植物学通报 2003，20（2）：137～143
Chinese Bulletin of Botany

基因定点整合技术及其在苔藓研究中的进展①

王先平 胡 勇 何奕昆
（首都师范大学生物系 北京 100037）

摘要 基因定点整合技术是20世纪80年代后兴起的一种分子生物学技术，在精细研究基因功能，消除转基因沉默，基因治疗等有重要意义。基因定点整合技术是功能基因组学研究的重要手段。目前关于基因定点整合技术在酵母和鼠胚胎干细胞中的应用已经很成熟，高等植物由于同源重组频率较低而限制了它的应用，但小立碗藓（*Physcomitrella patens*）基因同源重组频率较高，基因定点整合技术得到了成功应用，可望成为一种新的分子生物学的模式植物。本文针对基因定点整合的原理、技术路线及进展作一综述。

关键词 基因定点整合，同源重组，小立碗藓

The Research of Gene Targeting in Moss

WANG Xian-Ping HU Yong HE Yi-Kun
（*Biology of Department*, *Capital Normal University*, Beijing 100037）

Abstract Gene targeting is a molecular technique that has been exited since 1980'. It can be applied in researching the gene function, removing gene silence and gene therapy. Gene targeting is an important method in functional genetics. Now, it has been successfully used in yeast and mouse embryonic stem cells (ES). But in plants it has been limited because of lower recombination frequency. The recombination frequency in *Physcomitrella patens* is high, so it is expected to be a new molecular biological model plant.

Key words Gene targeting, Homologous recombination, *Physcomitrella patens*

随着许多模式生物大规模基因组测序的完成，研究重点必然会转移到对这些基因的功能分析，即由"基因组学"研究阶段进入"功能基因组学"研究阶段。目前，在研究基因功能方面出现许多新型的技术方法。如DNA微列阵、基因标签、基因诱捕、基因（蛋白质）芯片以及与质谱仪联用解析系统等技术。在众多方法中，"基因定点整合技术"仍是一种研究基因功能的有效方法。它已经广泛的应用在酵母（*Saccharomyces cerevisiae*）和鼠的胚胎干细胞中。

1 基因定点整合的遗传学基础

基因定点整合技术又称为基因打靶（gene targeting），是指构建含有同源序列的DNA片段的整合载体，通过各种转化方法，使外源DNA序列与靶序列之间发生同源重组，从而将靶DNA序列定点破坏，通过一系列筛选手段，最终得到定向转化的细胞。所谓同源重组，

①2002年北京市自然科学基金重点项目（5021001）。
http://www.cjoinline.com/doi/10.3724/CBB-t22-070
作者简介：王先平，女，1978年生，首都师范大学生物系研究生。何奕昆，男，1956年生，首都师范大学生物系教授，博士生导师。研究方向：分子细胞遗传学，发表论文40余篇。E-mail：yhe@duke.edu
收稿日期：2002-05-20 接受日期：2002-08-29 责任编辑：孙冬花

是指在细胞内,具有相似的核苷酸排列顺序的 DNA 片段之间,可以根据碱基互补配对原理,相互配对并发生交换。同源重组可分为染色体内同源重组,染色体外同源重组和转染 DNA 分子与基因组染色体同源重组。后者即对某个感兴趣的基因进行修饰,插入一个选择基因,然后通过各种方法将构建好的重组载体转入到细胞中,这样可能通过同源重组将染色体上的靶基因与外源的 DNA 序列发生相互或不相互的交换,而将靶基因置换掉。研究表明在酵母中参与同源重组的酶有 RAD-52,RAD-50/ku-70 等蛋白家族(Paques and Haber, 1999)。而且,在酵母中同源重组是修复 DNA 双链断裂(double strand breaks DSB)的主要方式(Struhl, 1983)。目前,基因定点整合方式有三种解释。

