-RR) 并没有引起背景细胞荧光强度的增加[99]。值得注意的是, 所有利用高强度 ZFNickase 进行处理过的细胞克隆中的重复片段都会缩短。这些结果提示, ZFNickases 与发卡结构特异性的 ZFNs 相似, 有可能改变了 ZFNs 介导 CAF 重复片段发生 DSBs 的能力。

最近, 有一种 TALEN 被证实能够高效切割酵母染色体中的 CAG 重复片段[23]。TALEN 的一个组件被设计用于锚定一段独特的序列与重复序列相邻的序列, 而其余部分则直接作用于重复序列本身。当细胞表达开始时, TALEN 能够高效地切割重复片段, 产生完全性的片段缩短。而且, 只在所设计靶向位置能检测到片段的缩短。虽然 DSBs 能够造成人类细胞基因组的插入和扩增[22,95], TALENs (和 CRISPR/Cas 核酸酶) 也能被修饰用于产生缺刻。

DNA 直接性疗法的缺陷和挑战

能够引起疾病的 TNR 重复序列是编辑化核酸酶治疗的巨大挑战。无论疾病是隐形的还是显性的, 理想化的疗法都应该能够特异性的靶向于扩增的等位基因并在基因组水平上缩短 TNR。关键性的需求是如何特异性的对扩增的重复性等位基因进行操作。这种疗法没有必要一定将扩增重复片段完全缩短至正常水平。即使是小范围的缩减也具有治疗价值, 由于疾病的严重程度和发生和重复片段的起始长度不具有简单的功能相关性, 但是一种长度依赖性的体细胞的功能不稳定性的发展也能够对组织产生影响, 例如大脑皮层和纹状体[82,103,104]。由于体细胞扩增的比率依赖于 TNR 的长度, 重复片段的减少能够轻微降低体细胞扩增的频率, 从而能够延缓疾病的发生并降低其严重程度。永久性的缩短扩增的重复等位基因长度能够显著地改善 TNR 异常所引起的症状。

一个关键性的 TNR 疾病所独有的问题是引起疾病应答的同一个重复位点也可能存在于人类基因组的其他位置, 并且具有不同的长度和纯度; 没有一例疾病相关的 TNR 序列是独特的[3]。因此, 治疗时作为标靶的 TNR 序列的破坏对其他的基因位点来讲必须是友好安全的。在基因组中最常见的 TNRs, 其长度经常不足以进行有效的靶向。在人类基因组中有低于 20%的 CAG 重复序列不超过 9 个单位, 这是 zfGCA/zfGCT 能够切割的极限, 只有 6 个片段长度超过了 20 个单位[105]。虽然对于这些 TNR 靶向性的编辑化核酸酶, 在自然状态下室友

很少的相关位点，理解这些位点突变的产生对基因组的影响仍是必要的。即使是进行独特基因组位点修饰过的核酸酶也无法避免脱靶 DNA 损伤的产生[106-108]。

新一代测序技术为核酸酶在靶位点介导的序列的改变提供了迅速和性价比高的测量方法，并可对大范围的脱靶事件发生的频率进行定量并探究随后的突变修复。目前这种新的成熟的测序平台具有很多，也有一些平台正处于发展中。今天，测序手段在成本、速度、通量、测序长度、犯错率和偏差方面都有了许多改进[109]。对 TNR 分析来说，最重要的因素就是读值的长度，如果测序的范围没有完全覆盖 TNR 的长度，那么就很难精确定义位点的基因表型。除此之外，在过去的几年中，数据分析方法也得到了极大的提高，能够精确检测点突变和小片段的插入[110]，以及例如微卫星变异等更为复杂的改变[111-113]，和更加大型的结构改变[114]。新一代基因技术的实力在最近有关于 HeLa 基因组的不稳定性测序和重构中得到了证明，在多个基因组位点都发现了大比例的基因突变[115-117]。而且，一些高水平扩增的三碱基重复序列，即使是应用传统 Sanger 测序法无法检测到的一些序列，在目前水平下也可进行研究，例如 PacBio 长片段测序平台[118]。类似的平台为重复序列的扩增和缩短提供了快速测量的可能性，但是仍然无法解决外源片段插入和目标片段之外的基因的删除。

总结和展望

对疾病相关的 TNRs 直接性的攻击为人工设计的重复片段靶向的核酸酶发展提供了巨大的可能。虽然直到目前为止，只有 ZFNs 被用于人类细胞，且只靶向于 CAG 重复序列，但是这类研究为 TNRs 改造提供了一种高效的方法，尤其是在重复片段缩短方面。目前，利用人工编辑核酸酶解决 CAG 重复扩增有三种方法：介导冲服位点产生 DSBs；在 CAG 链或 CTG 链产生缺刻；切除重复片段处所产生的 CAG 或 CTG 的发卡结构[22,23,95,99]。虽然每一种核酸酶所产生的全部事件还无法很细致的进行研究，可以肯定的是只在一条链上发挥作用的 ZFNs，例如 ZFNickases 和发卡特异性的 ZFNs，可以在缩短重复片段的前提下不引入额外复杂的片段删除和插入。因此，有研究发现经人工编辑过的核酸酶能够产生所需的有效的治疗作用——即缩短重复片段。

然而和独特的靶位点不同，TNRs 在基因组中具有多个拷贝，使得基因组会受到脱靶的 TNR 特异性的编辑核酸酶的影响。对于 CAG 特异性的 ZFNs，切割活性所依赖的长度范围会限制脱靶效应的产生[22,95]，因为大多数 CAG 重复片段都很短[105]。除此之外，有大约四分之三的 CAG 重复片段位于外显子之外[105]；因此，即使发生意外性的改变也大多不会影响细胞的存活和功能。与此不同，TALENs 和 CRISPR/Cas 核酸酶能够被用于设计识别靶向区域内独特的片段，并在目标基因位点进行选择性切割[23]。这一策略对于 ZFNs 有些困难，因为锌指结构域目前还无法被设计用于识别结合几乎所有的 DNA 重复片段。除了已有的策略之外，有必要对人工设计的核酸酶对基因组范围内的 TNRs 的敏感性进行确证，这一任务可通过新一代的测序技术解决。只有在这之后才有可能对于人工编辑的核酸酶介导重复片段缩短的 TNR 疾病治疗方法是否可作为一种有效的治疗手段进行正确的评价。

在这篇回顾中我们的着眼点主要在于人工设计核酸酶和利用这一工具进行疾病相关 TNRs 片段缩短的能力和限制。然而我们忽略了一个更大的挑战：即如何将这一技术用于人体。假设我们有能力构建具有理想性能的人共涉及核酸酶——具有疾病相关 TNR 序列高效的切割活力、具有片段缩短的倾向性、以及在基因组非目标为点最小化的切割活性，但是我们要如何才能将这一系统导入人体呢？这一技术在细胞系、精子、卵细胞、受精卵或早期胚胎中的应用能够在子代及后辈中排除疾病的发生，但是还需要面临无法克服的伦理和社会问题。另一种途径是治疗患病的个人。大多数 TNR 疾病只与有限的细胞或组织有关。因此，有可能利用一些局部给药途径在患病细胞中利用细胞或组织特异性的载体或细胞种类特异性的核酸酶进行表达。在原则上来讲，这一方法能够保护处于危险中的细胞并缓解疾病的症

状。这种技术也不会改变生殖细胞系使得患者的后代产生风险。然而更严峻的是，在第二代细胞群落中 TNR 的效应，由于在初代细胞中所显现出的缺陷，有可能造成新的疾病，并需要进行额外的治疗。毫无疑问的是，这一疗法的发展还需要克服很多困难，但是编辑化核酸酶本身也开始显示出其巨大的应用前景。

致谢 作者感谢 Shondra Preuett-Miller 博士和 Matthew Porteus 为构建 ZFN 核酸酶所做出的努力；Lynn Zechiedrich 博士、Jonathan Fogg 和 Vincent Dion 为构建 DNA 微环所付出的努力；以及 Beatriz Santillan 博士、Yi Zhang 和 Ramon Roman-Sanchez 在多个实验上的帮助。这一工作由 NIH 基金的 JHW (GM38219 和 EY11731) 项目和 T-32 博士后培训基金的 CM (EY07001) 项目提供支持。

参考文献

1. Richard GF, Kerrest A, Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol Mol Biol Rev.* 2008;72(4):686 – 727.
2. Rando OJ, Verstrepen KJ. Timescales of genetic and epigenetic inheritance. *Cell.* 2007;128(4):655 – 68.
3. Gemayel R, Vincens MD, Legendre M, Verstrepen KJ. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet.* 2010;44:445 – 77.
4. Legendre M, Pochet N, Pak T, Verstrepen KJ. Sequence-based estimation of minisatellite and microsatellite repeat variability. *Genome Res.* 2007;17(12):1787 – 96.
5. Subramanian S, Mishra RK, Singh L. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol.* 2003;4(2):R13.
6. Verstrepen KJ, Jansen A, Lewitter F, Fink GR. Intragenic tandem repeats generate functional variability. *Nat Genet.* 2005;37(9):986 – 90.
7. Kashi Y, King DG. Simple sequence repeats as advantageous mutators in evolution. *Trends Genet.* 2006;22(5):253 – 9.
8. Mittelman D, Wilson JH. Stress, genomes, and evolution. *Cell Stress Chaperones.* 2010;15(5):463 – 6.
9. Orr HT, Zoghbi HY. Trinucleotide repeat disorders. *Annu Rev Neurosci.* 2007;30:575 – 621.
10. Brouwer JR, Willemsen R, Oostra BA. Microsatellite repeat instability and neurological disease. *Bioessays.* 2009;31(1):71 – 83.
11. Pearson CE, Edamura KN, Cleary JD. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet.* 2005;6(10):729 – 42.
12. Albrecht A, Mundlos S. The other trinucleotide repeat: polyalanine expansion disorders. *Curr Opin Genet Dev.* 2005;15(3):285 – 93.
13. La Spada AR, Taylor JP. Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat Rev Genet.* 2010;11(4):247 – 58.
14. Gomes-Pereira M, Monckton DG. Chemical modifiers of unstable expanded simple sequence repeats: what goes up, could come down. *Mutat Res.* 2006;598(1-2):15 – 34.
15. Lin Y, Hubert Jr L, Wilson JH. Transcription destabilizes triplet repeats. *Mol Carcinog.* 2009;48(4):350 – 61.
16. Mirkin SM. Expandable DNA, repeats and human disease. *Nature.* 2007;447(7147):932 – 40.

17. Dion V, Lin Y, Hubert Jr L, Waterland RA, Wilson JH. Dnmt1 deficiency promotes CAG repeat expansion in the mouse germline. *Hum Mol Genet.* 2008;17(9):1306 – 17.
18. Tome S, Panigrahi GB, Lopez Castel A, Foiry L, Melton DW, Gourdon G, et al. Maternal germline-specific effect of DNA ligase I on CTG/CAG instability. *Hum Mol Genet.* 2011;20(11):2131 – 43.
19. Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, McMurray CT. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature.* 2007;447(7143):447 – 52.
20. Kovtun IV, Johnson KO, McMurray CT. Cockayne syndrome B protein antagonizes OGG1 in modulating CAG repeat length in vivo. *Aging (Albany NY).* 2011;3:509 – 14.
21. Hubert Jr L, Lin Y, Dion V, Wilson JH. Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. *Hum Mol Genet.* 2011;20(24):4822 – 30.
22. Mittelman D, Moye C, Morton J, Sykoudis K, Lin Y, Carroll D, et al. Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells. *Proc Natl Acad Sci U S A.* 2009;106(24):9607 – 12.
23. Richard GF, Viterbo D, Khanna V, Mosbach V, Castelain L, Dujon B. Highly specific contractions of a single CAG/CTG trinucleotide repeat by TALEN in yeast. *PLoS One.* 2014;9(4), e95611.
24. Nithianantharajah J, Hannan AJ. Dynamic mutations as digital genetic modulators of brain development, function and dysfunction. *Bioessays.* 2007;29(6):525 – 35.
25. Freudenreich CH, Kantrow SM, Zakian VA. Expansion and length-dependent fragility of CTG repeats in yeast. *Science.* 1998;279(5352):853 – 6.
26. Kim N, Jinks-Robertson S. Transcription as a source of genome instability. *Nat Rev Genet.* 2011;13(3):204 – 14.
27. Lin Y, Dion V, Wilson JH. Transcription promotes contraction of CAG repeat tracts in human cells. *Nat Struct Mol Biol.* 2006;13(2):179 – 80.
28. Kim N, Jinks-Robertson S. Guanine repeat-containing sequences confer transcription-dependent instability in an orientation-specific manner in yeast. *DNA Repair (Amst).* 2012;10(9):953 – 60.
29. Wierdl M, Greene CN, Datta A, Jinks-Robertson S, Petes TD. Destabilization of simple repetitive DNA sequences by transcription in yeast. *Genetics.* 1996;143(2):713 – 21.
30. Mochmann LH, Wells RD. Transcription influences the types of deletion and expansion products in an orientation-dependent manner from GAC*GTC repeats. *Nucleic Acids Res.* 2004;32(15):4469 – 79.
31. Lin Y, Leng M, Wan M, Wilson JH. Convergent transcription through a long CAG tract destabilizes repeats and induces apoptosis. *Mol Cell Biol.* 2010;30(18):4435 – 51.
32. Nakamori M, Pearson CE, Thornton CA. Bidirectional transcription stimulates expansion and contraction of expanded (CTG)ⁿ(CAG) repeats. *Hum Mol Genet.* 2011;20(3):580 – 8.
33. Jung J, Bonini N. CREB-binding protein modulates repeat instability in a Drosophila model for polyQ disease. *Science.* 2007;315(5820):1857 – 9.
34. Lin Y, Dent SY, Wilson JH, Wells RD, Napierala M. R loops stimulate genetic instability of CTG.CAG repeats. *Proc Natl Acad Sci U S A.* 2010;107(2):692 – 7.
35. Lin Y, Wilson JH. Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. *Mol Cell Biol.* 2007;27(17):6209 –

- 17.
36. Lopez Castel A, Cleary JD, Pearson CE. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol.* 2010;11(3):165 – 70.
37. Schweitzer JK, Livingston DM. The effect of DNA replication mutations on CAG tract stability in yeast. *Genetics.* 1999;152(3):953 – 63.
38. Rosche W, Jaworski A, Kang S, Kramer S, Larson J, Geidroc D, et al. Single-stranded DNA- binding protein enhances the stability of CTG triplet repeats in *Escherichia coli* [published erratum *J Bacteriol* 1996;178(23):7024]. *J Bacteriol.* 1996;178(16):5042-4.
39. Lopez Castel A, Tomkinson AE, Pearson CE. CTG/CAG repeat instability is modulated by the levels of human DNA ligase I and its interaction with proliferating cell nuclear antigen: a distinction between replication and slipped-DNA repair. *J Biol Chem.* 2009;284(39):26631 – 45.
40. Hou C, Chan NL, Gu L, Li GM. Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. *Nat Struct Mol Biol.* 2009;16(8):869 – 75.
41. Pelletier R, Krasilnikova MM, Samadashwily GM, Lahue R, Mirkin SM. Replication and expansion of trinucleotide repeats in yeast. *Mol Cell Biol.* 2003;23(4):1349 – 57.
42. Razidlo DF, Lahue RS. Mrc1, Tof1 and Csm3 inhibit CAG/CTG repeat instability by at least two mechanisms. *DNA Repair (Amst).* 2008;7(4):633 – 40.
43. Collins NS, Bhattacharyya S, Lahue RS. Rev1 enhances CAG/CTG repeat stability in *Saccharomyces cerevisiae*. *DNA Repair (Amst).* 2007;6(1):38 – 44.
44. Napierala M, Bacolla A, Wells RD. Increased negative superhelical density in vivo enhances the genetic instability of triplet repeat sequences. *J Biol Chem.* 2005;280(45):37366 – 76.
45. Hubert Jr L, Lin Y, Dion V, Wilson JH. Topoisomerase 1 and single-strand break repair modulate transcription-induced CAG repeat contraction in human cells. *Mol Cell Biol.* 2011;31(15):3105 – 12.
46. Bhattacharyya S, Lahue RS. *Saccharomyces cerevisiae* Srs2 DNA helicase selectively blocks expansions of trinucleotide repeats. *Mol Cell Biol.* 2004;24(17):7324-30.
47. Dhar A, Lahue RS. Rapid unwinding of triplet repeat hairpins by Srs2 helicase of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 2008;36(10):3366 – 73.
48. Kerrest A, Anand RP, Sundararajan R, Bermejo R, Liberi G, Dujon B, et al. SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. *Nat Struct Mol Biol.* 2009;16(2):159 – 67.
49. Pearson CE, Ewel A, Acharya S, Fishel RA, Sinden RR. Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum Mol Genet.* 1997;6(7):1117 – 23.
50. Manley K, Shirley TL, Flaherty L, Messer A. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat Genet.* 1999;23(4):471 – 3.
51. Schweitzer JK, Livingston DM. Destabilization of CAG trinucleotide repeat tracts by mismatch repair mutations in yeast. *Hum Mol Genet.* 1997;6(3):349 – 55.
52. Jaworski A, Rosche WA, Gellibolian R, Kang S, Shimizu M, Bowater RP, et al. Mismatch repair in *Escherichia coli* enhances instability of (CTG)_n triplet repeats from human hereditary diseases. *Proc Natl Acad Sci U S A.* 1995;92(24):11019 – 23.
53. Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP, Monckton DG. Pms2 is a genetic enhancer of trinucleotide CAG{middle dot}CTG repeat somatic mosaicism: implications for

- the mechanism of triplet repeat expansion. *Hum Mol Genet.* 2004;13(16):1815-25.
54. Jarem DA, Wilson NR, Delaney S. Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry.* 2009;48(28):6655 – 63.
55. Parniewski P, Bacolla A, Jaworski A, Wells RD. Nucleotide excision repair affects the stability of long transcribed (CTGⁿCAG) tracts in an orientation-dependent manner in *Escherichia coli*. *Nucleic Acids Res.* 1999;27(2):616 – 23.
56. Richard GF, Dujon B, Haber JE. Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. *Mol Gen Genet.* 1999;261(4-5):871 – 82.
57. Richard GF, Goellner GM, McMurray CT, Haber JE. Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex. *EMBO J.* 2000;19(10):2381 – 90.
58. Mittelman D, Sykoudis K, Hersh M, Lin Y, Wilson JH. Hsp90 modulates CAG repeat instability in human cells. *Cell Stress Chaperones.* 2010;15(5):753 – 9.
59. Hebert ML, Spitz LA, Wells RD. DNA double-strand breaks induce deletion of CTG.CAG repeats in an orientation-dependent manner in *Escherichia coli*. *J Mol Biol.* 2004;336(3):655 – 72.
60. Bowater RP, Jaworski A, Larson JE, Parniewski P, Wells RD. Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in *Escherichia coli*. *Nucleic Acids Res.* 1997;25(14):2861 – 8.
61. Reddy K, Tam M, Bowater RP, Barber M, Tomlinson M, Nichol Edamura K, et al. Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Res.* 2011;39(5):1749 – 62.
62. Gorbunova V, Seluanov A, Mittelman D, Wilson JH. Genome-wide demethylation destabilizes CTG.CAG trinucleotide repeats in mammalian cells. *Hum Mol Genet.* 2004;13(23):2979 – 89.
63. Debacker K, Frizzell A, Gleeson O, Kirkham-McCarthy L, Mertz T, Lahue RS. Histone deacetylase complexes promote trinucleotide repeat expansions. *PLoS Biol.* 2012;10(2), e1001257.
64. Lahiri M, Gustafson TL, Majors ER, Freudenreich CH. Expanded CAG repeats activate the DNA damage checkpoint pathway. *Mol Cell.* 2004 2004/7/23;15(2):287-93.
65. Pieretti M, Zhang FP, Fu YH, Warren ST, Oostra BA, Caskey CT, et al. Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell.* 1991;66(4):817 – 22.
66. Bidichandani SI, Ashizawa T, Patel PI. The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am J Hum Genet.* 1998;62(1):111 – 21.
67. Grabczyk E, Usdin K. The GAAⁿTTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner. *Nucleic Acids Res.* 2000;28(14):2815 – 22.
68. Soragni E, Herman D, Dent SY, Gottesfeld JM, Wells RD, Napierala M. Long intronic GAAⁿTTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia. *Nucleic Acids Res.* 2008;36(19):6056 – 65.
69. Osborne RJ, Thornton CA. RNA-dominant diseases. *Hum Mol Genet.* 2006;15 Spec No 2:R162-9.
70. Lee JE, Cooper TA. Pathogenic mechanisms of myotonic dystrophy. *Biochem Soc Trans.*

- 2009;37(Pt 6):1281 – 6.
71. Li Y, Jin P. RNA-mediated neurodegeneration in fragile X-associated tremor/ataxia syndrome. *Brain Res.* 2012;1462:112 – 7.
72. Willemsen R, Levenga J, Oostra BA. CGG repeat in the FMR1 gene: size matters. *Clin Genet.* 2011;80(3):214 – 25.
73. Lin Y, Wilson JH. Transcription-induced DNA toxicity at trinucleotide repeats: double bubble is trouble. *Cell Cycle.* 2011;10(4):611 – 8.
74. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al. Antisense transcription in the mammalian transcriptome. *Science.* 2005;309(5740):1564 – 6.
75. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The antisense transcriptomes of human cells. *Science.* 2008;322(5909):1855 – 7.
76. Wells RD. Non-B DNA, conformations, mutagenesis and disease. *Trends Biochem Sci.* 2007;32(6):271 – 8.
77. Sinden RR, Potaman VN, Oussatcheva EA, Pearson CE, Lyubchenko YL, Shlyakhtenko LS. Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J Biosci.* 2002;27(1 Suppl 1):53 – 65.
78. Yoon SR, Dubeau L, de Young M, Wexler NS, Arnheim N. Huntington disease expansion mutations in humans can occur before meiosis is completed. *Proc Natl Acad Sci U S A.* 2003;100(15):8834 – 8.
79. Shelbourne PF, Keller-McGandy C, Bi WL, Yoon SR, Dubeau L, Veitch NJ, et al. Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum Mol Genet.* 2007;16(10):1133 – 42.
80. Reyniers E, Martin JJ, Cras P, Van Marck E, Handig I, Jorens HZ, et al. Postmortem examination of two fragile X brothers with an FMR1 full mutation. *Am J Med Genet.* 1999;84(3):245 – 9.
81. De Biase I, Rasmussen A, Endres D, Al-Mahdawi S, Monticelli A, Coccozza S, et al. Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients. *Ann Neurol.* 2007;61(1):55 – 60.
82. Swami M, Hendricks AE, Gillis T, Massood T, Mysore J, Myers RH, et al. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum Mol Genet.* 2009;18(16):3039 – 47.
83. Lin Y, Dion V, Wilson JH. Transcription and triplet repeat instability. In: Wells RD, Ashizawa T, editors. *Genetic instabilities and neurological diseases.* 2nd ed. San Diego: Academic Press; 2006. p. 691-704.
84. De Biase I, Rasmussen A, Monticelli A, Al-Mahdawi S, Pook M, Coccozza S, et al. Somatic instability of the expanded GAA triplet-repeat sequence in Friedreich ataxia progresses throughout life. *Genomics.* 2007;90(1):1 – 5.
85. Tassone F, Hagerman RJ, Gane LW, Taylor AK. Strong similarities of the FMR1 mutation in multiple tissues: postmortem studies of a male with a full mutation and a male carrier of a premutation. *Am J Med Genet.* 1999;84(3):240 – 4.
86. Taylor AK, Tassone F, Dyer PN, Hersch SM, Harris JB, Greenough WT, et al. Tissue heterogeneity of the FMR1 mutation in a high-functioning male with fragile X syndrome. *Am J Med Genet.* 1999;84(3):233 – 9.
87. Wohrle D, Salat U, Glaser D, Mucke J, Meisel-Stosiek M, Schindler D, et al. Unusual

- mutations in high functioning fragile X males: apparent instability of expanded unmethylated CGG repeats. *J Med Genet.* 1998;35(2):103 – 11.
88. Kaytor MD, Burrignt EN, Duvick LA, Zoghbi HY, Orr HT. Increased trinucleotide repeat instability with advanced maternal age. *Hum Mol Genet.* 1997;6(12):2135 – 9.
89. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol.* 2001;21(1):289 – 97.
90. Handel EM, Alwin S, Cathomen T. Expanding or restricting the target site repertoire of zinc finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol Ther.* 2009;17(1):104 – 11.
91. Liu Q, Xia Z, Zhong X, Case CC. Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J Biol Chem.* 2002;277(6):3850 – 6.
92. Segal DJ, Dreier B, Beerli RR, Barbas 3rd CF. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5' -GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A.* 1999;96(6):2758 – 63.
93. Carroll D, Morton JJ, Beumer KJ, Segal DJ. Design, construction and in vitro testing of zinc finger nucleases. *Nat Protoc.* 2006;1(3):1329 – 41.
94. Gorbunova V, Seluanov A, Dion V, Sandor Z, Meservy JL, Wilson JH. Selectable system for monitoring the instability of CTG/CAG triplet repeats in mammalian cells. *Mol Cell Biol.* 2003;23(13):4485 – 93.
95. Liu G, Chen X, Bissler JJ, Sinden RR, Leffak M. Replication-dependent instability at (CTG) × (CAG) repeat hairpins in human cells. *Nat Chem Biol.* 2010;6(9):652 – 9.
96. Pearson CE, Wang YH, Griffith JD, Sinden RR. Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)_n. (CAG)_n repeats from the myotonic dystrophy locus. *Nucleic Acids Res.* 1998;26(3):816 – 23.
97. Petruska J, Arnheim N, Goodman MF. Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res.* 1996;24(11):1992 – 8.
98. McMurray CT. DNA secondary structure: a common and causative factor for expansion in human disease. *Proc Natl Acad Sci U S A.* 1999;96(5):1823 – 5.
99. Moyer C. Generation of Instability within Trinucleotide Repeats. PhD Thesis, Baylor College of Medicine. 2013.
100. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol.* 2007;25(7):786 – 93.
101. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol.* 2007;25(7):778 – 85.
102. Santillan BA, Moyer C, Mittelman D, Wilson JH. GFP-Based Fluorescence Assay for CAG Repeat Instability in Cultured Human Cells. *PLoS One.* 2014;9(11), e113952.
103. Dragileva E, Hendricks A, Teed A, Gillis T, Lopez ET, Friedberg EC, et al. Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol Dis.* 2009;33(1):37 – 47.
104. Wheeler VC, Lebel L-A, Vrbancic V, Teed A, Riele H, MacDonald ME. Mismatch repair

- gene Msh2 modifies the timing of early disease in HdhQ111 striatum. *Hum Mol Genet.* 2003;12(3):273-81.
105. Kozlowski P, de Mezer M, Krzyzosiak WJ. Trinucleotide repeats in human genome and exome. *Nucleic Acids Res.* 2010;38(12):4027 – 39.
106. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppl C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol.* 2011;29(9):816 – 23.
107. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol.* 2013;31(9):822 – 6.
108. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods.* 2011;8(9):765 – 70.
109. Ward RM, Schmieder R, Highnam G, Mittelman D. Big data challenges and opportunities in high-throughput sequencing. *Syst Biomed.* 2013;1:29 – 34.
110. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491 – 8.
111. Fondon 3rd JW, Martin A, Richards S, Gibbs RA, Mittelman D. Analysis of microsatellite variation in *Drosophila melanogaster* with population-scale genome sequencing. *PLoS One.* 2012;7(3), e33036.
112. Gymrek M, Golan D, Rosset S, Erlich Y. lobSTR: A short tandem repeat profiler for personal genomes. *Genome Res.* 2012;22(6):1154 – 62.
113. Highnam G, Franck C, Martin A, Stephens C, Puthige A, Mittelman D. Accurate human microsatellite genotypes from high-throughput resequencing data using informed error profiles. *Nucleic Acids Res.* 2013;41(1), e32.
114. Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics.* 2012;28(18):i333 – i9.
115. Landry JJ, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stutz AM, et al. The genomic and transcriptomic landscape of a HeLa cell line. *G3 (Bethesda).* 2013;3(8):1213-24.
116. Mittelman D, Wilson JH. The fractured genome of HeLa cells. *Genome Biol.* 2013;14(4):111.
117. Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, Martin BK, et al. The haplotypes-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature.* 2013;500(7461):207 – 11.
118. Loomis EW, Eid JS, Peluso P, Yin J, Hickey L, Rank D, et al. Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res.* 2013;23(1):121 – 8.
119. Paques F, Haber JE. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev.* 1999;63(2):349 – 404.
120. Owen BA, Yang Z, Lai M, Gajec M, Badger 2nd JD, Hayes JJ, et al. (CAG)_n-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat Struct Mol Biol.* 2005;12(8):663 – 70.
121. Salinas-Rios V, Belotserkovskii BP, Hanawalt PC. DNA slip-outs cause RNA polymerase II arrest in vitro: potential implications for genetic instability. *Nucleic Acids Res.* 2011;39(17):7444 – 54.

应用编辑核酸酶产生 HIV 抵抗细胞

George Nicholas Llewellyn , Colin M. Exline , Nathalia Holt ,and Paula M. Cannon

摘要 在基因和细胞基础疗法的发展中 HIV-1/AIDS 经常被认为是一种优先开发疾病，因为在当代治疗中这类疾病具有很高的负荷，患者需要终生服用抗逆转录病毒药物。编辑化核酸酶能够破坏特定的基因、也能够进行精确的基因编辑或是在靶向位点进行敲入。作为核酸酶基因破坏能力应用的开发，HIV-1 感染提供了一种理想化靶标 CCR5 基因。这是病毒进入 CD4+ T 细胞的最为常见的共受体。重要的是，CCR5 缺失是被发现能够被很好的耐受，因为有很高比例的人群都是自然状态下的 CCR5 等位基因缺失纯合子表型。因此，利用锌指核酸酶对自体 T 细胞的 CCR5 进行破坏是编辑化核酸酶在人体应用的首例。这一疗法的进一步应用包括前造血干细胞 CCR5 的破坏，对于 T 细胞中另一种 HIV-1 共受体——CXCR4 的破坏，或者在破坏后的 CCR5 位点引入抗 HIV 基因进行联合治疗。最后，编辑后核酸酶的基因破坏活性也可用于对患者细胞中整合的 HIV-1 基因组中的药物耐受相关基因的破坏，这一疗法将会提升药物治疗的效果从而完全杀灭病毒。

关键词 同源重组修复（HDR）• 造血干细胞（HSCs）• 非同源末端修复（NHEJ）• TAL 效应器核酸酶（TALENs）• 锌指核酸酶（ZFNs）

G. N. Llewellyn , Ph.D. • C. M. Exline , Ph.D. • N. Holt , Ph.D. • P. M. Cannon , Ph.D. (*)
Department of Molecular Microbiology and Immunology , Keck School
of Medicine, University of Southern California , 2011 Zonal Avenue,
HMR502 , Los Angeles , CA 90033 , USA
e-mail: gllewell@usc.edu; colinexline@gmail.com; nat@nathaliaholt.com; pcannon@usc.edu

© American Society of Gene and Cell Therapy 2016
T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental
Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_10

介绍

人类免疫缺陷病毒（HIV-1）能够引起严重的，终身性的感染，在未治疗人群中具有很高的致死率。虽然目前抗逆转录病毒联合疗法能够有效地抑制 HIV-1 的复制，但是无法最终治愈患者。这也意味着患者需要终生接受抗逆转录病毒药物的治疗，并需要承担巨额的经济负担及药物毒性、病毒耐受或“治疗疲劳”的风险。实际上，在美国有数据显示只有 25% 的 HIV 感染人群成功的接受了治疗并通过抗逆转录病毒药物实现了病毒的完全抑制[1]。因此，必须想办法改变控制策略并尝试真正治愈感染，包括运用细胞或基因疗法。

HIV-1/AIDs 在很长一段时间内都被认为是可进行基因疗法的干预的候选疾病，早期的研究试图通过细胞遗传学的手段产生“细胞内免疫”从而抑制 HIV-1 的复制[2]。一些最早期的基因治疗实验大多是和 HIV-1 相关的，传统上是采用整合性逆转录病毒载体使得抗 HIV 基因进行长时程的表达，例如反式显性 RevM10 蛋白和多种以 RNA 为基础的抑制剂（见 Peterson 等人的综述[3]）。最近，HIV-1 感染其本身被证实适合用于首次人体的编辑核酸酶的导入，可用于 CCR5 敲除锌指核酸酶（ZFNs）的导入[4]。除了 CCR5 的破坏之外，在抗 HIV 疗法中人工编辑核酸酶被进一步用于在 CCR5 损伤位点处插入抗 HIV 的基因从而起到联合治疗的作用。除此之外，核酸酶的基因破坏能力也可被用于灭活宿主细胞中整合入的 HIV-1 基因组，这也是移除潜伏性感染细胞的一个好办法。人工编辑核酸酶进行 HIV-1 治疗的应用演化将在本综述中进行讨论并总结于表 1。

利用 ZFNs 破坏 CCR5 共受体

人工编辑核酸酶和 DNA 修复

例如锌指核酸酶、TAL 效应器核酸酶（TALENs）、归巢内切酶和 CRISPR/Cas9 系统等人工编辑核酸酶都通过相同的原理发挥功能，即通过在 DNA 序列中靶向产生双链损伤（DSB），并通过 DNA 修复途径完成编辑。归巢内切酶和 Cas9 具有自然状态下的核酸内切酶活性，而 ZFNs 和 TALENs 则是通过分子设计与 DNA 结合域相连并通过同型二聚化的非特异性 II S 型核酸内切酶对目标区域进行切割，例如 FokI 限制性内切酶。并由此引发不同的细胞途径的 DSB 修复，包括非同源末端重组修复途径（NHEJ），最常见的修复结果是在基因损伤区域产生随机性的插入和删除（多态性），从而引发基因敲除的产生（图 2）。除此之外，DSBs 能够通过同源片段进行更为精确地修复，例如姊妹染色体。类似的直接性同源修复（HDR）能够将同源的“外来模板”序列，在人工编辑核酸酶起作用的同时，将相同的遗传信息拷贝到目标位点。这一过程的结果可以是小范围的基因编辑，例如在缺陷基因部位产生一个点突变，或是在 DSB 损伤位点引入更为特异性的长片段。通过这种方式，人工编辑的核酸酶能够被用于产生三种结果：基因破坏，基因编辑和基因敲入（图 2）。

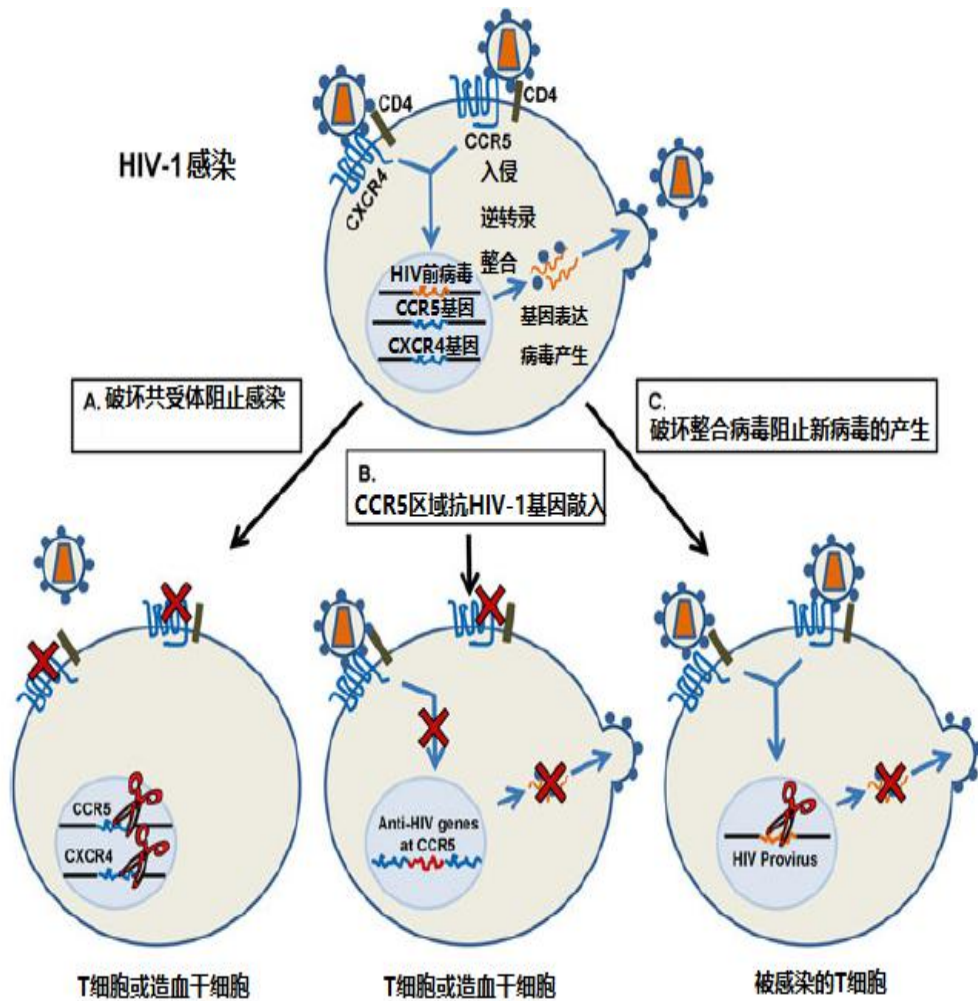


图1.抗HIV-1疗法中人工编辑核酸酶的潜在性应用。主要的HIV-1的主要生命周期如图所示。HIV-1通过和CD4及共受体蛋白——例如CCR5或CXCR4连接入侵宿主细胞，宿主的选择是由HIV-1表面的糖蛋白决定的。在入侵宿主细胞后，病毒的RNA基因组逆转录成DNA，并整合进入宿主细胞基因组。从这之后，整合后的前病毒作为一种细胞内基因起作用，转录形成mRNA并产生HIV-1蛋白，形成新的RNA基因组病毒，组装成病毒颗粒后通过细胞出芽的方式离开细胞表面。(a)人工编辑核酸酶被用于破坏CD4+ T细胞或它们的造血干细胞(HSC)祖细胞(只有CCR5)的共受体CCR5或CXCR4，从而阻止HIV-1的入侵。(b)额外的抗HIV-1基因被插入到损伤的CCR5位点，阻止HIV-1生命周期中其它阶段的进行。(c)人工编辑的核酸酶靶向破坏HIV-1序列，从而失活宿主细胞内的HIV-1整合基因组。

在早期，DSB的修复被认为是一种可预知的细胞通路。如果暴露的DNA末端被Ku70/80复合体所保护，那么就会启动NHEJ通路[5]。如果暴露的损伤末端为单链DNA，则HDR过程就会被启动[6]。到底是启动HDR还是NHEJ，这取决于细胞周期，HDR只发生于DNA复制后短暂的S期和G2期[6,7]，即修复所需的因子都处于激活(磷酸化)的状态[8]，而此时姊妹染色体相距位置最近，并更有可能成为同源性模板[9]。与此相反的是，NHEJ主要发生于G1期，虽然其在整个细胞周期都被激活。

造血干细胞中(HSC)是多种基因治疗应用的重要靶细胞，NHEJ相较于HDR发生更为

普遍[10,11]，在核酸酶的编辑结果中也更为常见。除此之外，在 HDR 介导的基因编辑中，DNA 外源片段的与核酸酶的共同导入是必须的，这一操作能够引起显著的细胞毒性。由于这些因素，在 HSC 中 NHEJ 介导的基因破坏过程更为容易。然而，基因破坏的治疗学应用收到了限制。在这一方面，CCR5 是人工编辑核酸酶在 HIV-1 疾病中的应用中的一个非常合适的靶标基因（图 2）。

