

利用修饰的随机引物构建非洲爪蟾酵母双杂交定向 cDNA 文库

武景阳^{1,2}, 李朝翠¹, 孔清华^{1,2}, 毛炳宇^{1,*}

(1. 中国科学院昆明动物研究所 遗传资源与进化国家重点实验室, 云南 昆明 650223; 2. 中国科学院研究生院, 北京 100049)

摘要: 描述了一种新的构建 cDNA 文库的方法, 其中用来合成 cDNA 第一链的随机引物 5' -端被加上碱基 d (AC), 在 cDNA 双链合成后所添加的连接头的末端含有碱基 d (GTCG)。在 cDNA 双链 3' -端, 连接头连接上后会形成一个完整的 *Sall* 酶切位点 d (GTCGAC), 而在 5' -端几乎不会形成 *Sall* 位点。在 *Sall* 酶切后利用形成的 3' -粘性末端与 5' -*EcoRI* 粘性末端一起将 cDNA 双链定向导入线性化的质粒载体, 再通过转化细菌获得 cDNA 文库。利用此方法构建了一个不同发育时期非洲爪蟾胚胎酵母双杂交 cDNA 文库。检测了文库中空载体的比例, 插入片段大小和不同基因的表达水平几个指标, 都符合预期, 但是同时发现定位于 mRNA 的 3' -端的插入片段比例比较低。这种倾向性符合文献报道, 应该是实验系统的倾向性, 不影响进一步的酵母双杂交实验。这些数据证明成功地构建了定向非洲爪蟾酵母双杂交 cDNA 文库。

关键词: 非洲爪蟾; 酵母双杂交; 随机引物; 定向 cDNA 文库

中图分类号: Q343.1; Q78; Q959.5 文献标识码: A 文章编号: 0254-5853-(2008)04-0368-05

Construction and Analysis of A Directional *Xenopus laevis* Embryonic cDNA Library for Yeast Two-hybrid Using Modified Random Primers

WU Jing-yang^{1,2}, LI Chao-cui¹, KONG Qing-hua^{1,2}, MAO Bing-yu^{1,*}

(1. State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming 650223, China;

2. Graduate University of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Here we describe a new strategy for cDNA library construction, in which the random primers directing the first strand cDNA synthesis contain additional d (AC) at their 5'-ends, and the linkers which will be added to the double strand cDNA contain d(GTCG) at the 5'-ends. When the linker is added to the 3'-end of the cDNA, a complete *Sall* site d(GTCGAC) will form at the 3'-end but not at the 5'-end of the cDNA fragment. After *Sall* digestion, the 3'-cohesive and the pre-existing 5'-*EcoRI* cohesive ends are used to introduce the double stranded cDNA into the linearized plasmid vectors. The ligation products were used to transform *E.coli* to produce a cDNA library. Using this method, we constructed a directional *Xenopus laevis* embryonic cDNA library for yeast two-hybrid. We checked the ratio of empty vectors, the size of inserts and existence of several genes, the results showed that cDNA library construction was successful.

Key words: *Xenopus laevis*; Yeast two-hybrid; Random primer; Directional cDNA library

构建 cDNA 文库的第一步是逆转录 mRNA 成为它的 cDNA 拷贝。其中起催化作用的逆转录酶需要一个短片段的 DNA 或者 RNA 作为引物来添加核苷酸。目前构建 cDNA 文库所用的引物有两种, 一种也是最常见的是含有限制性酶切位点的 Oligo (dT) 的引物, 另外一种是 6 聚或 9 聚脱氧核糖核

苷酸 (dN₆ 或者 dN₉) 引物 (Sambrook et al, 2001)。用含有限制性酶切位点的 Oligo (dT) 的引物所构建的 cDNA 文库是单方向的, 同时具有比较好的代表性, 真实地反应了组织中 mRNA 的表达水平。这种方法的缺点是具有序列的倾向性, 所克隆的基因片段往往位于基因的 3'-端, 原因是 mRNA 的二级

收稿日期: 2008-05-14; 接受日期: 2008-06-07

基金项目: 国家杰出青年科学基金 (30425011); 国家自然科学基金重点项目 (30530380)