1) 普遍接受的观点是基因定点整合是由于转化 DNA 片段两侧同源序列与靶 DNA 序列之间发生双交换而产生(Rothstein, 1983)(图 1A)。

2) 有证据表明还存在着另外一种整合方式,Leung 等(1997)观察到在酵母的一类转化事件中有扇形突变菌落(sectored colonies)产生,这种现象在修复系统突变的个体中更加明显。用上述观点无法解释这种现象,因为在双交换中外源 DNA 两侧同源序列都应该与靶 DNA 整合在一起,不会产生扇形突变菌落。故此,Leung 等(1997)提出另一种解释,这种现象的发生,是由于外源 DNA 一侧的同源序列与靶 DNA 序列退火,形成一种异源双链的中间体,之后再进行复制。该异源双链结构往往会被细胞内的修复体系删除或纠正,因为体内的错配修复体系更倾向于使用自己本身已有的且没有断裂的 DNA 链,使用外源的 DNA 链的几率仅有 5%。Negritto 等(1997)的实验也支持这个说法,当插入片断与靶位点完全同源,并且只多了一个选择标记基因时,他观察到在 *msh2*(*msh2* 是一个与错配修复有关的基因)突变的酵母株中基因取代频率增加了 40 倍(图 1B)。

3) 外源 DNA 同源序列与靶 DNA 配对并作为引物诱导互补链的合成。实验表明在含有着丝粒的外源 DNA 序列两侧分别加上染色体某一位点的正向序列与反向序列,即构建成等臂染色体片段载体(isochromosome fragmetation vector IsoCFV),将此载体转化酵母,转化后筛选阳性克隆,得到含等臂染色体(isochromosome IsoCF)的转化株,即在转化株的外源转化 DNA 片段两端是同一染色体臂的两个拷贝。这是由于所转化的 DNA 两侧的 3′端作为引物侵入染色体中,引起新链的合成(Morrow *et al*,1997)(图 1C)。

2 基因定点整合技术流程

2.1 载体的构建

基因定点整合中最关键的是载体的构建,依据整合方式不同载体构建的策略不同。载体可以分为插入型载体和替换型载体。插入型载体通常在外源同源序列外显子的上游插入一个正选择标记基因,并在外源同源序列中构建特异的酶切位点。将载体酶切线性化后进行转化,通过一次交换便可以将外源序列及质粒骨架插入到靶序列中,从而导致靶 DNA 序列功能丧失,进而研究该基因的功能。但这种方法的缺点是不能直接区别和筛选外源 DNA 定点整合的细胞克隆和随机插入的细胞克隆,产生的串联重复区有时不稳定,容易形成第二次同源重组。

http://www.sciengine.com/doi/10.3724/CBB-t22-070

另一种载体为置换型载体,目前这类载体是常用的基因定点整合载体。置换型载体的构建和插入型载体构建基本相同,但它不需要在同源序列中构建特异的酶切位点,酶切

位点设计在同源序列之外，不需要在转化之前进行酶切线性化。它是通过与靶 DNA 序列的双交换而将带有选择标记的外源同源序列整合到染色体上，而将靶 DNA 序列置换下来。一般来说用于同源重组的 DNA 片段长度在 30 ~ 45 bp 之间。同时为了达到富集定向重组的转化株，在同源序列外又设计了负选择标记，这样，通过正、负选择标记筛选就可以富集定向重组子了。

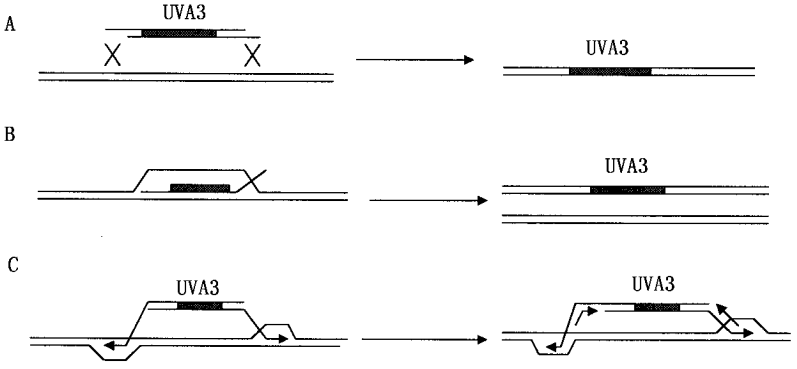

图 1　基因定点整合的三种方式(阴影部分为选择标记基因)
A. 第一种方式，外源同源序列的两端与靶序列发生双交换；B. 第二种方式，外源同源序列的一端侵入靶序列并发生重组；C. 第三种方式，外源同源序列的两端同时与靶序列退火，并作为引物，引起新链的合成。

Fig.1　The models of gene targeting (the shaded indicate the selectable marker gene)
A. The first model suggested the gene targeting is a double-crossover event; B. The second model showed gene targeting could occur by the assimilation of one strand of the transformed DNA; C. The third model suggested the gene targeting is envisions of the 3′-ends of the transformed fragment as a primer for extensive DNA synthesis.