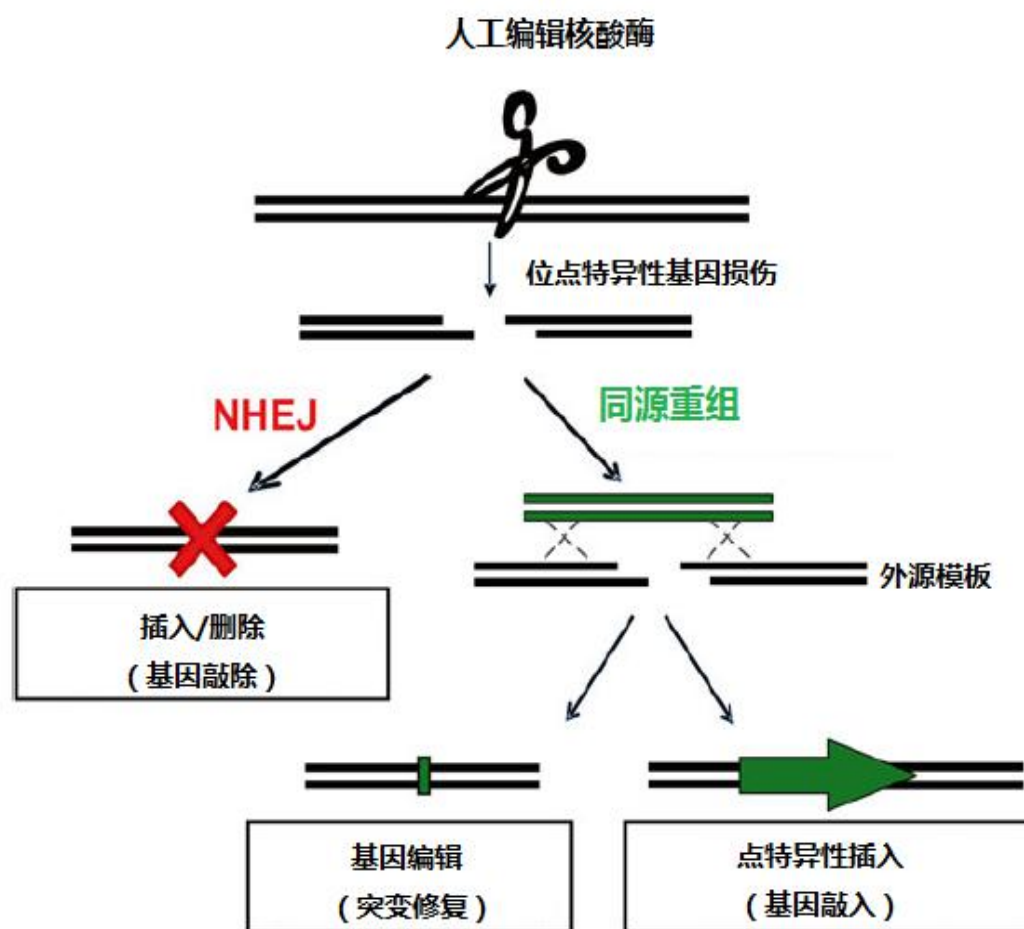


图2.人工编辑核酸酶激活触发的不同结果。人工编辑核酸酶在靶位点产生双链损伤（DSB）。如果非同源末端（NHEJ）修复被触发，最常见的结果是基因破坏。除此之外，DSBs能够通过同源重组进行修复，外源模板向细胞内的导入能够引导这一通路产生直接性的基因编辑，或促进DSB位点新的遗传基因的特异性插入。

CCR5 的理性化破坏——一种抗 HIV 疗法

CCR5 是一种趋化性受体，也是 HIV-1 进行感染时主要的共受体，与主要的受体——CD4 共同发挥作用[12]（图 1）。然而，这一功能在人类中并不是必要的，因为有很高比例的人群（~1%）是 CCR5 $\Delta 32$ 等位基因缺陷型纯合子。这一类人群对于 HIV-1 的感染完全耐受[13,14]且并未表现出显著性的表型缺失[13,15]。这些特征鼓励了一系列 CCR5 靶向性抗 HIV 药物的发展，其中包括了后来被用于 CCR5 治疗的 Maraviroc[16]。

除此之外，有人体数据表明 CCR5 阴性的细胞具有抵抗 HIV 感染的能力，也就是所谓的“柏林患者”。这一个体是一个 HIV 阳性白血病患者，并接受了 HSC 移植作为他白血病疗法的一部分，供体样本具有 CCR5 Δ 32 纯和突变。在接受治疗后，柏林患者保持了 HIV-1 阴性近 8 年的时间，也认为是第一例治愈了 HIV-1 的案例[17-18]。虽然在治疗过程中包含有激烈的侵蚀方法破坏白血病细胞，并显著减少了受感染细胞，但是供体的 CCR5 Δ 32 表型被认为具有关键性的作用，因为有一些其他的 HIV 患者接受了无 CCR5 Δ 32 供体的骨髓移植后疾病并未得到简单性的治愈[20,21]。这些发现提示在患者自身细胞内通过遗传学手段阻止了 CCR5 表达也能够控制 HIV-1 的复制。

在此之前拮抗 CCR5 的方法包括抗 CCR5 的 shRNAs 或核酶，传递系统包括整合逆转录病毒或慢病毒[22-27]。这种方法能够提供显著性的、但是不完全的抑制，需要被编辑细胞中抗 CCR5 元件终生性的表达。由于整合性载体的使用提供了这一可能，类似的载体基因表达在一段时间内会被沉默，整合性载体也会面临未知的突变产生的风险[28]。由于这些因素，利用人工编辑核酸酶进行 CCR5 的失活具有很大的价值。由于核酸酶的表达只需要是瞬时性的，这一方法具有潜在的安全性，不需要永久性或整合性的载体系统，但仍能引起基因的永久性改变[29-35]。无论是能够被 HIV-1 感染的 CD4+细胞还是 HSC 祖细胞，都是能够进行 CCR5 编辑的理想型细胞。

第一次临床实践：利用 ZFNs 在 T 细胞中进行 CCR5 基因的敲除

CD4+ T 细胞是 HIV-1 的主要感染靶标，虽然巨噬细胞等其他的 CD4+细胞也能够被病毒感染。T 细胞是能够进行基因操作的理想临床靶标细胞，在抗肿瘤领域中它们的应用已经积累了丰富的经验，已经建立了成熟的传代、扩增、基因操作、贮存和患者再灌注的流程[30,36]。而且，和最初的预期不同，经遗传标记的 T 细胞能够在体内存在至少 11 年，提示每个类似的修饰后 T 细胞都能够产生持久性的效果[37]。这些因素都使得经 ZFN 破坏 T 细胞的 CCR5 成为一种抗 HIV 的疗法，并被首先用于人工编辑核酸酶的临床实践。

利用 ZFNs 拮抗 CCR5 基因的想法首先由 Mani 等人提出[38]，他们分别建立了两对不同的 ZFN 靶向于 CCR5 的第二第七跨膜结构域（图 3）。每一对 ZFN 具有三个锌指单体，能够识别 DSB 位点的 9-bp 的目标序列。利用这些 ZFNs，作者能够在体外系统中破坏 CCR5 质粒 DNA。从这之后，多种研究显示 ZFNs 在多种人类细胞中能够起到很好的效果。很多研究利用 ZFN 对在近 160 核苷酸处开放阅读框产生了一个 DSB，该位点处于 CCR5 蛋白的第一个跨膜结构域，在每一个单体中包含有四个锌指结构[29-35]。很偶然的是，在一次基因敲除实验中，这一条件下最为常见的结果是发生 5bp 等位基因的插入，发生频率在 10%-30%之间，并在靶位点产生两个阅读框内终止子[31,33,35]。

ZFNs 引起的 CCR5 的破坏从而影响了 HIV-1 的复制这一现象在一系列不断升级的 T 细胞系[29,31]、主要的离体 T 细胞系[29-31,39]、移植了人 T 细胞的免疫缺陷鼠[29,31,39]中都被证实。在最早期的研究中，Perez 等人[31]发现在感染了 CCR5 依赖性的 HIV-1 之后 PM1 T 细胞系中 CCR5 的破坏率从 2.4%提升到 73%。作者进一步研究发现利用 Ad5/F35 慢病毒载体在主要的 T 细胞中导入 ZFNs 时 CCR5 的破坏率在 28%和 33%之间[40]。当修饰后的 T 细胞移植到 NOG 小鼠中并导入 HIV-1 时，与进行未编辑 T 细胞融合的小鼠相比，血浆病毒血症下降了 7.2 倍，在 HIV-1 感染后 CCR5 破坏的比率也提升了三倍（8.5-27.5%）。这些前临床试验证明利用人工编辑的核酸酶在正常及 HIV 感染后的人类 T 细胞中破坏 CCR5 基因具有可行性。多项基于 T 细胞的 HIV 感染者的临床试验也采用了 Ad5/F35 载体联用 160ZFN 对的方案（表 1）。

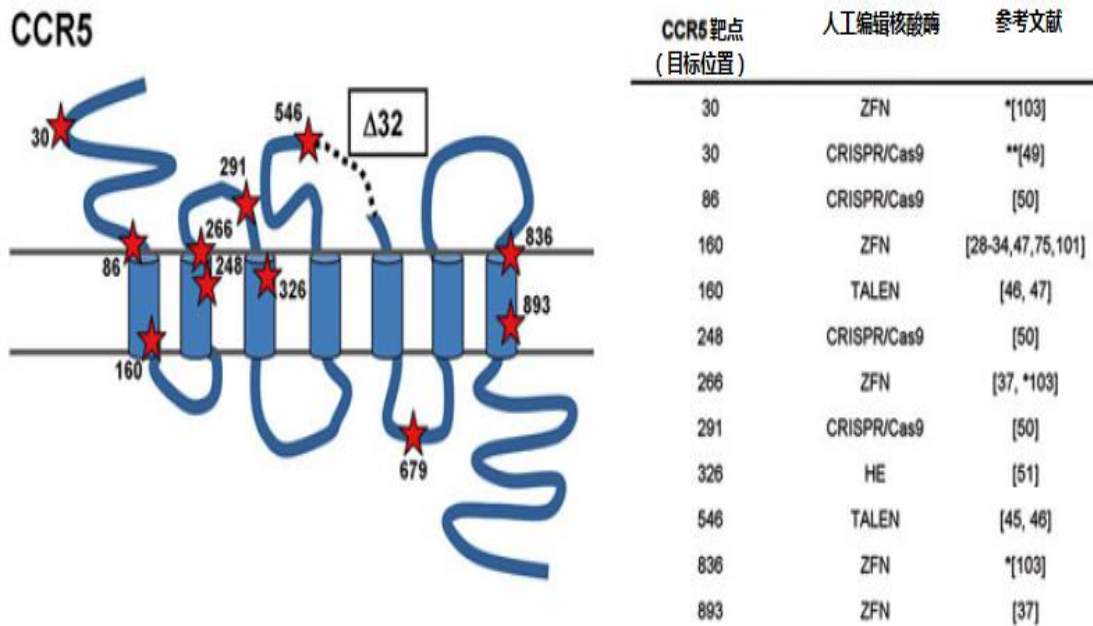


图3. 人工编辑核酸酶在CCR5的靶位点。在CCR5蛋白简图中一系列人工编辑核酸酶的作用位点的大致位置，更为详细信息请见图表。图中也显示出在32bp处的早期截短和CCR5 Δ 32等位基因的缺陷。*Kim等人[119]检测了315对ZFN在33个位点处的活性。其中有三处显示出高活性和低脱靶率。**Cho等人[53]检测了10种CRISPRs对CCR5的切割活性，其中一种被发现具有很高的活性和较低的脱靶率。HE，归巢内切酶。

ZFN 介导的 CCR5 破坏的早期临床试验主要集中于将 CCR5 ZFNs 导入患者自身的（自体同源）T 细胞，随后利用 CD3/CD28 进行刺激扩增。这一方法被报道能够产生最多 3×10^{10} 个 T 细胞，在给药 10 天后 CCR5 的破坏比率在 30-36%[30]。ZFNs 最初用 Ad5/35 载体进行转导，但是最近实验大多采用 mRNA 电穿孔的方式。第一个完整的临床试验——NCT00842634，12 名患者每人移植了 10^{10} 个 CCR5 编辑后的 T 细胞，编辑比例在 10.9%到 27.7%不等[4]。这一过程能够被很好的耐受，只有一例患者被报道在融合过程中产生了不良反应。CCR5 编辑 T 细胞能够持续存在数月，生命半衰期在 48 周左右。有 6 名患者在治疗后 4 周进行分析治疗中断（ATI），相关的逆转录病毒疗法（ART）在 12 周后中断，在剩下的 6 名患者中有 4 名患者完成了这一疗法。ART 的回撤 ATI 经常会引起 HIV-1 病毒血症的快速增高[41]，这也会对 CCR5 阴性 T 细胞产生选择性压力。除此之外，由于 HIV-1 的回升是一种典型的特征模式，ATI 为正确评估 CCR5 损伤的细胞对 HIV-1 的影响提供了机会。

在 ATI 期间，有一些抗 HIV 效应的症状显现。虽然在 ATI 期间，CCR5 修饰过和未修饰的 CD4 T 细胞数量均有下降，CCR5 修饰后 T 细胞下降的比率显著低于未经修饰的细胞（-1.81 细胞/平方毫米/天相对于 -7.25 细胞/平方毫米/天， $p=0.02$ ），虽然这一结果相较于平均值并不具有显著性（ $p=0.08$ ）。除此之外，有一名患者在 ATI 期间，HIV-1 RNA 的水平在 ATI 起始 6 周后也未见回升，甚至是在下一次治疗开始前就下降至检测限以下。由于这名患者是 CCR5 Δ 32 等位基因杂合子，很有可能是因为这“一半”的遗传背景促进了 ZFNs 介导的 CCR5 纯和阴性细胞的产生。这一假设在随后的 NCT01044654 临床试验中得到了进一步的验证，有 10 名 CCR5 Δ 32 杂合型患者接受了治疗（图 1）。为了进行更为深入的研究，最近的 T 细胞临床试验使用了环磷酰胺进行治疗（Cytosine）。这一前处理期望能在病人 ZFN 编辑后 T 细

胞融合前减少患者体内的 T 细胞，从而促使编辑后 T 细胞占有更大的比例。

表 1. 利用 ZFN 修饰后的自体 T 细胞治疗 HIV 患者的临床试验^a

临床试验	状态	队列/研究人群	细胞剂量 ^b ($\times 10^9$)	ZFN 递送方式, 细胞类型	其余治疗	登记项目
NCT008426 34	完成 ^c	1.多药耐药, 治疗失败	5-10	慢病毒, CD4 细胞	ATI (队列 2)	0
		2.ART 治疗, 病毒血症, CD4 统计 >459	5-10			6
		3.ART 治疗, 病毒血症, CD 统计 <500	5-10			6
NCT010446 54	完成	1.ART 治疗, 病毒血症, CD4 统计在 200-500	5-10	慢病毒, CD4 细胞	ATI (队列 5)	3
		2.ART 治疗, 病毒血症, CD4 统计在 200-500	20			3
		3.ART 治疗, 病毒血症, CD4 统计在 200-500	30			3
		4.多药耐药, 治疗失败	5-30			0
		5.CCR5 Δ 32 杂合型, ART 治疗	5-30			10
NCT012526 41	完成	未接受 ART, CD4 统计数 >500	5-30	慢病毒, CD4 细胞		N/A
NCT015431 52	招募中	1.ART 治疗, 病毒血症, CD4 统计数 >500	5-30	慢病毒, CD4 细胞 (除 3* 队列外, CD4/CD8 细胞)	200mg 环磷酰胺	3
		2.ART 治疗, 病毒血症, CD4 统计数 >500	5-30		0.5g/平方米环磷酰胺	6
		3.ART 治疗, 病毒血症, CD4 统计数 >500	5-30		1.0g/平方米环磷酰胺	3
		4.ART 治疗, 病毒血症, CD4 统计数 >500	5-30		2.0g/平方米环磷酰胺	3
		5.ART 治疗, 病毒血症, CD4 统计数 >500	5-30		1.5g/平方米环磷酰胺	3
		3*.ART 治疗, 病毒血症, CD4 统计数 >500	5-30		1.0g/平方米环磷酰胺 ATI (所有队列)	达到 8
NCT022256 65	招募中	1.ART 治疗, 病毒血症, CD4 统计数 >500	上限为 40, 两个剂量	mRNA , CD4/CD8 细胞	1.0g/平方米环磷酰胺 (所有两个队列)	3-6
		2.ART 治疗, 病毒血症, CD4 统计数 >500	上限为 40, 三个		ATI (所有两个队列)	3-8

			剂量			
NCT02388594	招募中	1.ART 治疗, 病毒血症, CD4 统计数 > 450	5-30 (预 计)	mRNA, CD4 细胞	环磷酰胺处 理 (队列 2 和 3)	最 大 至 15
		2.ART 治疗, 病毒血症, CD4 统计数 > 450				
		3.ART 治疗, 病毒血症, CD4 统计数 > 450			ATI (所有队 列)	

a 从网站 Clinicaltrials.gov 获得的信息, 经赞助公开报道

b 所有的 ZFN 修饰后细胞单剂量给药融合, 除非特意标注

c[4]

ART 抗逆转录病毒治疗, ATI 分析治疗中断, N/A 无法获得

利用 ZFNs 在 HSC 细胞中进行 CCR5 的破坏

人工编辑核酸酶也能够被用于在 HSC 细胞中进行 CCR5 基因的破坏。这些祖细胞能够分化成为多种类型的免疫系统细胞, 包括 CD4+ 细胞。虽然 HSC 相较于成熟 T 细胞来讲操作难度更大, 但是由于这类细胞生命周期较长能够起到更好的治疗效果, 由于 HSC 既能分化成为骨髓细胞也能成为髓性细胞系, 这也意味着能够产生 HIV-1 的非 T 细胞靶标, 例如巨噬细胞。

我们之前曾经报道 160 ZFN 对具有编辑人类 HSC 细胞的能力, 操作对象为分离自脐带血的 CD34+ 细胞[35]。通过 DNA 电穿孔转染 (核转染) 能够将 ZFNs 导入 HSC, 能够破坏 17% 的 CCR5 等位基因。编辑后的 HSC 能够被移植到免疫缺陷的 NSG 小鼠中, 对照组采用未经编辑的 HSC 分化诱导形成包括 CD4+ T 细胞在内的人免疫系统进行对照。利用 R5 趋向性的 HIV-1 进行感染后, 利用 ZFN 处理过的血液和淋巴系统的 HSC 相较于对照 HSC 能够更有力的抵御 HIV-1 的感染, 对照组的 CD4+ T 细胞快速减少。对感染后的 8-12 周小鼠的血液和淋巴系统进行分析发现 HIV-1 对 CCR5 阴性细胞具有很强的选择性。在 HIV-1 感染期间, 将未处理和 ZFN 处理过的 HSC 移植小鼠用相同剂量的 HIV-1 进行感染, 感染部位包括血液和组织。这一实验试图说明即使个体具有很少比例的 CCR5 阴性细胞, 也能够 HIV-1 感染中被选择, 并最终被选择使占比提高至足以影响病毒复制的频率。

虽然实验结果显示利用 ZFN 对 HSC 进行处理是对成熟 T 细胞进行编辑的一种可行方式, 但是仍需要方法以保证将核酸酶导入 HSC 细胞后能够最小程度的影响祖细胞进行分化的功能。除此之外, 虽然在前临床研究的 HSC 来自于人的脐带血或胎儿肝组织, 但自体 HSC 基因治疗的临床靶向细胞位骨髓残留的 HSC, 这类细胞与其它来源细胞仍会具有一些完全不同的特征[42,43]。成人的 HSC 分离主要通过外周血的细胞动员, 包括利用细胞因子 (G-CSF)、化疗药物 (环磷酰胺) 或小分子, 例如 CXCR4 拮抗剂——AMD3100[44]。我们之前曾经报道过在成人的 CD34+ HSC 细胞中利用 G-CSF 和与当今 T 细胞实验中相同的 Ad5/F35 位点的 160 ZFN 载体进行处理后, CCR5 等位基因的编辑比率能够达到 25% (表 1) [33]。然而, HSC 的慢病毒转染被证明相较于 T 细胞更具有挑战性, 利用 PMA 对细胞进行瞬时的低剂量处理能够将 CCR5 的破坏比例提升 5% 以上。由于更高水平的基因破坏比率经常伴有治疗毒性, 因而慢病毒载体也未必就是这类细胞最为优化的处理方式。

与之相反, 我们最近发现 mRNA 电穿孔转染是一种相对毒性较低、效率较高的将 ZFNs、TALENs 转入到 HSC 细胞中的方法, 能够在最低毒性下将 CCR5 的破坏比率提高到 50%[45]。这一方法也能被大规模应用, 能够一次满足病人治疗所需的全细胞剂量, 并因此提高这一过程的临床功用。

利用其它的人工编辑核酸酶进行 CCR5 的破坏

最近以 ZFN 对 CCR5 的编辑为基础的临床试验所取得的进步促进了基于其他的人工编辑核酸酶的 CCR5 特异性试剂的发展。由于和 ZFNs 相比 TALENs 的设计更为简易,使得 CCR5 靶向 TALENs 也得到了飞速的发展,这些核酸酶之间的差异包括 CCR5 靶向位点的不同以及 TALEN 蛋白的基础设计不同(例如,C-末端的长度)。Miller 等人[46]最早报道在 K562 细胞系中利用 C28 或 C63 骨架的 TALENs 能够分别引起 15-20%的 CCR5 的破坏。对应每种骨架结构,测试了多种 TALEN 的左右组合,在 DNA 识别区的每一个单体间设计了多个长度间隙。所有被评估的组合体都被设计靶向于 CCR5 的二次跨膜外区域片段,与自然状态下的 $\Delta 32$ 突变位点较近(图 3)。最近,我们发现 C17 骨架的 TALEN 也能够靶向于 $\Delta 32$ 附近位点,在 HSC 细胞中的破坏率能够达到 60%[47]。与之相似的是,一种 C17 骨架靶向于 157 位点的 TALEN,能够覆盖 160 ZFNs 的识别区域,经 Mussolino 等人的评估在 293T 细胞中能够达到 17%的 CCR5 基因破坏率[48]。有趣的是,这些研究者发现这一结构在同源性 CCR2 基因处的脱靶率低于 160 ZFN。在随后的研究中,对这一结构的毒性进行了更为严格的检测,作者也发现 TALEN 处理后的细胞具有更低的总毒性[49]。虽然 TALENs 和 ZFNs 都靶向于 CCR5 的相似位点,这两种试剂靶向识别的结构的不一致为这两种方法的比较带来了困难。而且尤其是在临床应用中,由于脱靶率和毒性分析对于候选核酸酶的筛选至关重要,这一分析也需要在临床靶细胞中进行,例如主要的人类 T 细胞或 HSC。最后,有报道称靶向于 CCR5 157 位点的 TALENs 能够通过 mRNA 电穿孔转染导入细胞,并可在 PM1 细胞系(最高可达 94%)和主要的 T 细胞(在 17%和 46.8%之间)中进行破坏[50]。

由不同种类的核酸酶引起的 DNA 损伤具有本质的不同,即使是用相同的 FokI 核酸酶介导的损伤。例如,在 43 项独立研究中在 122 个靶向位点对总共 1456 个突变序列进行了分析后, Kin 等人发现由 ZFNs 和 TALENs 造成的基因破坏有所不同。其中尤其显著的是,他们发现 ZFNs 在 DSB 位点能够产生更为频繁的插入和删除,而 TALENs 造成的删除经常更为纯粹[51]。造成这一差异的原因还未可知,作者推测在 ZFN 对的两个单体之间具有较小的空隔区域,因而所产生的较为一致的切割区域更易被插入,而 TALEN 两个单体之间的连接空隙更大,所产生的切割位点异质性更强,所造成的删除也就相对更为频繁。在早期实验中,广泛采用的 CCR5 位点 160 ZFN 对能够频繁的额引起损伤位点 5bp 的插入,在阅读框内形成两个终止子,从而中止 CCR5 蛋白的翻译。

归巢内切酶也被用于靶向 CCR5。这些核酸酶能够识别 18-34bp 的较长区域,对靶 DNA 的切割模式和限制性酶相近。有近百种不同特异性的归巢内切酶被报道,包括来源于 baker's 酵母的 I-Sce1,绿藻叶绿体的 I-Cre1。虽然它们在 DNA 链上的切割效率很高,但是它们很少具有像是 ZFNs 或 TALENs 那样易于进行识别编辑的特异性编码结构。这使得利用这些酶对 CCR5 进行重新定向更加具有挑战性。直到今天,只有一种归巢内切酶被报道能够用于 CCR5 的特异性切割,该种酶以 I-Cre1 为基础靶向于蛋白的第三段跨膜结构域(图 3)[52]。研究者利用已有的 I-Cre1 衍生物进行组装,使得原有 I-Cre1 的 24 个识别位点中的四个发生了改变。有趣的是,通过与一种 DNA 合成末期的酶——TREX2 核酸外切酶进行共表达后,能够将 HSC 中的 CCR5 破坏率从 5%提高到 37%,两种蛋白均采用慢病毒载体进行共表达。TREX2 的共表达也能够提高 ZFNs 和 TALENs 靶向 CCR5 的破坏率近三倍,提示这是一种提高基因破坏率的通用方法[52]。当前归巢内切酶的设计是限制其进行广泛应用的瓶颈,其中的小尺寸蛋白可通过病毒载体促进其转导,这也使得他们成为特定基因疗法中的一种颇具吸引力的工具。

最后,对于 CRISPR/Cas9 系统,由于其靶向位点的特异性是由互补 RNA 决定的,这类试剂的编辑、构建和评估都十分简易,这也使得它们可对于靶向位点进行高通量分析。例如,

Cho 等人[53]设计了 CRISPRs 靶向于 CCR5 的 10 个不同位点, 利用 CRISPR 对 K562 细胞 CCR5 第一跨膜结构域的破坏率达到了 60% (图 3)。与之相似的是, Cradick 等人发现 CRISPRs 在 293T 细胞中 CCR5 的破坏水平很高 (21-27%), 能够靶向于第二跨膜结构域和第二膜外结构域的 N-末端[54]。CRISPRs 被发现能够在 HSC 中起效。在最近的研究中, HSC 被转染了一个能够表达 GFP 的 Cas9 质粒。随后基于 GFP 的表达对 Cas9 表达细胞进行 FACS 分选, 并通过电穿孔转染导入 CCR5 向导 RNA 质粒。随后, 有 25.8-30% 的 HSCs 衍生克隆被发现具有 CCR5 等位基因的破坏[55]。CRISPR/Cas9 也能通过 Ad5/F35 或慢病毒载体靶向于 CCR5: Li 等人利用 Ad5/F35 载体在主要的 T 细胞中能够造成 CCR5 等位基因 30.3% 的破坏[56]。Wang 等人利用慢病毒导入系统在 CEMss-CCR5 T 细胞系中将 CCR5 阴性细胞的比率提高到了 43%[57]。

CXCR4 是基因敲除的另一靶标

基于人工编辑核酸酶的 CCR5 共受体失活通常有以下两个潜在限制: (1) 基因的两个拷贝都需要被破坏才能产生 CCR5 阴性细胞, 和 (2) HIV-1 感染特定的细胞株也可通过其它的共受体——例如 CXCR4, 进入细胞。X4 趋向性病毒和 R5 趋向性不同, 这类病毒出现于约 50% 的 HIV 晚期患者中, 而靶向于 CCR5 的疗法或药物能够加速这一选择过程。这一问题可通过 CXCR4 靶向性核酸酶和 CCR5 试剂进行联用得到解决, CXCR3 在 HSC 回归骨髓祖细胞中所起到的重要作用不可替代, 因而 CXCR4 的破坏只适用于 T 细胞而非 HSC。

在 T 细胞中靶向 CXCR4 被两个课题组发现。Wilen 等人发现经 CXCR4 ZFNs 处理的 T 细胞在体外能够免受 X4 趋向性 HIV-1 病毒的感染, 在移植有 CXCR4 ZFN 编辑的 T 细胞的 NSG 小鼠体内也具有部分的保护功能[58]。然而, 这以保护作用随着 R5 趋向性的 HIV-1 病毒的产生而逐渐失效。研究者也发现他们能够在纯合的 CCR5 Δ 32 个体中对 T 细胞进行进一步的编辑改造, 而这些细胞能够保护机体免受 R5 和 X4 趋向性的 HIV-1 细胞的感染。Yuan 等人进一步发现利用 ZFNs 对 CXCR4 进行编辑能够在体内和体外保护细胞, 而且在体内这种方法所产生的 X4 趋向性 HIV-1 的抵抗的效率要高于慢病毒载体导入 CXCR4 shRNA 的效率[59]。

Didigu 等人通过向单一 T 细胞系或主要的 T 细胞中导入两类靶向于 CCR5 和 CXCR4 的 ZFNs 将这一疗法向前推进了一步[29]。双重破坏的 T 细胞在经历了体外的 R5 和 X4 趋向性的病毒的感染后快速扩充, 证实了其保护性。与之前的实验相一致, 这以保护作用在体内实验中也被发现, 该活性通过移植了事先感染了 R5 和 X4 趋向性病毒 T 细胞的小鼠得到了证实。在该实验中, 在第 55 天时, 相较于未经处理的细胞或只接受过 CCR5 破坏的细胞, 接受了两种 ZFN 处理的 T 细胞的 CD4 数量提高了 200 倍以上[29], 这一结果进一步证实相较于只破坏 T 细胞中的 CCR5, 在 HSC 中破坏 CCR5 的同时破坏 T 细胞中的 CXCR4 能够进一步提高 CCR5 的破坏效果。然而值得注意的是, 虽然在这些模型系统中并未观察到毒副作用的产生, 在 T 细胞中破坏 CXCR4 的影响仍然未知, 在应用于人体之前这一方法仍需要进一步的测试。

除了基因敲除之外: 点特异性基因敲入与抗 HIV 疗法的联合

如上所述, HIV-1 适应除 CCR5 意外其它的共受体感染的能力会限制依赖于 CCR5 破坏的疗法的效果。在这里, 与 NIH 中进行 CCR5 抑制剂治疗药物相类似[60], 在治疗之前, 患者需要首先被筛选以确定没有感染 X4 趋向性病毒。除了这一限制外, CCR5 敲出的策略只在纯合破坏的细胞中有效, 虽然相对于野生型患者, CCR5 Δ 杂合型患者具有更好的预后[61-64]。由于这一限制, 利用人工编辑核酸酶在 CCR5 位点促进抗 HIV 基因插入的同源整合的策略能够提高单独破坏 CCR5 的效果 (图 1)。

有多种抗 HIV 分子, 在逆转录整合和慢病毒载体中被认为是方便 HIV-1 治疗基因, 也可被用于基因敲入的疗法。其中包括 HIV-1 的反式显性抑制结构域, 例如 RevM10 突变 [65,66] 及其修饰序列, 可作为细胞限制性因素的 HIV 抗性片段, 例如 TRIM5 α 和 APOBEC3G。

例如，虽然 HIV-1 在自然状态下对人源性的 TRIM5 α 会产生抗性，在恒河猴同源的 Rhesus 中受到抑制[67]，而人工编辑的人源化 rhesus TRIM5 α 的衍生品能够保持这一抗病毒活性[68,69]。与此类似，虽然 HIV-1 Vif 蛋白在正常情况下能够降解 APOBEC3G，这一胞嘧啶脱氨酶也能够在病毒基因组逆转录的过程中产生 G->A 的突变，而 D128K 点突变能够耐受 Vif，并因此抑制 HIV-1 的复制[70,71]。可作为 CCR5 位点插入候选片段的包括以 RNA 为基础的治疗工具，包括核酶或小片段干扰 RNAs (siRNA)，能够直接性的抵抗不同 HIV-1 的感染，例如 Rev[72]、Vif[73]、Nef[74]、Gag[75]、Env[76,77]或 Tat[72,73,78]，以及具有拮抗 HIV-1 融合潜力的 C46 肽抑制剂[79]。在 CCR5 位点插入三种不同的抗 HIV 基因（人源化 rhesus TRIM5 α ，APOBEC3G D128K 和 Rev M10）在细胞系模型中被证明具有可行性[80]。

对于 HIV-1 疗法等，临床中在早期人源化细胞中进行点特异性的基因敲入还需要进行效率的提高，这是一项极具挑战性的工作。在一项毒性实验中，Lombardo 等人报道利用三种非整合性慢病毒载体 (IDLVs) 转导各自对应的 GFP 同源重组到 CCR5 位点的效率可达到 50%。然而，在人源性 HSC 中进行同样的 GFP 整合，只能达到 0.11% 的效率。随后，相同的课题组利用五种腺病毒载体转导 CCR5 ZFNs，并联用 IDLVs 转导外源性模板，能够在主要的人类 T 细胞的 160 位点达到 5% 的靶向敲入率[82]。

我们在前期研究了 160 ZFNs 以质粒的形式导入 HSC 中的能力，结果对 CCR5 等位基因造成了 17% 的破坏率[35]。当一个 CCR5 同源模板质粒具有一个内在性的 PGK-GFP 表达功能域，并与 ZFNs 共同导入细胞，虽然在 DNA 介导的毒性上有所增加，剩下的 HSC 群落能够被移植到 NSG 小鼠中，并能在多种成熟的人类细胞系中将 GFP 的表达频率稳定提升至 5% (图 4a)。当人源化的小鼠被 R5 趋向性的 HIV-1 感染时，HIV 介导的 CD4 T 细胞的破坏能够对核酸酶编辑后的 CCR5 阴性细胞进行选择[35]。在这种实验条件下，当 GFP 插入到 CCR5 位点的时候，HIV 感染也会同时选在 CD4 细胞中具有 GFP 表达的小部分细胞 (图 4a,b)。这一观察现象与 GFP 功能域特异性的整合到 CCR5 位点相一致，这一结果也得到了 PCR 分析的验证 (图 4c)。

除了这些激动人心的发展之外，HIV-1 其自身也会作为一种筛选试剂，提高 CCR5 基因特异性敲入的频率，目前在 HSC 细胞中靶向敲入能够达到的水平无法匹配更为标准化的整合病毒载体介导的基因整合，还需要更效率的，尤其是在早期 HSC 中的靶向整合。最近利用非整合性慢病毒[83,84]或 AAV 载体[85]或其它的能够提高人类细胞中 HDR 相对于 NHEJ 比率同源修复模板导入方法正在实验中[86,87]。

在 iPSC 中进行 CCR5 的破坏

HSC 在不损伤分化能力的情况下在体外进行培养和扩增具有很大的困难，这限制了编辑后细胞在进行患者融合之前的初步基因组筛选。在未来，有可能提取出患者特异性的全功能 HSC 并诱导成为多功能干细胞 (iPSC)，从而也就能够产生更多的编辑后细胞的同源重组克隆。通过这一目标的实现，有多个课题组目前报道了 iPSC 的编辑情况：Ye 等人利用 CRISPRs 和 TALENs 精确产生具有 CCR5 Δ 32 突变的多功能 iPSC[88]，Yao 等人破坏了人类胚胎祖细胞和 iPSC 中的 CCR5，并将其分化成为 CD34⁺ 细胞，和分化程度更高的造血细胞系[89]。

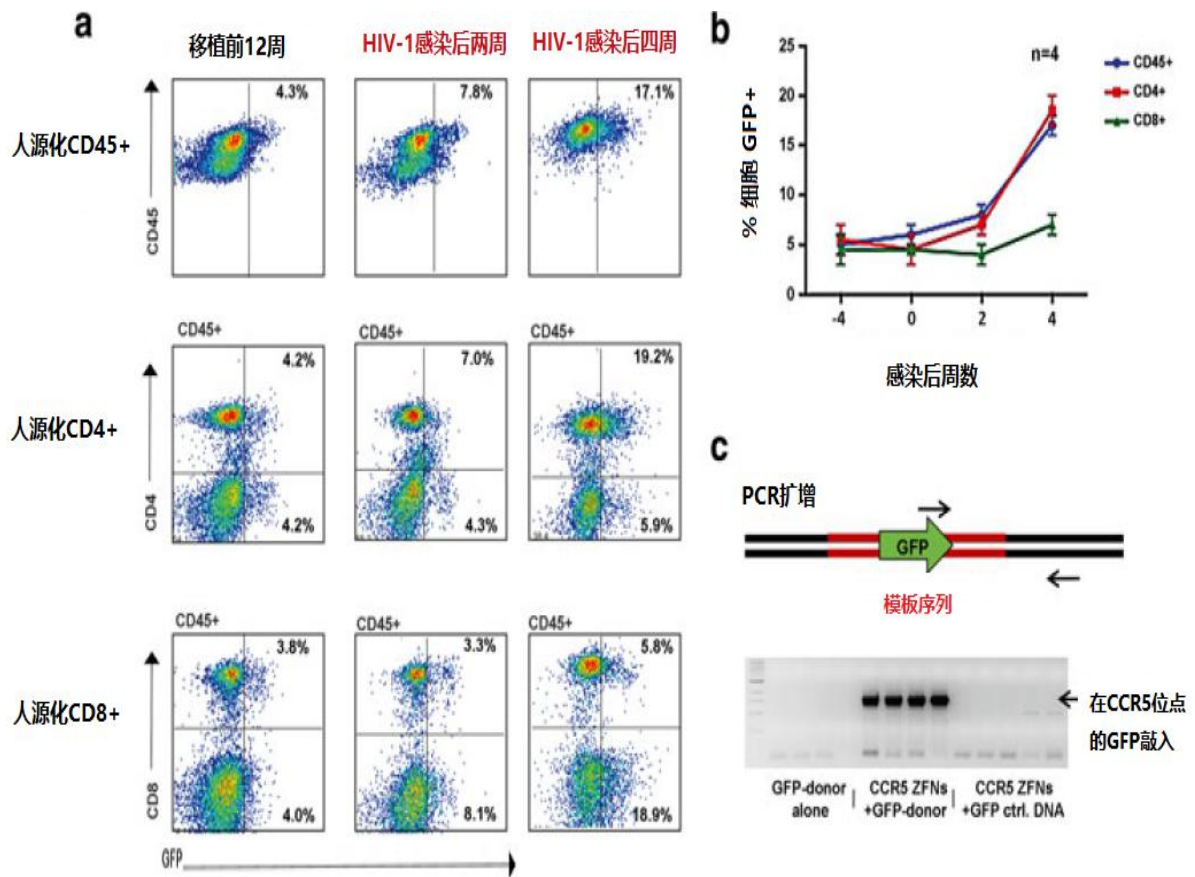


图4.在CCR5位点进行基因敲入。在人类脐带血HSC中电穿孔转染能够表达CCR5位点特异性的160 ZFNs和一个包含有PGK-GFP功能域且两侧具有CCR5同源序列的模板。将细胞如前述移植到新生的NSG小鼠中[35]。在12周龄时，使小鼠感染R5趋向性的HIV-1病毒。(a)在指定的时间点，利用FACS分析血液总人源性白血球(CD45细胞)、CD4和CD8群落中的GFP表达。(b)随着HIV-1的转染，在作为HIV-1宿主细胞的CD4和CD45中GFP表达提高，但是在CD8中没有这种现象。(c)在CCR5位点特异性敲入GFP功能域，并经过处理过后HSC的PCR扩增进行检验，PCR引物中的一条是GFP表达特异性的，一条是CCR5特异性的，但是超出了模板序列所在区域。

利用人工编辑核酸酶破坏 HIV-1 基因组

HIV 特异性的核酸酶发展过程

利用人工编辑核酸酶对抗 HIV-1 的另一项策略是利用人工编辑核酸酶去靶向 HIV-1 基因组本身(图 1)。虽然抗逆转录病毒药物具有很高的效率且能够抑制 HIV-1 至检测线以下，但是他们不能完全消除病毒，一旦停药反弹无可避免。这一来源的反弹病毒也就是所谓的潜在 HIV 存库，即在体内处于休眠期的病毒，无法被抗逆转录病毒真正杀灭，但是可以在今后的时间被激活。其中最为有名的存库是记忆 CD4 T 细胞[90-92]。真正在这些细胞中移除或失活这些病毒的方法应该能够引起药物非依赖型的 HIV-1 抑制，甚至是完全治愈。

有多项有关于不同人工编辑核酸酶的报道了能够识别 HIV-1 基因组的核酸酶[93-96]。在大多数的研究中，靶向序列为 HIV-1 的长末端序列(LTR)，存在于病毒整合性 DNA 片段的所有两个末端。LTRs 具有很多功能，5'序列驱动病毒基因组转录的进行，3'序列调控转录终止和聚腺苷酸化。完整的 LTRs 对于 RNA 的修正过程是必须的，能够成功地引起病毒生命周期中转录和整合的进行。除此之外，在 HIV-1 基因组中的两个 LTR 克隆为靶向序列的两端提

