

## 热点评析

## 体内基因组编辑的最新研究进展

黄芳 郭礼和

(复旦大学上海医学院, 上海 200032)

在本刊上一期的“热点评析”栏目中<sup>[1]</sup>, 我们介绍了一种新的基因治疗技术——体内原位基因组编辑。这项技术也可称作“基因组工程”, 是在基因工程和基因组学发展的基础上产生的一门新技术。它适用于体内原位基因组的修饰和改造, 为生命科学的基础研究、生物品种改良、创立新的物种和体内原位基因治疗开辟了新的方法和途径, 为生物经济的发展起到推波助澜的作用。

上一期的“热点评析”栏目中, 我们介绍了在血友病小鼠模型上利用锌指核酸酶(ZFN)技术结合具有亲肝性的腺相关病毒AAV8, 对模型小鼠肝脏细胞基因组进行在体编辑, 重建了疾病小鼠的凝血功能。但这些研究成果向临床转化之前, 仍有许多安全性和有效性的问题需要解决, 例如: 基因突变矫正的效率优化; 人类基因组编辑的脱靶效应(off-target)(非特异性)筛查; ZFNs的毒性和安全性等。今年9月 *Nature Methods* 杂志发表了两篇有关这方面的研究论文, 一篇是关于利用人工合成的单链寡脱氧核苷酸来代替基因打靶载体, 从而提高基因组编辑的效率<sup>[2]</sup>; 另一篇是关于如何通过离体筛选的手段分析ZFN的脱靶效应(非特异性打靶)<sup>[3]</sup>。本文对这两项基于ZFN技术的体内基因组编辑的最新进展作进一步的介绍。

## 1 利用单链寡脱氧核苷酸和锌指核酸酶实现高效基因组编辑

发表在今年9月 *Nature Methods* 杂志上的这一研究工作是由Sigma-Aldrich生物技术公司、哥本哈根大学和加州大学的研究人员共同完成的<sup>[2]</sup>。

利用锌指核酸酶对基因组进行特定的改造通常依赖于基因打靶载体质粒(donor plasmids)的选择, 它必须携带打靶位点两侧200~800 bp长度的同源臂。然而, 同经典的基因打靶载体相比, 质粒作为打靶载体要优异, 因为它小而简单, 但仍需要几周的

时间去构建。若是用人工合成的单链脱氧寡核苷酸(single-stranded oligodeoxynucleotides, ssODNs)作为基因打靶载体的话, 只要打靶效率得到保证, 它的优越性就会更加明显了, 因为它的设计与合成可在几天内完成, 从而能够加快和简化基于锌指核酸酶技术的基因打靶研究。在该论文中, 作者采用人工设计和合成的单链脱氧寡核苷酸代替传统的基因打靶载体来完成细胞内基因组编辑, 使基因组编辑变得更为灵活、高效和实用。

之前的研究表明, 双链DNA断裂介导的基因改造在断裂点附近最为有效。作者用单链脱氧寡核苷酸作为打靶载体进行了这样的尝试: 人工设计和合成相应AAVS1位点的ssODNs AAVS1-95 (数字95表示核苷酸序列的长度)寡核苷酸片段, 在AAVS1内的ZFN切割位点处引入限制性内切酶Hind III。此Hind III酶切位点的两侧各有约40 bp长的同源序列。用此单链寡核苷酸进行基因组打靶, 从转染的细胞中挑出7株, 作者检测到定点插入的效率为7%~57%, 插入效率与细胞类型有关。作者进一步确定了ssODNs中最小同源序列的长度, 当两侧的同源序列之和减少到30个碱基时, 同源重组的效率明显降低。此外, 作者还发现当ssODNs离断裂位点的距离超过0.1~1 Kb时, 检测不到任何的同源重组体。打靶载体ssODNs相比于打靶载体质粒, 具有更高的效率和整合序列的保真性。另外, 质粒DNA与mRNA都可用来表达ZFNs。

ZFNs很难做到在期望的突变位点上精确地切割。作者尝试了如何在RPS6KA3基因上产生一个单密码子变异。RPS6KA3基因编码激酶RSK2, 而RSK2同智力迟钝、精神运动与骨骼障碍(psychomotor and skeletal disorders)、癌症等有关。其酶活性中心区域一个位点发生半胱氨酸残基变成缬氨酸的突变, 使得RSK2对激酶抑制剂fmk不敏感。作者构建了一个人工ZFN锌指核酸酶, 该ZFN的切割