2.2　转化

在载体构建完成之后，通过各种方法将载体转化进细胞中，植物中常采用的转化方法有农杆菌介导的转化和 PEG 介导的原生质体的转化等。动物细胞中的转化方法有显微注射，脉冲电导等。

2.3　筛选

通过在培养基中添加合适的抗生素，转化的植株由于获得抗性标记可以正常生长，非转化的植株则无法生长。由于高等植物的同源重组效率较低，从而设计了正、负选择抗性标记，这样，那些非定点整合的植株由于是载体随机整合的，负选择标记也插入到基因组中，因而在含有负选择抗性标记的培养基上产生致死效应，从而达到富集定向转化株的目的。常用的正选择标记基因有 *neo*，*hprt* 等，负选择标记基因有 *codA* 等。

2.4　鉴定

一般通过 PCR 或采用 southern blot 来检测是否进行了基因打靶，PCR 的一条引物设计在外源同源序列的一端，另一条引物设计在紧靠靶 DNA 的染色体 DNA 上。通过扩增，如果有特异性的区带被扩增出来则说明成功地进行了基因打靶。采用 southern blot 方法，将转化的植株全基因组进行酶解，再通过探针杂交，从而得到定向的转化株。

https://www.sciengine.com/doi/10.3724/CBB-t22-070

3 基因定点整合技术在苔藓研究中的进展

目前植物中成功利用基因定点整合技术的有烟草、拟南芥及苔藓植物小立碗藓等。但是烟草基因打靶的频率仅为 10^{-6}(是指每个活细胞中基因打靶的频率)(Paszkowski *et al*,1988),而在拟南芥中最高时也只有 10^{-4}(Kempin *et al*,1997),可知在植物中基因打靶仍然很难以应用。推测其原因可能是高等植物在修复断裂双链的过程中倾向于非同源重组。然而,Schaefer 和 Zryd (1997)报道,小立碗藓基因打靶的频率在 $10^{-3} \sim 10^{-4}$,同源重组率比高等植物高出 100 ~ 1 000 倍,使基因打靶技术在植物功能基因研究中成为可能。

苔藓植物是一种配子体占主体,而孢子体寄生在配子体上的隐花植物。它的配子体有明显的两种发育形态,原丝体和茎叶体。原丝体是由单列细胞排列形成的丝状体,茎叶体则具有明显的三维结构。苔藓植物与高等植物具有相同的激素,对环境的反应也相似,但其发育模式相对简单,是研究植物的形态建成,向光性和极性的极好材料(Reski, 1998a)。目前常用来研究的苔藓植物主要有葫芦藓(*Funaria hygrometrica*)、角齿藓(*Cerafodon purpureus*)、小立碗藓(*Physcomitrella patens*)和墙藓(*Tortula ruralis*)。

Kammerer 和 Cove (1996)首次在小立碗藓中实现了基因的同源重组。他们首先将含 *hpt*, *amp* 基因的质粒载体通过 PEG 介导的方法转化小立碗藓原生质体,通过筛选得到单拷贝的转化株。之后用含 *amp nptII* 基因的质粒再次转化此转化株,用潮霉素和 G418 进行抗性筛选,最终得到含有 *hpt*, *npt* Ⅱ 基因的转化株。遗传学分析表明,第二次转化的转化效率较第一次的转化效率高出 1 倍,而且 87% 的被检测植株中两个抗性标记基因 *hpt* 和 *npt* Ⅱ 之间存在着明显的连锁关系。这些证据表明第二个质粒载体插入到第一个载体染色体位置的附近,由此得到初步的关于小立碗藓同源重组的证据。之后根据小立碗藓的基因序列构建重组载体进行转化。分析其结果发现在小立碗藓中基因打靶效率高达 90%(指的是基因打靶成功的植株占总转化株的比率)(Schaefer and Zryd ,1997),与酵母的基因打靶频率相当(酵母的基因打靶频率为 95%),这为基因打靶技术在植物中的应用开启了一扇门。