供了核酸酶，以及为双切除和重连接提供了可能，并能够导致 HIV-1 基因组永久性的切除。

最早的有关整合性 HIV-1 基因组破坏的报道是通过人工编辑的 cre 重组设计识别 HIV-1 的特异性序列[93]。这种酶常被用于识别 34bp 的 loxP 位点，随后被用来开发 HIV-1 病毒进化枝 A 的 LTRs 相似序列——TZB0003[97]。经人工改造过的重组酶，称为 Tre 重组酶，能够破坏 Hela 细胞中的整合性 HIV-1 基因组，并能够降低新的 HIV-1 颗粒的产生[93]。Tre 重组酶随后在体内模型中得到了发展[95]，转移了慢病毒载体的主要 T 细胞或 HSC 在 HIV 诱导启动子的作用下能够表达重组酶。实验所采用的 Tre 抗性的 HIV-1 LTR 启动子包含有串联的 HIV-1 Tar 组分，能够对 HIV 致病性的病毒反式激活子即 Tat 产生应答。随后，由于重组酶只会在感染了 HIV-1 的细胞中进行表达，因此能够降低潜在性的毒副作用。在这个概念验证实验中，转染后细胞有 30%到 60%移植到了 Rag2^{-/-}γc^{-/-}小鼠，随后用 R5 趋向性的 HIV-1 进行感染，使 Tre 能够识别改造后的 LTRs 序列。接受了 Tre 重组酶细胞移植的动物中的 CD45 和 CD4 细胞的水平稳定上升，与感染后 2 周相比，感染后 12 周的病毒水平下降了大约 1-log。接受了非转导细胞移植的小鼠与之相反，人源性 CD45 和 CD4 细胞水平有所降低，而同一时期的 HIV-1 水平趋于稳定或上升。然而更为有趣的事，技术的更为广泛的应用需要重组酶识别区域的发展，需要进行人工改进使得识别区域相对于自然状态下的 loxP 位点更为发散且在多种类型的 HIV-1 病毒中高度保守。类似的人工编辑的挑战能够促进新的设计工具的发展[98]。在这一方面，最近有报道另一种 Cre 重组酶衍生物（uTre）能够识别 HIV 主要的 A、B、C 亚型中具有 94%保守的 LTR 序列，而这种方法也更加接近临床[99]。

对于靶序列识别区域限制更少的人工编辑核酸酶（ZFNs、TALENs 和 CRISPR/Cas9），在 HIV-1 的基因组能够提供除 LTRs 之外的多个靶位点。图 5 显示出了 HIV-1 基因组中片段多样性的熵分析，尤其是其中的保守序列能够为这样的靶向位点提供来源。ZFNs 已经被证实能够靶向于 LTRs 位点，而在病人样本分离得到的病毒中具有高度的保守性[94]。这些试剂被报道在进行 DNA 电转之后能够破坏 Jurkat T 细胞系中 60%的 GFP 表达 HIV-1 基因组，能够减少感染的主要人源性 PBLs（降低 29%）或 CD4+ T 细胞（降低 31%）中病毒的产生。与之相似的是，TALENs 也被用于 HIV DNA 的靶向。Ebina 等人能够显著性的破坏 Jurkat 细胞系模型中的整合的 HIV 片段，这些病毒经 TNF α 刺激后可由 LTR 驱动产生 GFP 的表达。通过电转 mRNA 表达靶向于 LTR 的 TALEN 对，他们发现刺激引起的 GFP 阳性细胞的数量有所下降，从 63%下降到 4.3%，并报道在 53%的细胞中 HIV 基因组被完全移除。

最后，CRISPRs 也被用于整合性 HIV-1 DNA 的靶向。在这类研究中，所采用的试剂的异质性更高[101]能够靶向于不只一处的 HIV-1 序列，因而相对于其他疗法来说，能够更快地克服耐药性。Ebina 等人设计了两种 CRISPR gRNAs 靶向于 LTR 的 TAR 和 NF κ B 区域，通过测试发现它们能够拮抗具有 Tat 基因的整合性 LTR-GFP 报道基因的表达（JLAT 细胞）[96]。在进行了三轮 CRISPR/Cas9 电转之后，LTR 功能性的破坏显现了出来，在 25%-32%的激活细胞中 GFP 的表达减少到 25%-32%[96]。与之相似的是，Zhu 等人检测了 10 种不同的 gRNAs 在 JLAT 细胞中识别 HIV LTR 和 Po1 区域的活性，研究发现两者联合时能够降低 GFP+细胞的水平 24 倍以上[102]。Liao 等人也发现导入多种 CRISPR gRNAs 相较于只导入一种 gRNAs 能够更好的破坏 HIV 序列，他们也进一步地发现抗 HIV CRISPRs 的表达能够阻止 HIV 感染[103]。最后，Hu 等人报道了 CHME-5 细胞——一种具有 HIV 感染的小神经胶质细胞中，两种不同的 LTR 靶向性 gRNAs 的效果，在激活后 GFP 的表达水平从 76%降低至 17.1%或 3.9%[104]。

抗 HIV 核酸酶应用的挑战

以上所描述的研究展示了利用人工编辑核酸酶破坏 HIV-1 的迈出的重要的第一步，需要将能够阻止 HIV-1 基因组整合的这些试剂导入到患者细胞中，还有更多的工作需要去做，例如这些细胞也可能会影响染色体的凝集（图 5），已经有研究证实 HIV-1 能够诱导突变的产生以

及影响耐受性,有假设称这也是 HIV 特异性的人工编辑核酸酶的应用障碍之一。在这一方面,同时靶向于多条序列是具有优势的,包括所有已知的能够耐受的位点,虽然这可能会提升疗法的复杂性。而且,与核酸酶的所有治疗性应用一样,脱靶的潜在性效果需要进行描述。最后,除非有潜在性的保护措施进行应用,如上所述的 Tat 诱导的 Tre 重组酶系统中,还需要开发将试剂导入体内的 HIV 感染细胞的方法。这一需求对于潜在性的感染细胞尤其具有挑战,因为 HIV 基因表达的缺失意味着没有明显的指标能够指征携带有 HIV-1 基因组的细胞。因此,利用感染细胞的一些替代性指标进行所有 T 细胞的分散的、有选择性的导入[90]有可能是被需要的。

人工编辑核酸酶临床应用的思考

将人工编辑核酸酶导入到人类细胞的策略

将人工编辑核酸酶导入细胞是一个已知的巨大的挑战,包括不止一种 ZFN、TALEN 和 CRISPR/Cas9 核酸酶的共同导入,在需要 HDR 介导的基因编辑的情况下还需要修复模板的加入。除此之外,在 DNA 中还有大量的可被 ZFNs 和 TALEN 识别的重复序列区域。然而,由人工编辑核酸酶激活引起的永久性遗传信息的改变只需要靶细胞的瞬时性表达,有多种的传递系统也可被应用。

随着早期研发的进行以及多种临床方案的采用,ZFNs 已经成为了最广范围内活性得到评估的导入方法。采用的标准化病毒载体包括腺病毒、腺相关病毒(AAV)和 IDLVs[31,33,81,82,105-107]以及不常见的杆状病毒载体[32]。通过电转这种更为简单的导入方式,核酸酶的 DNA 和 mRNA 都能被导入到细胞中[32,35,45,47,81,82]。有趣的是,ZFN 蛋白也能够直接被细胞摄取,通过这种方式,在细胞系中 CCR5 的破坏率可达到 27%,在主要的人类 CD4 T 细胞中可达到 8%[34]。作者也报道了这种方法造成的 293T 细胞中的 CCR5 在九个位点的脱靶率低于 ZFN DNA 转染造成的脱靶,通过这种方式也会造成导入的 ZFNs 的浓度下降。通过相似的方法,Chen 等人将 ZFN 蛋白与转铁蛋白分子通过可裂解的二硫键进行连接构建了一种体系[108]。在该实验中,ZFN 通过转铁蛋白的内吞作用被摄入细胞,这一系统在多种细胞类型中都能发挥功能,提示这一受体具有很广泛的分布。这类的基于蛋白的导入系统相对于其它的以病毒为基础的系统较为简单,通过减少细胞中 ZFN 蛋白的暴露量,能够确定性的降低脱靶效应,这一功能与核酸酶的浓度和暴露时间均相关。

虽然 TALENs 和 ZFNs 具有很多共同的设计结构,TALEN 的导入相对于 ZFNs 更为复杂。最初,利用常见的 C63 骨架构建的 TALEN 相对于 ZFNs 大约长 1kb(每个单体),这种尺寸使得其在例如 AAV 这样的载体中的包装更为复杂。而且,利用 TALE 重复单位构建的 1 碱基识别编码与锌指分子的 3bp 识别片段不同,因此靶向于相同序列的 TALENs 具有大于 ZFNs 三倍的识别序列。这种设计也会引起确定类型载体的不稳定性,尤其是慢病毒载体[107],TALEN 亚单位之间高水平的同源重组能够引起逆转录期间两个包装 RNAs 之间的重排和删除。这一问题的一个可能性的解决办法是在 TALE 的每一个重复单位造成多个沉默突变降低同源的产生。或者,在相同的研究中发现 Ad5 载体能够很好的传导 TALENs,能够造成 HeLa 细胞系、永生化肌细胞和骨髓造血干细胞中 47-55%的基因破坏[107]。最近有一项研究利用慢病毒载体转导的 TALEN 被突变失去了作为逆转录酶的活性。最终产生的 RNA 基因组能够表达插入了 IRES 和 poly A 末端的 TALEN,类似于一个 mRNA[109]。

对于所有种类的人工编辑核酸酶,目前的技术还大多局限于体外转导水平,但是更为广泛的应用需要体内转导方法的开发,例如 siRNAs[110],这对于确定的病毒载体或许是可能的。例如,具有不同的体内趋向性的 AAV 亚型,AAV 载体被证实能够介导 ZFN 单体和外源模板通过趋肝性 AAV 的腹腔注射进入小鼠的肝细胞[105]。而且,最近在重新定位的腺病毒

载体中的进步显示，基于细胞表面蛋白 scFvs 或配体，在未来或能够将人工编辑核酸酶导入到特定的细胞类型中[111-117]。

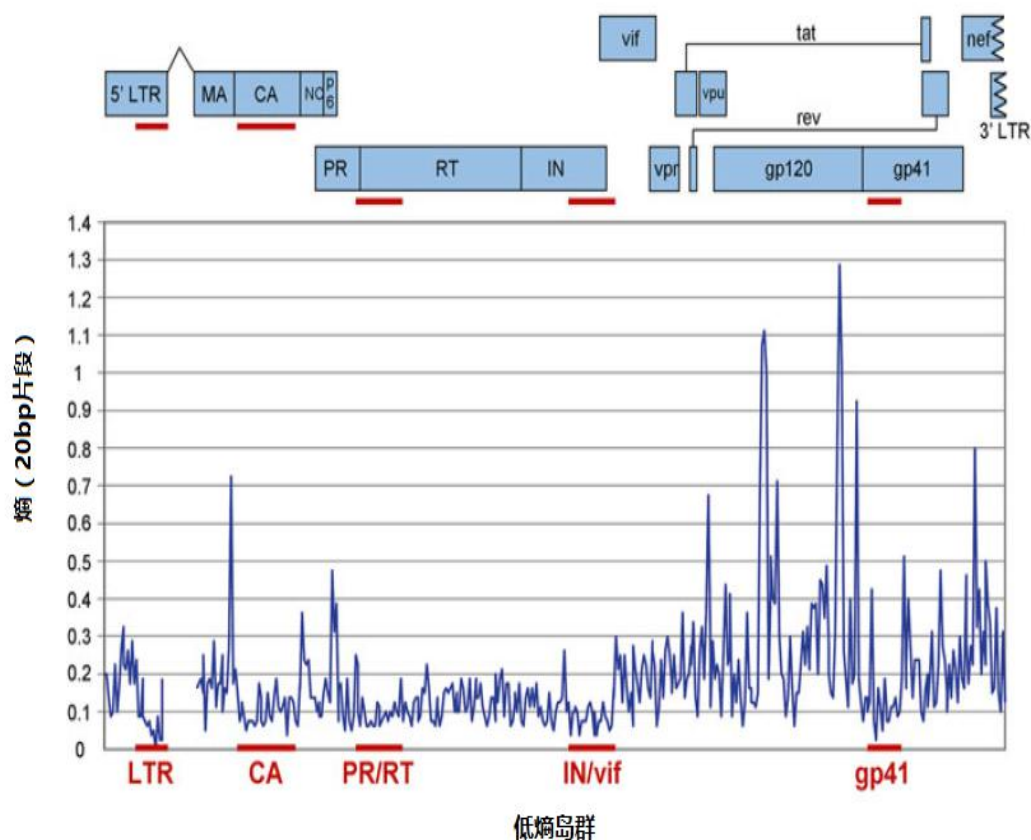


图5. 在HIV-1基因组中低熵岛群的鉴定。利用Entropy-one 工具 (www.hiv.lanl.gov) 对来源于Los Alamos HIV-1序列数据库的102 LTR和442编码区序列进行检索。对不同位点异质性的熵进行测量，低熵代表着高度的保守性。以每20bp的片段为单位进行平均熵值的统计和标注，对低熵（等于或低于0.1）岛群进行标注（红线）。

脱靶效应与毒性

人工编辑核酸酶进行临床应用的一个关键性问题是毒副作用的产生，包括细胞毒性[118]，或核酸酶在脱靶位点的作用[31,33,119-121]。毒性的产生可能是由于核酸酶自身的表达造成的，也可能是细胞对 DNA 损伤的应答反应，或者是介导方法等引起的细胞压力产生的综合性的作用。脱靶产生的位置可以利用与目标序列高度同源的生物信息方法进行预测，以及在体外实验筛选中发现的易脱靶位点[122]，或通过在 DSB 位点进行 NHEJ 介导的 IDLVs 的整合能力的预测[123]。由于类似的脱靶事件和毒性反应具有细胞类型的特异性，细胞系的研究也无法准确预测出主要的人类细胞的情况。因此前临床试验必须包括在预计脱靶位点中活性的评估以及对病人剂量的 ZFN 处理狗的 T 细胞或 HSC 细胞进行更为一般性的癌症形成能力的评估，需要同时进行体外细胞实验和随后的免疫缺陷鼠的移植[30,42]。

最近，有多个方法被用来检测脱靶事件的发生。例如，高通量的基因组范围的易位测序（HTGTS）被用于鉴定 I-sce 归巢内切酶处理过细胞的错位连接[124]，而且这一方法也被 Frock 等人改进，包括利用片段扩增 PCR（LAM-PCR）对 CRISPRs 和 TALENs 产生的脱靶位点进行鉴定[125]。这一研究也发现 CRISPR 介导的脱靶活性也能通过 CRISPR 缺刻酶降低脱靶活性[53]，而且也发现 TALEN 介导的脱靶效应通常是由于同源二聚化产生的[125]。

其它的基因组范围的脱靶效应的检测方法包括 GUIDE 测序（基因组范围，DSBs 能力的无偏测序鉴定），这一检测方法基于将双链寡脱氧核苷酸插入到 DSBs 序列[126]和 BLESS 实验（损伤标签，链霉素亲和素的富集和下一代基因测序技术），该技术利用生物素标记的寡核苷酸对 DSBs 进行标记[127,128]。还有另一种方法是 digenome 测序（降解基因测序），利用核酸酶处理体外的基因组 DNA，随后对产生的片段进行全基因组测序[129]。由于这些实验所产生的数据不依赖于生物信息的数据，能够对隐藏位点进行更为客观的评价。然而，HTGTS、GUIDE 测序和 digenome 测序等方法仍旧不完美，因为在一项 VEGF-A CRISPR 靶向实验中这些测序方法分别得到了一些不同的脱靶位点（回溯见[120]）。

对于常用的 160 CCR5 ZFNs，脱靶分布常见于高度同源的 CCR2 基因，相对于 CCR5 位点能够达到 9-10%的比率[31,33,49,119,122,123]，随着 ZFNs 的表达提高破坏率也有提高。通过这种方式，对核酸酶活性的鉴定就有了实际的标准，以靶向破坏率为最大值对可接受的脱靶活性进行量化。与此同时，降低核酸酶脱靶活性的方法也在开发中。例如，通过 Fok1 核酸内切酶的异二聚化连接，ZFNs 和 TALENs 都更加具有特异性，这也限制了独特的单体组件和目标异二聚体可能性的配对连接[130]。通过类似的方法，CRISPR/Cas9 的特异性能够通过两种 RNA 识别域连接的失活的 Cas9 蛋白和 Fok1 部分融合后提高[131]。最后，TALENs 中 DNA 连接 RVDs 的人工编辑能够提高靶向特异性并减弱脱靶效应[47]。

总结

HIV-1 的生命周期为基于人工编辑核酸酶的疗法干预提供了多个机会。病毒的感染需要借助细胞共受体——CCR5 的作用，而这一基因对于人类宿主来说是非必须的，为核酸酶进行基因破坏的应用提供了基础，在这一方面 ZFNs 的应用目前处于多种基因编辑工具的临床应用开发的最后期。除此之外，基因破坏伴发的 HDR 能够介导 CCR5 位点中抗 HIV 基因的插入。而且，针对于抗逆转录疗法无效的存在于患者细胞中的 HIV-1 基因组，也可作为基因破坏的靶标，但是将核酸酶导入潜在细胞的过程仍存在巨大的挑战。最后，和所有的产生 HIV 抗性细胞的基因疗法一样，HIV-1 能够被用于对 HIV 抗性细胞电的筛选从而加速其自身的灭亡。

致谢 我们由衷的感谢来自 Sangamo BioSciences 的合作者提供的宝贵的经验，包括 Michael Holmes、Jianbin Wang 和 Philip Gregory，并感谢 Liz Wolffe 在表 1 的编译中提供的帮助。这一工作得到了 James B. Pendleton 慈善信托基金、NIH 基金 HL073104、AI10149 和 HL129902 以及加利福尼亚 HIV/AIDS 研究项目资金 ID12-USC-245 的支持。

参考文献

1. <http://aids.gov/federal-resources/policies/care-continuum> , 2013.
2. Baltimore D. Gene therapy. Intracellular immunization. *Nature*. 1988;335(6189):395 – 6.
3. Peterson CW, et al. Combinatorial anti-HIV gene therapy: using a multipronged approach to reach beyond HAART. *Gene Ther*. 2013;20(7):695 – 702.
4. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med*. 2014;370:10.
5. Sun J, et al. Human Ku70/80 protein blocks exonuclease 1-mediated DNA resection in the presence of human Mre11 or Mre11/Rad50 protein complex. *J Biol Chem*. 2012;287(7):4936 – 45.
6. Symington LS, Gautier J. Double-strand break end resection and repair pathway choice. *Annu Rev Genet*. 2011;45:247 – 71.

7. Chapman JR, Taylor MR, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell*. 2012;47(4):497 – 510.
8. Escibano-Diaz C, et al. A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol Cell*. 2013;49(5):872 – 83.
9. Heyer WD, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet*. 2010;44:113 – 39.
10. Mohrin M, et al. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell*. 2010;7(2):174 – 85.
11. Rossi DJ, et al. Hematopoietic stem cell quiescence attenuates DNA damage response and permits DNA damage accumulation during aging. *Cell Cycle*. 2007;6(19):2371 – 6.
12. Deng H, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*. 1996;381(6584):661 – 6.
13. Samson M, et al. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*. 1996;382(6593):722 – 5.
14. Novembre J, Galvani AP, Slatkin M. The geographic spread of the CCR5 Delta32 HIVresistance allele. *PLoS Biol*. 2005;3(11), e339.
15. Dean M, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CCR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science*. 1996;273(5283):1856 – 62.
16. Wood A, Armour D. The discovery of the CCR5 receptor antagonist, UK-427,857, a new agent for the treatment of HIV infection and AIDS. *Prog Med Chem*. 2005;43: 239 – 71.
17. Allers K, et al. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. *Blood*. 2011;117(10):2791 – 9.
18. Yukl SA, et al. Challenges in detecting HIV persistence during potentially curative interventions: a study of the Berlin patient. *PLoS Pathog*. 2013;9(5), e1003347.
19. Hutter G, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med*. 2009;360(7):692 – 8.
20. Hutter G, Zaia JA. Allogeneic haematopoietic stem cell transplantation in patients with human immunodeficiency virus: the experiences of more than 25 years. *Clin Exp Immunol*. 2011;163(3):284 – 95.
21. Hayden EC. Hopes of HIV cure in ‘Boston patients’ dashed. *Nature News* 2013.
22. Ringpis GE, et al. Engineering HIV-1-resistant T-cells from short-hairpin RNA-expressing hematopoietic stem/progenitor cells in humanized BLT mice. *PLoS One*. 2012;7(12), e53492.
23. Lee MT, et al. Inhibition of human immunodeficiency virus type 1 replication in primary macrophages by using Tat- or CCR5-specific small interfering RNAs expressed from a lentivirus vector. *J Virol*. 2003;77(22):11964 – 72.
24. ter Brake O, et al. Lentiviral vector design for multiple shRNA expression and durable HIV-1 inhibition. *Mol Ther*. 2008;16(3):557 – 64.
25. Qin XF, et al. Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5. *Proc Natl Acad Sci U S A*. 2003;100(1):183 – 8.
26. An DS, et al. Optimization and functional effects of stable short hairpin RNA expression in

- primary human lymphocytes via lentiviral vectors. *Mol Ther.* 2006;14(4):494 – 504.
27. Li MJ, et al. Inhibition of HIV-1 infection by lentiviral vectors expressing Pol III-promoted anti-HIV RNAs. *Mol Ther.* 2003;8(2):196 – 206.
28. Wu C, Dunbar CE. Stem cell gene therapy: the risks of insertional mutagenesis and approaches to minimize genotoxicity. *Front Med.* 2011;5(4):356 – 71.
29. Didigu CA, et al. Simultaneous zinc-finger nuclease editing of the HIV coreceptors *ccr5* and *cxcr4* protects CD4+ T cells from HIV-1 infection. *Blood.* 2014;123(1):61 – 9.
30. Maier DA, et al. Efficient clinical scale gene modification via zinc finger nuclease-targeted disruption of the HIV co-receptor CCR5. *Hum Gene Ther.* 2013;24(3):245 – 58.
31. Perez EE, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol.* 2008;26(7):808 – 16.
32. Lei Y, et al. Gene editing of human embryonic stem cells via an engineered baculoviral vector carrying zinc-finger nucleases. *Mol Ther.* 2011;19(5):942 – 50.
33. Li L, et al. Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Mol Ther.* 2013;21(6):1259 – 69.
34. Gaj T, et al. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat Methods.* 2012;9(8):805 – 7.
35. Holt N, et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol.* 2010;28(8):839 – 47.
36. Kalos M, June CH. Adoptive T cell transfer for cancer immunotherapy in the era of synthetic biology. *Immunity.* 2013;39(1):49 – 60.
37. Scholler J, et al. Decade-long safety and function of retroviral-modified chimeric antigen receptor T cells. *Sci Transl Med.* 2012;4(132):132ra53.
38. Mani M, et al. Design, engineering, and characterization of zinc finger nucleases. *Biochem Biophys Res Commun.* 2005;335(2):447 – 57.
39. Yi G, et al. CCR5 gene editing of resting CD4(+) T cells by transient ZFN expression from HIV envelope pseudotyped nonintegrating lentivirus confers HIV-1 resistance in humanized mice. *Mol Ther Nucleic Acids.* 2014;3, e198.
40. Shayakhmetov DM, et al. Efficient gene transfer into human CD34(+) cells by a retargeted adenovirus vector. *J Virol.* 2000;74(6):2567 – 83.
41. Chun TW, et al. Relationship between pre-existing viral reservoirs and the re-emergence of plasma viremia after discontinuation of highly active anti-retroviral therapy. *Nat Med.* 2000;6(7):757 – 61.
42. Hofer U, et al. Pre-clinical modeling of CCR5 knockout in human hematopoietic stem cells by zinc finger nucleases using humanized mice. *J Infect Dis.* 2013;208 Suppl 2:S160 – 4.
43. Lepus CM, et al. Comparison of human fetal liver, umbilical cord blood, and adult blood hematopoietic stem cell engraftment in NOD-scid/gammac^{-/-}, Balb/c-Rag1^{-/-}-gammac^{-/-}, and C.B-17-scid/bg immunodeficient mice. *Hum Immunol.* 2009;70(10):790 – 802.
44. Lapid K, et al. Egress and mobilization of hematopoietic stem and progenitor cells: a dynamic multi-facet process. Cambridge: Harvard Stem Cell Institute; 2008. StemBook [Internet].
45. Cannon PM, et al. Electroporation of ZFN mRNA enables efficient CCR5 gene disruption in

- mobilized blood hematopoietic stem cells at clinical scale. *Mol Ther.* 2013;21:S71 – 2.
46. Miller JC, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol.* 2011;29(2):143 – 8.
47. Llewellyn N, et al. Next generation TALENs mediate efficient disruption of the CCR5 gene in human HSCs. *Mol Ther.* 2013;21:S72.
48. Mussolino C, et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 2011;39(21):9283 – 93.
49. Mussolino C, et al. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 2014;42(10):6762 – 73.
50. Mock U, et al. mRNA transfection of a novel TAL effector nuclease (TALEN) facilitates efficient knockout of HIV co-receptor CCR5. *Nucleic Acids Res.* 2015;43(11):5560 – 71.
51. Kim Y, Kweon J, Kim JS. TALENs and ZFNs are associated with different mutation signatures. *Nat Methods.* 2013;10(3):185.
52. Certo MT, et al. Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. *Nat Methods.* 2012;9(10):973 – 5.
53. Cho SW, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 2014;24(1):132 – 41.
54. Cradick TJ, et al. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* 2013;41(20):9584 – 92.
55. Mandal PK, et al. Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell.* 2014;15(5):643 – 52.
56. Li C, et al. Inhibition of HIV-1 infection of primary CD4+ T cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J Gen Virol.* 2015;96(8):2381 – 93.
57. Wang W, et al. CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. *PLoS One.* 2014;9(12), e115987.
58. Wilen CB, et al. Engineering HIV-resistant human CD4+ T cells with CXCR4-specific zinc-finger nucleases. *PLoS Pathog.* 2011;7(4), e1002020.
59. Yuan J, et al. Zinc-finger nuclease editing of human cxcr4 promotes HIV-1 CD4(+) T cell resistance and enrichment. *Mol Ther.* 2012;20(4):849 – 59.
60. <http://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv-guidelines/8/co-receptor-ortropism-assays> . Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. 2013.
61. Meyer L, et al. Early protective effect of CCR-5 delta 32 heterozygosity on HIV-1 disease progression: relationship with viral load. The SEROCO Study Group. *AIDS.* 1997; 11(11):F73 – 8.
62. de Roda Husman AM, et al. Association between CCR5 genotype and the clinical course of HIV-1 infection. *Ann Intern Med.* 1997;127(10):882 – 90.
63. Meyer L, et al. CCR5 delta32 deletion and reduced risk of toxoplasmosis in persons infected with human immunodeficiency virus type 1. The SEROCO-HEMOCO-SEROGEST Study Groups. *J Infect Dis.* 1999;180(3):920 – 4.
64. Ioannidis JP, et al. Effects of CCR5-Delta, 32, CCR2 – 64I, and SDF-1 3' A alleles on HIV-1

disease progression: an international meta-analysis of individual-patient data. *Ann Intern Med.* 2001;135(9):782 – 95.

65. Bevec D, et al. Inhibition of human immunodeficiency virus type 1 replication in human T cells by retroviral-mediated gene transfer of a dominant-negative Rev trans-activator. *Proc Natl Acad Sci U S A.* 1992;89(20):9870 – 4.

66. Bonyhadi ML, et al. RevM10-expressing T cells derived in vivo from transduced human hematopoietic stem-progenitor cells inhibit human immunodeficiency virus replication. *J Virol.* 1997;71(6):4707 – 16.

67. Stremlau M, et al. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature.* 2004;427(6977):848 – 53.

68. Sawyer SL, et al. Positive selection of primate TRIM5alpha identifies a critical species specific retroviral restriction domain. *Proc Natl Acad Sci U S A.* 2005;102(8):2832 – 7.

69. Anderson J, Akkina R. Human immunodeficiency virus type 1 restriction by human-rhesus chimeric tripartite motif 5alpha (TRIM 5alpha) in CD34(+) cell-derived macrophages in vitro and in T cells in vivo in severe combined immunodeficient (SCID-hu) mice transplanted with human fetal tissue. *Hum Gene Ther.* 2008;19(3):217 – 28.

70. Schrofelbauer B, Chen D, Landau NR. A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif). *Proc Natl Acad Sci U S A.* 2004;101(11):3927 – 32.

71. Xu H, et al. A single amino acid substitution in human APOBEC3G anti-retroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion. *Proc Natl Acad Sci U S A.* 2004;101(15):5652 – 7.

72. Anderson J, et al. Safety and efficacy of a lentiviral vector containing three anti-HIV genes--CCR5 ribozyme, tat-rev siRNA, and TAR decoy--in SCID-hu mouse-derived T cells. *Mol Ther.* 2007;15(6):1182 – 8.

73. Kumar P, et al. T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice. *Cell.* 2008;134(4):577 – 86.

74. ter Brake O, et al. Evaluation of safety and efficacy of RNAi against HIV-1 in the human immune system (Rag-2(-/-)gammac(-/-)) mouse model. *Gene Ther.* 2009;16(1):148 – 53.

75. Novina CD, et al. siRNA-directed inhibition of HIV-1 infection. *Nat Med.* 2002;8(7):681 – 6.

76. Kohn DB, et al. A clinical trial of retroviral-mediated transfer of a rev-responsive element decoy gene into CD34(+) cells from the bone marrow of human immunodeficiency virus-1-infected children. *Blood.* 1999;94(1):368 – 71.

77. Humeau LM, et al. Efficient lentiviral vector-mediated control of HIV-1 replication in CD4 lymphocytes from diverse HIV+ infected patients grouped according to CD4 count and viral load. *Mol Ther.* 2004;9(6):902 – 13.

78. Michienzi A, et al. A nucleolar TAR decoy inhibitor of HIV-1 replication. *Proc Natl Acad Sci U S A.* 2002;99(22):14047 – 52.

79. Zahn RC, et al. Efficient entry inhibition of human and nonhuman primate immunodeficiency virus by cell surface-expressed gp41-derived peptides. *Gene Ther.* 2008;15(17):1210 – 22.

80. Voit RA, et al. Generation of an HIV resistant T-cell line by targeted “stacking” of restriction factors. *Mol Ther.* 2013;21(4):786 – 95.

81. Lombardo A, et al. Gene editing in human stem cells using zinc finger nucleases and integrase- defective lentiviral vector delivery. *Nat Biotechnol.* 2007;25(11):1298 – 306.
82. Lombardo A, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods.* 2011;8(10):861 – 9.
83. Genovese P, et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature.* 2014;510(7504):235 – 40.
84. Hoban MD, et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood.* 2015;125(17):2597 – 604.
85. Wang J, Exline CM, et al. Highly efficient homology-driven genome editing in CD34+ hematopoietic stem/progenitor cells by combining zinc finger nuclease mRNA and AAV donor delivery. *Nat Biotechnol.* 2015;33(12):1256 – 1263.
86. Chu VT, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol.* 2015;33(5):543 – 8.
87. Maruyama T, et al. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol.* 2015;33(5):538 – 42.
88. Ye L, et al. Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. *Proc Natl Acad Sci U S A.* 2014;111(26):9591 – 6.
89. Yao Y, et al. Generation of CD34+ cells from CCR5-disrupted human embryonic and induced pluripotent stem cells. *Hum Gene Ther.* 2012;23(2):238 – 42.
90. Chomont N, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med.* 2009;15(8):893 – 900.
91. Chun TW, et al. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature.* 1997;387(6629):183 – 8.
92. Finzi D, et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med.* 1999;5(5):512 – 7.
93. Sarkar I, et al. HIV-1 proviral DNA excision using an evolved recombinase. *Science.* 2007;316(5833):1912 – 5.
94. Qu X, et al. Zinc-finger-nucleases mediate specific and efficient excision of HIV-1 proviral DNA from infected and latently infected human T cells. *Nucleic Acids Res.* 2013;41(16):7771 – 82.
95. Hauber I, et al. Highly significant antiviral activity of HIV-1 LTR-specific tre-recombinase in humanized mice. *PLoS Pathog.* 2013;9(9), e1003587.
96. Ebina H, et al. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep.* 2013;3:2510.
97. Blackard JT, et al. Transmission of human immunodeficiency type 1 viruses with intersubtype recombinant long terminal repeat sequences. *Virology.* 1999;254(2):220 – 5.
98. Surendranath V, et al. SeLOX—a locus of recombination site search tool for the detection and directed evolution of site-specific recombination systems. *Nucleic Acids Res.* 2010;38(Web Server issue):W293 – 8.
99. Karpinski J, et al. Universal Tre (uTre) recombinase specifically targets the majority of HIV-1

- isolates. *J Int AIDS Soc.* 2014;17(4 Suppl 3):19706.
100. Ebina H, et al. A high excision potential of TALENs for integrated DNA of HIV-based lentiviral vector. *PLoS One.* 2015;10(3), e0120047.
101. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339(6121):819 – 23.
102. Zhu W, et al. The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology.* 2015;12:22.
103. Liao HK, et al. Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat Commun.* 2015;6.
104. Hu W, et al. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc Natl Acad Sci U S A.* 2014;111(31):11461 – 6.
105. Li H, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature.* 2011;475(7355):217 – 21.
106. Joglekar AV, et al. Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol Ther.* 2013;21(9):1705 – 17.
107. Holkers M, et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* 2013;41(5), e63.
108. Chen Z, et al. Receptor-mediated delivery of engineered nucleases for genome modification. *Nucleic Acids Res.* 2013;41(19), e182.
109. Mock U, et al. Novel lentiviral vectors with mutated reverse transcriptase for mRNA delivery of TALE nucleases. *Sci Rep.* 2014;4:6409.
110. Khatri N, et al. In vivo delivery aspects of miRNA, shRNA and siRNA. *Crit Rev Ther Drug Carrier Syst.* 2012;29(6):487 – 527.
111. Morizono K, et al. Lentiviral vector retargeting to P-glycoprotein on metastatic melanoma through intravenous injection. *Nat Med.* 2005;11(3):346 – 52.
112. Lin AH, et al. Receptor-specific targeting mediated by the coexpression of a targeted murine leukemia virus envelope protein and a binding-defective influenza hemagglutinin protein. *Hum Gene Ther.* 2001;12(4):323 – 32.
113. Frecha C, et al. A novel lentiviral vector targets gene transfer into human hematopoietic stem cells in marrow from patients with bone marrow failure syndrome and in vivo in humanized mice. *Blood.* 2012;119(5):1139 – 50.
114. Anliker B, et al. Specific gene transfer to neurons, endothelial cells and hematopoietic progenitors with lentiviral vectors. *Nat Methods.* 2010;7(11):929 – 35.
115. Paraskevaku G, et al. Epidermal growth factor receptor (EGFR)-retargeted measles virus strains effectively target EGFR- or EGFRvIII expressing gliomas. *Mol Ther.* 2007;15(4):677 – 86.
116. Kneissl S, et al. Measles virus glycoprotein-based lentiviral targeting vectors that avoid neutralizing antibodies. *PLoS One.* 2012;7(10), e46667.

117. Kneissl S, et al. CD19 and CD20 targeted vectors induce minimal activation of resting B lymphocytes. *PLoS One*. 2013;8(11), e79047.
118. Alwin S, et al. Custom zinc-finger nucleases for use in human cells. *Mol Ther*. 2005;12(4):610 – 7.
119. Kim HJ, et al. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res*. 2009;19(7):1279 – 88.
120. Koo T, Lee J, Kim JS. Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9. *Mol Cells*. 2015;38(6):475 – 81.
121. Hendel A, et al. Quantifying on- and off-target genome editing. *Trends Biotechnol*. 2015;33(2):132 – 40.
122. Pattanayak V, et al. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods*. 2011;8(9):765 – 70.
123. Gabriel R, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol*. 2011;29(9):816 – 23.
124. Chiarle R, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell*. 2011;147(1):107 – 19.
125. Frock RL, et al. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol*. 2015;33(2):179 – 86.
126. Tsai SQ, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*. 2015;33(2):187 – 97.
127. Crosetto N, et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods*. 2013;10(4):361 – 5.
128. Ran FA, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. 2015;520(7546):186 – 98.
129. Kim D, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods*. 2015;12(3):237 – 43, 1 p following 243.
130. Doyon Y, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods*. 2011;8(1):74 – 9.
131. Tsai SQ, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32(6):569 – 76.