位点距离突变位点(Cys436)有27 bp的长度, 还设计了一个长度为125个核苷酸的打靶载体ssODN (ssODN RSK2-125) (图1)。在打靶载体ssODN中引入了一个 *Bam*H I位点, 用于对重组克隆的限制性片段长度多样性 (restriction fragment length polymorphism, RFLP) 的检测。最后, 作者非常缜密地在打靶载体ssODN内的ZFN结合位点处引入两个变异核苷酸, 以确保ZFNs不会对突变后的基因位点进行第二次切割。如果发生二次切割, 非同源末端连接修复(nonhomologous end-joining repair, NHEJ)可能会导致基因中微小的缺失或插入, 影响基因的功能。

将ssODN RSK2-125和RSK2 ZFN mRNA转入人K562细胞, 通过RFLP分析, 同源重组率高达22%~32%, 而单细胞克隆中携带双等位基因C436V位点突变的克隆约占3%。激酶活性分析实验表明随机挑选的三个携带C436V位点突变的细胞克隆都

对RSK2抑制剂fmk不敏感。

此外, 作者利用上述在体基因组编辑技术对基因或外显子的定点缺失作了研究。

目前, 利用ZFN介导的基因组内的缺失绝大多数是由非同源末端连接过程中的错误修复产生的, 一对ZFN引发的缺失长度往往<50 bp。而大片的缺失需要两对ZFNs来完成。作者采用ssODNs结合ZFN技术对*AAVS1*位点进行了大片段缺失的研究, 选择缺失的长度为0.1~100 Kb(图2)。在人K562细胞中, *AAVS1*位点为三倍体(triploid), 经过单次ZFN处理, 用PCR对细胞克隆进行筛选, 结果显示这三个等位基因分别完全缺失了1, 5, 10 Kb的几率为2.9%, 0.6%和0.2%, 而缺失较短长度(100 bp和500 bp)的三等位基因几率高达21%和10.8%。当缺失长度为100 Kb时, 仍有近0.2%的重组几率。对7种不同细胞株和4种不同基因, 作者做了技术验证。采用*AAVS1* ZFNs和位

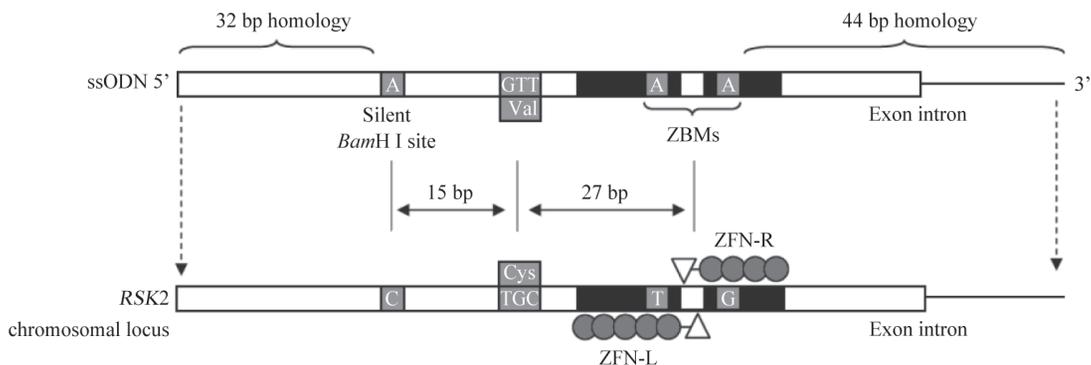


图1 ssODN设计(图的上半部)和人RSK2基因(图的下半部)及锌指核酸酶(ZFN)结合的左(L)右(R)臂示意图<sup>[2]</sup>

RSK2-125供体ssODN为125个核苷酸的聚单链脱氧核苷酸, 可以对人的RSK2基因进行编辑, 引入三种突变: (1) 将一个胞嘧啶替换为腺嘌呤, 产生一个*Bam*H I酶切位点; (2) 将编码436位半胱氨酸的密码子TGC转变为GTT, 即在蛋白水平上变异为缬氨酸; (3) ZFN左右臂中(T和G)各引入一个腺嘌呤核苷酸(A)突变, 防止ZFN的二次切割(ZFN-blocking mutation, ZBM)。