为何在小立碗藓有如此之高的同源重组效率?目前有三种解释(Schaefer, 2001)。首先,苔藓是单倍体占世代周期的主体,在其体内可能存在着某些特殊的生理特征使其在 DNA 断裂双链的修复过程中倾向于同源重组。对不同生物基因打靶频率进行比较,发现基因打靶频率大于 1% 的生物多集中在低等或单倍体真核生物,如原生动物、酵母、一些绿藻和真菌。推测在单倍体生物基因组中,为了保持基因组的完整性,不得不以同源重组的方式修复 DNA 断裂双链。第二种解释认为高效的同源重组效率与细胞周期有关 (Reski, 1998b)。对小立碗藓基因整合研究发现,成功进行了基因打靶的小立碗藓原生质体大多数集中在 G2/M 期。但也有实验表明基因打靶效率与细胞周期之间没有必然的关系,在酵母中,除了对由 HO 限制性内切酶诱导的 DNA 断裂双链修复在 G1 期可提高 100 倍之外,迄今为止没有发现细胞周期与基因打靶效率之间存在必然关系(Paques and Haber, 1999)。第三种解释认为与转化方法有关。在苔藓中 PEG 介导的转化方法似乎对基因打靶特别有效。

http://www.cnki.com.cn

最近在小立碗藓中分离到两个高度同源的重组相关基因 *RAD51A*, *RAD51B* (Mark-

mann-Mulisch *et al*，2002），令人惊奇的是这两个基因均不含有内含子，而含有内含子是小立碗藓属基因的一个特点，推测可能在其体内存在着不同的重组相关机理，但尚无明显的证据。

Strepp 等（1998）成功地敲除了小立碗藓的叶绿体分裂基因—*Fts Z*，得到的转化株细胞内只存在一个巨大的叶绿体，而在野生型植株中成熟细胞则含有 50 多个叶绿体。Girod 等（1999）对小立碗藓一个编码 26S 蛋白酶亚基的基因—*mcb*1（multi-ubiqutin chain binding gene）进行敲除，发现小立碗藓的生长受阻，形成异常的茎丝体（caulonemata），不能从单列细胞的原丝体向有三维结构的茎叶体转变。可见基因定点重组技术在小立碗藓中得到了有效的应用。

4　提高高等植物同源重组的策略

1）在某个特定位点引入 DNA 双链断裂（DSB）可使同源重组的效率提高。研究表明 DSB 是真核生物同源重组的一个关键步骤。I-Sce I 编码酵母线粒体的一种核酸内切酶，此酶能特异性地识别 18 bp 的碱基序列。通过 I-Sce I 基因的瞬时表达来诱导产生 DSB 可使烟草的同源重组效率增加到 $2.2 \sim 18.3 \times 10^{-3}$（Puchta *et al*，1996）。然而由于 I-Sce I 限制性核酸内切酶所识别的特异的核苷酸序列在基因组中分布较少，因而这种策略在实际应用上不太可行。

2）通过表达参与原核生物同源重组的酶来提高同源重组效率。将大肠杆菌参与同源重组的基因 *RecA* 转入烟草使烟草染色体内同源重组率提高了一倍，使 DSB 的修复能力提高了 33.3 倍，但并没有使得基因打靶的频率提高（Reiss *et al*，2000）。此外，Ruv C 参与大肠杆菌同源重组的中间体（Holliday junction）的形成。发现 Ruv C 的过表达使得烟草的转化 DNA 分子与染色体同源重组频率提高 56 倍，染色体间同源重组频率提高 12 倍以及染色体内的同源重组频率提高 11 倍（Shalev *et al*，1999）。但由于与宿主蛋白复杂的相互作用影响到宿主体内某些同源重组过程，很难预料外源同源重组蛋白在植物同源重组中的作用效果。