*通讯作者 (Corresponding author), E-mail: mao@mail.kiz.ac.cn

结构和逆转录酶易于从长片段上脱离 (Kimmel et al, 1987)。如果基因含有长的 3'-UTR, cDNA 往往不能包含基因的开放阅读框。用随机引物构建的 cDNA 文库也具有比较好的代表性, 一定的条件下 cDNA 位于基因的 5'-端、3'-端或中间的可能性是差不多的 (Haymerle et al, 1986)。它的缺点是插入的 cDNA 片段具有两个方向, 尤其是对于表达性的文库, 约 50% 的 cDNA 克隆含有反向插入片段所以是没有用的。尽管有各种各样的克隆全长 cDNA 的技术, 比如 SMART, 但是对于酵母双杂交 cDNA 文库来说, 全长未必有太大的意义。酵母双杂交用于探测蛋白与蛋白之间的相互的作用, 这是以蛋白的结构域 (domain) 或者结构元件 (motif) 为基础的, 一般长度从几个到几百个氨基酸。含有太长的 cDNA 插入片段的载体难以用常规的醋酸锂化学转化方法转化进入酵母, 或者所转录和翻译出来的大的蛋白因为对酵母有毒性而使酵母不能生长, 或者大的蛋白由于缺少高等真核细胞里的稳定性修饰最后被降解。常规的商业化的酵母双杂交 cDNA 文库的插入片段长度一般在 0.5—3 kb 之间, 峰值在 1.5kb 左右。

这里我们介绍一种以修饰的随机引物合成第一链为基础的单向非洲爪蟾酵母双杂交 cDNA 文库的构建方法, 克服了经典的随机引物文库中插入片段方向上的随机性, 插入的 cDNA 是单方向的, 使得有效克隆数量大大的增加, 并且能够很好的覆盖基因的读码框, 这对于探测蛋白-蛋白的相互作用非常有利。

1 研究材料与方法

1.1 材料

实验动物成体非洲爪蟾购自美国 Nasco 公司。爪蟾的培养、体外受精和胚胎培养按照 Gawantka et al (1995) 的方法进行。爪蟾胚胎发育的时期按照 Nieuwkoop & Faber (1967) 所描述的来确定。

1.2 总 RNA 和 mRNA 的提取

分别取不同时期的胚胎, 包括 10—11 期、20—22 期、25—28 期、38 期和 44 期, 合并在一起后加入 TRIzol 试剂 (Invitrogen) 并进行匀浆, 按照产品说明书操作直到得到总 RNA 并且溶于 DEPC 处理的纯净水中。然后用 Oligotex mRNA Mini Kit (Qiagen) 从总 RNA 中提取 mRNA。

1.3 双链 cDNA 的合成

首先进行 cDNA 第一链的合成, 将修饰的随机引物 5'-pACNNNNNN-3' 与定量后的 mRNA 按照引物/mRNA 质量比 0.25 混合 (Morris et al, 1995), 在 70℃ 放置 10 min 后置于冰上, 然后分别加入 dATP、dGTP、dTTP (Promega) 和 dm⁵CTP (Fermentas), 5×first strand buffer、0.1 mol/L DTT 和逆转录酶 SuperScriptII RT (Invitrogen), 混合好后先在 25℃ 放置 10 min, 然后升温到 37℃ 进行逆转录反应 40 min。第一链的合成结束后, 在反应液中加入 5×second strand buffer、dNTP、*E. coli* DNA ligase、*E. coli* DNA polymerase I 和 *E. coli* RNaseH (Invitrogen), 混合好后在 16℃ 放置 2h, 然后加入 T4 DNA polymerase (Invitrogen) 在 16℃ 继续反应 5 min。第二链的合成结束后, 加入 0.5 mol/L EDTA 来终止反应。产物用酚、氯仿抽提后, 加入 100% 乙醇、3 mol/L pH5.2 NaAc 和糖原进行沉淀。沉淀然后溶于纯净水, 并进行定量。

1.4 加接头, 限制性酶切与目的 cDNA 片段获得
接头共有 3 套: A、B 和 C, 分别比前一个在中间的序列上多出一个核苷酸 (如下划线所示)。接头单链的序列如下:

A 5'-AATTCCCACGCGTGTCG-3',
A' 5'-pCGACACGCGTGGG-3'
B 5'-AATTCCCAACGCGTGTCG-3',
B' 5'-pCGACACGCGTTGGG-3'
C 5'-AATTCCCAAACGCGTGTCG-3',
C' 5'-pCGACACGCGTTTGGG-3'