鉴定人工编辑核苷酸脱靶效应的策略

Eli J. Fine , Thomas James Cradick , and Gang Bao

摘要 利用人工编辑核酸酶对预选的 DNA 序列进行特异性的切割促进了基因编辑的发展。由核酸酶诱导的 DNA 损伤能够通过非同源末端修复 (NHEJ) 或直接同源修复 (HDR) 进行细胞修复, 并能够在范围内的组织和细胞系中进行基因编辑。然而, 如果核酸酶在基因组中的切割位点不在目标区域, 那就是所谓的“脱靶”, 这能够产生突变、染色体缺失或重组、产生功能的获得/缺失和细胞毒性。虽然锌指核酸酶 (ZFNs)、TAL 效应核酸酶 (TALENs) 和 CRISPR/Cas9 系统已经成功地用于在细胞中产生特异性的 DNA 损伤, 但是特异性的缺乏能够引起脱靶切割。已经开发了方法能够对脱靶事件进行预测和定量, 这对于优化核酸酶的设计和特异性的基因编辑方式的筛选至关重要。这些方法能够显著性地促进基因编辑应用领域中人工编辑核酸酶的设计。

关键词 基因编辑 • 核酸酶 • 脱靶 • 特异性 • TAL 效应器核酸酶 (TALEN) • 锌指核酸酶 (ZFN) • CRISPR/Cas9

E. J. Fine • T. J. Cradick • G. Bao , Ph.D. (*)

Department of Biomedical Engineering , Georgia Institute of Technology and Emory University , 313 Ferst Dr. , Atlanta , GA 30332 , USA

e-mail: ejfine@gmail.com; tj@alum.mit.edu; gang.bao@bme.gatech.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_11

缩略语

bp	碱基对
Cas9	CRISPR 相关蛋白 9，一种来源于链霉菌的核酸内切酶
CCR5	趋化因子（C-C 基序）受体 5 基因
CRISPR	群集定期间隔短回文重复
CRISPR/Cas9	DNA 的切割需要向导 RNA 和相关蛋白的联合
DSB	双链损伤
FokI	来源于黄杆菌属 <i>oceanokoites</i> 的一种核酸内切酶
gRNA	在 CRISPR/Cas9 系统中所使用的向导 RNA
HBB	血红蛋白 β 基因，也被称为 β -球蛋白
HDR	DNA 同源修复重组修复途径
IDLV	整合缺陷性慢病毒载体
Indel	短片段的插入、删除或 DNA 插入和删除的总称
NGS	下一代基因测序，例如 Illumina 属于这一范畴
NHEJ	DNA 非同源末端修复途径
RVD	重复性的可获得性双残基，具有 DNA 碱基特异性的 TAL 重复片段的两个氨基酸
SMRT	独立分子 real-time，第三代基因测序平台
TALEN	转录激活样效应器核酸酶
ZFN	锌指核酸酶

介绍

人工编辑核酸酶能够切割靶位点 DNA 促进基因编辑，这一过程通过细胞修复途径完成。DNA 切割能够借助所提供的模板 DNA 显著提高同源重组（HR）介导的基因编辑的频率[1,2]，高效的基因敲除能够在损伤位点通过非同源末端（NHEJ）修复途径引起的小片段的插入和删除（indel）完成。不可否认，保证全基因组中只有一个切割位点是存在巨大挑战的。人工编辑核酸酶需要一些办法最小化脱靶效应，以防由脱靶产生的突变、染色体缺失或重组，以及由其引发的基因调控和细胞毒性的改变。有很多的核酸酶被用于产生这些特异性的损伤，包括归巢内切酶[3,4]、锌指核酸酶（ZFNs）[5-7]、TAL 效应器核酸酶（TALENs）[8-11]和最近常用的 CRISPR/Cas9[12-14]。这些核酸酶中的每一个都没有完美的特异性，都被发现具有脱靶切割，即使是自然状态下的归巢内切酶——如 I-SCEI 也不例外[15]。

这些核酸酶家族的每一种都能在其靶向位点和脱靶位点同时产生切割。不同家族的核酸酶被证实具有不同的切割模式和不同的脱靶位点。在一个家族中，靶向位点的不同也会引起脱靶位点的显著不同。因此，在实际应用中进行多种优化以最好的平衡活性和脱靶就尤为重要。对于理解每种不同层面的差异也是非常重要的，因为这会影响全基因组脱靶实验的一些参数。

定位核酸酶的脱靶活性

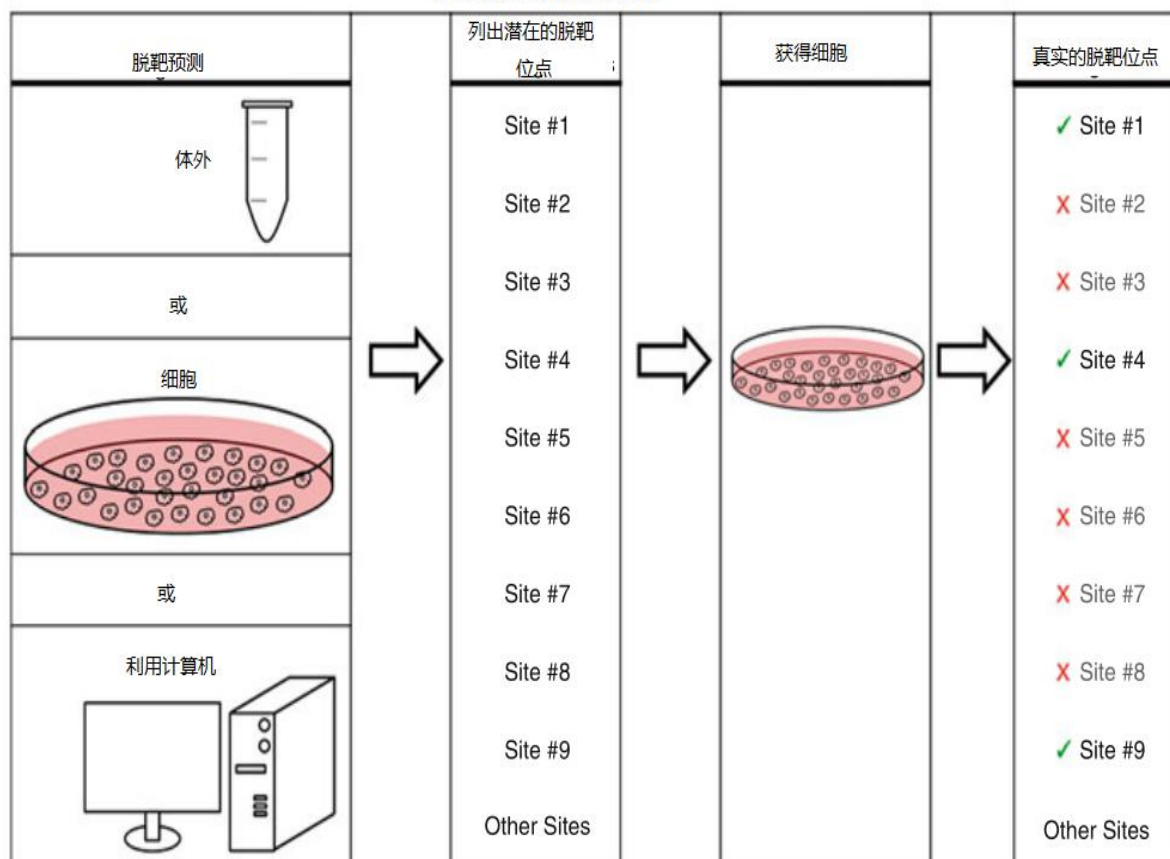


图1.核酸酶脱靶活性的定位。脱靶位点可通过体外实验、细胞实验或计算机技术进行预测，产生一系列的潜在脱靶位点。提取利用核酸酶处理后细胞的基因组，进行靶点扩增、检测并获取真实的脱靶位点信息。

目前具有多种策略能够检测核酸酶的脱靶活性。最简单的也是应用最为广泛的方式是通过观察细胞群落整体的脱靶活性，例如 DNA 修复位点的增加、细胞周期的调节异常或细胞死亡的增加。更为复杂的策略包括发现脱靶切割的特异性位点以及测量核酸在这些位点诱导产生的突变率。有很多种方法采用常见的框架来定位点特异性的脱靶活性（图 1）。首先，基于核酸酶的实验特征描述罗列出基因组中可能存在的脱靶位点（细胞或体外实验）或通过计算机模型运算。随后，利用核酸酶处理细胞，并对这些位点中核酸酶的脱靶活性进行分析以确定该脱靶位点的真实性。

影响脱靶活性的因素

当选择人工编辑核酸酶系统时，评价项目的可接受风险和脱靶作用的限度（图 1）是非常重要的。每一个核酸酶家族，及其中具有差异的一些核酸酶，都需要在活性和特异性方面进行权衡。而且，当采用策略来判定脱靶效率时，了解何种类型脱靶更易发生对于调整计算机模型参数、体外分析试验中寡核苷酸库的构建或体内基因区域内潜在位点的定位至关重要。

锌指核酸酶

C₂H₂ 型锌指是人类基因组中应用最为广泛的 DNA 连接基序[16]。锌指 DNA 识别区域由独立的锌指分子序列组成，每个包含约 30aa，能够通过疏水作用和锌离子的螯合作用形成一个 β β α 稳定结构[17,18]。每个锌指结构域能够识别 3-4 个核苷酸，但是有一些锌指结构的特

异性不是很强。锌指是分子模块化的结构，能够通过组装靶向于全新的序列，但是存在干扰成功率的位置效应和序列限制[19,20]。将锌指 DNA 识别区和细菌核酸内切酶 FokI 的催化区域相连产生新的人工的限制酶（ZFNs）[21,22]。这一工作的发展源于 FokI 的催化区域是非特异性的，其识别结构可被其它的 DNA 识别区域所取代[21]。在将其与两个 ZFNs 的半位点识别区域进行正确的方向和空隙的连接之后，其特异性显著性的提高。这使得两个 FokI 结构域能够发生二聚化并对于介于两者之间的区域进行切割[23]。这一独特的锌指结构能够被用于改造，或者在保持框架恒定的前提下对所连接的残基进行改造使得 ZFNs 能够直接靶向于全新的序列[6,24]。ZFNs 被发现能够产生脱靶切割和细胞毒性[25-30]。研究进一步通过对每种 ZFNs 进行修饰和进一步纯化能够提高特异性，描述如下。

蛋白和 DNA 之间的相互作用

有多种方法试图被用于降低 ZFN 的脱靶切割和毒性。很多这一类的研究都通过对蛋白质和 DNA 连接更为深刻的了解并尝试利用特异性更强的核酸酶直接提高特异性[31]。分选实验利用噬菌体展示技术[32-35]和细菌系统对锌指结构域结合的紧密程度进行鉴定，通过与细菌染色体竞争对其特异性进行评估[36-38]。有多种筛选平台能够被用于进行多种锌指结构的分选，但是这些方法仍存在一些困难[39,40]。靶向于非鸟嘌呤富集区段对于 ZFNs 仍十分困难。总之，试图提高 ZFNs 特异性的方法还未取得显著性成效。

多聚锌指序列

在人工转录因子或 ZFNs 中提高锌指数量能够导致 DNA 亲和力的提高。四个和六个锌指结构域经常用来构建两个指状单位的系列连接[41]。有很多研究组通过提高锌指结构域的数量来提高 DNA 连接区段的亲和力和特异性[42,43]，虽然很少有工作通过比较靶向于同一区段不同数量指状结构的来验证其可行性。有一种能够和抑制区域结合的六指锌指能够高特异性的降低靶向基因调控[44]。由于这是抑制结构域，在沉默基因中的脱靶无法观测到，也无法在类似的靶向抑制实验中比较其与更少的指状结构的活性。最近有一项靶向与脱靶切割的对比试验比较了在重叠位点三指和四指 ZFN 对的活性。其中四指锌指对具有更高的靶向效率和更低的脱靶率[45]。然而，指状结构域的加入并不总会引起特异性的提高；在人工培养的大鼠细胞中，由 Sigma-Aldrich 设计的 CompoZr ZFNs 被发现具有大量的脱靶[46]。在今后的工作中，还需要敏感的测序方法用来比较不同数量的 ZFNs 之间的活性。

蛋白连接区域

也有研究试图优化 ZFR 或 ZFN 中独立的锌指结构域之间的连接结构[41,47]，但是没有通用性的规则能够适用于所有的情况。有一些方法能够通过降低连接区域和 DNA 之间的非特异性相互作用提升整体的特异性[48]。当锌指结构域和 FokI 结构域之间的连接发生改变时，半位点间隔需求也随之改变，并会对目标位点的靶向和脱靶产生影响。

转录激活样效应核酸酶

转录激活样效应分子（TALENs）是 DNA 连接蛋白家族中的一员，发现于植物噬菌体 Xanthomonas[10,11,50,51]。将 TALE DNA 连接区段与 ZFNs 中所采用的来自细菌核酸内切酶 FokI 的相同区段的非特异性催化结构域进行连接后，能够产生 TALENs[52]。与 ZFNs 相同，特异性的重要组件需要按照正确的方向和间隔安装在两个 TALENs 的两个半位点上[23]。

蛋白与 DNA 的相互作用

每个 TALE DNA 连接结构域都包括 33-35 个连接结构域，差异首先发生于每个重复片段的第 12 和 13 位点，即重复差异性双残基（RVDs）[50]。TALEN 的特异性源自自然产生的 TALE 靶向位点 RVDs 中 DNA 识别编码的不同[53]。TALE 或 TALEN DNA 连接区域能够被很方便的设计，因为在每个 RVD 和所靶向的核苷酸之间存在简单性的一对一的关系[10,11,54]。虽然腺嘌呤、胸腺嘧啶和胞嘧啶之间的连接是直接性的，但是有多种 RVDs 能够与鸟嘌呤连接；其中最为

常见的 RVD, “NN” (天冬酰胺-天冬酰胺), 能够以近乎相同亲和力与鸟嘌呤和腺嘌呤连接。包含有“NN” RVD 的 TALEN 具有更加良好的活性, 但是相较于具有“NK” (天冬酰胺-赖氨酸) 的 TALENs 也具有更强的脱靶效应[45]。

蛋白质连接区域

目前还没有针对于 TALENs 的 DNA 连接区域内的非经典连接的研究, 但是有多种 TAL 重复序列和 FokI 结构域之间的连接区域正在测试中。不同的连接区段具有不同长度的自然状态下的具有保守的 TAL 重复片段 C 末端的 AvrBs4 TALE 骨架, 例如, 常见的+63 连接在进行 AvrBs4 骨架截短和 FokI 结构域融合之后保留了 63 个氨基酸的序列。虽然有些研究文章宣称存在确定性的连接能够提高 TALEN 在目标靶点的切割活性, 但是结果却不尽相同[54,55], 除了已经有足够证据能够证明完整性的 AvrBs4 骨架 (也就是+231 连接) 相较于截短骨架活性更低[52,56]。然而, 目前已知的是不同的连接结构对于两个 TALEN 半位点之间有效切割的范围具有显著性的影响。

在切割空隙的长短设计上需要做出权衡。切割空隙变宽能够减小 TALEN 靶位点的限制。然而, 这种设计也会提高脱靶的可能, 因为空隙越大 TALEN 切割所造成的半位点的错配也会更加频繁。+63 连接是最为常见的连接形式, 能够切割的片段长度在 10 到 30bp 之间, 虽然在一些实验中所观测到的切割长度只有 6bp 那么短[55]或是 39bp 那么长[54]。有一种+47 的连接结构被发现相对于+63 的切割范围更为严格, 但是也具有一定的范围, 从 12 到 21bp 不等[55]。有一种+18 连接被报道切割范围更为精确, 在 13 到 17bp 之间 (范围急剧缩短, 但是也有研究能在 24bp 处检测到切割) [54]。一种+17 的连接结构被报道切割的长度很相近, 能够切割 9-18bp 长度序列和 21bp 长度[55]。如果把限制脱靶活性作为应用的一个重要考量, 那么更短的 (+17 或+18) 的连接结构更为合适。

有一个研究组试图通过对 C-末端连接的重编程从而提升特异性[57]。他们将三或七个阳离子氨基酸替换成谷氨酸从而限制 C-末端结构域和 DNA 之间的非特异性作用。体外实验的结果很成功 (靶向率略有提高, 脱靶率大幅下降), 在培养细胞中所展现出来的靶向活性也很令人振奋。然而, 进一步研究发现只在脱靶极低的情形下才会发挥作用。

对 ZFNs 和 TALENs 的 FokI 结构域进行改造

无论是 ZFNs 还是 TALENs 能够和 FokI 核酸内切酶的催化基团连接切割 DNA。虽然曾经试图改变 ZFNs 或 TALENs 的催化基团, 例如 PvuII[58]或 I-TevI[59], FokI 仍是被最广泛采用的。在 FokI 结构域中已经进行过多种改造和升级使其能够胜任不同应用领域的基因编辑。

专一的异源二聚体

ZFNs 最初被设计, 是由“左” ZFN 和连有 FokI 切割结构域的“右” ZFN 组成的异源二聚体。如果一种脱靶位点由两个“左”或两个“右”半位点组成, 有可能使得两个 ZFNs 的 FokI 结构域发生二聚化从而产生 DSB, 能够将潜在的脱靶连接区域数提高一倍。通常情况下, 同源二聚化通常占据位点脱靶原因的大部分比率[28,45], 因为相对于其他形式单独的 ZFN 缺少特异性。为了解决这一问题, 对 FokI 的二聚化表面进行重编辑, 从而抑制同源二聚化的产生[60,61]。虽然实验的结果减少甚至是消除了同源二聚化所产生的脱靶效应[60,61], 在早期的一些研究中被发现靶向效率也有所下降[29]。异源专一性的构建极大地降低了同源二聚化位点脱靶事件的产生, 但是一些更为彻底的研究发现被认为是异源专一性的 ZFNs 也会在同源二聚化位点发生比较罕见的脱靶切割和突变诱导[27,28]。最目前的研究中, 最为流行的专一性的异源二聚化 FokI 对为“ELD/KKR”组合, 在每个 FokI 结构域都包括有三个突变, 无论是对于 ZFNs[62]还是 TALENs[63]来说, 相对于野生型的 FokI 结构域, 靶向性都有所提高。

Sharkey 增强

在 FokI 结构域进行过多种改造用以增强 ZFN 的切割活性[64]。有多种突变被发现能够提高哺乳动物细胞中的切割活性十倍以上,也有一些突变被发现能够与专一性的异源二聚化的 FokI 结构兼容。其中效率最高的突变 (S418P, K441E) 被称为“Sharkey”。虽然靶向活性有所提高, Sharkey 也被发现能够提高 bulk 细胞实验中的脱靶活性,即使是连接了专一性异源二聚的结构也无法避免[25],所以当脱靶成为了主要的考虑因素时,其应用需要慎重。一项更为彻底的对比试验即 Sharkey vs. 野生型 TALENs 的试验还未见报道。

只有刻痕

如果某种实验中所期望的 DNA 修复途径为直接性同源修复,那么与发生突变的非同源末端不同 (mutNHEJ),有一种可以降低脱靶活性的选择是利用“只产生刻痕”的酶。具有 D450A 突变的 FokI 能够失活 ZFN 中的催化结构域使其无法切割 DNA 链。因此,当一个具有活性和一个不具有活性 ZFN 相连的时候,不能产生完整的双链损伤,但是能够在 DNA 的一条链上产生损伤,产生一个“缺刻”。DNA 缺刻的修复不大会通过 NHEJ 通路,也很少会产生突变[65]。ZFNikases 因此被用来在人类细胞中诱导靶位点 HDR 的产生,相较于标准的 ZFN 核酸酶只会产生很低水平的 mutNHEJ 事件[65]。然而相较于正常的核酸酶,缺刻酶所产生的 HDR 的绝对比率也大幅度下降。虽然还未被证明,ZFN 缺刻酶和专一性异源二聚化结构的联用在理论上能够显著性的降低或完全阻止脱靶的发生。TALEN 缺刻酶目前还没有报道,但是已有未发表数据指出具有半失活 (D50A) FokI 结构域的 TALEN 对在一些情况下仍会在靶位点和脱靶位点产生 mutNHEJ 事件。

CRISPR/Cas9 系统

有一类人工编辑核酸酶不依赖于蛋白质和 DNA 之间的相互作用,而是基于一种细菌免疫系统的 RNA 介导的 DNA 切割酶而且是成簇的、间隔调节的,短回文重复序列 (CRISPR) [66-70]。相较于 ZFNs 和 TALENs, CRISPR 具有较大的改变,在对不同基因进行靶向时 CRISPR 相关 (Cas) 蛋白都是相同的;只有单链的向导 RNA (sgRNA) 能够被改变重新定位切割位点[12,14]。每一次构建都只需将一段退火后的寡核苷酸片段进行退火。这与 ZFNs 和 TALENs 完全不同,这些酶还需要对蛋白进行重新设计,对质粒进行重新构建,以对全新的序列进行靶向[71]。

gRNA-DNA 和 PAM 之间的相互作用

在 CRISPR/Cas9 系统中的早期研究显示:在远离前空格连接基序 (PAM) 区域由于向导链与 DNA 序列的错配能够产生高活性的脱靶[12,13,72,73]。CRISPR 系统也被发现能够和 PAM 序列附近的片段发生错配从而切割片段[55,65,74-76]。对于调控 CRISPR 脱靶切割的明确“规则”还未可知,而且特异性片段和向导 RNA (gRNA) 之间的相互作用还会受到强烈的未知因素的影响。然而,目前有趋势发现发生于接近于 PAM 的“种子”区的错配相较于 PAM 远距离错配具有更低的容忍度。PAM 是决定特异性的重要区域,即使是单碱基的突变在大多情况下也很可能造成切割的完全失效[77,78]。对于最为常见的 Cas9 (来自于链霉菌菌属,也称为 SpCas9) 差异识别的 PAM 序列,为连带有两个鸟嘌呤的任意序列 (基序 NGG),但是在 NAG PAMs 中的脱靶也能够被允许[79-81]。不同的 CRISPR 系统具有不同的 PAM 序列,只是在长度和序列偏好上有所不同,在 Cas 分子中进一步的工作从而产生更长的 PAM 有可能会提高特异性[82]。最近有工作发现相对于单独采用 gRNA (sgRNA) 的方式,分别采用转录激活 CRISPR RNA (tracrRNA) 和 CRISPR RNA (crRNA) 的方式有可能产生更低的脱靶活性[83]。

对 Cas9 进行改造

虽然在 Cas9 切割的结构学[84]和机制[85]上的研究飞速发展,但是 Cas9 还远未达到像是 FokI 那样可以随心所欲编辑的程度,然而,目前已经发现了一些改造方式能够降低脱靶活性。与 FokI 相近,只产生“缺刻”突变的 Cas9 也通过在一个切割活性结构域造成 D10A 或 H840A

突变的方式进行半失活[86,87]。这些 CRISPR 缺刻酶仍然会起到在目标位点诱导 HDR 但同时降低脱靶的 NHEJ 水平的作用。用来产生抵偿缺刻的 CRISPR 缺刻酶对诱导 NHEJ 的发生率机制 DSBs 相似[86,87]。由于单独在脱靶位点产生的缺刻相较于 DSBs 产生突变的概率低很多(低 50-1500 倍)，利用两个脱靶位点正在 100bp 之内的缺刻酶对所产生的双链损伤相较于标准的 Cas9 脱靶活性脱靶活性有所降低。将 FokI 和催化失活的 Cas9 进行融合所产生的融合蛋白需要将两个 FokI 结构域进行二聚化，切割模式与 ZFNs 和 TALENs 类似[88,89]。由于单缺刻 Cas9 在脱靶位点也能诱导低比率的 NHEJ，这些 RNA 导向的 FokI 核酸酶 (RFNs) 被发现具有更强的特异性，通过靶向位点的深度测序并未检测到 NHEJ 的产生，结果提示 Cas9 缺刻酶所产生的脱靶活性很低 (利用相同的 gRNA)。

整体实验的脱靶活性

如果核酸酶能够作用于目标靶向位点且并未在目标细胞中体现出总的细胞毒性，bulk 实验能够被用于检验脱靶水平的第一步。这些实验为核酸酶在基因组中是否会对大部分细胞群落产生不可逆的影响提供信息。这个实验的一个常见的阴性对照是大范围核酸酶 I-SceI，可在 AddGene (#21299) 获得哺乳动物表达质粒。I-SceI 高度的特异性在正常情况下并不会诱导在大部分细胞发生脱靶。

γ H2AX 位点

在基因组产生的双链损伤能够引起 DNA 损伤位点 H2AX (γ H2AX) 组蛋白磷酸化的快速产生[64]。大量的核酸酶脱靶活性能够引起显微镜或流式细胞术可见的多个位点的脱靶活性。Guo 等人报道有一款 Milipore (登记号#17-344) 公司的流式分析商业试剂盒能够被用于检测核酸酶的脱靶活性[64]。

细胞周期异常

由核酸酶脱靶活性产生的基因组中的 DSBs 能够阻止不同细胞周期之间的过渡，直到修复的完成。HeLa FuCCI 细胞 (最初由 Sakaue-Sawano 等人发现[90]，获得自 Amalgaam) 是一种简易的可转染系统，能够通过荧光信号判断细胞所处的三个复制时期： G_1 ， G_1/S 或 $S/G_2/M$ 。利用这一实验检测核酸酶的脱靶活性的方法由 Mussolino 等人提出[30]。

细胞生存和凋亡

如果细胞无法修复由核苷酸脱靶活性产生的 DSBs，细胞通常会发生凋亡或通过其他方式降低生存率。随着凋亡比率的提高 Annexin V 阳性细胞也被观测到[30]。或者，能够摄取 7-氨基酸放射菌素 D 染料 (7-AAD) 的为死细胞。虽然 7-AAD 并不能区分凋亡和坏死，但是相较于碘化丙啶染料 (PI) 来讲具有成本低、简易等优点，因为 7-AAD 的荧光谱和常用的转染对照的荧光蛋白并不存在重叠[91]。

荧光的缺失

如果核酸酶作为 DNA 的形式导入细胞，最常见也是最为直接的方式是将例如 eGFP 这样的荧光蛋白进行共转染[74]。在两个时间点 (通常是转染后的第 2 天和第 5 天) 具有荧光的细胞比例进行检测。如果在较早的时间点检测到了荧光，那么代表核酸酶成功的导入了细胞。因此，荧光的缺失 (例如第 5 天的 40%对比第 2 天的 80%，有约 50%的荧光缺失) 能够被用于指征核酸酶的毒性。这一方法能够比较测量接受了核酸酶转染的细胞群落相较于未转染核酸酶的同细胞的群落生存率的下降。

基于脱靶位点的实验预测方法

很多之前的有关于核酸酶脱靶活性的研究利用特异性的核酸酶的一些实验性的特征对潜在的脱靶位点进行预测。虽然基于预测方法的实验一般十分有效（几乎所有采用这些方法的已知研究中都会发现至少一个真实的脱靶位点），这些方法很具有技术挑战性、成本很高、而且浪费时间。由于这些技术在实际应用时会存在困难，有很多从未在原有实验室以外的地方进行重复过。

SELEX

通过指数富集法（SELEX）进行配体的系统化升级是一种已经成型的技术，能够对与靶向分子具有高亲和力的核酸序列进行筛选。这一方法被 Sangamo Biosciences 的实验室用于确证核酸酶的特异性。SELEX（图 2）已经被用于检测 ZFNs[92,93]和 TALENs[52,75,76]的连接偏好以及随后通过导向生物信息实验预判潜在的基因脱靶位点。操作流程具体为：（1）对核酸酶进行亲和分子标记，例如血球凝集素（HA），（2）在体外表达核酸酶，（3）与寡核苷酸半随机库进行孵育（通常偏向于核酸酶所期望的目标连接位点），（4）利用抗体捕捉核酸酶蛋白（5）对 DNA 片段进行 PCR 扩增。步骤（3）-（5）随后重复多轮，并用步骤（5）富集得到的富集产物替换步骤（3）中起始的半随机库。在进行过足够的筛选循环后，对 PCR 扩增产物进行测序和鉴定。SELEX 通常情况下能够扩增得到 20-50 个独特的 DNA 连接序列 [52]；如果发现了过多或者过少的片段，可以在最初的扩增循环后进行测序验证或增加筛选的循环数。这些序列随后会产生位点权重矩阵（PWMs）指定每个位点的连接偏好（图 2）。

一旦每个核酸酶半位点的 PWMs 建立后，就可以对整个基因组进行有效的生物信息检索并对每个位点进行评分。每个潜在的复式核酸酶的脱靶位点，其中的一半能够铜鼓近似长度的间隔序列进行筛选，而另一半位点，能够通过 PWM 对每个具有潜在脱靶位点的核苷酸进行评分（注意：来自于 Sangamo 的研究者并未公布 PWMs 评分的公式，但是所有的位点产物计算结果都大致相似[94]）。在基因组中的全部位点随后能够进行排序，并可从中选择合适的一组进行进一步的研究。

虽然这一技术还面临很多质疑，但是已经有实验验证实在 ZFNs 和 TALENs 中能产生稳定可靠的结果。这一技术的缺陷包括它只能提供每个核酸酶半位点连接偏好的信息，因此会忽略核酸酶切割中所需的两个半位点之间的相互作用。另一个限制是这一实验完全在体外进行，因此会忽略细胞环境下蛋白的改变以及细胞中有可能影响 DNA 基因组脱靶位点的因素，例如染色体的结构、可操作性和甲基化状态。最后，由于起始的寡核苷酸库是半随机的，这一方法更偏向于寻找目标位点相关高度同源的位点。然而，这也是目前为止报道最为广泛的能够成功的找到核苷酸酶脱靶位点的技术，而且也是仅有的两项被报道的能够预测 TALEN 真实脱靶位点的技术之一[75,76]。

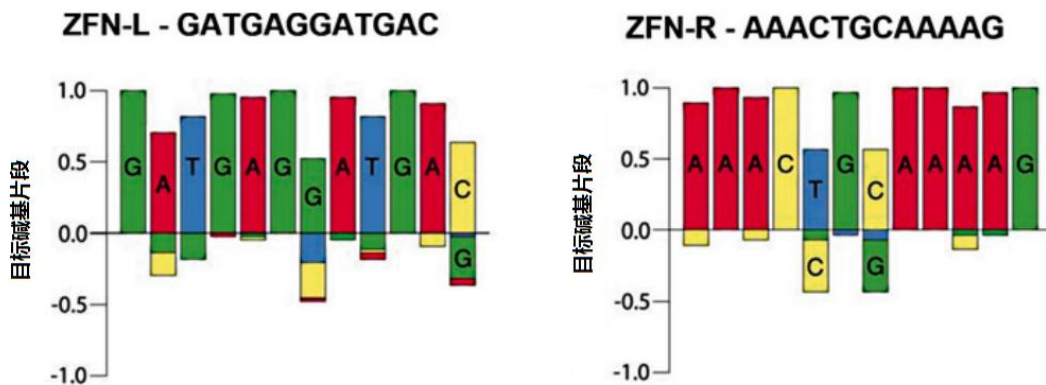


图2.ZFN靶向特异性。一种位点权重矩阵，能够给予SELEX分析对CCR5靶向的ZFNs的特异性进行鉴定。来源于Perez, EE 等人。(2008)利用锌指核酸酶通过基因编辑在CD4+ T细胞中构建HIV-1抗性。Nat Biotechnol, 26, 808-816。经自然出版组授权。

细菌单杂交

细菌单杂交法 (BIH) 和 SELEX 相似，能够对核酸酶单体的偏好进行分析。首先，需要构建一个报告基因上游的半随机（偏向于目标核酸酶位点）区域的报道质粒库。这一数据库随后整合到核酸酶 DNA 连接片段片吗的质粒上，与转录激活因子进行融合[29]。E.coli 克隆能够表达报道基因，是源于核酸酶 DNA 识别区域具有充分的亲和力，能够激活基因，从而被选择并通过对质粒测序的方式对半随机连接位点进行鉴定。所有的片段发现于 E.coli 克隆，能够通过编译产生 PWM 用于基因组潜在的脱靶位点的筛选，方式与 SELEX 相同。Scot Wolfe 的实验室利用 BIH 进行核酸酶的脱靶预测，也是到目前为止唯一使用该方法的课题组，且这一技术只被用于 ZFNs 脱靶活性的预测[29]。在单一单体分析方面这一方法与 SELEX 面临同样的质疑，但是由于该技术在细胞（虽然是非真核的细菌中）环境下进行，具有一定的优势，相较于完全在体外进行的试验能够更好的模拟蛋白质和 DNA 之间的相互作用。

体外切割

与前述的两种分别独立检测每个单体的预测方法不同，体外切割试验能够明确的指出在随机混合池中哪一种 DNA 能够被核酸酶切割[27]。这一方法已经被应用到 ZFNs[27]、TALENs[57] 和 CRISPRs[95]领域，但是只在 David Liu 的实验室中进行过应用。通过这种方式，构建了一种半随机寡核苷酸库，具有全核酸酶识别位点。对于成对的核酸酶（例如 ZFNs、TALENs 或配对的 CRISPR 缺刻酶），所有两个单体的半位点都在检测范围内，并通过近似空格长度的片段进行分选。这种核酸酶随后在体外表达，与寡核苷酸库进行孵育。多次酶学和凝胶分离步骤能够对核酸酶切割的序列进行筛选。在核酸酶孵育之前和之后对库进行深度测序，对核酸酶切割的片段进行鉴定。一项生物信息学的实验随后进行，从而确定在体外的一些位点的切割是否也同样发生于目标基因组中。这些位点的脱靶活性随后通过实验进行验证。

这一技术也具有很多的优势和限制。通过检测核酸酶切割而非单纯的连接，在原实验室中[27]提出了 ZFN 二聚化“能量补偿”假说模型，在一个半位点产生的大量的错配能够通过另一个半位点的很少或几乎没有错配得到补偿。然而，由于这个技术几乎不在体外进行，无法对细胞环境下核酸酶和基因组 DNA 之间的效果进行统计。而且，由于寡核苷酸库是半随机的，分析结果偏向于目标核酸酶靶点中更高同源水平的位点。

最近这一方法得到的改进在于对大数据更好的应用[26]。这一方法原有的应用是通过基因组发现与准确的配对切割片段[27,95]，但是这些与基因组配对的片段中能被核酸酶切割的只是总片段中的一小部分。通过利用一种 Bayesian 机器运行算法穷举出原实验中能够被 ZFNs 切割的 CCR5 和 VEGF 片段，随后对每种核酸酶在给定片段中发生切割的可能性进行评分。随后利用生物信息学筛选全基因组，这一过程和 SELEX 和 BIH 中 PWM 筛选相似，并得到评分较高的位点。通过这种方法对脱靶位点进行分析发现其能够定位低同源性的真实脱靶位点及低活性的位点。令人印象深刻的是，这一方法的犯错率也很低，并在 ZFNs 靶向 CCR5 和 VEGF 实验中[26]和两项 TALENs 靶向 CCR5 和 ATM 实验中[57]发现了大量新的真实脱靶位点。不幸的是，这种方法令人难以置信的难度和在每个核酸酶中所投入的巨大的时间消耗还需要进行进一步的探究，这也意味着体外切割实验和随后建立的 Bayesian 分级器机器算法限制

了应用这种方法研究核酸酶的数量。

IDLV LAM-PCR

由聚合酶链式反应（IDLV LAM-PCR）介导的整合缺陷型慢病毒载体连接扩增是两种完全在细胞环境中进行的脱靶检测方法中的一种。这一方法由 Christof von Kalle 的实验室研发[28]。在这一方法中，导入 IDLV 的细胞编码一种选择性标志物，例如绿色荧光蛋白（GFP）。由于这一病毒具有整合缺陷性，其整合进入基因组的能力受到了限制。因此，在数周后培养分裂的细胞中的 IDLV 基因序列迅速减少，在细胞分裂中也不发生复制。如果加入了核酸酶，所产生的 DSB 能够引起 IDLV 在细胞基因组中更效率的整合。随后在这一实验中，在培养细胞分裂数周后，与对照组细胞相比有大部分的核酸酶处理细胞能够表达选择标记的基因。随后对这些细胞进行分选和病毒整合位点的分析。简单来讲，这种方法在基因组 DNA 上进行 LAM-PCR 所采用的引物能够结合于 IDLV 区域的长尾重复末端（LTR）。由 LAM-PCR 所产生的扩增片段包括了基因组中与 LTR 相连接的部分，因此能够对 IDLV 的整合位点的扩增片段进行高通量的测序。整合位点簇（CLIS）分析通过限制两个独立性的整合位点必须在 500bp 之内过滤掉了许多与核酸酶活性无关的随机性整合（虽然在基因组中有一些脆弱的为为点也被证实是整合位点簇）。下一步是对 CLIS 位点周围的空间进行检索，寻找出一段和目标靶向位点同源的有可能作为核酸酶切割的真实位点的序列；随机片段空间相较于核酸酶靶点具有约 45%的同源性[28]，则测序存在大于 60%的核酸酶位点同源性。被预测出的脱靶位点随后可通过无 IDLV 的核酸酶处理过的细胞进行验证。

这一方法的主要限制在于缺乏敏感性。这一缺陷是此方法的固有属性，因为这一过程依赖于 DSB 修复中 IDLV 被获取的罕见事件。因此，很多脱靶位点，尤其是低活性的位点会被忽略；异源专一性的 CCR5 ZFNs 的 IDLV LAM-PCR 分析[28]仅发现了 38 个已知位点中的四个[26,45]。这一方法成功的被运用于 ZFN[28]，TALEN[96]和 CRISPR[96]脱靶位点并能够在细胞内环境下，在细胞基因组的原有结构内，对高活性的脱靶位点进行无偏测算。由于在寡核苷酸筛选库中没有偏向性，这一方法能够定位目标核酸酶靶向中极地同源性（66%）的真实脱靶位点[28]。由于这一方法缺少敏感性，它可能不是最优化的作为患者治疗中的核酸酶检测方式，也无法被用于检测能够造成不良反应的罕见脱靶的发生，但它仍然是一种经常被采用的检测手段，因为它没有偏向性的特点使其能够发现不适宜作为核酸酶特异性的标准模型中指导寡核苷酸酶建库或计算机研究的位点。

染色体免疫共沉淀测序

染色体免疫沉淀反应加深度测序（ChIP-Seq）是一种用来确定与特定蛋白相连的基因组片段的成熟方法。ChIP-Seq 能够通过一种亲和表位（通常是血球凝集素）对核酸酶进行基因标记并催化核酸酶失活（因此 DNA 能够结合但不能被切割），能够在细胞中表达改造后的核酸酶，将蛋白和 DNA 交联在一起，修剪基因组 DNA 成为更小的片段，并对核酸酶进行纯化（及与之交联的 DNA），利用抗体（免疫沉淀反应），对与核酸酶相连的 DNA 片段进行测序，随后在基因组中对这些序列进行定位。早在 2013 年，有研究发现 ChIP-Seq 能够用于在活细胞中进行无偏向的基因组范围内的检测[71]，但是截止到目前为止还未有实验达到最初预计的结果。

二聚化的核酸酶，例如 ZFNs、TALENs、RFNs 以及成对的 Cas9 缺刻酶，都对于 ChIP-Seq 有特别的要求。如同在 2013 年晚期利用 ChIP-Seq 对 CCR5 ZFNs 进行鉴定的一些不成功的尝试：“在基因组中可能存在上千的高亲和力单体靶向位点，然而单体并不足以产生损伤。而且，二聚化的 ZFN 位点与所有的两个单体之间连接较弱，DNA 切割效率很低，但是结合不足以达到 ChIP 的检测限” [26]。

由于 Cas9 能够作为单体核酸酶发挥作用，有三个课题组试图在 2014 年早期利用 ChIP-Seq 定位 CRISPR/Cas9 的脱靶位点。由于 Cas9 连接发生于目标位点之外的基因组的很多位点（能达到上千个，取决于所采用的 gRNA），但是通过两个课题组的检索发现脱靶活性（NHEJ）只发生于有限的少部分位点[97,98]，这说明这一方法在检测核酸酶脱靶效率方面具有很高的假阳性率。其它的课题组（Kuscu 等人）宣称通过她们的 ChIP-Seq 实验预测到了 53% 的真实核酸酶活性脱靶位点[99]。然而，但是用不同的窗口尺寸去观察核酸酶的诱导效率时，会发现 Kuscu 等人的结果具有很大的差异，而且在检测 indels 方面采用了非常规的方法（利用了 Bowtie2 中的 CIGAR 输出值而非对阅读序列和模板序列进行比对分析），这也使得他们的结果与其它的两个课题组具有显著性的不同。

总之，ChIP-Seq 目前还不是可信用度很高的预测核酸酶脱靶活性的方法。虽然它在识别靶向序列时具有很高的精度，但是 ChIP-Seq 连接所检测到的脱靶位点有时对没有核酸酶活性的检测位点会给出很高的评分[99]，ChIP-Seq 可能在根本上就不适合鉴定核酸酶的脱靶效应。目前有越来越多的证据显示 Cas9 进行 DNA 切割需要经过多个步骤[97]：与染色体上的多个位点连接，并在能够与 gRNA 互补的靶位点停留，但是只对最佳配对处进行切割。ChIP-Seq 能够检测到所有的连接序列，造成了核酸酶很高的假阳性率。在切割和连接之间缺少区分的方法，ChIP-Seq 有可能无法检测到低脱靶切割频率的位点，因为 ChIP-Seq 在那些位点所产生的信号无法提升至背景噪音以上。

进行 CRISPR 脱靶位点鉴定的其他基因组范围工具

随着 ChIP-Seq 的失败，最近有很多其它的方法被开发用于在基因组范围内对脱靶位点进行一种“无偏差”（不受目标位点的同源区段所影响）的预测。基因组 DNA 的降解也被用于进行脱靶 1 的预测[100]。还有一种鉴定位点的方式是采用高通量的、基因组范围内的迁移测序（HTGTS），通过有可能发生的迁移来鉴定潜在的脱靶切割位点，该方法利用了 PCR 和 NGS 的线性化扩增[101]。然而，重要的是由 HTGTS 原本所预测到的潜在性的脱靶位点，永远都无法证实是 NHEJ 所必须的，从而无法证实其作为脱靶位点的真实性。虽然有正交数据显示至少是有一些预测位点是真实的，更为透彻的研究需要对这种方法所得到的确切的假阳性率进行探究。

这一类最后的方法称为 DSBs 测序评估的全基因组无偏鉴定（GUIDE-Seq）[102]。这一方法和 IDLV 类似，但是采用了短双链寡核苷酸代替的全尺寸病毒。虽然看起来只有很小的偏差，但是却对方法的敏感性具有显著性的影响（这也是 IDLV 的主要缺陷），能够对脱靶位点上产生 NHEJ 的频率精确到百分之一。GUIDE-Seq 具有很低的假阳性率，这是 IDLV 系统的一个优势，也是到目前为止唯一的将脱靶位点和 NHEJ 的发生关联起来的方法。另一方面，在 GUIDE 测序操作下，通过 COSMID 生物信息工具穷举出所有的位点进行比较并选出最好的一个 gRNA（VEGF#1），结果发现 92% 的总脱靶活性（通过 GUIDE 测序读取）发生在 COSMID 所预测的最有可能的 25 个位点之内（图 3）。虽然计算机对 CRISPR 脱靶活性预测水平的逐步提高，但是基因编辑的应用仍对以实验为基础的高花费的彻底性的脱靶活性分析（例如 GUIDE 测序）而非单纯的计算机预测方式（例如 COSMID）具有很大的需求。在同一样本中通过对这两种方法的对比能够最大量的提供信息。仍存在 COSMID 能够预测出但其它方法无法发现的真实位点。例如在 CCR5 靶向性的 ZFNs 中这一独特的方法被发现能够定位额外的真实脱靶位点[45,101]。