3）通过正、负选择富集转化株。因为高等植物基因打靶的频率很低，常在构建载体上采用正、负选择标记的方法来富集重组子。Thykjaer 等（1997）在载体的外源同源区内插入正选择标记基因 *NPT* Ⅱ，在同源区侧翼加上负选择标记基因 *codA* 来构建正负选择标记载体，*codA* 是编码胞嘧啶脱苷酶，此酶可使 5-氟胞嘧啶变为 5-氟尿嘧啶，从而抑制胸腺嘧啶的合成，最终导致植株的死亡。通过 G418 和 5-氟胞嘧啶来筛选重组子，使拟南芥筛选效率提高 1 000 倍。

4）调节内源基因表达，提高同源重组的效率。通过抑制参与非同源重组基因或过表达参与同源重组的基因来提高同源重组的效率。MIM 蛋白的过表达可以使拟南芥的同源重组效率提高（Hanin *et al*，2000）。Gherbi 等（2001）通过 RAD50 T-DNA 插入突变使得拟南芥染色体内同源重组效率增加了三倍。

5）构建有效的基因打靶载体。引入重组酶体系 Cre/lox。Cre 重组酶能催化两个 *lox* 位点发生特异性的重组。在载体的外源同源序列前设置 *lox* 位点，Cre 重组酶就可以催化载体上的 *lox* 位点与染色体上的 *lox* 位点发生特异性重组，从而将外源的 DNA 序列整合

到靶位点上。但由于 Cre 重组酶不可能很快失活,导致染色体上的两个 *lox* 位点再次发生重组而将已整合在上面的外源 DNA 序列又从染色体上脱落。有几种方法可以提高 Cre 酶的重组效率。Sauer 和 Henderson (1990)利用启动子融合的策略,将质粒载体上的 *lox* 位点连接在无启动子的 *tk* 基因上游,而在染色体靶位点上 *lox* 置于启动子 3′端,这样只有 *lox* 位点特异性地整合在靶位点上才可使 *tk* 基因表达活性,抗 HAT。Fukushige 和 Sauer (1992)则只是将质粒载体上的 *lox* 位点置于起始密码子 ATG 下游,CHO 细胞染色体上的 *lox* 位于无起始密码 ATG 的 *neo* 的 5′端,通过特异性的重组才可使启动区-ATG-*lox* 位点与 ATGless-*neo* 形成融合的 *lox-neo* 基因,抗 G418。Albert 分别将 *lox* 位点左侧 13 bp 的反向重复序列与右侧的 13 bp 的反向重复序列突变,形成 LE 突变 *lox* 位点(LE mutant lox site)与 RE 突变 *lox* 位点(RE mutant lox site),通过 Cre 酶的催化形成一个野生型的 *lox* 位点和一个 LE + RE 的突变 *lox* 位点。这样 Cre 酶就很难识别后者,从而形成稳定的整合体(Albert *et al*,1995)。Corneille 等(2001)成功地利用 Cre/lox 重组体系对玉米质体的负选择标记基因 *codA* 进行删除,获得了无选择标记的转化植株,Cre/lox 重组体系可能为获得无选择标记的转基因植物提供了一种方法。

基因定点整合技术自问世以来,在基因功能研究方面发挥了重要作用,虽然目前它在植物中的利用仍存在困难,但随着人们对基因同源重组原理有更深入了解,基因定点整合技术会得到相应的改进和提高。同时由于许多平行技术的使用,使得基因定点整合技术在基因功能研究、后基因组学、蛋白质工程、植物的品种改造以及消除转基因沉默等方面发挥更大的作用。

参 考 文 献

Albert H, Dale E C, Lee E, Ow D W, 1995. Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J*, **7**: 649 ~ 659

Corneille S, Lutz K, Svab Z, Maliga P, 2001. Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system. *Plant J*, **27**: 171 ~ 8

Fukushige S, Sauer B, 1992. Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *Proc Natl Acad Sci U S A*, **89**: 7905 ~ 7909

Gherbi H, Gallego M E, Jalut N, Lucht J M, Hohn B, White C I, 2001. Homologous recombination in planta is stimulated in the absence of Rad50. *EMBO Rep*, **2**: 287 ~ 291

Girod P A, Fu H, Zryd J P, Vierstra R D, 1999. Multiubiquitin chain binding subunit MCB1 (RPN10) of the 26S proteasome is essential for developmental progression in *Physcomitrella patens*. *Plant Cell*, **11**: 1457 ~ 72.