在 10×DNA oligo annealing buffer (100mmol/L Tris HCl, pH7.5, 1 mol/L NaCl, 10mmol/L EDTA) 存在的条件下, A 和 A', B 和 B' 以及 C 和 C' 以终浓度 100 μmol/L 进行混合。在 95℃ 变性 5 min, 然后在 65℃ 放置 20 min, 自然冷却。然后将 3 套接头等比例混合, 同双链 cDNA 按照摩尔比 120:1 的比例在 16℃ 进行连接。所用的连接酶是 Invitrogen 的 T4 DNA Ligase, 因为它的缓冲液中含有 PEG8000, 便于于平端的连接。连接产物用酚、氯仿抽提后, 加入 100% 乙醇, 3 mol/L pH5.2 NaAc 和糖原进行沉淀。沉淀溶于纯净水中后, 加入限制性酶 *Sall* (Fermentas) 进行酶切反应。反应结束后, 将反应液在含有 EB 的 1% 琼脂糖凝胶, 0.5×TBE 缓冲液中进行电泳, 回收大小在 500—3 000 bp 的 DNA, 然后用 MiniElute Gel Extraction Kit (Qiagen) 纯化双链 cDNA 片段, 并进行定量。

1.5 cDNA 文库的构建, 文库 DNA 的提取与克隆鉴定

将纯化的 cDNA 片段与用 *EcoRI* 和 *XhoI* 双酶切 (*XhoI* 与 *SalI* 所切出的粘性末端互相匹配) 的线性化载体 pGADT7 (Clontech), 按照摩尔比 7:1 的比例进行连接。连接产物用酚、氯仿抽提后沉淀, 溶于 15 μL 纯净水中, 然后转化 XL1-Blue Electroporation-Competent Cells (Stratagene)。转化后的细菌按照每块培养皿 150,000 克隆的密度涂布于 15cm 直径的培养皿中, 在 37 $^{\circ}\text{C}$ 恒温条件下培养 14h。然后用含有 25% 甘油的 LB 溶液将克隆刮下来, 用 Qiagen plasmid Midi kit 提取质粒 DNA, 将所提取的所有的 DNA 合并, 混匀后以小份分装并且冰冻储存。

1.6 cDNA 文库质量的鉴定

挑取 30 个细菌单克隆溶于 10 μL LB 溶液中, 取 1 μL 作为模板进行 PCR。所用的引物是 MATCHMAKER 5' 和 3' AD LD-Insert Screening Amplimer (Clontech), 专门用于扩增 pGADT7 系列载体的插入片段。将阳性的克隆进行测序, 然后与相应的 mRNA 进行序列对比以确定插入片段在基因中的位置。

最后取一定量的文库总 DNA, 通过 PCR 检测是否包含了部分已知基因。我们选择了 *Sox2*、*krox20*、*slug*、*Nkx6.3*、*Brunol2b*、*Brunol4*、*mcpA* 和 *mcpB* 共 8 个基因, 所扩增的片段长度分别为 214、354、262、367、240、233、180 和 360bp。

2 结果与讨论

2.1 总 RNA 与 mRNA 质量的鉴定

在总 RNA 提取后, 取少量的样品进行琼脂糖凝胶电泳, 电泳后在紫外线下观察到 18S 和 28S RNA 的条带清晰可见, 无降解现象。在进一步纯化得到 mRNA 后, 取少量的样品进行 RT-PCR 来检测一些基因的表达, 显示 mRNA 的代表性良好。

2.2 双链 cDNA 的合成

为了克服常规随机引物构建文库所导致的插入片段方向上的随机性, 我们在随机引物的 5'-端加上 *SalI* 酶切位点的最后两位, 所以合成的随机引物序列为 5'-pACNNNNNN-3' (Morris et al, 1995)。没有加上完整的酶切位点甚至更多的保护性碱基的原因, 是考虑到太多的碱基可能会产生序列上的倾向性。酶切位点的另外 4 个碱基, 是在连接头双

链的后四位, 只有当连接头与随机引物连接时才会形成一个完整的 *SalI* 酶切位点。连接头在 5'-端将保持一个粘性的 *EcoRI* 位点。双链 cDNA 在用 *SalI* 酶消化后, 将具有两个不同的粘性末端, 可以定向导入线性化的载体中。为了防止双链 cDNA 片段内部具有 *SalI* 位点而被切开, 我们用 dm^5CTP 来代替 dCTP, 产生甲基化的酶切位点 GTm^5CGAC , 从而不能被 *SalI* 切开 (图 1)。

