

细菌转录组分析样品制备方法

Sample Preparation For Bacterial Transcriptome Analysis

武迎春^{1, 2}, 郝光飞^{1*}, 韩东飞^{2*}

¹生命科学与食品工程学院, 河北工程大学, 邯郸, 河北; ²农业环境与可持续发展研究所, 中国农业科学院, 北京

*通讯作者邮箱: handongfei@caas.cn, wuyucool@126.com

摘要: 伴随着全球模式微生物基因组测序计划的进行, 微生物转录组研究在我国取得了快速发展。第二代、第三代测序技术的出现, 更使大规模应用转录组方法解决科学问题成为可能。转录组研究可以从整体水平研究基因功能和基因表达动态, 揭示特定生物学过程和微生物代谢调控的分子机制, 已广泛应用于基础研究、代谢工程和药物研发等领域。其中, 通过细菌转录组研究来揭示生命基本过程, 如生命形成、生物进化、基础代谢、疾病发生、药物靶点等, 成为生物学研究的重要手段。随着 RNA 测序价格的日益低廉, RNA 测序逐渐成为筛选分子生物学后续研究方向的最省时、省力、最经济的方法, 研究者往往使用 RNA 测序来推进项目的进展。然而, 许多新入行的研究者对细菌 RNA-Seq 测序样品准备相关步骤知之甚少, 基于此本文将对细菌转录组样本制备过程进行详细说明。

关键词: 细菌, 转录组, 样品制备, RNA-Seq

材料与试剂

表 1. 实验所需材料及试剂

Table 1. Materials and reagents

产品名称	生产公司	商品型号
50×TAE 缓冲液	aladdin 生物工程有限公司	T197242
DNA Ladder	Promega 生物工程有限公司	Z5300
Agarose	中科瑞泰生物科技有限公司	9012-36-6
RNAiso plus (Takara)	TaKaRa 公司	T9180
MICROB Express Kit (Ambion)	Ambion 公司	AM1905

产品名称	生产公司	商品型号
glycerol	aladdin 生物工程有限公司	G116205
PrimeScript™ 1st strand cDNA Synthesis	TaKaRa 生物工程有限公司	6110A
DNA 1000 LabChip	Caliper 生物科技有限公司	G2938
10 ×PCR Buffer (Mg ²⁺ plus)	TaKaRa 生物工程有限公司	9151A
dNTP Mixture	TaKaRa 生物工程有限公司	4030
T4 RNA Ligase	TaKaRa 生物工程有限公司	2050A
T4 RNA Ligase Buffer	TaKaRa 生物工程有限公司	2050A
0.1% BSA	TaKaRa 生物工程有限公司	2050A
Reverse Transcriptase XL (AMV)	TaKaRa 生物工程有限公司	2621
Glycogen	碧云天生物有限公司	D0812
AMV RT Buffer	BBI 生命科学有限公司	B610020
DreamTaq DNA Polymerase (EP0704)	赛默飞世尔科技有限公司	B300107
SYTO 9 Green	Invitrogen, Eugene, oregon,	s34854
mirVana miRNA Isolation Kit	赛默飞世尔科技有限公司	AM1561
EDC 粉末	BBI 生命科学有限公司	C600433
0.1mol/L MES buffer (pH4.8)	生工生物工程股份有限公司	C506052
DEPC-Treated Water	赛默飞世尔科技有限公司	AM9916
NaOH	BBI 生命科学有限公司	A100173

仪器设备

表 2. 实验所需仪器及设备

Table 2. Instruments and equipments

仪器名称	型号	生产公司
生化恒温摇床器	HZQ-QA	上海精宏实验设备有限公司
生化培养箱	SPX-250	上海一恒科学仪器有限公司
水平电泳槽	DYCP-31DN	北京六一生物科技有限公司
电泳仪	DYY-6D	北京六一生物科技有限公司
微波炉	M1-L213C	美的集团有限公司