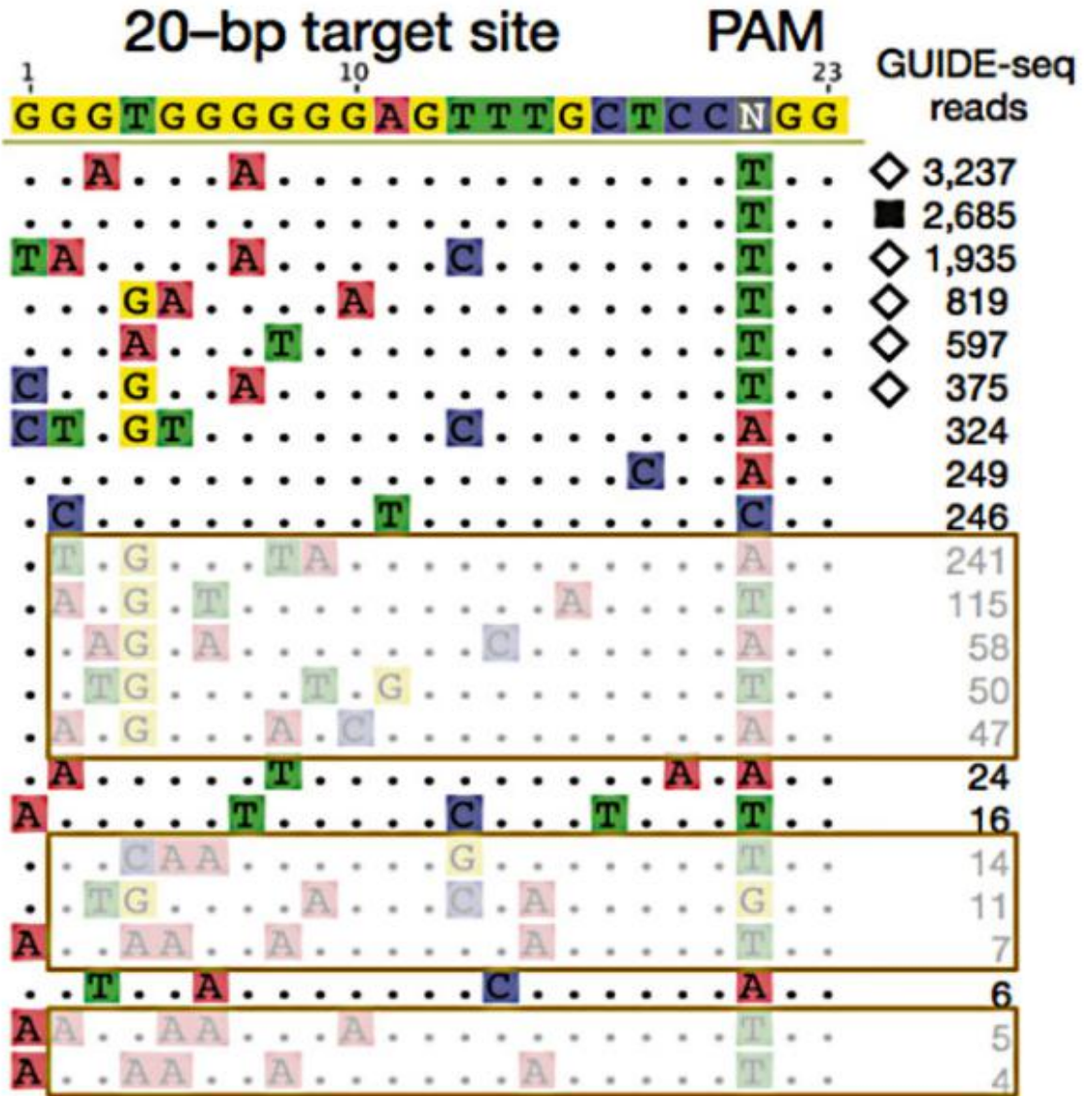


图3. 比较COSMID和GUIDE测序鉴定得到的脱靶位点。在VEGF gRNA#1中，通过COSMID所预测出的最有可能的25个潜在靶点占GUIDE测序得到的脱靶发生频率的92%。COSMID未检测到而GUIDE测序发现的位点用框圈出或进行淡化标记。这些位点未被COSMID发现是因为他们具有四个或在一个案例中具有五个，和gRNA错配的基因序列。由COSMID所列举出的第七位点具有gRNA的删除和两个错配。来源于Tsai SQ等人。(2014) GUIDE测序能够进行基因组范围内的CRISPR-Cas核酸酶脱靶切割的检测。Nat Biotechnol, 33, 187-197。数据的公开得到了Nature出版小组的批准。

在计算机中脱靶位点的预测方法

虽然以实验为基础的脱靶预测方式合理且准确，而且能够满足大部分核酸酶脱靶位点的检测需要，但是随着核酸酶设计的不断变化，其缺点也被不断放大。精确的计算机预测手段为脱靶分析通量的显著性提高提供了可能，相较于以实验为基础的方法，弥补了其所需的巨额花费和操作困难等缺陷。只有两个以实验为基础的预测方法（IDLV LAM-PCR 和 ChIP-测序）依赖于寡核苷酸的半随机数据库。然而，但是数据库的随机度和片段的长度直接相关；在存量有限的数据库中相较于长片段，短片段具有更高的随机度。大多数的以实验为基础的脱靶位点分析都以分别能够识别 9-10bp 或 12-13bp 序列的三指或四指 ZFNs 为对象。在过去的几年中有趋势发展这一技术用于检测更长的靶向序列；例如 Sigma-Aldrich 研发的每个单体能识别 18bp 序列的 6 指 ZFNs，也就是标准化的 CompoZr 核酸酶，靶向于 16-20bp 长度序列的 TALEN 单体，和能偶靶向识别 22bp 长度的 CRISPR/Cas9 系统（包括原空格序列和 PAM）。由于寡核苷酸库与目标位点同源性很高，这些以实验为基础的方法的优势被极大的拉低。诚然，利用 SELEX 实验所检测到的 TALEN 的三个脱靶位点中，基于片段重组的简单相关算法发现其中的两个都是真实的脱靶位点[45,75,76]。然而在体外的 CRISPR/Cas9 系统中[95]，相较于原始的基因组同源预测方法，所得到的真实的脱靶位点往往更少[78,103]。

目前的另一个趋势是将具有缺陷的（图 4）的 IDLV LAM-PCR 和 ChIP 测序开发成为具有更高核酸酶特异性的方法。虽然这些无偏预测的方法具有不依赖于寡核苷酸库的优点，但是它们在低频脱靶位点缺乏足够的敏感性。虽然在 ZFN 的脱靶位点预测中只需活性 >1% 的就能够被发现[26-30,45]，但是 TALENs 和配对的 CRISPR 缺刻酶则需要更高的特异性。直到今天，只有一个 TALEN 的目标同源性极低的（<87%）脱靶位点被发现具有 >1% 的脱靶活性[57]。目前配对的 CRISPR 缺刻酶相较于具有活性的单独的 CRISPR 核酸酶的脱靶活性低 50-1500 倍，在已知位点中的活性都远远低于 1%[86]。RNA 介导的 FokI 核酸酶（RFNs）是最近发现的另一种能够提高 CRISPR 特异性的体系[88,89]。

最后的一个主要趋势就是具有缺陷的基于实验的预测方法正面临着核酸酶发展过程中的瓶颈问题。对于大多种数据库来讲，鉴定 ZFN 所连接的靶位点高效切割螺旋是一个巨大挑战（图 4a）[20]。如果把相当多的时间和努力都用来开发一种独立具有活性的核酸酶，那么利用以实验为基础的脱靶预测方式确定核酸酶的脱靶位点就是一个明智性的选择。然而，基于目前的技术，高活性的 TALENs 和 CRISPR 能够被很简易的设计用于切割几乎所有的目标序列[12,104]。一个单独的研究者在数周之内就能得到十数种靶向于目标基因的高活性核酸酶，由此，基于实验方法对每个核酸酶的脱靶切割位点进行预测进而挑选特异性最高的核酸酶的方法就显得有些不切实际。这些限制能够引起大多数候选活性核酸酶产生任意的脱靶切割，无法对其所产生的脱靶结构进行正确的分析（图 4b）。然而，由于在计算机中脱靶位点的预测更为迅速，有越来越多的核酸酶通过采用这种方式评价脱靶切割的水平，从而能够对进行基因编辑应用的候选核酸酶进行更为理性化的选择（图 4c）

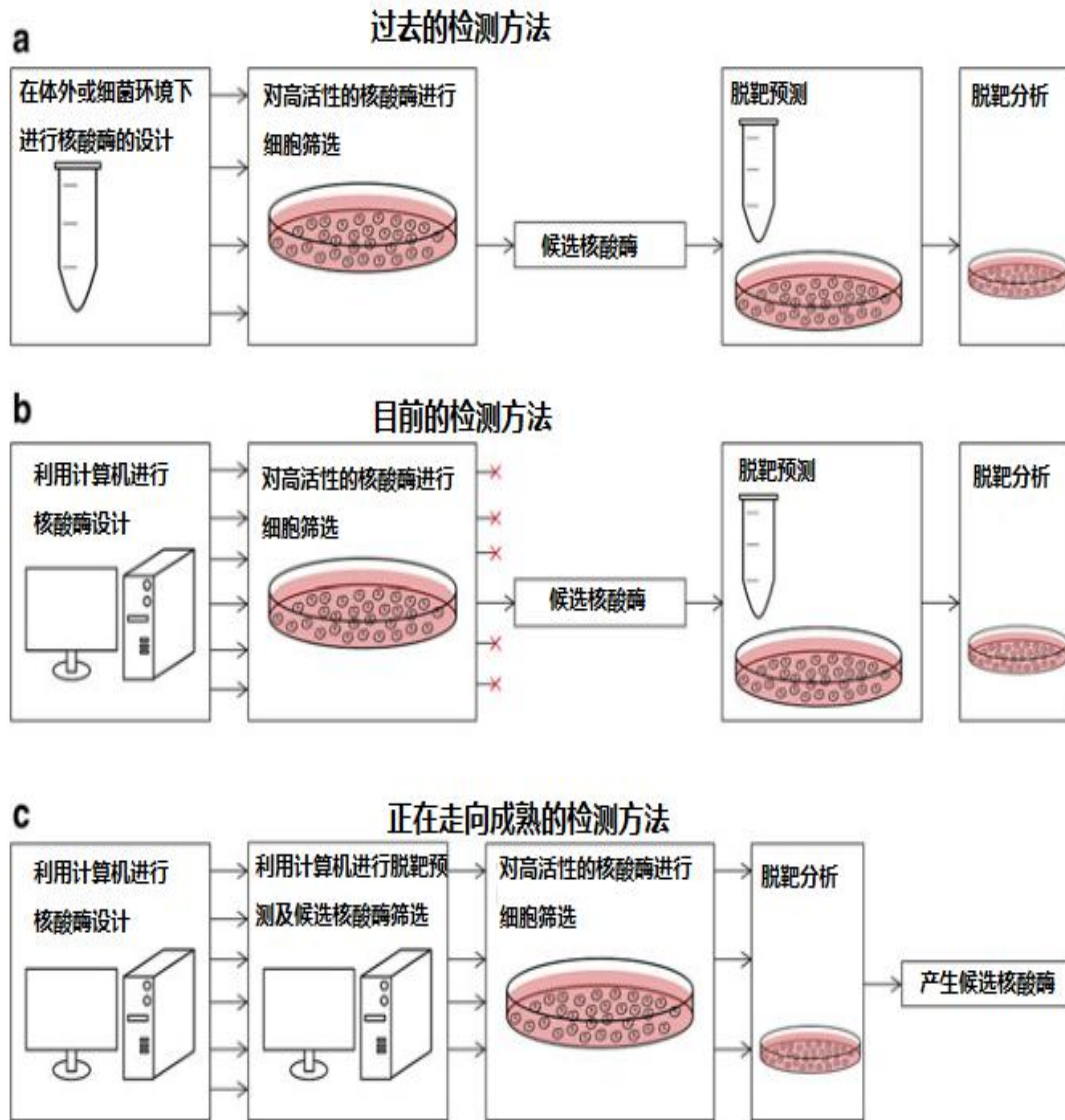


图4. 在核酸酶发展过程中检测方式的转变。(a) 在过去的研究中，TALEN和CRISPR的设计模式，核酸酶的筛选需要经历繁杂的过程，而在细胞检测中只有很少的具有代表性的一部分具有高活性。基于试验的脱靶效率的检测能够被用于少量的候选核酸酶特异性的筛选。(b) 计算机设计的简化导致了大量高活性的核酸酶的产生，这也对基于实验的资源密集型的脱靶预测方法造成了巨大的挑战，因此有很多具有活性的核酸酶在进一步的分析中被随机性的排除（红色的‘X’）计算机中脱靶预测方法的发展为所设计的核酸酶结构的初筛提供了可能，使得对高活性的核酸酶进行脱靶位点的发现及分析成为可能，因而可以基于最低的脱靶诱导效率的模拟预测对理想化的核酸酶进行更为理性的选择。

进行 ZFN 脱靶预测的计算机工具

PROGNOS

The Predicted Report Of Genomewide Nuclease Off-target Sites (PROGNOS)是一种在线工具，能够模拟 ZFNs 和 TALENs 的潜在脱靶活性（图 5a）[45]。对于 ZFNs，它能够进行多种因素的评估，包括序列的同源性、FokI 结构域造成的错配距离、鸟苷酸残基的存在率、ZFN 切割的二聚化能量补偿、独立的锌指亚单位的键能，从而对基因组中潜在的脱靶位点进行排序（图 5b）。PROGNOS 能够罗列出超过一半的已知的 ZFN 脱靶位点，并在经过了充分研究的 CCR5 ZFNs 中发现了一个全新的脱靶位点，并且也在新设计的 3 指 ZFNs[45]、4 指 ZFNs[30,45]和 Sigma-Aldrich 开发的 5 指 CompoZr ZFNs[105]也都发现了新的脱靶位点。为了辅助脱靶分析，PROGNOS 也能够产生不同的脱靶位点的 PCR 检测引物，能够被用于热循环条件下高成功率的高通量 PCR 扩增的进行（图 5c）。相对于其他的计算机预测工具，PROGNOS 的一个主要缺陷在于需要长时间的线上服务程序的运行，使得大数量的候选核酸酶的预筛选的灵活性很差。对于具备有一些编程知识的研究者来说，PROGNOS 也能够被下载并在本地通过命令行进行更加快速的运行。PROGNOS 的获得地址为

<http://baolab.bme.gatech.edu/Research/BioinformaticTools/prognos.html>

<http://bit.ly/PROGNOS>.

a **PROGNOS在线页面**

Nuclease Target: Genome to search:

ZFN:
 5' - -NNN...NNN- [] -3'
 3' - [] -NNN...NNN- -5'

TALEN:
 5' - [] -NNN...NNN- -3'
 3' - -NNN...NNN- [] -5'

Human (hg19)
 Mouse (mm9)
 Rat (rn4)
 O. sativa (MSU6)
 D. Melanogaster (dm3)
 Zebrafish (danRer7)
 ⋮

Maximum mismatches per half-site (1-10):

Allowed spacer lengths:

Include homodimeric nuclease sites:

Separation between Cel1 cleavage products (bp):

b **Sample Output**

23016 Potential off-target Sites		Rankings	Mismatches	PCR Product Size	...
Exons:	Promotors:	1-HBB	0	271 bp	...
1450	529	2-CANX	1	469 bp	...
Introns:	Intergenic:	3-ATG7	1	298 bp	...
9671	11366	⋮	⋮	⋮	⋮

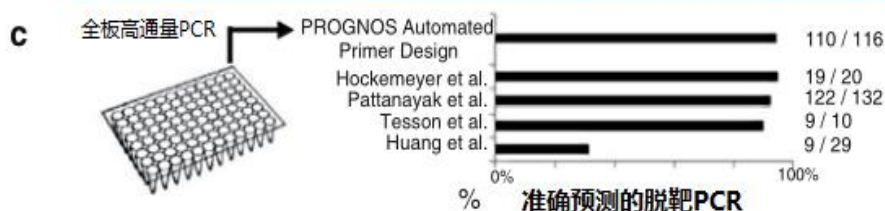


图5. 利用PROGNOS进行脱靶分析。(a) 将核酸酶靶向位点、目标基因组、可容许的间隔长度输入在线页面中。(b) PROGNOS根据基因组参数检索情况和根据检测结果的脱靶易发生率对所有的脱靶位点进行排序。(c) 根据输出结果自动设计产生PCR引物, PROGNOS所设计的引物相较于其它文献中的手动设计的引物具有较高的成功率。来源于Fine EJ等人(2014)一种在线的能够预测锌指和TALE核酸酶脱靶切割活性的生物信息学工具。Nucleic Acids Res, 42, c42。发表获得了来自Oxford大学出版社的允许。

ZFN-Site

ZFN-Site 是一个网站页面, 能够基于靶向序列和已知的核酸酶特异性对多个基因组的脱靶位点进行预测检索[94]。ZFN 位点采用 FetchGWI 搜索引擎, 能够提供快速的详尽的位点检索服务。位点的输出与基因组浏览器相关联, 能够促进脱靶位点的筛选。ZFN-site 的一个主要限制是在它的搜索算法中只能允许每个 ZFN 半位点最多两个错配的发生, 但是 ZFNs 发生脱靶的真实位点很可能存在四个或更多的锌指亚单位的错配。然而, 其快速检索的能力使其成为一种能够对潜在的 ZFN 目标位点进行快速筛选以保证在基因组中不存在除目标位点以外的两个或更少错配发生的额外位点。ZFN-site 的在线地址为 <http://ccg.vital-it.ch/tagger/targetsearch.html>。

TALEN 脱靶预测的计算机工具

PROGNOS

如 ZFN 部分所述, PROGNOS 是一种线上工具, 能对 ZFNs 和 TALENs[45]潜在的脱靶活性进行模拟。对于 TALENs, 其能对多个因素进行评估, 包括片段同源性、TALEN 二聚体的相互作用、N-末端错配的距离、具有错配连接的来自于“强”RVDs (例如 NN 和 HD) 的补偿效应, 以及通过对人工编辑的 TAL 结构域进行 SELEX 分析得到的 RVD-核苷酸连接偏好性, 从而提供基因组潜在的脱靶位点列表(表 5b)。利用 SELEX 预测方法得到的三个真实的 TALEN 脱靶位点中[75,76], 有两个位点通过 PROGNOS 方法检测位于列表排序的顶尖位置。对于七个新设计的 TALENs, PROGNOS 能够被用于定位八个额外的真实脱靶位点[30,45]。为了辅助分析, PROGNOS 也能够产生进行检测的不同脱靶位点的 PCR 引物测序片段, 能够被用于单组热循环条件下高通量高成功率的测序分析(图 5c)。

TALENoffer

TALENoffer 是另一种用于预测 TALEN 脱靶位点的网站工具[106]。它基于自然 TAL 效应器中 RVD 核苷酸的连接偏好进行建模预测潜在的脱靶位点。虽然它在脱靶位点定位的真实性方面要大大的优于原有的 TALE-NT 网络工具[106,107], 但是有研究发现其精确度不如 PROGNOS[45], 而且截止至目前也并未预测出任何全新的脱靶位点。然而, TALENoffer 进行全基因组检索的速度要快于 PROGNOS (以及快于 TALE-NT), 这使得其相对于其它网站能够更加迅速的筛选潜在的 TALEN 连接位点, 以确保基因组中不会存在脱靶活性评分较高的位点。TALENoffer 的获取地址为 <http://galaxy2.informatik.uni-halle.de:8976/>。

进行 CRISPR 脱靶预测的计算机工具

虽然 CRISPR/Cas 系统在全组织范围内被证实具有高效的切割活性, 脱靶切割仍然是一个主要的问题。可能的脱靶位点的检测对于向导序列的选择和处理后的彻底检查至关重要。可能性的基因脱靶切割位点可通过与特异性的 PAMs 相连的基因组中每个位点的向导链的比较进行鉴定。通常情况下使用者需要首先需要输入一些搜索条件, 脱靶程序能够输出与输入向导序列相似的与 PAM 相连的指定约束片段。目前已知的程序在搜索方法上具有很大的差别,

能够对输出结果进行评估和排序。排序结果正是在远离 PAM 的地方错配更有可能发生，然而切割可能发生在与 PAM 相近的错配向导链和目标位点之间。这一工具还需要进行进一步的完善，尤其是在排序方面，如下所述。

排序

由于进行脱靶活性筛选的位点数量是受限的，因此有必要对基因组位点进行排序从而匹配使用者所提供的条件。由于大量的真是脱靶的存在，且在向导 RNA 和基因组序列之间的错配碱基数甚至可大于 3 个，因此进行排序就显得尤为重要[77,78,103]。为了能够输出每一个可能的切割假定位点，很多程序将基因组互补序列和向导序列以及特定的 PAMs 错配的数量和定位进行考量。为了对给定的 gRNA 进行排序，一些程序能够进行基因区域的扫描并基于所有发现的脱靶位点或所采用的评分系统对靶向链进行排序。而通过这些工具所产生的数据能够更好的引导脱靶分析的进行，能够更好的对脱靶位点和向导链的选择进行排序。

Optimized CRISPR design tool 作为一种早期的程序能够基于人类细胞中不同位点上多种 sgRNAs 的检测实验对错配效果进行评估[78]。其他的工具能够提供更为简单的评分系统，输出结果观察得到的 5' 末端 gRNA 和基因组片段错配相较于 PAM 附近序列更为频繁的现象相一致。因此在多种工具中所采用的排序系统的构建基于多个因素的综合考量，尤其是 gRNA 和目标基因位点之间错配的总数和定位。在多种不同的工具中采用相同向导 RNA 的结果是相似的，但是排序结果存在偏差，如图 6 所示[108]。

在这种时候，所有的排序只基于序列信息，但是与基因组内容相关的因素并未纳入考量范围内。因此，所鉴定的序列都得到了相同的结果。我们和其他人观察到在很多已知序列的不同位点的切割水平具有显著性的差异[108]。在一个例子中，排序讲稿的脱靶位点（OT2）相对于一个靶位点具有更高的切割活性，而另一个位点在检测限之下（OT3）（图 6）。评分只能代表具有切割的可能或在不同细胞类型中或不同环境下有可能达到的最高切割水平。这一类的数据充分考量了基因组序列的内容和相似的 gRNA 和基因序列质检切割水平预测的困难性。为了模拟这一类的限制，需要对细胞特异性的数据进行整合，例如在高可能性的染色体中可能存在的位点，染色体甲基化以及转录连接位点；ENCODE 数据库的进一步成长有赖于这一信息的支持，未来的工具将会对这些因素进行整合。

Loci ID	COSMID (Rank)	Cas Online Designer (Rank)	ZiFit	CRISPR Design Tool	Cas Offender (Sorted)	Mutation Rate (%)	Gene
R01_OT2	2-6	2-7	15	-	18-139	43.6	None
R01	1	1	on	1	on	35.2	HBB
R01_OT10	7	2-7	3	-	3-17	23.4	None
R01_OT1	2-6	2-7	16	-	-	21.8	None
R01_OT5	2-6	2-7	5	-	3-17	15.9	None
R01_OT7	143-145	73-76	24	-	18-139	12.9	SECISBP2
R01_OT4	2-6	2-7	7	-	3-17	10.8	None
R01_OT8	355-357	238-241	34	-	18-139	6.6	VTI1A
R01_OT6	143-145	73-76	25	-	18-139	2.7	FSTL5

图6. 比较在线搜索工具的输出情况。通过减弱的T7E1酶活性对所观测到的具有两个错配位点的gRNA R-01的靶向及脱靶突变位点进行罗列。对匹配序列位点（第一碱基之外）的名字进行颜色标记。注释基因所对应的位点在右侧进行罗列。利用不同的在线检索工具进行脱靶位点分析。如果某一特异性的位点可通过已知工具得到，则列于最上端（例如Cas OFFinder），这一工具的位点排序输出结果（如果合适）展示如图。利用这一工具无法预测得到的位点用破折号或灰色框表示。

穷举发现所有的位点

最近有一项对比研究将其他工具输出得到的位点与真实的脱靶切割位点进行比较发现有一些工具根据用户输入的参数穷举出所有位点[108]。一些脱靶检索工具的差异在于在它们的非编码区域中包括CG 或其他的可能会高频出现的一些重复区域。以此必须注意程序中所采用的检测参数，以防在他们所设计的区域中不含有数以千计的脱靶位点。

我们最近发现能够被切割的基因组DNA对比RNA向导链或长（‘DNA突起’）或短（‘RNA突起’）的片段的切割水平要高于缺乏错配的互补位点[109]。由于在该研究中观测到了‘突起位点’的高水平脱靶活性，因而对COSMID进行了发展将这些位点加入到基因检索中，也被写入到COSMID工具的添加建议中（人际传播）。将突起的建议加入后，能够提高位点的输出数量，因此对多个位点进行脱靶活性的鉴定仍是必要的。

检索特征

新的CRISPR工具的产生或改进、以及对现有工具的提高正在稳步推进。目前脱靶检索工具（表1）的差距在于如上所述的根据使用者提供的线索穷举出的所有基因位点的能力以及它们的范围特征。一些工具首先扫描输入序列或基因组中的序列，随后根据预测得到的脱靶位点的情况确定向导链[110]，而另一些工具则是通过输入向导链穷举出所有可能的脱靶位点。由于在很多工具上都发现了大量的规律，虽然有一些还不能够称之为规律，利用应用E-CRISPR对用户所提供的序列进行脱靶检索时，类似的规律就被用于检测所敲入的基因是否能够被切割[111]。JackLin的CRISPR/Cas9 gRNA finder能够链接到一系列RNA二级结构网站工具，虽然它只能进行一些非穷举的BLAST和BLAT检索。如果一种工具缺乏目标基因组，可以根据需求进行多网站的联合。多种工具能够对两个正确性的间隔连接位点或许引起的CRISPR缺刻酶[86,112]对或CRISPR FokI[88,89]对[113-115]发挥切割作用所产生的脱靶位点进行定位，这橡胶油分别运行两个独立位点后进行整合更加简单，例如COSMID的表格输出结果就能实现独立性整合的功能[108]。

在一些CRISPR设计工具中的一些脱靶分析过程包括，检索特异性的基因区域，基于在所选基因组中的脱靶位点的预测数目划定潜在向导链的范围。虽然这一方法能够很简单的产生小脱靶概率事件的gRNA，范围的划定和微点的比较需要进行大量的假定工作从而充分评估可能的多种假定脱靶切割位点所产生的不良影响。这些CRISPR设计工具通过基于对脱靶事件发生率最低的位点的排序进行靶向位点的快速筛选[78,110,114-116]。这些工具在筛选脱靶位点的方式和位点排序的输出上具有显著性的不同，因此需要利用一种穷举检索工具进行重新扫描对所选定的靶向位点进行优化。

通过对给定数量位点的检测从而进行脱靶活性的测量与多种因素相关。如上所述，脱靶工具的差异在于如何进行匹配特定条件的所有位点的穷举。实验结果表明很多真实的脱靶位点并不能通过这种方式得到[102]。除此之外，新一代的测序经常需要提供足够数量的已知序列从而进行精确的测量，因为利用突变检测实验脱靶事件经常会被漏检，因此常常无法检测到低于1-2%的突变发生率[55,117]。虽然还存在很多限制，这种类型的比较能够提供每个独特的向导链的特异性的比较性读数。利用多种生物信息学和实验性驱动脱靶分析技术对一

种锌指核酸酶 (ZFN) 对进行研究发现, 每一种方法都能够输出已知脱靶位点中的一些区段, 而这些区段也可以通过其它方法获得[45], 因此想要定位所有的 CRISPR 脱靶位点需要采用多种方法联用的方式, 但是这仍然需要生物信息学和实验学方法的提高。

脱靶活性检测方法

如上所述的所有方法只有预测的功能, 而所获得的任何潜在脱靶位点不经过分选实验验证该位点中核酸酶活性的真实性。虽然一些预测方法相较于其他方法具有较低的错误发现率, 但是所有方法所得出的预测位点经过大量的验证实验后仍会发现一些没有脱靶活性的情况。如果利用特定的分析没有在该位点发现脱靶活性, 在可能的情况下可以选用敏感度更高的实验方法。有多种提升脱靶活性检测方法敏感性和难度的方式将在下面的内容中进行概述。每种检测方法的第一步是对潜在脱靶位点周围的目标区域进行 PCR 扩增。不同的检测方法对 PCR 扩增具有不同的要求, 尤其是在长度方面, 所以必须注意不同方法中 PCR 引物的设计。为了限制 PCR 反应期间突变的发生, 经常会用到高保真酶。在所有的检测实验中, 并行分析对照处理的细胞从而对实验背景中的脱靶位点信号水平进行对比监测是重要的。然而这个需求有一个可能的例外, TOPO 测序具有极地的犯错率。TOPO 测序也可被用来对对照处理细胞进行检测从而排除该位点的遗传多态性。

错配检测酶

这是一种常见的酶学实验检测技术, 能够对特定位点的脱靶活性进行快速的应答。然而这种方法的特异性很低, 通常情况下检测限度在约 1% 左右。在这种方法中, 扩增子能够通过变性、重退火形成异源二聚双链核酸分子, 其中一条链来源于野生型的扩增子, 另一条链来源于具有核酸酶诱导突变的扩增子。加入能够选择性切割这些异源二聚双链核苷酸的酶, 例如 Surveyor Nuclease (也被称为细胞酶, 来自 Transgenomic) 或 T7 核酸内切酶 (也被称为 T7EI, 来源于 New England BioLabs), 能够通过琼脂糖或聚丙烯酰胺凝胶电泳的谱带强度进行直接的定量。通过凝胶中条带的强度能够计算出样品中具有核酸酶诱导突变等位基因的百分比 [117,118]。这一实验在扩增子长度范围在 300 到 500bp 之间的时候效率最好。由于需要对同一块凝胶中的三条不同的条带进行定量, 包括一条原有条带和两条更小尺寸的在错配位点被酶切割开的两条反应产物, 扩增子长度的选择必须谨慎以保证三条带在切开的时候能够通过电泳充分的分开; 确保核酸酶发生切割的一般性原则是核酸酶的切割位点距离扩增子中点 >45bp, 距离扩增子每个尾端的距离 >45bp。目前能够获得利用 T7EI 酶检测核酸酶脱靶活性的详细分步方法说明 [119]。

TOPO 测序

TOPO 测序相对于错配酶检测的优势在于这种方法是真实的序列进行读取和编译。这一方法的全过程包括: 利用 TOPO TA kit (来自 Invitrogen) 对扩增子进行克隆并连入质粒, 通过 E.coli 进行转化, 随后从多个独立的 E.coli 克隆中提取质粒, 对质粒进行 Sanger 测序, 最后对测序数据中核酸酶诱导突变的证据的分析。由于 Sanger 测序能用来检测相对较长的序列 (通常 >750bp), 当所设计的 450-550bp 的扩增子的核酸酶切割位点位于扩增子中点附近时, 在 NGS 方法中有可能被忽略的长片段 indels 可以通过短片段阅读进行检测。由于 Sanger 测序非常准确的, TOPO 测序过程也能够克服 NGS 中易发生错读的典型序列 (例如 GC 富集延长或同聚体片段) 以及定位复杂的背景噪音 (图 6), 例如 CRISPR/Cas9 经常发生的小片段的 indels [77]。由于 TOPO 测序低通量的自然属性, 敏感性通常很低, 但是可以通过预热和对多个克隆进行测序而提高。

然而, 对于导入了核酸酶的胚胎分化组织或细胞克隆来说, TOPO 测序由于不需要较低

的检测限而且能够进行序列中全部的等位基因检测，因而是一种非常有用的工具。为了对具有两种等位基因的细胞群落进行分析，将扩增产物连入质粒并转化进入 *E.coli*。至少进行七个克隆质粒的测序以确保有足够的置信区间 ($p < 0.01$) 使至少一个测序结果包含每一种等位基因。在进行了胚胎的显微注射后，核酸酶仍然能够在胚胎生长的两个细胞阶段引起突变，有可能在全基因组中造成四种等位基因差异。因此，取自全组织的样本基因组 DNA 中，至少对 17 个 *E.coli* 群落进行测序以确保所有的死忠等位基因变异均有发生 ($p < 0.01$)。

SMRT 测序

单分子实时定量 (SMRT) 测序是目前发展起来的“第三代”测序平台，相较于 TOPO 测序和错配酶检测方面敏感性具有很大的提高。利用一种价格在约 \$ 250 范围内的试剂通过一次测序大约能获得 2,500 深度的测序数据(虽然不同的机器使用费用会造成测序评分的差异)。这也使得 24 个潜在脱靶位点的分析的敏感度提升至约 0.1%的精度，这也使得这一平台成为了实验室对有限数量的核酸酶进行分析时可被采用的一种理想手段。对于初次涉足高通量测序的实验室来说，SMRT 平台的另一优势是过程简单。与类似 Illumina 这样的平台相比，在扩增子的设计上具有最小的约束条件：扩增子长度在 270 到 400bp 之间，核酸酶靶向位点距离每侧尾端具有 >40bp 的长度。只要满足了上述两个条件，就能够将来自于多个脱靶位点的扩增子进行混合(大约保持相近的比例)进行单次的测序。对于引物序列没有特殊的要求，对于标准化的 PCR 反应所需的过程都可以进行。如果可能的话，如果扩增子位于较短末端范围内(测序读数将会高于平均质量)且核酸酶位点位于扩增子中点附近(能够观测到长片段删除事件)，能够得到更好的结果。利用 SMRT 测序进行核酸酶脱靶活性分析的描述可见于 Fine 等人的作品[45]。

Illumina 测序

对于需要高敏感度的情况，Illumina 测序能够做到最高的通量和每次反应最低的成本。然而如果每次反应需要覆盖 2,500 的长片段时，SMRT 能够提供更高质量的读数。如果将不同实验中的样品进行混合，在相同的敏感性条件下相对于 SMRT 能够节省大量的花费。然而，在进行 Illumina 测序扩增时必须注意。在两侧引物序列中必须存在特殊的适应性片段，而且必须经过第二轮的 PCR 筛选，对系列引物进行评估。重组性的扩增子长度根据所采用 Illumina 化学特性类型的不同具有较大的差异(单链末端 vs 双链末端，100vs250vs300bp 阅读长度)。当第一次尝试进行 Illumina 测序时，强烈推荐向测序中心机构或其它的相关公司进行咨询，对 Illumina 测序的过程和易犯错产生错误结果的关键性步骤进行深入的理解。

结论

人工编辑核酸酶是具有价值的研究工具，尤其是在治疗领域具有较高的发展潜力。基因编辑领域正受到越来越多的关注，而且在多类新开发的人工编辑核酸酶中也取得了振奋人心的结果；然而这种核酸酶都在一定程度上缺乏特异性。由脱靶切割引起的可能的灾难性的结果需要新的方法和实验进行准确预测并对脱靶事件进行定量，并寄期望于通过以上信息能够挑选出特异性最强、最安全的核酸酶进行应用，并保证经核酸酶处理后的细胞具有最小的脱靶活性。对脱靶切割活性进行检测和定量的方法在未来也可能辅助提升核酸酶的特异性。

致谢 这一工作得到了国立健康研究所的 NIH 纳米药物发展中心资金[GB 的 PN2EY018244]的支持。E.J.F 还额外得到了国家科学基金会研究生研究奖学金的支持[DGE-1148903]。

参考文献

1. Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS. Insertion of DNA-sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature*. 1985;317:230 – 4.
2. Capecchi MR. Altering the genome by homologous recombination. *Science*. 1989;244:1288 – 92.
3. Marcaida MJ, Muñoz IG, Blanco FJ, Prieto J, Montoya G. Homing endonucleases: from basics to therapeutic applications. *Cell Mol Life Sci*. 2010;67:727 – 48.
4. Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, Pâques F. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther*. 2011;11:11 – 27.
5. Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. *Nat Biotechnol*. 2005;23:967 – 73.
6. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005;435:646 – 51.
7. Chandrasegaran S, Smith J. Chimeric restriction enzymes: what is next? *Biol Chem*. 1999;380:841 – 8.
8. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res*. 2011;39:359 – 72.
9. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. 2010;186:757 – 61.
10. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*. 2009;326:1509 – 12.
11. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*. 2009;326:1501.
12. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819 – 23.
13. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013;31:233 – 9.
14. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337:816 – 21.
15. Petek LM, Russell DW, Miller DG. Frequent endonuclease cleavage at off-target locations in vivo. *Mol Ther*. 2010;18:983 – 6.
16. Tadepally HD, Burger G, Aubry M. Evolution of C2H2-zinc finger genes and subfamilies in mammals: species-specific duplication and loss of clusters, genes and effector domains. *BMC Evol Biol*. 2008;8:176.
17. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcrip-

- tion factor IIIA from *Xenopus* oocytes. *EMBO J.* 1985;4:1609 – 14.
18. Frankel AD, Berg JM, Pabo CO. Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc Natl Acad Sci U S A.* 1987;84:4841 – 5.
 19. Wolfe SA, Grant RA, Elrod-Erickson M, Pabo CO. Beyond the "recognition code": structures of two Cys2His2 zinc finger/TATA box complexes. *Structure.* 2001;9:717 – 23.
 20. Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods.* 2008;5:374 – 5.
 21. Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci U S A.* 1992;89:4275 – 9.
 22. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A.* 1996;93:1156 – 60.
 23. Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A.* 1998;95:10570 – 5.
 24. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science.* 2003;300:763.
 25. Gaj T, Guo J, Kato Y, Sirk SJ, Li CFB. Targeted gene knockout by direct delivery of zinc finger nuclease proteins. *Nat Methods.* 2012;9:805 – 7.
 26. Sander JD, Ramirez CL, Linder SJ, Pattanayak V, Shores N, Ku M, Foden JA, Reyon D, Bernstein BE, Liu DR, et al. In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* 2013;41, e181.
 27. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods.* 2011;8:765 – 70.
 28. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppl C, Nowrouzi A, Bartholomae CC, Wang J, Friedman G, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol.* 2011;29:816 – 23.
 29. Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA. Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Res.* 2011;39:381 – 92.
 30. Mussolino C, Alzubi J, Fine EJ, Morbitzer R, Cradick TJ, Lahaye T, Bao G, Cathomen T. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 2014;42:6762 – 73.
 31. Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure.* 1996;4:1171 – 80.
 32. Choo Y, Klug A. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc Natl Acad Sci U S A.* 1994;91:11168 – 72.
 33. Jamieson AC, Kim SH, Wells JA. In vitro selection of zinc fingers with altered DNA-binding specificity. *Biochemistry.* 1994;33:5689 – 95.
 34. Rebar EJ, Pabo CO. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science.* 1994;263:671 – 3.
 35. Choo Y, Klug A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc Natl Acad Sci U S A.* 1994;91:11163 – 7.
 36. Joung JK, Ramm EI, Pabo CO. A bacterial two-hybrid selection system for studying protein-

- DNA and protein-protein interactions. *Proc Natl Acad Sci U S A*. 2000;97:7382 – 7.
37. Joung JK. Identifying and modifying protein-DNA and protein-protein interactions using a bacterial two-hybrid selection system. *J Cell Biochem Suppl*. 2001;37:53 – 7.
 38. Durai S, Bosley A, Abulencia AB, Chandrasegaran S, Ostermeier M. A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions. *Comb Chem High Throughput Screen*. 2006;9:301 – 11.
 39. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell*. 2008;31:294 – 301.
 40. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods*. 2011;8:67 – 9.
 41. Moore M, Choo Y, Klug A. Design of polyzinc finger peptides with structured linkers. *Proc Natl Acad Sci U S A*. 2001;98:1432 – 6.
 42. Wolfe SA, Ramm EI, Pabo CO. Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers. *Structure*. 2000;8:739 – 50.
 43. Imanishi M, Hori Y, Nagaoka M, Sugiura Y. DNA-bending finger: artificial design of 6-zinc finger peptides with polyglycine linker and induction of DNA bending. *Biochemistry*. 2000;39:4383 – 90.
 44. Zhang HS, Liu D, Huang Y, Schmidt S, Hickey R, Guschin D, Su H, Jovin IS, Kunis M, Hinkley S, et al. A designed zinc-finger transcriptional repressor of phospholamban improves function of the failing heart. *Mol Ther*. 2012;20:1508 – 15.
 45. Fine EJ, Cradick TJ, Zhao CL, Lin Y, Bao G. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. *Nucleic Acids Res*. 2014;42, e42.
 46. Abarregui-Pontes C, Creneguy A, Thinard R, Fine EJ, Thepenier V, Fournier LRL, Cradick TJ, Bao G, Tesson L, Podevin G, et al. Codon swapping of zinc finger nucleases confers expression in primary cells and in vivo from a single lentiviral vector. *Curr Gene Ther*. 2014;14:12.
 47. Choo Y, Klug A. A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Res*. 1993;21:3341 – 6.
 48. Jantz D, Berg JM. Reduction in DNA-binding affinity of Cys2His2 zinc finger proteins by linker phosphorylation. *Proc Natl Acad Sci U S A*. 2004;101:7589 – 93.
 49. Handel EM, Alwin S, Cathomen T. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol Ther*. 2009;17:104 – 11.
 50. Schornack S, Meyer A, Römer P, Jordan T, Lahaye T. Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. *J Plant Physiol*. 2006;163:256 – 72.
 51. Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol*. 2009;48:419 – 36.
 52. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*. 2011;29:143 – 8.
 53. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. Efficient design and assembly of custom TALEN and other TAL

- effector-based constructs for DNA targeting. *Nucleic Acids Res.* 2011;39, e82.
54. Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable Di-residues. *PLoS One.* 2012;7, e45383.
55. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 2011;39:9283 – 93.
56. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug li RG, Tan W, Penheiter SG, Ma AC, Leung AYH, et al. In vivo genome editing using a high-efficiency TALEN system. *Nature.* 2012;491:114 – 8.
57. Guilinger JP, Pattanayak V, Reyon D, Tsai SQ, Sander JD, Joung JK, Liu DR. Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat Methods.* 2014;11:429 – 35.
58. Schierling B, Dannemann N, Gabsalilow L, Wende W, Cathomen T, Pingoud A. A novel zinc-finger nuclease platform with a sequence-specific cleavage module. *Nucleic Acids Res.* 2012;40:2623 – 38.
59. Beurdeley M, Bietz F, Li J, Thomas S, Stoddard T, Juillerat A, Zhang F, Voytas DF, Duchateau P, Silva GH. Compact designer TALENs for efficient genome engineering. *Nat Commun.* 2013;4:1762.
60. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol.* 2007;25:786 – 93.
61. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol.* 2007;25:778 – 85.
62. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods.* 2011;8:74 – 9.
63. Cade L, Reyon D, Hwang WY, Tsai SQ, Patel S, Khayter C, Joung JK, Sander JD, Peterson RT, Yeh J-RJ. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.* 2012;40:8001 – 10.
64. Guo J, Gaj T, Barbas 3rd CF. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J Mol Biol.* 2010;400:96 – 107.
65. Ramirez CL, Certo MT, Mussolino C, Goodwin MJ, Cradick TJ, McCaffrey AP, Cathomen T, Scharenberg AM, Joung JK. Engineered zinc finger nucleases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.* 2012;40:5560 – 8.
66. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindromic repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology.* 2005;151:2551 – 61.
67. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science.* 2010;327:167 – 70.
68. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet.* 2010;11:181 – 90.