与 Oligo (dT) 引物为基础构建文库的方法不同, 在使用随机引物的情况下, 可以通过调节随机引物和 RNA 模板的比例来控制所产生的双链 cDNA 的大小。我们构建文库的目的是用于筛选蛋白-蛋白相互作用, 这是以蛋白的结构域或者结构元件为基础的, 不需要 cDNA 全长。我们用 36 期的胚胎总 RNA 为模板, 加入与模板质量比分别 0.1、0.15、0.25、0.5 和 1 的随机引物来合成双链 DNA。我们发现随机引物和 RNA 模板的质量比小于 0.25 时, 产生的双链 cDNA 的量比较少; 当质量比大于 0.25 的时候, 产生的 cDNA 的量比较多, 但合成的片段比较小; 比例在 0.25 的时候, 合成的 cDNA 片段大小在 500—3000bp 之间, 峰值约在 1kb 左右, 产生的 cDNA 的量也合适, 符合我们的需要 (图 2A)。我们利用 2.5 μg 起始 mRNA, 采用的随机引物与 mRNA 的质量比为 0.25, 合成的双链 cDNA 电泳结果如图 2B 所示。我们同时使用了相差一个碱基的 3 套连接头, 以期获得尽量多的含有正确读码框的克隆。在加上连接头后, 我们采用胶回收的方法, 获得大小在 500—3000bp 之间的 cDNA 片段, 然后用 Qiagen 的 MinElute gel extraction kit 进行纯化, 获得双链 cDNA 约 1.2 μg 。

2.3 cDNA 文库的构建与质量的鉴定

纯化的双链 cDNA 与 *EcoRI/XhoI* 双酶切的 pGADT7 载体进行连接。连接产物纯化后, 先取少量来确定电转化细菌的效率, 然后全部进行转化, 按照每块培养皿 (直径 15cm) 约 150,000 个克隆涂平板, 总共获得约 7×10^6 个独立克隆, 每块培养皿提取出约 150 μg DNA, 我们最后得到了约 7mg 文库总 DNA。

对于 cDNA 文库质量的鉴定, 首先是检测不含有插入片段的空克隆比例。我们挑出 30 个单克隆进行 PCR, 发现有两个不含有插入片段, 空克隆比例为 6.67%, 小于预期的 10%。其它的阳性克隆所含有插入片段, 大小基本上都在 500—1000bp 之间,