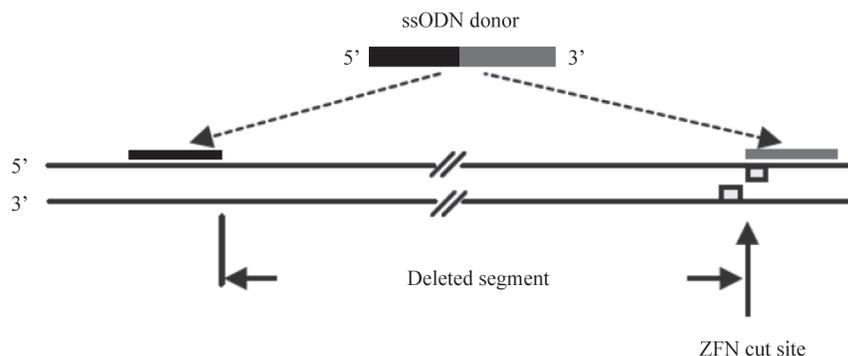


图2 利用ssODNs和ZFNs可以实现染色体片段的缺失<sup>[3]</sup>

设计的缺失位于ZFN切割部位。ssODN由两部分构成, 远离ZFN切割位点的远端序列(5'部分)和携带ZFN半序列的近端序列(3'部分)。

于两侧的两个ssODNs, 可实现切点处5'和3'两个方向上的缺失。作者还做到了在大片段缺失的同时, 插入诸如*loxP*重组位点这样的短序列, 以便对基因组进行再一次改造。

文章的最后, 作者针对之前报道的ssODNs对细胞的不利影响进行了分析。以往文献报道的研究有一个共同的地方, 就是使用的ssODNs携带磷硫酰键(phosphorothioate linkages)。近来, 研究表明磷硫酰修饰(phosphorothioate modification)的ssODNs对细胞生理活性具有毒性, 而ssODNs本身对细胞几乎没有影响。在该研究中, 作者正是采用了未修饰的ssODNs进行的在体基因组编辑技术操作, 故而这种技术方法对细胞是安全的。

作者的最终结论是, 利用未经修饰的单链脱氧寡核苷酸和精确定位的锌指核酸酶可以实现三大类在体基因组编辑技术操作: (1) 定点突变; (2) 长短序列的定点缺失; (3) 在进行长序列的定点缺失技术操作的同时, 可以插入短的遗传元件。在没有抗生素筛选的条件下, 基因组编辑的效率为1%~30%。这篇论文证明: 利用单链脱氧寡核苷酸(ssODNs)代替传统的打靶载体结合ZFN技术进行在体基因组编辑是可行而有效的, 能够更为方便地对哺乳动物细胞乃至真核系统进行基因组修饰和改造。

## 2 通过离体筛选的手段揭示ZFN的脱靶效应

ZFN技术的发明和应用, 为开展动植物基因组改造带来了前所未有的方便, 并为基因治疗和动植物品种改造提供了新技术、新方法和新思路。有关ZFN的专一性研究是这一技术应用的另一个关注问题。通常真核细胞基因组很庞大, 含有ZFN切点的序列可能很多, 虽然对该酶识别的旁侧序列延长到20多个碱基对, 但在庞大的基因组序列中还会有相似的序列出现, 这样就会引起脱靶效应产生。研究脱靶效应自然就会成为大家关心的热点。在今年9

月份出版的*Nature Methods*上刊登了由哈佛大学、麻省总医院、哈佛医学院的科学家联合发表的文章<sup>[3]</sup>, 对CCR5-224和VF2468两个ZFN的脱靶情况进行了系统的研究。

CCR5-224识别人CCR-5基因序列上的一个24 bp的位点, 其对HIV病人的治疗靶点已在进行临床研究; VF2468识别人*VEGFR*基因中的一个18 bp的位点, 可以用于抗血管生成的相关研究和临床应用。研究组首先根据这两个ZFN识别序列各自建立了一个容量为 $10^{11}$ 的靶点DNA序列突变库, 经和不同浓度ZFN作用后, 对切点序列进行扩增和高通量测序, 获得了ZFN在体外的识别专一性和可切割位点的特征。研究表明, 数以千计的识别位点的突变序列可以被上述两种ZFN切割, 其脱靶效应随ZFN浓度的升高而上升。发生一个突变的识别序列几乎100%都能被相应的ZFN识别并切割, 在所发现的可切割位点中有37个CCR5-224脱靶位点和2 652个VF2468脱靶位点存在于人类基因组中。研究者进一步在人K562细胞中发现其中的9种CCR5-224脱靶位点和31种VF2468脱靶位点可以实际发生, 并且CCR5-224的一个脱靶位点恰好在肿瘤抑制基因*BTBD10*的启动子内。研究者通过对切点序列的比较分析发现, 过剩的结合能的补偿和替代是ZFN发生脱靶反应的一个重要机制。设计过长的识别序列, 会导致过剩的结合能, 因此未必能够带来剪切专一性, 这一观点为设计高专一性的ZFN提出了一个新的思路。

### 参考文献 (References)

- 1 黄 芳, 郭礼和. 体内原位基因组编辑—介绍一种新的基因治疗技术. 中国细胞生物学学报 2011; 33(9): 1062-5.
- 2 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.
- 3 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.