69. Garneau JE, Dupuis M, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 2010;468:67 – 71.
70. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*. 2009;139:945 – 56.
71. Segal DJ, Meckler JF. Genome engineering at the dawn of the golden age. *Annu Rev Genomics Hum Genet*. 2013;14:135 – 58.
72. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*. 2012;109:E2579 – 86.
73. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. *Elife*. 2013;2, e00471.
74. Cornu TI, Cathomen T. Quantification of zinc finger nuclease-associated toxicity. *Methods Mol Biol*. 2010;649:237 – 45.
75. Tesson L, Usal C, Ménoret S, Leung E, Niles BJ, Remy S, Santiago Y, Vincent AI, Meng X, Zhang L, et al. Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol*. 2011;29:695 – 6.
76. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol*. 2011;29:731 – 4.
77. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*. 2013;41:9584 – 92.
78. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31:827 – 32.
79. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 2009;155:733 – 40.
80. Shah SA, Erdmann S, Mojica FJ, Garrett RA. Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol*. 2013;10:891 – 9.
81. Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol*. 2008;190:1401 – 12.
82. Fischer S, Maier LK, Stoll B, Brendel J, Fischer E, Pfeiffer F, Dyall-Smith M, Marchfelder A. An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J Biol Chem*. 2012;287:33351 – 63.
83. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim J-S. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2013;24:132 – 41.
84. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014;156:935 – 49.
85. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014;507:62 – 7.

86. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154:1380 – 9.
87. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013;31:833 – 8.
88. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol*. 2014;32:577 – 82.
89. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32:569 – 76.
90. Sakaue-Sawano A, Kurokawa H, Morimura T, Hanyu A, Hama H, Osawa H, Kashiwagi S, Fukami K, Miyata T, Miyoshi H, et al. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*. 2008;132:487 – 98.
91. Zembruski NC, Stache V, Haefeli WE, Weiss J. 7-Aminoactinomycin D for apoptosis staining in flow cytometry. *Anal Biochem*. 2012;429:79 – 81.
92. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKolver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol*. 2009;27:851 – 7.
93. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee Y-L, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol*. 2008;26:808 – 16.
94. Cradick TJ, Ambrosini G, Iseli C, Bucher P, McCaffrey AP. ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites. *BMC Bioinformatics*. 2011;12:152.
95. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013;31:839 – 43.
96. Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, Chang T, Huang H, Lin R-J, Yee J-K. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol*. 2015;33:175 – 8.
97. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*. 2014;32:670 – 6.
98. O'Geen H, Henry IM, Bhakta MS, Meckler JF, Segal DJ. A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res*. 2014;43:3389 – 404.
99. Kucsu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. 2014;32:677 – 83.
100. Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim JI, Kim JS. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods*. 2015;12:237 – 43.
101. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol*. 2015;33:179 – 86.
102. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C,

- lafrate AJ, Le LP, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol.* 2014;33:187 – 97.
103. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol.* 2013;31:822 – 6.
104. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol.* 2012;30:460 – 5.
105. Abarrategui-Pontes C, Creneguy A, Thinard R, Fine EJ, Thepenier V, Fournier LRL, Cradick TJ, Bao G, Tesson L, Podevin G, et al. Codon swapping of zinc finger nucleases confers expression in primary cells and in vivo from a single lentiviral vector. *Curr Gene Ther.* 2014;14:365 – 76.
106. Grau J, Boch J, Posch S. TALENoffer: genome-wide TALEN off-target prediction. *Bioinformatics.* 2013;29:2931 – 2.
107. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ. TAL effector-nucleotide targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res.* 2012;40:W117 – 22.
108. Cradick TJ, Qui P, Lee CM, Fine EJ, Bao G. COSMID: a Web-based tool for identifying and validating CRISPR/Cas Off-target sites. *Mol Ther Nucleic Acids.* 2014;3:e214.
109. Lin Y, Cradick TJ, Brown MT, Deshmukh H, Ranjan P, Sarode N, Wile BM, Vertino PM, Stewart FJ, Bao G. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* 2014;42:7473 – 85.
110. Xie S, Shen B, Zhang C, Huang X, Zhang Y. sgRNACas9: A software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLoS One.* 2014;9:e100448.
111. Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, Gagarinova A, Pogoutse O, Brown G, Binkowski A, et al. A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Mol Microbiol.* 2011;79:484 – 502.
112. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X, et al. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods.* 2014;11:399 – 402.
113. Xiao A, Cheng Z, Kong L, Zhu Z, Lin S, Gao G, Zhang B. CasOT: a genome-wide Cas9/gRNA off-target searching tool. *Bioinformatics.* 2014;21:2014.
114. Sander JD, Maeder ML, Reyon D, Voytas DF, Joung JK, Dobbs D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res.* 2010;38(Suppl):W462 – 8.
115. Ma M, Ye AY, Zheng W, Kong L. A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes. *Biomed Res Int.* 2013;2013:270805.
116. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* 2014;42:W401 – 7.
117. Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ. A rapid and general assay for monitoring endogenous gene modification. *Methods Mol Biol.* 2010;649:247 – 56.
118. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protocols.* 2013;8:2281 – 308.
119. Fine EJ, Cradick TJ, Bao G. Identification of Off-Target Cleavage Sites of Zinc Finger Nucleases and TAL Effector Nucleases Using Predictive Models. In: Storici F, editor, *Gene correction*. Totowa: Humana Press; 2014. Vol. 1114, pp. 371 – 83.

120. Zhu LJ, Holmes BR, Aronin N, Brodsky MH. CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-CAS9 genome-editing systems. *PLoS One*. 2014;9, e108424.

121. Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30:1473 – 5.

通过基因编辑核苷酸酶进行细胞编辑和疾病建模

Mark J. Osborn and Jakub Tolar

摘要 有两种相关技术推动着医学未来的发展：一是如何诱导和保持细胞的全能性，二是利用编辑核酸酶精确操纵基因组。总的来讲，这两种技术在新一代的细胞和基因治疗中具有巨大的潜力。这篇综述将会讨论人类和动物的干细胞模型和利用编辑核酸酶进行精确基因靶向和控制。在动物的研究和建模中，核酸酶被获准进行更为复杂和广泛的应用。早期的在小鼠基因组中不可靶向的序列通过对核酸酶的人工改造目前也能进行编辑。在发现基因编辑蛋白之前，大鼠全基因组的基因靶向和操作都是非常困难的。进行更大动物的基因编辑的能力能够降低器官移植的难度并能扩大食物来源的范围。最后，对造血干细胞、胚胎干细胞或体细胞进行编辑的能力能够被用于构建疾病相关模型，或用于针对性更强、效率更高的疗法中。总之，基因编辑核酸酶的效率和利用它们在不同细胞不同种属中进行基因编辑的能力为研究的创新提供了机遇，并为基因操作提供了更为复杂、精确性更高的合适模型。

关键词 间隔空隙调控短片段回文重复序列簇（CRISPR）/Cas9 • 胚胎干细胞（ESC）• 造血干细胞（HSC）• 同源重组（HR）• 插入/删除（indels）• 诱导多功能干细胞（iPSC）• 大范围核酸酶（MN）• 非同源末端插入（NHEJ）• 寡核苷酸模板（ODN）• 重编程 • 体细胞核移植（SCNT）• 转录激活样效应器核酸酶（TALEN）• 锌指核酸酶（ZFN）

M. J. Osborn , Ph.D.

Department of Pediatrics, Division of Blood and Marrow Transplantation, Masonic Cancer Center, Stem Cell Institute, and Center for Genome Engineering , University of Minnesota , Minneapolis , MN , USA

e-mail: osbor026@umn.edu

J. Tolar , M.D., Ph.D. (*)

Department of Pediatrics, Division of Blood and Marrow Transplantation , Stem Cell Institute, University of Minnesota , MMC 366, 420 Delaware Street SE , Minneapolis , MN 55455 , USA

e-mail: tolar003@umn.edu

© American Society of Gene and Cell Therapy 2016

T. Cathomen et al. (eds.), Genome Editing, Advances in Experimental Medicine and Biology 895, DOI 10.1007/978-1-4939-3509-3_12

从定义上来说，干细胞是一类具有自我更新和分化成为多种类型细胞能力的细胞。其类型包括胚胎干细胞（ESCs）、诱导多功能干细胞（iPSC）和造血干细胞（HSCs）。ESCs 最早在 1998 年被发现[1]，来源于胚芽分化和体细胞系形成之前囊胚阶段的内胚层细胞[2]。根据 August Weismann 的种质学说，这一特征在正常情况下能够赋予这些细胞可能的生存优势[3]。这一学说宣称来源于体细胞的细胞系和干细胞的获得性突变能够对躯干产生影响，但是反之不尽亦然[4]。例如，细胞系的基因组受到保护，自然选择倾向于对抗致病突变的产生。由于这些细胞具有作为再生和移植药物研究工具的可能，但是这些细胞的自然来源受到了严格的法律和道德的约束，这限制了它们的广泛性应用。

在最近的 60 年内，Weismann 学说中的部分内容被推翻，部分内容得到了证实。在二十世纪五十年代到六十年代之间，Robert Briggs、Thomas King 和 John Gurdon 发现当将不同分化程度的细胞内的细胞核导入到无核卵母细胞中时，能够产生一个完整生命个体（一只青蛙）[5,6]。在 1997 年，Ian Gurdon 及其同事利用这一类的体细胞核移植技术（SCNT）产生了 Dolly 绵羊[7]。这一过程依赖于体细胞核移植到无核卵母细胞中。然而，在 2006 年 Yamanaka 实验室发现通过导入特定的转录因子，体细胞能够进行重编程回到全能性的状态，这一类的细胞也被称之为诱导型多功能干细胞（iPSC）[8]。这些研究推翻了 Weismann 学说中有关于体细胞无法逆转产生“干性”的说法，但是证实了其中有关于体细胞基因组异常能够影响新细胞类型的表型[9]。

在干细胞领域中所产生的第二种技术赋予了使用者精确操作基因组的自由。这一技术采用多种形式和结构的具有统一的识别位点和独特的 DNA 连接片段的基因组编辑类核酸酶。大多数的研究将这些蛋白和核酸酶结构域相连并利用其固有属性切割 DNA。一旦 DNA 发生损伤，有两种主要的损伤途径被用于基因组的编辑：非同源末端连接（NHEJ）和同源重组（HR）。NHEJ 是易错途径，在缺乏外源性模板的情况下，在修复 DNA 损伤的同时经常伴有小片段的插入或删除（‘indels’），能够对 DNA 编码序列造成永久性的破坏。基因的修复依赖于无错的 HR 途径。当同时导入了核酸酶和外源性模板时，DNA 损伤的修复采用外源性的模板作为部分的“切割和替换”策略进行精确性的用户目标序列的插入。通过调控核酸酶进行 NHEJ 和 HR，能够完成复杂的精确性的基因敲除、敲入和修复策略，并能够使研究者对细胞进行改造构建模型用于疾病研究、药物发现或者用于基因纠正。这些用于基因编辑的核酸酶包括大范围核酸酶（MNs）、锌指核酸酶（ZFNs）、转录激活样效应器核酸酶（TALENs）和间隔空隙调控性短回文重复片段簇（CRISPR）/Cas9 系统。

大范围核酸内切酶

MN 的家族成员有很多，而其中的一种具有 LAGLIDADG 蛋白基序，被认为是最适合进行特殊位点的基因编辑的 MN[10]。MNs 曾用于人类细胞的编辑，但并不是直接作用于多功能干细胞[11,12]。在啮齿类动物中，有概念验证实验将绒毛蛋白基因敲入到一类 I-SceI 的靶向位点中，随后瞬时表达 I-SceI MN 能够引起 HR 提高 100 倍[13]。Ménoret 等人也进行了精确靶向实验，利用 I-CreI 归巢内切酶经人工改造后靶向于 Rag1 基因[14]。将质粒 DNA 注射到受精卵中能够引起 0.6% 的兔子和 3.4% 的老鼠在该位点产生突变，并能引起 T 细胞和 B 细胞水平显著下降，且保持 NK 细胞水平不变[14]。除了破坏基因的能力，MN 由于其高昂的成本和长时间的生产周期限制了其更为广泛的基因编辑的应用；然而，它们的单分子构造的、高活性的自然属性使它们在未来的基因编辑疗法中具有独特的优势。

锌指核酸酶

Zif268 转录因子能够作为定制化编辑 ZFNs 的功能性模板连接并调控多种细胞基因的表达。通过改变 Zif268 中与 DNA 残基相连的特异性，有多种能够识别用户期望序列的基因特异性

ZFNs 已经被生产出来。能够识别特定片段的典型的 ZFNs 具有一段与 DNA 片段相连的特殊组织阵列。阵列中每一个独立的单位、或是指状结构由大约 30 个氨基酸组成，对应三对 DNA 碱基对[15]。多个指状结构相互连接，形成了一个 ZFN 单体串联阵列，每一个类似的阵列都能识别连接特定的序列，也称之为“半位点”。因此，具有完整识别复杂目标位点的结构由两个半位点及两者之间的一段‘空白序列’构成，这一复合体能够介导双链损伤（DSB）的发生（图 1）[15,16]。直到今天，ZFNs 仍被广泛应用于哺乳动物干细胞领域。

常见的基因编辑啮齿类动物是研究者进行基因功能研究的核心工具，其中基因编辑的小鼠是最为常见的动物模型系统[17]。由于功能性的 ESC 细胞系只存在于小鼠而非其他物种，受精卵的显微注射能够被用于代替直接性的啮齿类 ESCs 的应用。

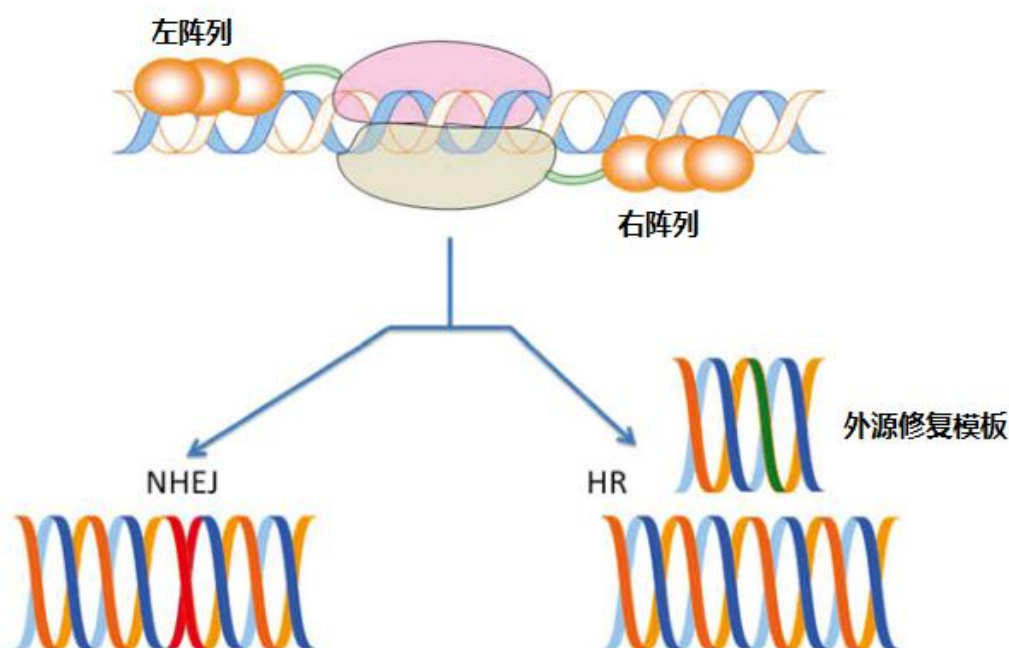


图1. 锌指核酸酶结构。上图所示为一类三指ZFN，具有完整的左和右阵列的异源二聚化复合体。每个独立的指状结构能够识别3bp的DNA，这一阵列与FokI核酸酶结构域进行连接（粉红色和棕色的椭圆），能够介导空隙区域DSB的产生。DNA损伤能够通过 易错的产生突变的非同源末端连接（NHEJ）或同源重组（HR）进行修复。红色指代NHEJ介导的插入/删除和由外源性模板介导的HDR（右）。图片利用Motifolio Inc.Sykesville,MD软件产生。

啮齿类动物

在 2009 年，第一个被证实的啮齿类基因组编辑实验由 Geurts 等人利用 ZFNs 在大鼠胚胎细胞中完成。他们靶向了一个外源整合的报告基因和内源性的 IgM 和 Rab38 基因，并发现在代际遗传中具有 25-100%的稳定基因破坏率[18]。第二项实验在大鼠模型中的 Il2rg 位点达到了约 20%的破坏率[19]。通过类似方法，这些研究开发了大鼠的高血压[18]和免疫缺陷[19]模型，并建立了单克隆抗体生产平台[18]。Cui 等人也进行了大鼠基因组的靶向整合，将外源的模板序列通过 HR 整合到 Mdr1a 和 PXR 基因中[20]。这些研究高效的解决了之前研究中大鼠基因不易被靶向的困难[18]，且在一些神经行为学的研究中大鼠相对于小鼠是更为具有优势的研究模型[21]。

ZFNs 在 2010 年成功地用于小鼠的目标基因敲除和敲入[22,23]。敲除实验包括在 C57BL/6 或 FVB/N 种小鼠中 *Mdr1a*、*Jag1* 和 *Notch3* 基因的靶向[22]。在刚出生的动物中进行编辑的效率在 20%到 75%之间，等位基因的破坏能够高效的通过生殖细胞系进行代际遗传[22]。通过成功的运用 ZFNs 和外源性模板对 *ROSA26* 位点进行靶向改造，Meyer 等人相对于一项早期的研究将 HR 的频率提高了十倍以上[23,24]。这些具有改革意义的基因编辑研究进展飞速（最少四个星期），目前基因编辑能够用于啮齿类动物的任何位点[22]。这些技术上的突破能够促进更为流行的意义更为深远的疾病模型的产生。例如，非肥胖糖尿病（NOD）小鼠已经成为了 1 型肥胖研究中的重要模型。Chen 等人利用 ZFNs 在 NOD 胚胎中进行操作，能够敲除编码 *CD137* 基因的 *Tnfrsf9* 基因[25]。他们发现 *CD137* 在胰岛炎的发展中并不重要，但是是促进过度肥胖的重要因素[25]。因此 ZFNs 和种属特异性的胚胎能够构造出纯种的 NOD 背景的疾病模型，并可移除潜在引起 ESCs 遗传背景污染的混杂的靶标或者是疾病标志因素，并移植到宿主中。小鼠中更扩大化的疾病模型也可用 ZFNs 产生。Meyer 等人发现将 ZFNs 和单链寡核苷酸导入小鼠胚胎细胞能够引起特定定位点的错义突变[26]。这一方法能够简化相关动物模型产生的模式并能够促进更为广泛的体内应用和疾病建模、基因研究和疗法探索。能够证实这一概念的第二篇文献来源于 Osiak 等人，设计了一种用于产生基因敲除鼠 ESCs 的无需筛选的方法[27]。在缺少选择性的标志的前提下，精确靶向于基因组中特定定位点的能力使构建的过程更为流畅，并最大程度的减少了疾病模型中潜在的异源转基因/外源插入片段的影响。

在更大的动物模型中应用 ZFN 介导基因编辑寄期望于提高食物的产量或通过减少免疫原性或不必要的特征使其在相关疗法中更加安全。这些产生模型用于最终人类间接体内疗法的动物研究主要集中于将皮肤细胞进行重编程产生多功能干细胞。在人类中这依赖于 iPSC 衍生技术。而在动物中，这主要集中于 SCNT，一种用于将细胞中的细胞核移植都缺乏细胞核的卵母细胞中的克隆技术[28,29]。

猪

在 2011 年，Whyte 等人利用 ZFNs 在猪的纤维母细胞中敲除了 GFP 报告基因，而 Hauschild 等人成功的破坏了猪的 α 1,3-半乳糖基转移酶（GGTA1）基因的全部等位基因[30,31]。对酶进行 GGTA1 编码需要产生 Gal 表位[31]，具有高度的免疫原性，能够引起异种移植模型急性的排斥反应[32]。删除或破坏免疫排斥的抗原应答为异种移植器官的基因编辑提供了巨大的可能性，每年能够帮助解决近 7,0000 人免疫器官来源不足的矛盾[33,34]。

牛

对牛基因组的编辑有可能提高每日食品产出的数量和质量。Yu 等人利用 ZFNs 介导牛基因组中 β -乳球蛋白（BLG）基因处的 NHEJ 破坏，能够破坏牛奶中作为主要抗原的该基因编码蛋白的产生[35]。Liu 等人利用 ZFNs 形式的缺刻酶即存在一个不具备 DNA 链切割活性单体的 ZFNs 二聚体。作用结果是造成了 DNA 单链损伤，极大的提升了靶向基因敲入的准确性[36]。利用所设计的 ZFN 缺刻酶靶向牛 *CSN2* 基因位点中的第二内含子，它们敲入了拟葡萄球菌溶葡萄球菌酶基因。乳腺炎是奶牛中的一种常见病，而预防性的抗生素疗法可见于约 90%的牧群中[37]，因为更健康的奶牛能够产生更多产量质量更好的奶制品[37]。Yu 及其同事发现，利用 ZFNs 和 SCNT 能够表达溶葡萄球菌酶和其他类似的基因能够对畜牧业产生有利的影响，因为由拟葡萄球菌产生的溶葡萄球菌酶能够控制拟葡萄球菌引起的乳腺炎[38]。这一策略的主要缺陷在于饮用级别的牛乳中抗生素的沉溺在能够使人获得对抗生素的耐受性[39]。耐药性是公共健康的一个重要的威胁，因为抗生素在人类中的应用被认为能够延长人类的预期寿命 2-10 年[40,,41]。

人类

可用于基因编辑的人类干细胞包括 HSCs、ESCs 和 iPSCs。HSCs 由于其具有分化成所有血细胞系的能力是非常理想的目标细胞群落，在全球的移植中心被用于多种恶性肿瘤疾病和非肿瘤疾病的治疗（2）。对于人类来说，只有少量的 ESC 细胞系能够被无限制的应用。然而有多种野生型和疾病模型的 iPSCs，能够绕开 ESCs 相关的潜在限制约束。

在人类细胞中对 HSC 基因组进行编辑的方法主要集中于利用 ZFNs 触发 NHEJ 和 HR 的 DNA 修复途径。HR 代表未来基因治疗的金标准，能够特异性的产生无缝突变修复。Lombardo 等人利用 ZFNs 和具有 GFP 或嘌呤霉素抗性基因的外源性模板在 CCR5 基因位点造成最大的基因靶向率为 0.11%[42]。在 2014 年，新一轮的突破性研究中成功利用 ZFNs 将基因插入 AAVS1 或 IL2RG 基因位点[43]。这一研究最近在 HSC 细胞系中取得了进展，并得到了更高比率的 HDR[44-46]。通过拮抗芳香烃受体 StemRegenin 1 (SR1) 驱动相关的 CD34+ HSC 的扩增（HSCs 间接体内疗法自发分化速度降低的结果）能够提高异种移植人鼠嵌合模型中移植的比率[47]。16,16-二甲基前列腺素 E2 (dmPGE2) 是一种稳定的 PGE2 的类似物，被发现能够调控造血作用并提高移植[48,49]。SR1 和 dsPGE2 的应用能够保留造血干细胞的表型，并通过整合缺陷型的慢病毒 (IDLV) 将 ZFN mRNA 和外源模板进行基因编辑过的人类细胞以更高的比率移植到小鼠中[43]。随后，被植入的细胞很可能进行爆发式增长并发挥对抗同种异体肿瘤细胞的功能。这一研究很快的就进行了临床转化：GMP 水平的 mRNA 和 IDLV 流程已经存在，IL2RG 基因外源性模板的 5-8 外显子和内源性的 1-4 外显子发生错位；这种技术随后也将广泛的用于 SCID。而且，在 AAVS1 HDR 的实验中所构建的方法也能够用于其它疾病，也为目前的造血细胞移植疗法的相关疾病提供了一种可被广泛采用的新策略。

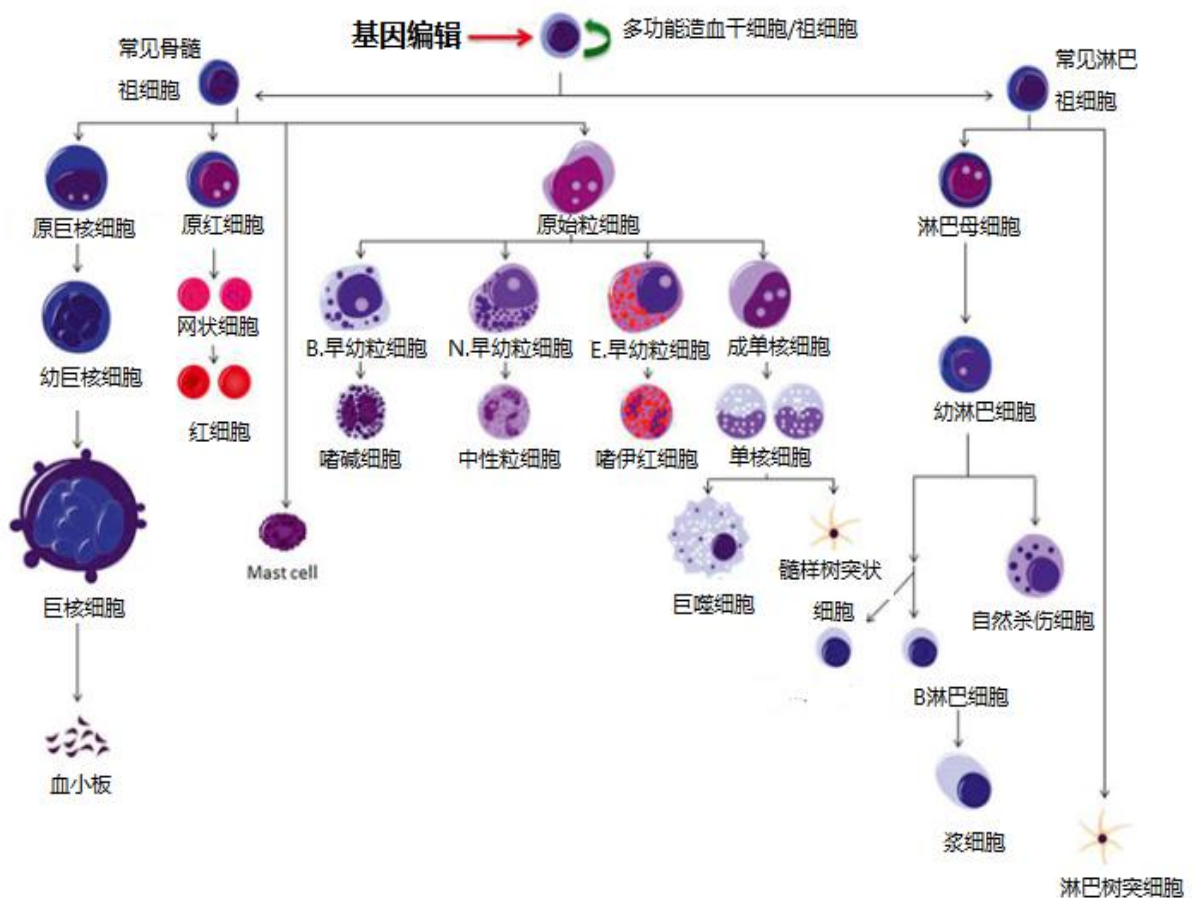


图2. 人类造血细胞。造血干细胞/祖细胞拥有自我更新的能力（绿色箭头）和形成常见淋巴细胞（CLP）和常见骨髓祖细胞（CMP）的能力。CLPs能够引起淋巴间细胞的产生：T-，B-，和NK细胞（右）。CMP能够产生有限的祖细胞（原始粒细胞，原始红细胞和原巨核细胞），其自身能够分化成为成熟的造血骨髓和红系效应细胞。红箭头表明作为基因编辑靶标的细胞。图像利用Motifolio Inc.Sykesville，MD软件绘制。

在DNA修复中的NHEJ在HSC基因编辑中具有治疗潜力，目前已经建成了能达到最大基因破坏效率的多基因导入平台。在这些细胞的操作中由病毒的包被的、非病毒转染的和基于蛋白转染的HSC基因的导入是关键性的问题，理想的遗传物质导入策略是不只进行简单的基因/蛋白表达，同时也进行核酸酶表达功能域本身低风险的基因组微干扰。Lombardo和Genovese研究组利用IDLV分别导入ZFNs和模板物质[42,43]。被包裹的DNA整合了D64V突变，能够产生线性或环状的物质，作为游离基因跨过细胞膜和核膜，并在细胞分裂过程中被遗失[50-54]。IDLV的一个潜在的缺陷是能够整合到基因组中由核酸酶切割产生的损伤位点或是基因中一些脆弱的内源性损伤位点[55,56]。具有ZFNs编码的腺病毒功能域也被导入HSCs，并能引起CCR5的接近25%的NHEJ破坏，能够抵消掉潜在的毒性[57]。这一相对高的比率是由于有丝分裂激活蛋白激酶/细胞外信号调控激酶通路的模式决定的[57]。其它研究发现药物诱导核酸酶基因转移的提高和在科研中是一种非常有兴趣的最大化核酸酶活性的方法[58]。由于在293腺病毒宿主细胞中ZFN的表达有可能产生毒性并影响滴度。Saydaminova等人设计了一种方法通过microRNA（MiRNA）序列等手段在293细胞中抑制ZFN基因的表达，但在造血干细胞中则不存在抑制表达的作用[59]。通过这种方法，他们在移植后的HSCs中检测到了约12%的基因编辑率。Li和Saydaminova研究组都发现了基因编辑后的细胞在鼠模型中的移植能力下降；然而，他们成功的在临床上达到的显著的CCR5的破坏率显示出腺病毒基因导入的方式在HSCs中具有很高的实用价值[57,59-61]。

核酸酶导入的非病毒方式包括蛋白或信使RNA（mRNA）基础平台，由于其瞬时性表达不会引起潜在性的基因整合是一种非常理想的方式。基于蛋白质的导入策略利用ZFN-转铁蛋白耦合物通过转铁蛋白受体介导的胞吞作用导入ZFNs[62]。这些复合物导入人类HSCs并在报告实验中介导整合细胞系中NHEJ和HR的产生[62]。利用Amaya Nucleofector装备的电穿孔导入，mRNA在HSCs中能够引起接近15%的CCR5基因的破坏[63]。而且，这一研究组还在前临床人源化小鼠模型中利用HSCs进行基因编辑对抗HIV-1的感染[63,64]。利用ZFNs破坏HIV感染途径的能力被用于最近的临床患者中，即将HIV感染途径中的CCR5基因进行纯和的CCR5 Δ 32突变并将这种造血细胞移植回患者体内阻止HIV的感染[65]。这一治疗手段最初被用于一例HIV/AIDS阳性的患有急性白血病的患者（也就是所谓的‘柏林患者’），试图一次性的治愈恶性肿瘤和HIV感染[65,66]。由于缺少CCR5 Δ 32/CCR5 Δ 32模板，ZFN编辑过的HSCs被认为是一种可被广泛施行的理想疗法。进一步的核酸酶基础基因破坏疗法的研究发现，在HIV中CXCR4基因也是HIV进行细胞感染的一种共受体，有很多HIV阳性的患者具有不同程度的CXCR4趋向性的HIV病毒[67]。由于这个原因，最近有实验室提出了ZFN的联合疗法。Didigu等人利用CXCR4和CCR5靶向的腺病毒骨架ZFNs，发现在人类T细胞中具有移除所有的两个HIV共受体的作用[68]。这一方法的潜在性临床限制是CXCR4是HSCs中的一个重要的调控分子[68]，该基因的破坏有可能扰乱HSC在体内正常的平衡。

HSCs的血细胞可塑性和广阔的临床应用使其成为了一种理想的可被人工设计核酸酶进行基因编辑的细胞类型；然而，它们在体外受限的造血组织形成能力限制了它们在综合性疾病模型和淋巴造血系统之外的再生药物中的应用。ESCs和iPSCs是填补这一空白和进行多个干细胞基因组核酸酶编辑的有力工具（图3和4）。

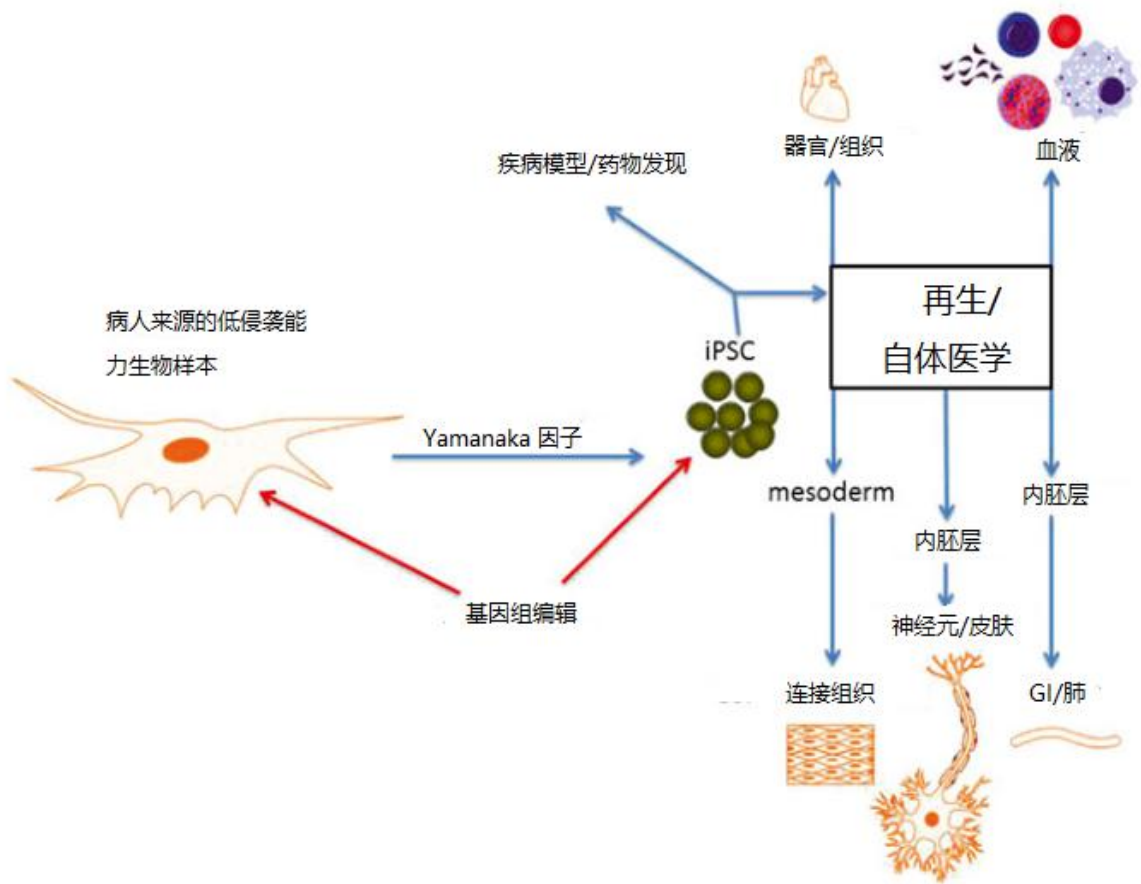


图3. 诱导型多功能干细胞编辑。基因编辑核酸酶能够在进行重编程前导入体细胞或直接导入iRSCs中。体细胞能够通过多种来源获得，来自患者的最小侵袭性的生物样本组织。将Yamanaka因子 (Oct3/4 , Sox2 , Klf4 , c-Myc) 导入目标细胞能诱导产生全能性。这些细胞能够作为疾病模型、药物发现、自体疗法/再生医学的平台发挥其广泛性的分化能力。图像的绘制利用Motifolio Inc.Sykesville , MD 软件。

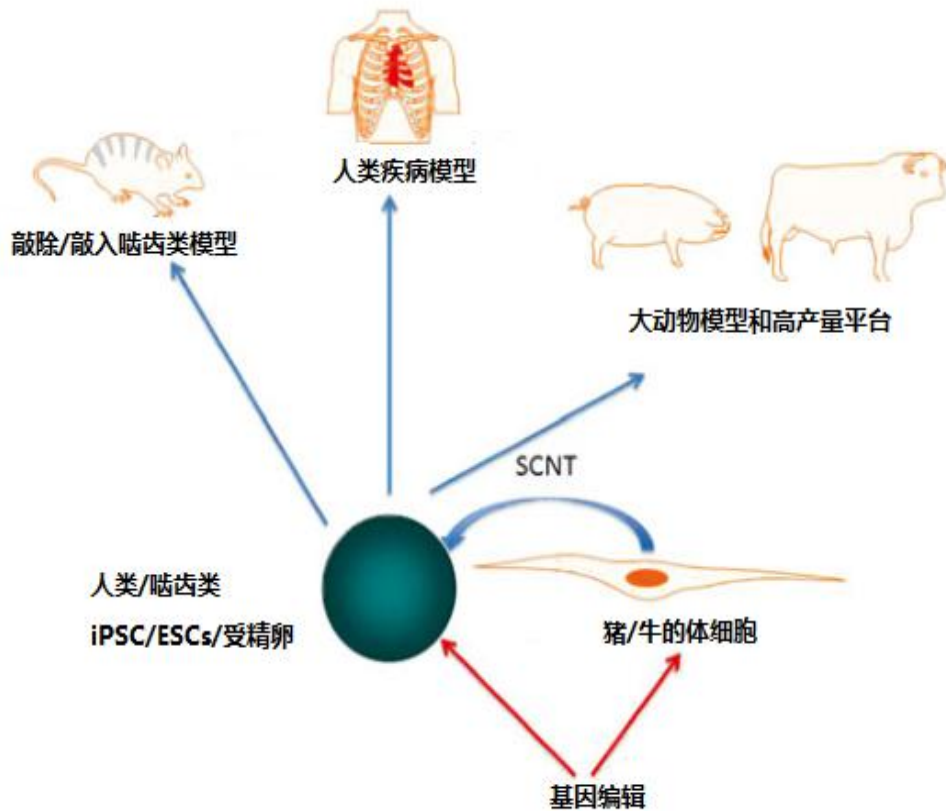


图4. 多物种胚胎干细胞编辑。在家畜模型中，基因编辑能够被用于体细胞，并随后通过核移植（SCNT）导入无核卵细胞中。小动物模型的基因编辑起始于早期的ESCs或者是早期胚胎的建立。人类干细胞的基因编辑发生于体外主要的细胞系。基因编辑能够提供完整的科研用动物模型，和人类ESCs（或iPSCs）作为一种体外基因/药物发现模型或谱系遗传研究。图片的绘制采用Motifolio Inc.Sykesville,MD软件。