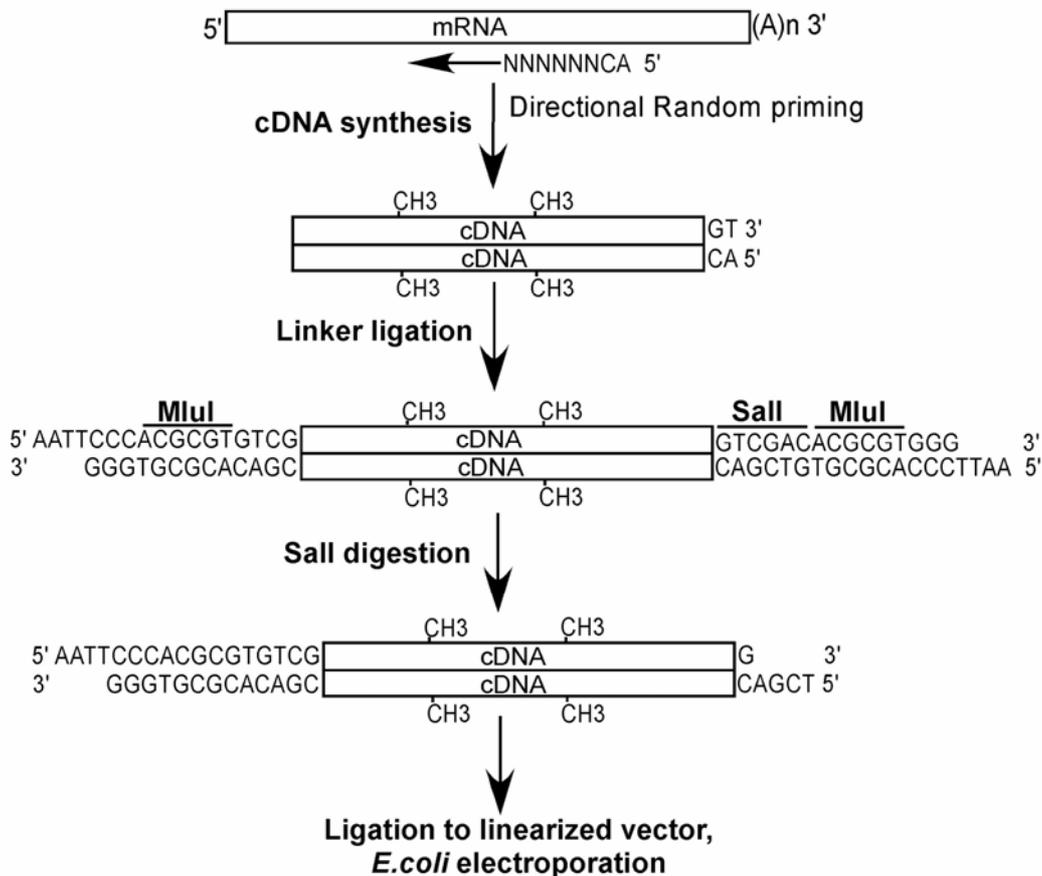


图 1 利用修饰的随机引物构建定向 cDNA 文库的设计思路与流程

Fig. 1 Strategies for the construction of directional cDNA libraries using modified random primers

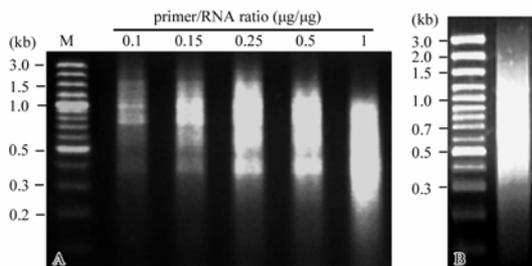


图 2 双链 cDNA 的合成

Fig. 2 Synthesis of the double-stranded cDNA

A: 不同随机引物/RNA 比例所获得的双链 cDNA 的片段大小与产量有所不同; B: 实际实验中所获的双链 cDNA 的电泳图。

A: amount and fragment size of the cDNA products produced using different random primer/template ratios; B: the cDNA products for the library construction.

M, DNA 分子量标记 (DNA molecular weight marker); kb, 千碱基对 (kilobase pairs)。

也符合预期值 (图 3)。为了检测文库中插入序列的方向性及代表性, 上述 28 个阳性克隆被测序, 其中 20 个克隆检索到了对应的全长 mRNA, 其插入

方向都是正确的。我们分析了克隆片段在相应 mRNA 中的位置。如果将全长 mRNA 平均分为 3 份, 插入片段位于 mRNA 的 5'-端、中间和 3'-端的比例为 14 : 14 : 5 (包括重叠的部分)。如果仅仅计算 mRNA 长于 3 kb 的 7 个克隆, 比例为 4 : 4 : 1 (包括重叠的部分)。表明我们构建的 cDNA 文库含有的插入片段位于基因 3'-端的倾向性较低。Morris et al(1995)所报告的插入片段位于基因的 5'-和 3'-端的百分比分别为 62%和 29%, 与我们的结果相似。这种倾向性的原因尚不明确。由于我们建库的目的是检测蛋白之间的相互作用 (用于酵母双杂交), 这种序列上的倾向性应该对将来筛选文库没有太大的影响。首先, 基因尤其是较长基因的 3'-端通常是较长的非翻译区而不是编码蛋白的读码框; 其次可以通过增加文库筛选的克隆数量来弥补这种倾向性。