Hanin M, Mengiste T, Bogucki A, Paszkowski J, 2000. Elevated levels of intrachromosomal homologous recombination in Arabidopisis overexpressing the MIM gene. *Plant J*, **24**: 183 ~ 189

Kammerer W, Cove D J, 1996. Genetic analysis of the result of retransformation of transgenic lines of the moss, *Physcomitrella patens*. *Mol Gen Genet*, **250**: 380 ~ 382

Kempin S A, Liljegre S J, Block L M, Roundsley S D, Yanofsky M F, Lam E, 1997. Targeted disruption in *Arabidopsis*. *Nature*, **389**: 802 ~ 803

Leung W, Malkova A, Haber J E, 1997. Gene targeting by linear duplex DNA frequently occurs by assimilation of a single strand that is subject to preferential mismatch correction. *Proc Natl Acad Sci. USA*, **94**: 6851 ~ 6856

Markmann-Mulisch U, Hadi M Z, Koepchen K, Alonso J C, Russo Vincenzo E A, Schell J, Reiss B, 2002. The organization of physcomitrella patens RAD51 gene is unique among eukaryotic organisms. *Proc Natl Acad Sci. USA*, **99**: 2959 ~ 2964

Morrow D M, Connelly C, Hieter P, 1997. "Break copy" duplication: a model for chromosome fragment formation in *Saccharomyces cerevisiae*. *Genetics*, **147**: 371 ~ 382

Negritto M T, wu X, Kuo T, Chu S, Bailis A M, 1997. Influence of DNA sequence identify on efficiency of targeted gene replace-

https://www.scientiasinica.com/cn/SIAS/372/EC0E9E2E2E2E7C0.pdf

ment. *Mol Cell boil*, **17**: 278 ~ 286

Paszkowski J, Baur M, Bogucki A, Potrykus I, 1988. Gene targeting in plants. *EMBO J*, **7**: 4021 ~ 4026

Paques F, Haber J E, 1999. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*, **63**: 349 ~ 404

Puchta H, Dujon B, Hohn B, 1996. Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci USA*, **93**: 5055 ~ 5060

Reiss B, Schubert I, Kopchen K, Wendeler E, Schell J, Puchta H, 2000. RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting in pants transformed by *Agrobacterium*. *Proc Natl Acad Sci USA*, **97**: 3358 ~ 3363

Reski R, 1998a. Development. Genetics and molecular biology of mosses. *Bot Acta*, **111**: 1 ~ 15

Reski R, 1998b. *Physcomitrella* and *Arabidopsis*: the david and goliath of reverse genetics. *Trends Plant Sci*, **3**: 209 ~ 210

Rothstein R J, 1983. One-step gene disruption in yeast. *Methods Enzymol*, **101**: 202 ~ 211

Sauer B, Henderson N, 1990. Targeted insertion of exogenous DNA into the eukaryotic genome by the Crerecombinase. *New Biol*, **2**: 441 ~ 449

Schaefer D G, Zryd J P, 1997. Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J*, **11**: 1195 ~ 1206

Schaefer D G, 2001. Gene targeting in physcomitrella patens. *Curr Opin Plant Biol*, **4**: 143 ~ 150

Shalev G, Sitrit Y, Avivi-Ragolski N, Lichtenstein C, Levy A A, 1999. Stimulation of homologous recombination in plants by expression of the bacterial resolvase RuvC. *Proc Natl Acad Sci USA*, **96**: 7398 ~ 7402

Strepp R, Scholz S, Kruse S, Speth V, Reski R, 1998. Plant nuclear gene knockout reveals a role in plastid division for homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc Natl Acad Sci USA*, **95**: 4368 ~ 4373

Struhl K, 1983. The new yeast genetics. *Nature*, **305**: 391 ~ 397

Thykjaer T, Finnemann J, Schausser L, Christensen L, Poulsen C, Stougaard J, 1997. Gene targeting approaches using positive-negative selection and large flanking regions. *Plant Mol Biol*, **35**: 523 ~ 530

https://www.sciengine.com/doi/10.3724/CBB-t22-070