Lombardo 等人最早利用 ZFNs 进行了人类 ESC 的基因编辑，即利用 ZFNs 对 HUES-3 和 HUES-1 细胞系中的 CCR5 基因进行靶向[42]。他们将 ZFNs 和一个 3 号外显子具有 GFP 的 CCR5 基因模板序列导入细胞中得到了约 5%的靶向整合率。重要的是，这类细胞具有全能性和自我更新的能力[42]。这一研究开启了利用 HR 在 ESC 的目标位点特异性的插入整合基因的先例。Hockemeyer 等人采用这一技术在所谓的‘安全位点’ AAVS1 置换插入了一段诱导表达功能域。第一次在 19 号染色体中 PPP1R12C 基因的第一号染色体中利用特异性的 ZFNs，他们导入了一段独立的表达结构域，包括启动子、嘌呤霉素基因和具有 2A-嘌呤霉素受体基因错配的多聚腺苷酸化信号（或基因诱捕靶向载体），依赖于和 PPP1R12C 基因第一外显子的适当靶向和错配[69]。与之类似，在 AAV 位点的基因靶向能够引起基因启动子的置换，驱动基因表达达到理想化的水平或由原有的 PPP1R12C 启动子调控引起持续性的活化[70]。利用这一策略并不会明显改变 iPSCs 或 ESCs 的多功能自然属性[69-71]。Lombardo 及其同事发现在 AAVS1 位点能够稳定表达 GFP 基因的 iPSCs 也能够表达内源性的 TRA-1、Nanog、Sox2 和 Oct4 多能性标记[71]。Wang 等人利用 ZFNs 介导 PPP1R12C 的 1 号外显子中 2A-嘌呤霉素抗性基因通过 HR 途径插入，随后构建一种泛素化启动子驱动的 tricistronic，包括有红色荧光蛋白、萤火虫荧光素酶和单纯性疱疹病毒胸腺激酶基因[72]。这一策略用到了荧光成像、生

物发光、和 ESCs、iPSCs 的正电子断层造影术和干细胞所衍生的内皮细胞和心肌细胞[72]。由于在 AAVS1 位点 HR 的直接性整合并未明显干扰 PPP1R12C1 基因或邻近基因的转录模式，然而目前关于这一点是否是真正的‘安全位点’还存在争议[71,73]。在干细胞中利用 ZFNs 很有可能会构建出特异性更强的靶向方法。在他们的 AAVS1 靶向方法构建的相同研究中，Hockemeyer 等人也设计了靶向 OCT4 (POU5F1) 位点的 ZFNs，能够介导全新的 OCT4-EGFP 报告基因的插入，从而能够通过实时成像对 hESCs 的多能性状态进行监控[69]。除此之外，他们也发现在 ESCs 和 iPSCs 中常被沉默的 PITX3 基因也能够被靶向，这也说明在 ESCs/iPSCs 中能够被靶向的序列与他们的转录状态无关[69]。

在 ESCs 和 iPSCs 中基因编辑的能力对于体外疾病建模非常重要，这也变成了加速转化研究的一类新功能。为了构建基因编辑 ZFNs 工具对人类疾病进行限制，Zou 等人在 ESCs 和 iPSCs 中对磷脂酰肌醇 N-乙酰葡萄糖苷氨基转移酶的 A (PIG-A) 亚类基因进行靶向，能够产生和严重血液异常阵发性夜间血红蛋白尿 (PNH) 患者细胞类似表型的细胞[74]。他们的策略依赖于 H1 ESCs 中 NHEJ 介导的基因删除和 HR 诱导的突变发生。利用 ZFNs 和外源靶向模板破坏 6 号外显子，他们能够分离出具有多能性的 iPSCs 和 ESCs 单细胞克隆[74]。这些研究将 ZFN 介导的基因编辑作为干细胞建模的有力工具。这一方法的前景天生取决于研究中所采用的统一性的工具对疾病表型的模拟，不会介导产生混杂的差异也不会作为实验材料的任意两组（或两株）细胞之间的多个不同的位点产生显著性的差异。

在干细胞的基因编辑中显著性差异的潜在来源是细胞产生/增殖的过程。被批准用于广泛研究的 ESC 细胞系在其分化能力和敏感性方面、表观遗传和基因稳定性方面具有差异[75]。这些因素有助于遗传多样性并且在体外无需精确性的比较[75]。对于 iPSCs，由于还牵扯到了细胞的产生，因而变得更为复杂。通常情况下，重编程细胞产生 iPSCs 可利用化学的、蛋白的、病毒的和非病毒的平台完成[8,76-78]。当采用整合性载体时，上色效果、异味转基因表达和拷贝数的差异在重编程的时候在疾病建模时都是潜在的混杂因素[79,80]。为了解决这一问题，Soldner 等人利用 ZFNs 产生同基因调控的帕金森疾病 (PD) 细胞系[79]。这一工作主要是在非患病的 ESCs 中对 A53T 或 E46K PD 突变进行编辑或在 PD 患者来源的 iPSCs 中修复突变[79]。通过这一方式，他们提出了解决多个遗传差异的良好方案并设计独立用于 ESC 和 iPSC 克隆的编辑器。最近有研究发现，利用所构建的 ZFNs 能够促进直接性的、统一性的同基因细胞系的产生[81]。

进一步的工作也研究了基因编辑逆转疾病突变回到野生型的能力。已经有研究证实可利用 ZFNs 修复镰状细胞贫血疾病突变[82-84]，Zou 等人在该研究中修复了具有红系细胞分化能力的 iPSC[83]。Yusa 等人进一步扩大了 ZFN 修复应用的范围，包括在 α 1-抗胰蛋白酶中造成蛋白酶缺陷的 Glu342Lys 突变的修复、修复后的肝样细胞在体外也显示出了正常的酶活性[85]。Rahman 等人在纤维母细胞和 iPSCs 中利用 ZFNs 修复了能够引起严重的免疫缺陷的 DNA 依赖性的蛋白激酶缺陷并成功的产生了具有正常表型的 T 淋巴细胞[86]。

Jiang 等人利用 ZFNs 进行了全基因组基因编辑，开发了一种全新的对唐氏综合症患者的成纤维母细胞进行重编程产生 iPSCs，随后利用 ZFNs 在 21 号染色体的 DYRK1A 位点插入了整合了诱导性的 X-失活 (XIST) 转基因[87]。检测了六个克隆。利用强力四环素对系统进行诱导产生的 21 号染色体上的 XIST 转基因的表达能够引起稳定的异染色体的改变和基因沉默形成一种“21 号染色体巴氏小体”[87]。全基因组转录检测发现目标染色体上的 95% 的表达基因被沉默[87]。这些研究将 ZFNs 应用于 21 三体的基因治疗以及体外唐氏综合症模型的构建，在模型中能够检测到 21 三体诱导产生的基因组/转录异常调节[87]。

除了多篇文献中有介绍过 ZFN 能够介导多组织的中高效的基因编辑，但是 ZFN 技术仍存在很多限制。构建出的 ZFNs 的活性并不足以进行广泛性的应用。虽然这一平台的模块性得到了提升，所连接的序列特异性的指状结构仍需要更强的亲和力，而其产生仍是一个专业

化的漫长的过程[88,89]。这些缺陷随着高度自由模块化化的 TALEN 的到来，极大的改变了干细胞和基因组编辑的整体情况。

TALEN

来源于黄杆菌属植物致病菌的 TAL 效应器蛋白。一段具有 14-24 组串联的 34 氨基酸重复中部序列具有 DNA 识别的能力，并可受简单的编码调控进行快速的组装[90-92]。与 ZFNs 相似，TALENs 以异二聚体的形式发挥作用（图 5a）。与 ZFNs 不同的是，每个 TALEN 单体都能被编辑识别 12 到 32bp 长度的 DNA。因此 TALEN 序列的特异性识别被认为十分精准，而在给定的基因组中，这一序列长度的变化可能很低。TALENs 已经被用于除 HSCs 之外的多个和 ZFNs 相似的模式模式中。

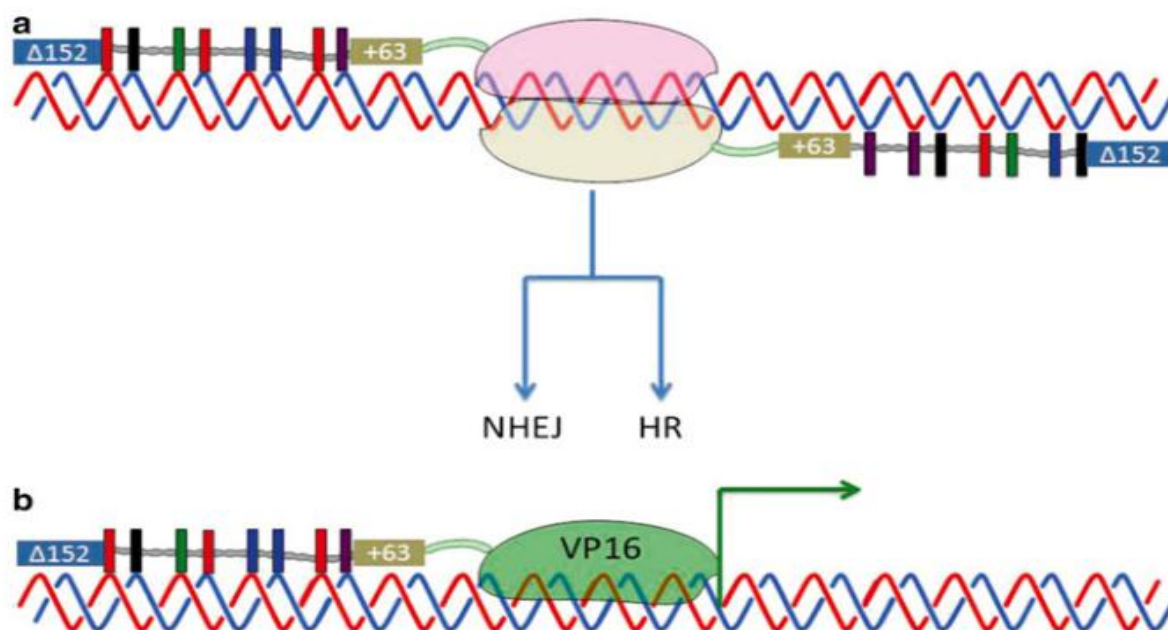


图5. TALEN核酸酶和转录激活样结构。 (A) TALE 核酸酶。一个TAL由剪短的152个氨基酸的N-末端和C-末端63氨基酸的野生型TAL序列以及能够简单识别4个DNA碱基编码的具有两个重复差异性的双残基（彩色框）。这一编码能够进行人工编辑产生用户定制化的TALEN，并和FokI酶的核酸酶区域进行融合（粉色和棕色椭圆）。TALE核酸酶是异源二聚化核酸酶，具有目标位点共定位的左片段和右片段，能够介导双链损伤并触发NHEJ或HR的产生。(B) TALE 激活子，FokI核酸酶功能域被单纯性疱疹病毒编码（或其他组分）转录激活子来源的多拷贝的VP16结构域所置换。TALEN单体与激活子融合用于靶向目标区域上调转录过程。图像利用Motifolio Inc. Sykesville,MD软件绘制。

啮齿动物

在 2011 年，Tesson 等人利用 TALENs 靶向于大鼠 IgM 位点以破坏其功能[93]。他们发现了高比率的 DNA 编码 TALENs 的基因编辑，但是并非纯合子修饰。与之相反，通过 mRNA 导入的方式，虽然切割比率很低，但是在更高浓度的 mRNA 条件下能够引起 50%的等位基因纯和编辑[93]。TALENs 也被用于大鼠 TLR4 基因的敲除，这一基因被认为和乙醇诱导的行为模式相关。Ferguson 等人设计了 TLR4 基因 1 号外显子靶向的 TALENs 并将其注射到单细胞胚胎中

[21]。在受精卵所产生的 13 个胚胎中有一只大鼠成功的产生了突变[21]。在大鼠中进一步的基因编辑的进展是 Mashimo 等人将靶向 Tyr 基因和靶向具有 5'-3'核酸外切酶切割活性的 (Exo1) 基因的 TALENs mRNA 导入大鼠并发现基因敲除的数量大鼠显著提高[94]。重要的是, TAL 和 Exo1 mRNA 的同时导入并未引起毒性, 这也说明在 DNA 修复途径中基因编辑的异位操作能够进一步的提高效率[95]。

在 2013 年, Sung 等人第一次将 TALENs 用于小鼠胚胎[96]。他们靶向了 Pibf1 和 Sepw1 基因, 并发现了高剂量的 TALEN mRNA 能够引起更高比例的等位基因纯和修饰。而低剂量 mRNA 能够使很多的动物状态变差。后一种效应与 Pibf1 TALEN 的脱靶毒性有关, 和其它研究中所认为的高剂量核酸酶能够引起高比例的编辑不同, 在他们的研究中发现低剂量的 TALENs 反而能够引起更高比率的编辑[96]。

在利用 TALENs 的多能性和适用性产生不同的基因敲除小鼠模型实验中, Davies 等人在三个不同种小鼠卵母细胞中利用 TALENs 失活 Zic2 基因产生了前脑无裂畸形的小鼠, 是人类常见的一种先天异常疾病[97,98]。通过注射 TALEN mRNA, 他们在 C3H/HeH 背景动物中观测到了 25%的基因编辑率, 在 C57BL/6J 动物中观测到了 10%的基因编辑率[92]。在相似的研究中, Qiu 等人在 C57BL/6 和 FVB/N 受精卵中获得了相似的基因编辑效率[99]。基于以上研究, 在多种品系的小鼠中更为复杂更为纯粹性的突变诱导试验还在进行。因此, 有可能通过构建具有不同遗传背景的小鼠来模拟人类种群中相似的疾病, 通过本质上的基因改变来影响生理应答[100]。

Wang 等人在一项研究中利用 TALEN 靶向于 Y 染色体上的 Sry 和 Uty 基因[101]进一步的推进了小鼠基因编辑领域的进展, 尤其是 TALENs 基因编辑领域。在之前的研究中, 出现过一种采用传统方法构建的具有缺陷基因 Y 染色体的不完全敲除小鼠[101]。TALENs 能够同时在 Sry 和 Uty 靶点达到很高的效率[101]。这也有利于对 Y 染色体上的功能性基因更为详尽的研究。Takada 等人还报道了利用 miR-10a 和 miR-10b 进行 MiRNA 基因靶向的方法。这一多功能性、简易性和高效的应用促进了种属特异性基因发现的研究[102]。

靶向性的敲入也能够产生人源化的小鼠疾病突变模型。Wefers 等人通过导入 TALEN mRNA 和一段具有 RAB38 基因错配突变的单链寡核苷酸模板(ODN)模拟 Hermansky - Pudlak 综合征的产生[103]。在第二组研究中, 研究组利用具有质粒编码多聚(A)尾端的全新的 TALEN 骨架又花了这一过程, 从而在体外移除了聚腺苷酸化的步骤[104]。在他们的首次研究中他们发现了 ODN 介导的 <2%的 HR 诱导率, 但是优化后的 TALEN mRNA 介导的肌萎缩侧索硬化症的 Fus 基因位点中 ODN 导入突变率为 6.8%[104]。他们也发现了 HR 和 NHEJ 相伴出现的情况。相似的结果也出现在 Low 等人的研究中, 研究利用 TALENs 和 ODN 模板修复 Crb1rd8 突变, 这一突变出现在很多的视网膜疾病研究的小鼠模型种系中[105]。ODN 模板的应用非常常见, 因为更易被合成, 并且在很小体积中能够达到很高的浓度。这一类的报道中显示其能够整合进入基因组产生全新的疾病多态性突变[106,107]。然而, 由 ODN 引起的 NHEJ 和 HR 事件的同时出现使得在一些情况下这一方法并不是最优的而且缺乏效率, 因为缺少药物耐受标记对多克隆进行筛选。

经典的基因靶向依赖于双链 DNA 模板的应用, 通常将靶位点整合到质粒中, 通常包含有同源臂并通常携带有药物耐受基因可用于筛选。这一方法的缺陷在于质粒的尺寸会限制导入的浓度。优势在于质粒冗长的同源臂能够降低不完整或短交联产生基因截短的风险, 而且筛选标记(例如药物耐受基因)的存在能够有助于目标基因的适当的恢复。有一项重要的研究, 利用类似的模板质粒, 构建了可进行基因组编辑的最适宜的条件: 在一个临床相关模型中对来自患者的多功能干细胞进行基因修复。在 2007 年, Hanna 等人在镰状细胞贫血的人源化小鼠模型中分离出一种成纤维细胞系, 对细胞进行重编程产生 iPSC, 利用质粒模板进行修复后, 诱导这些细胞分化形成造血细胞祖细胞, 并将这些细胞移植回镰状细胞小鼠中重新

产生正常的红细胞[108]。虽然在该方法中并未采用人工编辑的核酸酶，但是依旧取得了足够高的效率，而特定基因靶向的核酸酶的设计和引入使得这一疗法进入临床成为了可能。

猪

TALENs 已经被用于猪基因组的改造进行基因敲除从而产生大动物疾病模型。Carlson 等人利用 TALENs 在成纤维母细胞低密度脂蛋白受体位点产生小片段的插入或删除构建了活体高胆固醇血症模型，并随后进行了 SCNT 操作[107]。这一课题组也利用相似的方法产生了不孕症和黑色素瘤的猪模型。TALEN 编辑产生的成纤维母细胞通过 SCNT 产生独立的具有 DAZL 和 APC 等位基因敲除的个体动物模型[109]。

将靶向于猪基因组 GGTA1 位点的 TALEN mRNA 注入猪胚胎能够引起 >70% 的编辑比率[110]。后续研究利用具有新霉素抗性基因的模板在纤维母细胞中经过药物筛选后具有 >85% 的整合率，其中有 >25% 的细胞为纯合子基因编辑[110]。一项利用 ZFNs 的一致性研究的纯和等位基因破坏率约为 1%[31]。随后的研究并为使用药物筛选过程，所以进行直接性的比较很困难；然而，在基因编辑中如何获得高比率的编辑细胞是一个需要被重要考虑的问题。

牛

在 2012 年，在大动物胚胎模型中 TALENs 被报道具有很高的效率。这些研究利用了早期产生的具有很多原始细菌序列的 TALEN 骨架对 ACAN11 或 ACAN12 基因进行靶向编辑[91,111]。显微注射所能取得的整合频率最高，但是也会产生发育迟滞的现象，这一现象促使他们利用截短的 N-和 C-末端重新构建 TALENs，优化 TALENs 的活性降低其在哺乳动物细胞中的毒性[112,113]。当采用设计用于 GT-GDF83.1 基因靶向的 TALENs 时，他们在体外观察到了更高的 indels 比率却没有显著的发育迟滞的产生。

非人灵长类（NHP）

在 2014 年，Liu 等人利用 TALENs 在恒河猴和猕猴的细胞和组织中进行基因编辑[114]。通过注射靶向于 MECP2 基因的 TALEN 和促进 NHEJ 的第三编码 RAD51 的质粒，他们观察到了符合 Rett 症状相关雄性特征死胎的流产。通过这一过程，他们得到了一个 MECP2 indels 的单独雌性个体，试验显示了 NHPs 中基因编辑的能力以及相对于 mRNA 质粒 DNA 具有更高的基因编辑的效率[114]。

人类

在 2011 年，Hockemeyer 等人设计了 TALENs，并在相同的基因位点检测了其与之前研究中所采用的 ZFNs 的活性[69,115]。在 ESCs 和 iPSCs 的所有的五个靶向位点都观察到了模板序列的整合：OCT4（5'和 3'主要靶标），AAVS1 和 PITX3（启动子和终止子编码靶标），每个位点都保留了其多功能的特征[65]。这些主要的发现，随着 TALEN 模块技术的逐渐成熟，缩短了其产生的时间，提高了靶向编辑的能力并使其操作更具有效率为干细胞和基因编辑领域带来了无与伦比的变革与机遇。很多研究者开始通过干细胞编辑进行疾病的模拟或者治疗。Ding 等人发现了干细胞和 TALENs 在 15 基因靶向和突变产生中具有协同能力[116]。他们对干细胞中三种疾病基因进行表型研究：SORT1、AKT2 和 PLIN1[116]。SORT1 具有调节血液中的胰岛素、胆固醇水平和神经元存活率的功能。他们利用 TALENs 在两个不同的 ESC 细胞系中破坏了 SORT1 的 2 号或 3 号外显子。分化成为肝样细胞说明 SORT1 能够降低血液中 apoB 的水平，并因此降低胆固醇防止动脉粥样硬化[116]。而且，这些研究发现 SORT1 对于胰岛素应答的葡萄糖摄取十分必要，提示 SORT1 在人类胰岛素敏感性中具有作用[116]。最后，在 SORT1 缺失细胞中神经元的分化验证了之前数据中有关 SORT1 在脑内神经营养因子（proBDNF）介

导的神经元凋亡过程中具有重要作用的说法[116]。这些数据与之前的人类运动神经元中 proBDNF 介导的程序性细胞死亡过程中, SORT1 具有主要作用的结果相一致[116]。这些研究者也在体外利用 TALEN 构建了 AKT2 和 PLIN1 疾病细胞编辑模型。在 AKT2 基因中的 E17K 错义突变与胰岛素抵抗和体重增加有关;然而,这一差异的蛋白质研究缺乏,因为无法获得生理学相关的细胞类型[116,117]。为了解决这一问题,利用 TALENs 对 HUES 9 细胞中的 AKT 2 号外显子进行敲除,达到了 17/192 克隆的 indels 效率,随后进行第二轮的基因靶向获得了 2/96 频率的纯和突变[116]。为了更精确的模拟 E17K 突变,将 TALENs 和一个 67 核苷酸的 ODN 导入 HUES 9 细胞产生一种 AKT2^{E17K} 纯和细胞系用于研究[116]。通过将每一类的细胞群落诱导分化形成肝样细胞和脂肪细胞的过程显示 E17K 突变能够引起患者产生糖尿病类似的状况[116,117]。通过相似的策略,这些研究者利用 TALENs 在 HUES 9 细胞中的 PLIN1 基因的 8 号外显子产生了移码突变。他们得到了 70/293 克隆比例的 indels,其中一个与自然产生的延长型突变极其相似 (Val398fs 突变),并第二次引起的 PLIN1 的缩短[116]。通过诱导这些细胞分化产生脂肪细胞,他们还发现延长突变蛋白能够改变脂肪细胞中的贮存和脂滴形成的情况,这一模式与脂肪代谢障碍的患者类似[116]。

这些研究显示了基因编辑在干细胞体外疾病建模和研究中的强大功能。通过进行单细胞系的分析,他们能够减小患者之间或细胞克隆之间的基因差异,而这些差异在之前的研究中可能影响同基因疾病模型中疾病基因的表达[79]。而且,即使在之前的人类研究中他们取得的结果具有一致性[118],但是这与基因敲除鼠模型仍具有差异[118],这一点尤其是在种属特异性的疾病模型中具有突出的价值[116]。而且,TALENs 所具有的简易性扩大了其可被应用的范围,例如在 15 基因靶向策略中多细胞系的分化和该研究中基于三个基因的疾病表达模型的构建[116]。

基因编辑技术是疾病基因模型和干细胞研究中遗传背景保持一致的能力随着 DICE (双整合功能域置换) 系统的建立进一步提高[119]。这一系统依赖于 ESCs 或 iPSCs 的 H11 位点中具有 phiC31 和 Bxb1 attP 位点的整合片段的插入[119]。非重叠性的 attP 位点可被 phiC31 和 Bxb1 整合酶识别,并能够导致外来基因高度精确的置换,在单独的拷贝中具有相同的方向。不同的外源性整合片段在基因组中的随机性位点整合能够引起拷贝数的减少和表达的差异[119]。在整合片段中插入 TALENs 能够将 ESCs、两个 iPSC 细胞系和来自于帕金森患者的 1754 iPSCs 细胞系中的基因编辑的效率提高 8 倍[119]。一个相似的方法依赖于 AAVS1 位点中 loxp-Cre 转基因置换[120]。如上所述,疾病相关或治疗相关的基因能够被导入到多个干细胞平台的多个位点中 (例如 AAVS1、H11),能够减少经典转基因方式引起的不可预测性的结果。

TALEN 的基因研究也拓展到了 miRNA 领域。在人类成纤维母细胞中设计删除 miR-302/367 簇的 TALENs 也证明了在 iPSCs 的重编程中需要这些 miRNA 的存在[121]。在这一研究中也通过将 TALEN 阵列与一种 Kruppel 相关区域 (KRAB) 转录抑制子进行融合将 TALEN 的功能性研究拓展到了表观修饰领域[121]。利用 TAL-KRAS 抑制融合体,这些研究者在 miR 簇敲低的情况下观察到了重编程活性 4 倍的下降。TALEN 转录激活因子的 TAL DNA 识别区域被报道能和 VP16 的激活区域进行融合 (图 5b)。这一复合体能够被用于靶向 Oct4 位点并适度上调 ESCs 细胞中的转录[122]。将 TALE 重复序列与 TET1 羟化酶催化结构域融合也能够对甲基化 CpG 残基进行修饰,并伴有基因表达的上调[123]。与此类似,TALENs 作为基因编辑核酸酶的应用也能够引起基因组永久性的改变,或刺激基因组瞬时性的激活或短暂抑制基因的功能,这些功能的开发能够极大扩展这一技术的应用领域。

TALEN 基因纠正功能也在 iPSCs 中有所报道。利用串联连接有 α -1 抗胰蛋白酶 (AAT) 同源臂、嘌呤霉素抗性基因和截短的胸苷激酶基因融合产生的抗性基因所组成的质粒模板,Choi 等人完成了 AAT 基因的修复[124]。通过嘌呤霉素筛选,他们得到了 66/66 比率的靶向

iPSCs 编辑细胞，其中 25-33%为纯合子修复[124]。最后，利用一种 piggyBac 转位子系统，他们随后移除了抗性基因，只留下了模板序列，在 iPSC 分化得到的肝样细胞中观察到了功能的恢复[124]。Sun 等人利用相似的外源重组策略修复引起镰状细胞疾病的血红蛋白 β (HBB) 基因中 E6V 位点的修复[125]。有超过 60%的嘌呤霉素抗性克隆具有靶向基因的修复，而且 iPSCs 仍会保持正常的核型和多功能干性[125]。Osborn 等人利用 TALENs 在严重的隐性营养不良大疱性表皮松弛 (RDEB) 患者来源的主要纤维细胞进行了基因修复[56]。细胞随后进行重编程产生 iPSCs，并注射到免疫缺陷小鼠中，所产生的皮肤瀑样畸形瘤具有 RDEB 患者中所缺乏的 VII 型胶原蛋白沉积[56]。综上所述，这些研究为患者在重编程前和重编程后的间接体内细胞治疗提供了基础，并提供了一种基因个性化疗法和再生药物。

TALENs 的靶向能力优于 ZFNs，而且更容易产生[111,126]。TALENs 也能和多种结构域融合 (例如转录激活因子和抑制因子) 从而产生点特异性的新活性。虽然 TALENs 的研究很热门，但是这一技术仍相对较新且无法在一些组织中进行应用。TALENs 的导入效率受限于它们的尺寸，这也使得只有腺病毒适合该类核酸酶的导入而其余平台 (逆转录病毒、慢病毒和腺相关病毒载体) 不适于这项技术[127]。第二，目前 TALENs 的多路运输需要分别产生独立的多点特异性蛋白并保证能够高效的导入到目标位点。因为这些原因，目前更多的关注点集中于 CRISPR/Cas9 系统，这一系统能够提供给用户更大的便利性并能够引起干细胞更为复杂的基因编辑。

CRISPR/Cas9

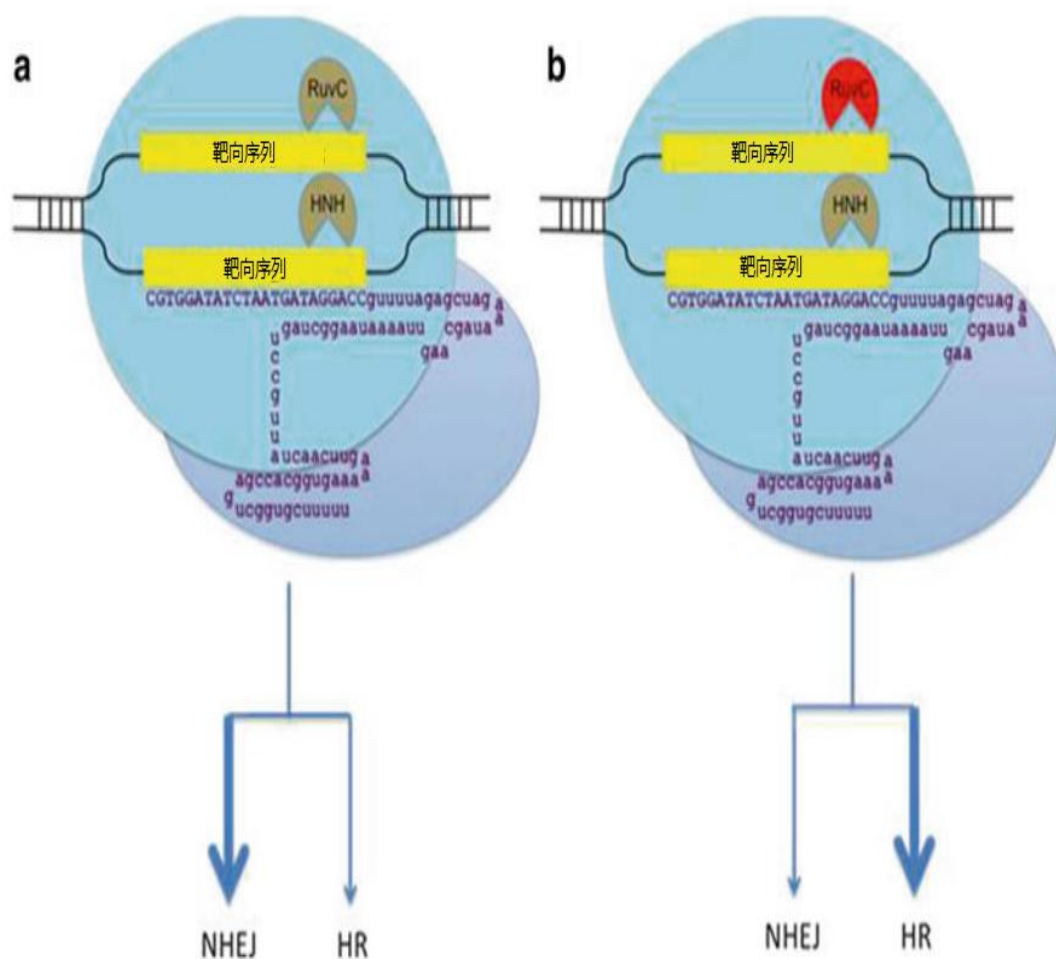
CRISPR/Cas9 存在与古生菌和细菌中，是用来抵御外来病毒噬菌体入侵的适应性免疫系统的一部分[128]。这一系统由 Cas9 核酸酶和一段具有前空白连接基序 (PAM) 相关特异性的向导 RNA (gRNA) 序列组成 (图 6) [129,130]。这一系统适用于哺乳动物的多平台应用并且具有高度的友好性。

啮齿动物

Mashiko 等人利用简单的导入方法能够将具有 Cas9 核酸酶和 gRNA 序列的单独质粒注射到小鼠受精卵中靶向于鼠源的 *Cetn1* 基因或 *Prm1* 基因[131]。更高剂量的质粒和更高效率的敲除相关，只有 2/46 比例的被编辑动物中发现了 Cas9 表达功能域的整合[131]。这一相对低的整合率简化了平台的导入/注射过程，因为其排除了 RNA 产生的需要。然而，RNA 的导入与高比率的基因编辑有关。Li 等人无法在 FVB/N 小鼠中利用低浓度的 DNA 完成 *Th* 基因的修饰，而将 DNA 浓度提高 2.5 倍时基因编辑的效率只提高了 0.3%[132]。与此相反，RNA Cas9 和 gRNA 在 C57B1/6 小鼠中的应用提高了 *Th* 位点的基因靶向效率提高到了 6.7%[132]。Fujii 等人发现具有两种独立靶向的 80 核苷酸长度 (与经典的约 40 核苷酸长度形成对比) gRNA 的 CRISPR/Cas 能够引起长片段的删除 (约 10kb)，并可遗传给子代小鼠[133]。这些数据显示多种 gRNA 能够同时在小鼠受精卵中发挥作用。将这一现象扩大到不同基因，Zhou 等人试图靶向 *IL2rg*、*Prf1*、*Prkdc* 和 *Rag1* 基因产生免疫缺陷鼠模型 (*IL2rg* 和 *Prkdc* 缺陷的 NSG 模型；*Rag1* 和 *IL2rg* 敲除的 BRG 模型和 *B2m*、*IL2rg* 和 *Prkdc* 纯和阴性的 NSG *B2m* 模型) [134]。而且，他们在三种动物中都发现了所有的 5 个靶向基因的破坏[134]。Wang 等人也对 5 个基因进行了靶向 (*Tet1*、*3*、*Uty* 和 *Sry*)，而在八个同源位点的五种基因中有 10%的筛选后 ES 克隆产生了突变[135]。在该研究组和一些其它的研究组中也发现在目标位点能够进行小片段的双链或单链的模板序列的整合[135,136]。这些整合包括点突变、*loxP* 位点的插入或 V5 抗原表位标签的引入[135,136]。通过这种方式能够一步产生具有点突变的小鼠，并可在同源位点高效引入条件性突变、荧光标记或 V5 抗原表位。这些策略为基因发现提供了路径，也为利用具有可能不存在抗体的候选序列标签 (例如 V5 抗原) 进行基础蛋白研究提供方法。

CRISPR/Cas9 系统中的免疫功能在全基因组发现研究中的多种应用正在被充分的挖掘 [137]。Koike-Yusa 及其同事导入了 87,897 个向导 RNA 靶向于持续性表达 Cas9 的小鼠的 ESCs 中 19,150 个小鼠蛋白编码基因 [137]。他们随后利用 6-巯鸟嘌呤或梭状芽孢杆菌 α -毒素对其进行筛选并鉴定和这些表型相关的 31 个基因 [137]。

更高级的哺乳动物胚胎细胞也被用于 CRISPR 的基因编辑。注射了 Cas9 和靶向于 Mc3r 或 Mc4r 基因的 gRNA RNA 的大鼠受精卵表现出了一定的基因破坏频率，分别具有 0.8% 和 10.6% 的基因代际遗传 [132]。CRISPR/Cas9 也可用于大鼠中的条件性基因靶向，能够引起瞬时的组织特异性的基因激活的调控。Ma 等人利用独立的环状模板质粒，靶向了 Dnmt1、Dnmt3a 和 Dnmt3b 三个基因，并产生了具有特定等位基因的后代 [138]。CRISPR/Cas9 的一个强大的应用是对单细胞大鼠 ESCs 进行高效靶向产生纯和的基因破坏或插入获得受精卵和稳定遗传的后代 [139]。因此，CRISPR 对大鼠基因组高效的调控由于其在鼠模型中所具有的优势为研究提供了新的思路，尤其是在毒理学和药理学研究中 [138,140]。



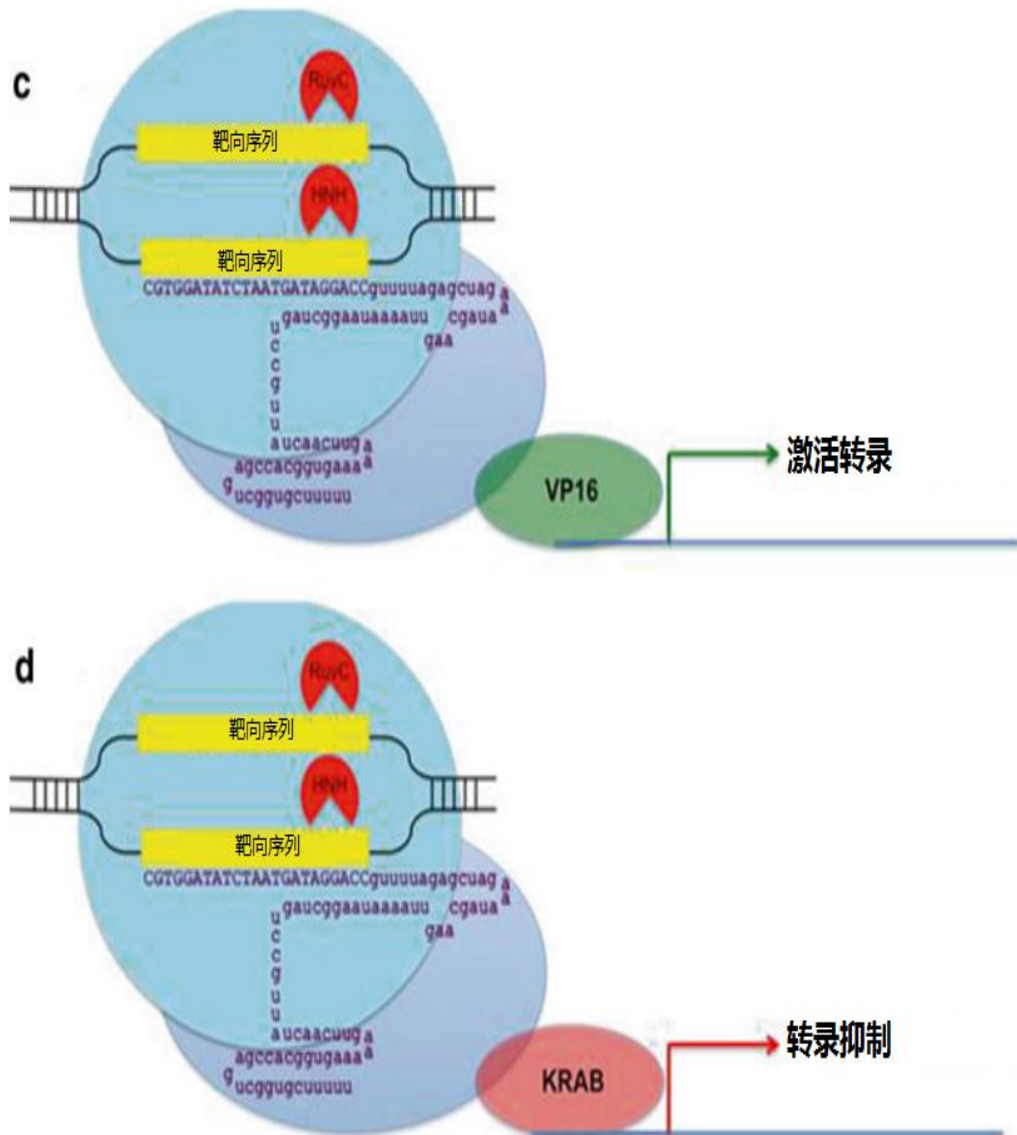


图6. CRISPR核酸酶和缺刻酶。 CRISPR/Cas9系统是由一段具有二级结构的向导RNA (紫色线段) 和与之相互作用的Cas9核酸酶所组成。复合体能够和目标序列连接, Cas9的两个结构域能够介导DNA正义链和反义链的损伤; HNH能够切割互补链, 而RuvC切割非互补链。(A) CRISPR/Cas9核酸酶。gRNA和Cas9共定位于目标位点, DNA双链的切割能够促进NHEJ和部分HR的进行。(B) CRISPR/Cas9缺刻酶。D10A突变能够使RuvC结构域失活 (用红色进行标记) 并能够切割DNA单链螺旋, 相对于NHEJ能够引起更频率的HR。(C) CRISPR/Cas9激活和 (D) CRISPR/Cas9抑制。失活HNH和RuvC能够移除Cas9的核酸酶活性。将这种Cas9与转录激活 (例如VP16) 或抑制结构 (例如Krueppel相关功能域) 进行融合能够产生具有基因表达调控功能的复合体。