为了检测 cDNA 文库是否具有原始总 RNA 材料的代表性, 我们选择了 *Sox2*、*krox20*、*slug*、*Nkx6.3*、

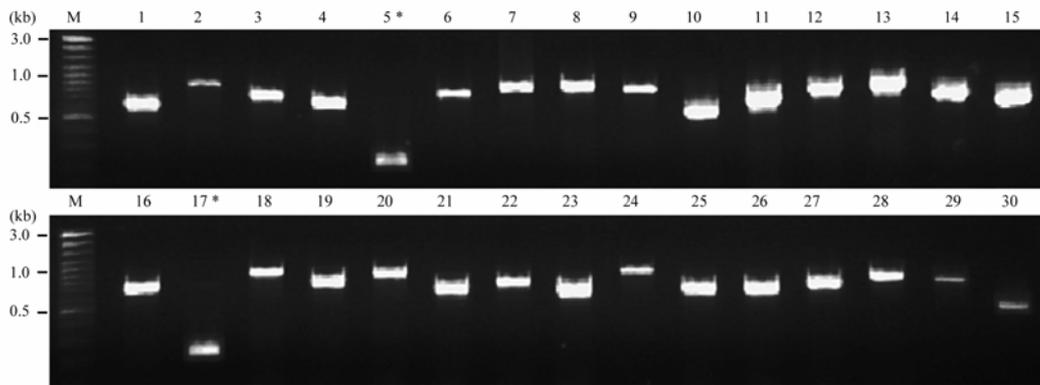


图 3 cDNA 文库空载率和插入片段大小的 PCR 检测
Fig. 3 Detection of the ratio of empty clones and the size of the inserts by PCR

*. 含有空载体的克隆 (clones with empty vectors)。

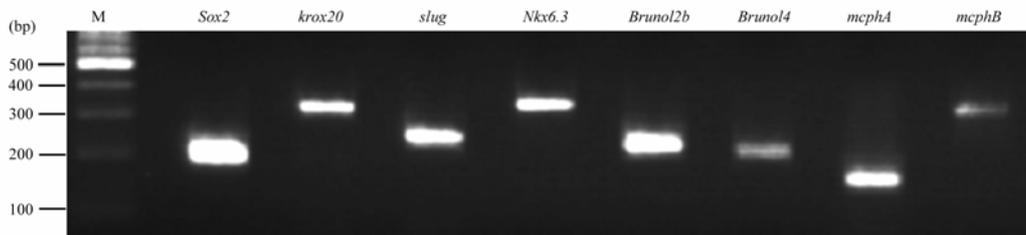


图 4 从 cDNA 文库中扩增不同基因片段的的结果
Fig. 4 Detection of the presence of 8 known genes in the cDNA library
M, DNA 分子量标记(DNA molecular weight marker); bp, 碱基对(base pairs)。

Brunol2b、*Brunol4*、*mcphA* 和 *mcphB* 8 个基因进行 PCR 检测。这 8 个基因表达量有高有低，其中 *Sox2* 和 *Brunol2b* 表达水平比较高，*Brunol4* 和 *mcphA* 比较低，*mcphB* 尤其低。*Brunol4* 的 mRNA 长度有 6kb，所用的引物约位于 5'-端的 1.2kb 处。PCR 的结果表明，即使是表达量很低同时也位于长的 mRNA 的 5'-端的基因片段仍然可以从文库中扩增出来，证明文

库能够真实地反映总 RNA 材料的代表性 (图 4)。

根据以上的分析，我们认为此非洲爪蟾酵母双杂交 cDNA 文库的构建是成功的。这种以随机引物为基础的定向 cDNA 文库的构建方法应该适用于以检测蛋白之间的相互作用为目的的文库构建，比如酵母双杂交和噬菌体展示文库等。

参考文献:

- Gawantka V, Delius H, Hirschfeld K, Blumenstock C, Niehrs C. 1995. Antagonizing the Spemann organizer: role of the homeobox gene *Xvent-1* [J]. *EMBO J*, **14** (24): 6268-6279.
- Haymerle H, Herz J, Bressan GM, Frank R, and Stanley KK. 1986. Efficient construction of cDNA libraries in plasmid expression vectors using an adaptor strategy [J]. *Nucleic Acids Res.* **14** (21): 8615-8624.
- Kimmel AR, Berger SL. 1987. Preparation of cDNA and the generation of cDNA libraries: Overview [J]. *Methods Enzymol*, **152**: 307-316.
- Morris B, Hammer B, Mierendorf R. 1995. A novel strategy for directional cloning of random primed cDNA [J]. *Novagen inNovations Newsletter*, **3**: 1-4.
- Nieuwkoop PD, Faber J. 1967. Normal Table of *Xenopus laevis* (Daudin) [M]. Amsterdam: North-Holland Publishing Co.
- Sambrook J, Russell DW. 2001. Molecular cloning: A Laboratory Manual, 3rd ed [M]. New York: Cold Spring Harbor Laboratory Press.