非人灵长类（NHP）

作为一种系统化模型，NHP 相对于其它物种更接近人类。合适的疾病模型的产生曾经十分困难，而且，生殖成熟和妊娠产生少量的后代需要一段很长的时期[141]。为了解决这一问题，Wan 及其同时利用 CRISPR/Cas9 制造了 p53 纯和基因敲除动物。他们的研究发现高剂量的 gRNA 和 Cas9 能够对桑椹胚/囊胚阶段的胚胎发育产生负面的影响[141]。然而，当将低剂量的 Cas9 和 gRNA 注入到 108 个受精卵中时，有 62 个形态正常的发育完全的胚胎被移植到 13 个代孕动物中，其中有四只成功怀孕[141]。其中两只受孕产生了 3 个互贷，其中有两个具有 p53 基因修饰[141]。这一研究也证实 CRISPR/Cas9 有可能会介导胚胎产生 HDR，为产生和人类异常疾病更接近的 NHP 模型提供了新的思路。

人类

最早的在哺乳动物细胞中 CRISPR/Cas9 系统的应用的两篇报道指出该技术在人类干细胞中具有重大的潜力[129,130]。Mali 等人最早发现 CRISPR/Cas9 在 PGP1 iPSC 细胞系的 19 号染色体的 AAVS1 位点具有诱导 NHEJ 产生 indels 的能力[129]。Zhang 实验室第一个证明了该系统的多功能性[130]并对 64,751 个不同的 gRNAs 进行阳性和阴性生存能力筛选后对 HUES62 ESC 细胞系的 18,080 个基因进行靶向扩大了这一效应[142]。这一方法相对于只能靶向转录组的 RNA 干扰（RNAi）方法更进了一步，而且 CRISPR 技术能够对于全基因组进行靶向（例如启动子和增强子）[142]。

虽然所设计 gRNA 能够对近 50% 的基因组进行靶向，但是野生型的酿脓链球菌 Cas9 只能对 G (N₂₀) GG 结构的基序进行靶向[129,130]。不同 Cas9 之间的差异能够提高靶向性，利用正交实验差异能够同时对多个独立目标位点进行靶向。脑膜炎双球菌的 Cas9/gRNA 平台能够识别 5'-NNNNGATT-3' PAM，和酿脓链球菌的靶向系统具有不同的识别区域[143,144]。利用奈瑟氏属菌，Hou 等人能够在 H9 人类 ESCs 中破坏基因并在 DNMT3b 位点整合一段报告基因[143]。而且，通过嘌呤霉素筛选策略，他们能够提高 H1 和 H9 ESCs 以及 iPSCs[145]细胞中基因编辑的比例，实验结果显示有 60% 的目标位点正确的进行了基因整合[143]。来源于金黄色葡萄球菌的 Cas9 能够靶向于 NNAGAAW 序列，其主要的优势在于相对于其它的 Cas9 cDNAs 尺寸约小 1kb，因此能够整合进入 AAV 病毒促进高特异性的体内体外基因编辑[146]。理性的基因编辑也促进了靶向性的 Cas9 的广泛性应用。Kleinstiver 等人通过联合设计、结构分析和基于进化系统的细菌筛选过程对酿脓链球菌、酿脓嗜热和金黄色葡萄球菌的 Cas9 蛋白进行了改造，促进了其对新的 PAM 区域的差异化识别，并在斑马鱼和人类细胞中发挥了功能[147]。因此，新的差异化的 Cas9 的产生和应用进一步促进并提高该技术的应用。

为了达到目的，直接性的 CRISPR 基因修复也可见于鼠和人类的干细胞。Schwank 等人 CFTR 基因 11 号外显子纯和 F508 缺失纤维化囊肿患者来源的大小肠干细胞中研究了 CRISPR/Cas9 的效应[148]。经无性增殖得到的类器官恢复了完整性的功能并对 CRISPR/Cas9 修复人类疾病突变进行了概念性的验证[148]。在具有致白内障的 Crygc 基因突变中，Wu 等人发现将 gRNA/Cas9 RNA 注射进入受精卵并结合外源性的寡核苷酸模板或内源性的正常等位基因能够修复治病突变[149]。目前已经有多个课题组报道了人源性 iPSC 疾病模型中多个疾病（ β -地中海贫血病，肌肉萎缩症，慢性肉芽性囊肿）相关突变的修复[150-155]，其中很多都需要选择性标记的导入，并通过一些方法提高其在整合细胞群落中的比率。通常这还需要第二步操作利用 cre-lox 或转座子删除选择标志物。Grobarczyk 等人介绍了一种无需筛选的过程，经过对细胞的机械操作和胰酶分解使 HDR 的比例达到约 2%，NHEJ 的比例约 15% [156]。Miyaoaka 等人阐述了利用一种核酸酶和一种外源性寡核苷酸模板处理 iPSC 混合池，随后运用液滴数字化的 PCR 对经无限稀释法获得的单克隆细胞株进行筛选的方法[157]与之

类似，配对的核酸酶介导的基因修饰与非筛选技术联用能够催生微创编辑法的产生，被编辑的干细胞能够用于下游实验的操作。而且，由于 Cas9 蛋白具有 RuvC 和 HNH 结构域，每个结构域都能在所对应的单链 DNA 螺旋产生损伤（‘缺刻’），对 Cas9 这两种结构域的其中一个进行失活能够只对一条 DNA 单链进行切割产生 DNA 缺刻（图 6b）[129,130,158]。单链上的缺刻更易于产生无错的 HDR，然而核酸酶经常会同时触发 HDR 和 NHEJ 事件[159,160]。而且，配对的缺刻酶被证实具有更高的特异性[161-163]。与之类似，Cas9 的广泛性编辑能够引起不同目的的（例如 HDR vs NHEJ）的基因编辑达到最大的效率。作为这一功能的延伸，Cas9 能够作为融合蛋白进行更为微妙的，瞬时性的基因调控。失活 RuvC 和 HNH 其中的一个能够得到催化失活的蛋白但仍具有 gRNA 连接活性，并能够和其他的功能性结构进行融合获得新的功能。在干细胞中这一策略被用于将 Cas9 和 VP48（最小化 VP16 序列的三个拷贝）的活性结构域进行融合（图 6c）[164]。SOX2、OCT4 和 IL1RN 基因能够被用于高特异性的瞬时性激活，并可通过全基因组微阵列表达分析进行验证[164]。

在人类干细胞的基因调控可通过 SOX17 基因 Cas9-VP16 激活子或一个 OCT4A Cas9-KRAB 抑制子寻找分化调控基因，并显示出了 CRISPRs 介导不同效应的广泛性功能（图 6d）[165]。这一系统为特异性的分化调控因子的发现提供了平台，这随后也可用于细胞分化下游通路的研究。这一策略也可被扩大应用在 iPSCs 中，利用 VP64-p65-Rta 与失活的 Cas9 进行融合产生的“三方面活化剂”诱导更高水平的转录激活的产生[166]。gRNA 其本身也可用于编辑，Konermann 等人也找到了提高活性的方法，即将 gRNA 与适配器序列项链促进非 Cas9 作用范围内的修复的进行[167]。通过更大范围内的靶向增强子（近端和远端）表观遗传修饰，进一步的提高的靶向能力[168]。

高级的 CRISPR/Cas9 系统能够通过反复迭代进行积累或是通过直接性的改造赋予新的功能产生更为敏感的平台。其在干细胞中可用于疾病模型的构建、产生和修复，以及可用于表观调控的操作，可作为工具发挥瞬时性和永久性的动态效应。

总结

基因编辑和干细胞生物学的前景是光明的，在分子和细胞研究中已经改变了生物医学的研究模式。最早在 iPSC[8]中进行的实验具有极其重要的价值，因为其发现了一个可用于包括发展中的干细胞生物学等在内的多个领域的强大工具。除此之外，iPSCs 平台还可用于体外疾病模型的构建和药物发现，也是自体移植细胞和再生性药物的一个潜在性来源。然而由出生后细胞所带来的限制 ESC 应用的道德障碍必须被移除。在 iPSCs 和 ESCs 之间差异性的进一步研究也需要进行，从而对新一代的疗法平台进行鉴定。ESCs 在未来的主要性应用将主要集中在于家畜的操作和人类疾病啮齿类动物模型的建立。

第二类可用于干细胞改造的工具包括基因特异性的核酸酶。在过去的十年之内，这一领域得到了飞速的发展，从只在很少的实验室中构建的 ZFNs 到 TALENs 和 CRISPRs 可被任意的研究者只通过基础的生物学技巧就能熟练掌握。重要的是，核酸酶平台的每一次迭代都拓展了其应用领域且并未对已有的功能造成冲击。每种形式的设计和应用都交互性的强化了全领域的发展。ZFNs、TALENs 和 CRISPRs 都在干细胞领域中得到了应用，目前还不知哪一种会取得主要的地位。其中驱动性的因素如表 1 所示。

表 1 核酸酶对比矩阵

平台	费用	特异性	编辑难易度	应用领域	应用能力	多位点靶向	导入能力
大范围核酸酶	\$ 50,000	低-高	困难	大鼠, 小鼠	核酸酶	无	多途径
ZFN	\$ 5000	低-高	困难-中等	人, 小鼠, 大鼠, 牛, 猪/体细胞, ESC, iPSC	核酸酶, 缺刻酶	无	多途径
TALEN	\$ 500	低-高	简单	人, 小鼠, 大鼠, 牛, 猪, 非人灵长类/体细胞, ESC, iPSC	核酸酶, 缺刻酶, 激活因子, 表观调控	无	很少途径
CRISPR	\$ 5	低-高	最低	人, 小鼠, 大鼠, 牛, 猪, 非人灵长类/体细胞, ESC, iPSC	核酸酶, 缺刻酶, 激活因子, 抑制因子, 表观基因调控	大范围	多途径

基因特异性的核酸酶构建费用如左侧数据所示。特异性主要与目标位点的复杂性和序列的异质性相关。低复杂性序列核酸酶的设计被认为能够产生高频率的脱靶。更为严格性的结果具有更高的特异性。构建的难易取决于成本、构建过程的复杂性、和候选酶筛选所需核心试剂获得的难易程度决定的。应用能力是指这一类蛋白除核酸酶外进一步作为缺刻酶、激活子、抑制子所发挥功能的特征。多喂点靶向和倒入能力代表同时进行多个位点靶向的能力, 以及在大范围的病毒或非病毒表达功能域进行包装和导入的核心平台效应蛋白。MN, ZFN 和 TALEN 能同时靶向有限数量的目标位点, 然而 CRISPR/Cas9 能在全基因组中发挥多种功能。TALENs 的载体导入最为受限。

核酸酶和多功能干细胞都具有潜在的能够限制其活性的有害方面。对于多功能干细胞, 这和重编程之前的基因和表观修饰的存在和积累相关。在体外进行过基因编辑的 ESCs 或 iPSCs, 可能是同一细胞系的细胞甚至是同一培养皿中正进行增殖的细胞[9]。iPSCs 及其亲本前体细胞和 ESCs 中被报道出现过异倍体。国际干细胞研究组织发现核型畸形在每三个细胞系中就会出现一例[169]。12 三体是人类 ESCs 和 iPSCs 中最常见的异常类型, 而 17 三体常见于鼠的 ESCs[169,170]。有趣的是, NANOG 基因作为全能性的主要调控者, 位于 12 号染色体上, 因此 12 染色体三体的细胞由于基因含量升高能够在体外阳性筛选中获得优势[171,172]。拷贝数和单核苷酸的差异是来源细胞最主要的差异, 在重编程和扩增过程中并没有显著性的毒害作用[173]。这些发现也说明对目标细胞进行严格的筛选能够最小化下游不良事件的发生, 而通过转录激活进行重编程的过程不大可能会造成突变的产生。

然而，重编程本省能够引起表观遗传背景显著性的改变。在该过程中，在 X 染色体和多个分散位点中发生了表观遗传背景的全面重置[174]。例如，在人类女性细胞中，X 染色体拷贝会发生随机性失活，而这一模式会在子代细胞中有所保留，尤其是在重编程过程中及之前[175]。然而，随着培养过程的进行，X 的失活会逐渐在 ESCs 和 iPSCs 细胞中蔓延、扩增甚至是占据培养种群的大多数[176,177]。同时也会发生定点的表观遗传修饰。而且，亲本类型细胞甲基化的保持（例如表皮细胞和造血细胞）也存在于重编程过程中（‘表观记忆’）。为了支持这一观点，细胞系特异性的 DNA 甲基化模式被报道在细胞分化成特定类型细胞系的过程中对细胞表型具有显著性的影响[173,178,179]。

细胞系的可塑性是多功能细胞的标志之一。最为常见的检测可塑性的实验是在免疫缺陷的小鼠中进行的致畸瘤实验。未分化的细胞更易产生致畸瘤，这也是临床应用的重点。在诱导的代际遗传细胞中移除基因背景受污染的 iPSCs 的方法已被发现，而这类细胞高效率的移除甚至是完全的清除对临床应用具有重大的意义[180,181]。异常化的甲基化状态和微妙的基因改变也能够通过对生理机能的干扰或免疫原性的产生影响 iPSC 分化细胞的临床应用[182]。在技术实施之前必须对质量进行严格性的保障和控制。由于全基因组、外显子组和表观测序的成本逐渐降低，这些方法能够对干细胞操作和核酸酶诱导的基因有害性改变进行多重质量评价。

特定类型的人工编辑核酸酶的选择和效率取决于多个因素：成本、用户进行特定基因靶向时进行设计的能力（例如存活性），应用的广泛性、所发挥的功能（例如核酸酶、缺刻酶等），以及进行高效传导的能力（影响因素见表 1）。而其中最重要的因素可能就是安全性。通过设计能够使核酸酶对特定的 DNA 序列具有识别能力；然而，同于重复性片段和低识别率 DNA 序列的存在，在主要的靶位点和 OT 位点之间还会产生脱靶效应。目前 OT 位点的预测和鉴定方法包括体外建模、通过指数富集法构建的配体进化系统（SELEX）、和无偏基因组筛选。SELEX 方法可用于 ZFNs[183]和 TALENs[115]，但是无法完全的替代真实的体外体内靶位点[184]。有一种新的方法被证实对 ZFNs 具有更好的预测性[185]，进行 CRISPR OT 位点预测的数据库也可得到[186,187]。

无偏方法包括整合缺陷型慢病毒诱捕 (IDLV) 和利用线性化扩增介导 (LAM) PCR[55,188]、GUIDE 测序[189]、双基因组测序[190]，高通量 LAM PCR、全基因组转录测序进行定位 (HTGTS) [191]，和直接在损伤原位点标记、进行链霉素富集和新一代的基因测序 (BLESS) [146,192]。这些高敏感性方法对于 OT 位点的定位和发现非常有效，然而，无法预知的 OT 效应并未史无前例，这些方法会与临床应用也并非必须[55,193]。而且，目前还没有一类确定性的、主要的核酸酶完全不存在 OT 潜力，因此每个靶位点和候选核酸酶都必须进行独立性的考量。目前，产生特异性最为重要的因素为被靶向位点的异质性/复杂性。为了获得最大化的效能，对核酸酶进行重新编辑产生特异性的异二聚体[194]或缺刻酶/配对化的缺刻酶版本[162,163]，能够显著降低 OT 效应[195]。对基因组水平 OT 的效应、靶向位点的异质性（例如第二/OT 位点的最小重叠）、以及编辑组分中更为严格的核酸酶结构进行评估能力的提高将会使临床应用的基因编辑试剂更加的安全。

干细胞和核酸酶在遗传疾病动物模型、基因发现和功能性基因分析中产生了巨大的影响。通过核酸酶在多功能干细胞极其直接分化形成的终末端效应细胞中介导的精确性的基因编辑也具有良好的应用前景。在概念验证性实验中，非经基因组编辑的 iPSCs 第一次被用于年龄相关的黄斑变性患者的治疗[196]。这一基础性的研究为未来基因编辑干细胞的应用奠定了基础。然而，想要完全开发这一技术所具有的潜力，在重编程阶段和核酸酶导入后安全性问题必须被严格的考虑。高通量测序技术在基因组、外显子组和表观遗传水平和应用生物信息学分析进行的大数据评估中所取得的进步[197]促进了核酸酶编辑细胞的临床应用。

临床数据的分析也必须涵盖社会学和伦理学问题的考量，这将对干细胞治疗的可接受性

产生深远的影响。*Primum non nocere*（首要问题是不要伤害病人）[196]是医学研究、伦理学研究和 iPSCs 研究的核心问题，由于治疗工具直接取自于成人组织，这一点相较于 ESCs 具有更大的可接受性。基因组编辑的能力使得即使是在叫停了生殖系细胞编辑的前提下，伦理学的讨论也变得更为复杂[198]。在 2015 年，Liang 等人利用 CRISPR/Cas9 对人类受精卵的血红蛋白 B 位点进行基因编辑但是效率极其低下[199]。因此，必须推进复杂性的生物学的发展、优化、评估技术、干细胞的利用和基因编辑模型的构建以及具有复杂性的社会学和生物学领域交互作用的技术的进步，并时刻警惕其可能出现的“脱靶”效应。

致谢 M.J.O. 得到了 Lindahl 家族和 Corrigan 家族基金的资助。J.T. 得到了国家健康研究所（R01 AR063070 和 R01AR059947）项目、国防部（USAMRAA/DOD 军队部门 W81XWH-12-1-0609 和 USAMRAA/DOD W81XWH-10-1-0874）项目、DebRA 国际性、Sohana 研究资金、Richard M. Schulze 家族基金，Jackson Gabriel Silver 资金、Epidermolysis Bullosa Medical Research 资金和宾夕法尼亚大学 MPS I-11-009-01 奖学金的资助。M..J.O 和 J.T 也得到了明尼苏达州儿童肿瘤研究资金的资助。这部出版物中的研究报告得到了国家中心的先进科学转化部门国家健康研究奖金的支持，奖金代码为 UL1TR000114（M.J.O.）。内容由作者本人负责，不代表国家健康研究所的官方观点。

参考文献

1. Thomson JA, et al. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998;282:1145 – 7.
2. Puri MC, Nagy A. Concise review: embryonic stem cells versus induced pluripotent stem cells: the game is on. *Stem Cells*. 2012;30:10 – 4.
3. Weismann A, Poulton EB, Schönland S, Shipley AE. *Essays upon heredity and kindred biological problems*. Oxford: Clarendon; 1891.
4. Solana J. Closing the circle of germline and stem cells: the primordial stem cell hypothesis. *Evodevo*. 2013;4:2.
5. Gurdon JB. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J Embryol Exp Morphol*. 1962;10:622 – 40.
6. Briggs R, King TJ. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc Natl Acad Sci U S A*. 1952;38:455 – 63.
7. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature*. 1997;385:810 – 3.
8. Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007;131:861 – 72.
9. Liang G, Zhang Y. Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell*. 2013;13:149 – 59.
10. Silva G, et al. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther*. 2011;11:11 – 27.
11. Popplewell L, et al. Gene correction of a duchenne muscular dystrophy mutation by meganuclease-enhanced exon knock-in. *Hum Gene Ther*. 2013;24:692 – 701.
12. Dupuy A, et al. Targeted gene therapy of xeroderma pigmentosum cells using meganuclease and TALEN. *PLoS One*. 2013;8, e78678.
13. Cohen-Tannoudji M, et al. I-SceI-induced gene replacement at a natural locus in embryonic

- stem cells. *Mol Cell Biol.* 1998;18:1444 – 8.
14. Menoret S, et al. Generation of Rag1-knockout immunodeficient rats and mice using engineered meganucleases. *FASEB J.* 2013;27:703 – 11.
 15. Durai S, et al. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* 2005;33:5978 – 90.
 16. Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. *Nat Biotechnol.* 2005;23:967 – 73.
 17. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet.* 2005;6:507 – 12.
 18. Geurts AM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science.* 2009;325:433.
 19. Mashimo T, et al. Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS One.* 2010;5, e8870.
 20. Cui X, et al. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol.* 2011;29:64 – 7.
 21. Ferguson C, McKay M, Harris RA, Homanics GE. Toll-like receptor 4 (Tlr4) knockout rats produced by transcriptional activator-like effector nuclease (TALEN)-mediated gene inactivation. *Alcohol.* 2013;47:595 – 9.
 22. Carbery ID, et al. Targeted genome modification in mice using zinc-finger nucleases. *Genetics.* 2010;186:451 – 9.
 23. Meyer M, de Angelis MH, Wurst W, Kuhn R. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci U S A.* 2010;107:15022 – 6.
 24. Brinster RL, et al. Targeted correction of a major histocompatibility class II E alpha gene by DNA microinjected into mouse eggs. *Proc Natl Acad Sci U S A.* 1989;86:7087 – 91.
 25. Chen YG, et al. Gene Targeting in NOD mouse embryos using zinc-finger nucleases. *Diabetes.* 2014;63:68 – 74.
 26. Meyer M, Ortiz O, Hrabe de Angelis M, Wurst W, Kuhn R. Modeling disease mutations by gene targeting in one-cell mouse embryos. *Proc Natl Acad Sci U S A.* 2012;109:9354 – 9.
 27. Osiak A, et al. Selection-independent generation of gene knockout mouse embryonic stem cells using zinc-finger nucleases. *PLoS One.* 2011;6, e28911.
 28. Campbell KH, McWhir J, Ritchie WA, Wilmut I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature.* 1996;380:64 – 6.
 29. Ogura A, Inoue K, Wakayama T. Recent advancements in cloning by somatic cell nuclear transfer. *Philos Trans R Soc Lond B Biol Sci.* 2013;368:20110329.
 30. Whyte JJ, et al. Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs. *Mol Reprod Dev.* 2011;78:2.
 31. Hauschild J, et al. Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proc Natl Acad Sci U S A.* 2011;108:12013 – 7.
 32. Platt JL, Lin SS. The future promises of xenotransplantation. *Ann N Y Acad Sci.* 1998;862:5 – 18.
 33. DuBose J, Salim A. Aggressive organ donor management protocol. *J Intensive Care Med.* 2008;23:367 – 75.
 34. Salim A, et al. The effect of a protocol of aggressive donor management: implications for

- the national organ donor shortage. *J Trauma*. 2006;61:429 – 33; discussion 433-5.
35. Yu S, et al. Highly efficient modification of beta-lactoglobulin (BLG) gene via zinc-finger nucleases in cattle. *Cell Res*. 2011;21:1638 – 40.
36. Liu X, et al. Zinc-finger nickase-mediated insertion of the lysostaphin gene into the beta-casein locus in cloned cows. *Nat Commun*. 2013;4:2565.
37. Oliver SP, Murinda SE. Antimicrobial resistance of mastitis pathogens. *Vet Clin North Am Food Anim Pract*. 2012;28:165 – 85.
38. Schuhardt VT, Schindler CA. Lysostaphin therapy in mice infected with *Staphylococcus aureus*. *J Bacteriol*. 1964;88:815 – 6.
39. van den Bogaard AE, Stobberingh EE. Epidemiology of resistance to antibiotics. Links between animals and humans. *Int J Antimicrob Agents*. 2000;14:327 – 35.
40. Hollis A, Ahmed Z. Preserving antibiotics, rationally. *N Engl J Med*. 2013;369:2474 – 6.
41. McDermott W, Rogers DE. Social ramifications of control of microbial disease. *Johns Hopkins Med J*. 1982;151:302 – 12.
42. Lombardo A, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol*. 2007;25:1298 – 306.
43. Genovese P, et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature*. 2014;510:235 – 40.
44. Watts KL, et al. CD34(+) expansion with Delta-1 and HOXB4 promotes rapid engraftment and transfusion independence in a *Macaca nemestrina* cord blood transplant model. *Mol Ther*. 2013;21:1270 – 8.
45. Boitano AE, et al. Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. *Science*. 2010;329:1345 – 8.
46. Delaney C, Varnum-Finney B, Aoyama K, Brashem-Stein C, Bernstein ID. Dose-dependent effects of the notch ligand Delta1 on ex vivo differentiation and in vivo marrow repopulating ability of cord blood cells. *Blood*. 2005;106:2693 – 9.
47. Mohrin M, et al. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell*. 2010;7:174 – 85.
48. North TE, et al. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature*. 2007;447:1007 – 11.
49. Goessling W, et al. Prostaglandin E2 enhances human cord blood stem cell xenotransplants and shows long-term safety in preclinical nonhuman primate transplant models. *Cell Stem Cell*. 2011;8:445 – 58.
50. Vargas Jr J, Gusella GL, Najfeld V, Klotman ME, Cara A. Novel integrase-defective lentiviral episomal vectors for gene transfer. *Hum Gene Ther*. 2004;15:361 – 72.
51. Nightingale SJ, et al. Transient gene expression by nonintegrating lentiviral vectors. *Mol Ther*. 2006;13:1121 – 32.
52. Yanez-Munoz RJ, et al. Effective gene therapy with nonintegrating lentiviral vectors. *Nat Med*. 2006;12:348 – 53.
53. Philippe S, et al. Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo. *Proc Natl Acad Sci U S A*. 2006;103:17684 – 9.
54. Leavitt AD, Robles G, Alesandro N, Varmus HE. Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. *J Virol*. 1996;70:721 – 8.

55. Gabriel R, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol.* 2011;29:816 – 23.
56. Osborn MJ, et al. TALEN-based gene correction for epidermolysis bullosa. *Mol Ther.* 2013;21(6):1151 – 9.
57. Li L, et al. Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Mol Ther.* 2013;21:1259 – 69.
58. Rahman SH, et al. The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. *Hum Gene Ther.* 2013;24:67 – 77.
59. Saydaminova K, et al. Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. *Mol Ther Methods Clin Dev.* 2015;1:14057.
60. Stephen SL, et al. Chromosomal integration of adenoviral vector DNA in vivo. *J Virol.* 2010;84:9987 – 94.
61. Harui A, Suzuki S, Kochanek S, Mitani K. Frequency and stability of chromosomal integration of adenovirus vectors. *J Virol.* 1999;73:6141 – 6.
62. Chen Z, et al. Receptor-mediated delivery of engineered nucleases for genome modification. *Nucleic Acids Res.* 2013;41, e182.
63. Holt N, et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol.* 2010;28:839 – 47.
64. Hofer U, et al. Pre-clinical modeling of CCR5 knockout in human hematopoietic stem cells by zinc finger nucleases using humanized mice. *J Infect Dis.* 2013;208 Suppl 2:S160 – 4.
65. Allers K, et al. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. *Blood.* 2011;117:2791 – 9.
66. Hutter G, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med.* 2009;360:692 – 8.
67. de Mendoza C, et al. Prevalence of X4 tropic viruses in patients recently infected with HIV-1 and lack of association with transmission of drug resistance. *J Antimicrob Chemother.* 2007;59:698 – 704.
68. Didigu CA, et al. Simultaneous zinc-finger nuclease editing of the HIV coreceptors *ccr5* and *cxcr4* protects CD4+ T cells from HIV-1 infection. *Blood.* 2013.
69. Hockemeyer D, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol.* 2009;27:851 – 7.
70. DeKolver RC, et al. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.* 2010;20:1133 – 42.
71. Lombardo A, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods.* 2011;8:861 – 9.
72. Wang Y, et al. Genome editing of human embryonic stem cells and induced pluripotent stem cells with zinc finger nucleases for cellular imaging. *Circ Res.* 2012;111:1494 – 503.
73. Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer.* 2012;12:51 – 8.
74. Zou J, et al. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell.* 2009;5:97 – 110.

75. Lengner CJ, et al. Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell*. 2010;141:872 – 83.
76. Narsinh KH, et al. Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors. *Nat Protoc*. 2011;6:78 – 88.
77. Lin T, et al. A chemical platform for improved induction of human iPSCs. *Nat Methods*. 2009;6:805 – 8.
78. Kim D, et al. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell*. 2009;4:472 – 6.
79. Soldner F, et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*. 2011;146:318 – 31.
80. Hussein SM, et al. Copy number variation and selection during reprogramming to pluripotency. *Nature*. 2011;471:58 – 62.
81. Dreyer AK, Cathomen T. Zinc-finger nucleases-based genome engineering to generate isogenic human cell lines. *Methods Mol Biol*. 2012;813:145 – 56.
82. Sebastiano V, et al. In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells*. 2011;29:1717 – 26.
83. Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPSC cells derived from an adult patient with sickle cell disease. *Blood*. 2011;118:4599 – 608.
84. Voit RA, Hendel A, Pruett-Miller SM, Porteus MH. Nuclease-mediated gene editing by homologous recombination of the human globin locus. *Nucleic Acids Res*. 2013.
85. Yusa K, et al. Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. *Nature*. 2011;478:391 – 4.
86. Rahman SH, et al. Rescue of DNA-PK signaling and T-cell differentiation by targeted genome editing in a *prkdc* deficient iPSC disease model. *PLoS Genet*. 2015;11, e1005239.
87. Jiang J, et al. Translating dosage compensation to trisomy 21. *Nature*. 2013;500:296 – 300.
88. Maeder ML, et al. Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell*. 2008;31:294 – 301.
89. Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK. Oligomerized pool engineering (OPEN): an ‘open-source’ protocol for making customized zinc-finger arrays. *Nat Protoc*. 2009;4:1471 – 501.
90. Christian ML, et al. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One*. 2012;7, e45383.
91. Cermak T, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011;39, e82.
92. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*. 2009;326:1501.
93. Tesson L, et al. Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol*. 2011;29:695 – 6.
94. Mashimo T, et al. Efficient gene targeting by TAL effector nucleases coinjected with exonucleases in zygotes. *Sci Rep*. 2013;3:1253.
95. Certo MT, et al. Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. *Nat Methods*. 2012;9:973 – 5.

96. Sung YH, et al. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol.* 2013;31:23 – 4.
97. Davies B, et al. Site specific mutation of the *Zic2* locus by microinjection of TALEN mRNA in mouse CD1, C3H and C57BL/6J oocytes. *PLoS One.* 2013;8, e60216.
98. Brown SA, et al. Holoprosencephaly due to mutations in *ZIC2*, a homologue of *Drosophila* odd-paired. *Nat Genet.* 1998;20:180 – 3.
99. Qiu Z, et al. High-efficiency and heritable gene targeting in mouse by transcription activator-like effector nucleases. *Nucleic Acids Res.* 2013;41, e120.
100. Kahle M, et al. Phenotypic comparison of common mouse strains developing high-fat diet-induced hepatosteatosis. *Mol Metab.* 2013;2:435 – 46.
101. Wang H, et al. TALEN-mediated editing of the mouse Y chromosome. *Nat Biotechnol.* 2013;31:530 – 2.
102. Takada S, et al. Targeted gene deletion of miRNAs in mice by TALEN system. *PLoS One.* 2013;8, e76004.
103. Wefers B, et al. Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc Natl Acad Sci U S A.* 2013;110:3782 – 7.
104. Panda SK, et al. Highly efficient targeted mutagenesis in mice using TALENs. *Genetics.* 2013;195:703 – 13.
105. Low BE, et al. Correction of *Crb1rd8* allele and retinal phenotype in C57BL/6N mice via TALEN-mediated homology-directed repair. *Invest Ophthalmol Vis Sci.* 2013.
106. Orlando SJ, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res.* 2010;38, e152.
107. Carlson DF, et al. Efficient TALEN-mediated gene knockout in livestock. *Proc Natl Acad Sci U S A.* 2012;109:17382 – 7.
108. Hanna J, et al. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science.* 2007;318:1920 – 3.
109. Tan W, et al. Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *Proc Natl Acad Sci U S A.* 2013;110:16526 – 31.
110. Xin J, et al. Highly efficient generation of GGTA1 biallelic knockout inbred mini-pigs with TALENs. *PLoS One.* 2013;8, e84250.
111. Christian M, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics.* 2010;186:757 – 61.
112. Miller JC, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol.* 2011;29:143 – 8.
113. Mussolino C, et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 2011;39:9283 – 93.
114. Liu H, et al. TALEN-mediated gene mutagenesis in rhesus and cynomolgus monkeys. *Cell Stem Cell.* 2014;14:323 – 8.
115. Hockemeyer D, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol.* 2011;29:731 – 4.
116. Ding Q, et al. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell.* 2013;12:238 – 51.
117. Hussain K, et al. An activating mutation of *AKT2* and human hypoglycemia. *Science.* 2011;334:474.

118. Musunuru K, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010;466:714 – 9.
119. Zhu F, et al. DICE, an efficient system for iterative genomic editing in human pluripotent stem cells. *Nucleic Acids Res*. 2013.
120. Zhu H, et al. Baculoviral transduction facilitates TALEN-mediated targeted transgene integration and Cre/LoxP cassette exchange in human-induced pluripotent stem cells. *Nucleic Acids Res*. 2013;41, e180.
121. Zhang Z, et al. Dissecting the roles of miR-302/367 cluster in cellular reprogramming using TALE-based repressor and TALEN. *Stem Cell Reports*. 2013;1:218 – 25.
122. Bultmann S, et al. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res*. 2012;40:5368 – 77.
123. Maeder ML, et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol*. 2013;31:1137 – 42.
124. Choi SM, et al. Efficient drug screening and gene correction for treating liver disease using patient-specific stem cells. *Hepatology*. 2013;57:2458 – 68.
125. Sun N, Zhao H. Seamless correction of the sickle cell disease mutation of the HBB gene in human induced pluripotent stem cells using TALENs. *Biotechnol Bioeng*. 2013.
126. Cermak T, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011;39(17):7879.
127. Holkers M, et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res*. 2013;41, e63.
128. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337:816 – 21.
129. Mali P, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823 – 6.
130. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819 – 23.
131. Mashiko D, et al. Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. *Sci Rep*. 2013;3:3355.
132. Li D, et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31:681 – 3.
133. Fujii W, Kawasaki K, Sugiura K, Naito K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res*. 2013;41, e187.
134. Zhou J, et al. One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. *Int J Biochem Cell Biol*. 2014;46:49 – 55.
135. Wang H, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;153:910 – 8.
136. Yang H, et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;154:1370 – 9.
137. Koike-Yusa H, Li Y, Tan E-P, Velasco-Herrera MDC, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol*. 2014;32(3):267 – 73.

138. Ma Y, et al. Generating rats with conditional alleles using CRISPR/Cas9. *Cell Res.* 2014;24:122 – 5.
139. Li W, et al. Genetic modification and screening in Rat using haploid embryonic stem cells. *Cell Stem Cell.* 2013.
140. Jacob HJ, Kwitek AE. Rat genetics: attaching physiology and pharmacology to the genome. *Nat Rev Genet.* 2002;3:33 – 42.
141. Wan H, et al. One-step generation of p53 gene biallelic mutant Cynomolgus monkey via the CRISPR/Cas system. *Cell Res.* 2015;25:258 – 61.
142. Shalem O, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* 2014;343:84 – 7.
143. Hou Z, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci U S A.* 2013;110:15644 – 9.
144. Esvelt KM, et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods.* 2013;10:1116 – 21.
145. Chen G, et al. Chemically defined conditions for human iPSC derivation and culture. *Nat Methods.* 2011;8:424 – 9.
146. Ran FA, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature.* 2015;520:186 – 91.
147. Kleinstiver BP, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* 2015.
148. Schwank G, et al. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell.* 2013;13:653 – 8.
149. Wu Y, et al. Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell.* 2013;13:659 – 62.
150. Sim X, Cardenas-Diaz FL, French DL, Gadue P. A doxycycline-inducible system for genetic correction of iPSC disease models. *Methods Mol Biol.* 2016;1353:13 – 23.
151. Song B, et al. Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. *Stem Cells Dev.* 2015;24:1053 – 65.
152. Li HL, et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports.* 2015;4:143 – 54.
153. Xie F, et al. Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* 2014;24:1526 – 33.
154. Flynn R, et al. CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPSCs. *Exp Hematol.* 2015;43(10):838 – 48.e3.
155. Huang X, et al. Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient iPSCs after genome editing of the sickle point mutation. *Stem Cells.* 2015;33:1470 – 9.
156. Grobarczyk B, Franco B, Hanon K, Malgrange B. Generation of isogenic human iPSC cell line precisely corrected by genome editing using the CRISPR/Cas9 system. *Stem Cell Rev.* 2015;11(5):774 – 87.
157. Miyaoka Y, et al. Isolation of single-base genome-edited human iPSCs without antibiotic selection. *Nat Methods.* 2014;11:291 – 3.

158. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*. 2012;109(39):E2579 – 86.
159. Davis L, Maizels N. Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc Natl Acad Sci U S A*. 2014;111:E924 – 32.
160. Osborn M, Gabriel R, Webber BR, DeFeo AP, McElroy AN, Jarjour J, Starker CG, Wagner JE, Joung JK, Voytas DF, von Kalle C, Schmidt M, Blazar BR, Tolar J. Fanconi anemia gene editing by the CRISPR/Cas9 system. *Hum Gene Ther*. 2014;26(2):114 – 26.
161. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol*. 2014;32:577 – 82.
162. Wyvekens N, Topkar VV, Khayter C, Joung JK, Tsai SQ. Dimeric CRISPR RNA-guided FokI-dCas9 nucleases (RFNs) directed by truncated gRNAs for highly specific genome editing. *Hum Gene Ther*. 2015.
163. Tsai SQ, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32:569 – 76.
164. Cheng AW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res*. 2013;23:1163 – 71.
165. Kearns NA, et al. Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development*. 2014;141:219 – 23.
166. Chavez A, et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*. 2015;12:326 – 8.
167. Konermann S, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015;517:583 – 8.
168. Hilton IB, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol*. 2015;33:510 – 7.
169. Amps K, et al. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol*. 2011;29:1132 – 44.
170. Ben-David U, Benvenisty N. High prevalence of evolutionarily conserved and species-specific genomic aberrations in mouse pluripotent stem cells. *Stem Cells*. 2012;30:612 – 22.
171. Mayshar Y, et al. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell*. 2010;7:521 – 31.
172. Mayshar Y, Yanuka O, Benvenisty N. Teratogen screening using transcriptome profiling of differentiating human embryonic stem cells. *J Cell Mol Med*. 2011;15:1393 – 401.
173. Ruiz S, et al. Analysis of protein-coding mutations in hiPSCs and their possible role during somatic cell reprogramming. *Nat Commun*. 2013;4:1382.
174. Liang G, Zhang Y. Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective. *Cell Res*. 2013;23:49 – 69.
175. Tchiew J, et al. Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell*. 2010;7:329 – 42.
176. Mekhoubad S, et al. Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell*. 2012;10:595 – 609.
177. Anguera MC, et al. Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. *Cell Stem Cell*. 2012;11:75 – 90.

178. Lister R, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*. 2011;471:68 – 73.
179. Ruiz S, et al. Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proc Natl Acad Sci U S A*. 2012;109:16196 – 201.
180. Lee MO, et al. Inhibition of pluripotent stem cell-derived teratoma formation by small molecules. *Proc Natl Acad Sci U S A*. 2013;110:E3281 – 90.
181. Tang C, et al. An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nat Biotechnol*. 2011;29:829 – 34.
182. Okita K, Nagata N, Yamanaka S. Immunogenicity of induced pluripotent stem cells. *Circ Res*. 2011;109:720 – 1.
183. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods*. 2011;8(9):765 – 70.
184. Cheng L, Blazar B, High K, Porteus M. Zinc fingers hit off target. *Nat Med*. 2011;17(10):1192 – 3.
185. Sander JD, et al. In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res*. 2013;41, e181.
186. Ran FA, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154:1380 – 9.
187. Ran FA, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8:2281 – 308.
188. Paruzynski A, et al. Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nat Protoc*. 2010;5(8):1379 – 95.
189. Tsai SQ, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*. 2015;33:187 – 97.
190. Kim D, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods*. 2015;12:237 – 43, 1 p following 243.
191. Frock RL, et al. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol*. 2015;33:179 – 86.
192. Crosetto N, et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods*. 2013;10:361 – 5.
193. Tebas P, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med*. 2014;370:901 – 10.
194. Doyon Y, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods*. 2011;8(1):74 – 9.
195. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 2014;32:279 – 84.
196. Nakano-Okuno M, Borah BR, Nakano I. Ethics of iPSC-based clinical research for age-related macular degeneration: patient-centered risk-benefit analysis. *Stem Cell Rev*. 2014;10:743 – 52.
197. Mardis ER. The \$1,000 genome, the \$100,000 analysis? *Genome Med*. 2010;2:84.
198. Lanphier E, Urnov F, Haecker SE, Werner M, Smolenski J. Don't edit the human germ line. *Nature*. 2015;519:410 – 1.
199. Liang P, et al. CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein Cell*. 2015;6:363 – 72.