仪器名称	型号	生产公司
冰箱	HYC-390	青岛海尔股份有限公司
微量分光光度计	2000c	Nanodrop 科技有限责任公司
台式冷冻离心机	TG16B	Eppendorf 中国有限公司
酶标分析仪	iMark	美国 Bio-Rad (伯乐) 有限公司
Bioanalyzer	2100	安捷伦科技 (中国) 有限公司
PCR 核酸扩增仪	C1000 Touch	美国 Bio-Rad (伯乐) 有限公司
凝胶成像分析系统	GelDoc XR+	美国 Bio-Rad (伯乐) 有限公司
高压型电泳电源	PowerPac HV	美国 Bio-Rad (伯乐) 有限公司
电热恒温水浴锅	HH-8	苏州江东精密仪器有限公司
制冰机	XD50KG	北京长流仪器有限公司
液相色谱	1200	安捷伦科技中国有限公司
紫外分光光度计	JH-2100	上海元析仪器有限公司
超净工作台	MKSD-CJ	无锡一净净化设备有限公司
立式压力蒸汽灭菌器	LS-35LD	上海博迅医疗生物仪器有限公司
震荡仪	SI-0246	海门市其林贝尔仪器制造有限公司
微型离心机	KST110	海门市其林贝尔仪器制造有限公司
扫描仪	LuxScan 10K/A	CapitalBio 仪器设备有限公司
金属浴	TCS10	卡尤迪生物科技有限公司
超声仪	KQ-3200	上海远淮化工科技有限公司
精密电子天平	UW820H	上海菁海仪器有限公司
超低温冰箱	DW-86L959BP	青岛海尔股份有限公司
热循环仪	PTC-220, PTC-225	MJ Research, Watertown, MA
涡旋振荡器	Vortex-Genie 2	奥然科学技术有限公司
超纯水系统	Milli-Q	Millipore 公司
微量移液器	各种量程	吉尔森公司

实验步骤

1、细菌培养与收集。

在相应抗性（目的细菌菌株所含抗性）的培养基中培养单一的目的菌种到对数中后期。须严格控制培养条件。在 4℃ 预冷的离心机中 10000-12000×g 离心 3-5 min 沉降菌体于 50 ml 离心管中。留少量菌液，将菌体重悬（全程需无菌操作）。对于厌氧细菌，除按如上操作外，在培养与菌体收集过程中需严格控制厌氧条件。

2、细菌总 RNA 的提取。

目前，常用的细菌 RNA 提取方法主要有变性法、CTAB 法、SDS-Phenol 苯酚法和热硼酸法^[1]。在商业化的今天，国内外生物试剂公司开发了许多细菌 RNA 提取试剂盒。例如，Takara 的 RNAiso plus；Invitrogen 的 Trizol 总 RNA 提取试剂盒；Purelink™ RNA Mini 试剂盒；上海生工的 UNIQ-10 色谱柱；Omega 的细菌 RNA 试剂盒；HP total RNA Kit 和 Qiagen 公司的 RNeasy Protect Bacteria Mini and Midi Kits 等。能够快速有效地从细胞或培养细胞中提取高质量和高纯度的 RNA。经文献参考对比发现，Takara 公司的 RNAiso plus 总 RNA 提取效果相对较好^[2]。该试剂盒的具体步骤如下：

（1）取新鲜或者 -80℃ 冻存的菌体细胞，约 100 mg，加入 1 mL RNAiso plus 裂解，振荡混匀后室温静置 5 min；

（2）以每毫升 RNAiso 对应 0.2 mL 氯仿的比例加入氯仿，振荡 15 s，室温静置 3 min；

（3）12000×g，4℃，离心 15 min，取上清液至新的 Eppendorf 管中（上层：无色水相为 RNA，中层：白色为 DNA，下层：红色为蛋白）；

（4）在得到的水相中加入等体积异丙醇，混匀后 -20℃ 放置 20~30 min；（提取 dsRNA 时，加入等体积的异丙醇和 3 M NaCl 混合液，异丙醇：3 M NaCl = 1：1）

（5）12000×g，4℃，离心 10 min，移除上清；

（6）加入 1 mL 75% 乙醇，该试剂由 DEPC（焦碳酸二乙酯）和无水乙醇配置，-20℃ 预冷，洗涤沉淀；

(7) $8000\times g$, 4°C , 离心 5 min 去除上清, 干燥 RNA 沉淀;

(8) 用 20~30 μL DEPC 水溶解 RNA;

(<https://www.takarabiomed.com.cn/ProductShow.aspx?m=20141220151857153056&productID=20141226160653343219> 以上为 RNAiso plus (Takara) 试剂盒说明书下载地址。)

注意: 所有的试管、移液器吸头等都要要求无核酸酶级别, 或经 DEPC 处理。

3、总 RNA 质量检测

(1) 总量。微量分光光度计测 260 nm 吸收值计算。

(2) 纯度。微量分光光度计测 260 nm/230 nm (>1.8) 吸收值的比值, 用于评估有机溶剂残留情况; 260 nm/280 nm 吸收值的比值 (1.8-2.1), 用于评估蛋白质污染比例情况^[3]。

(3) 完整性。在进行 bioanalyzer 分析之前, 可先进行琼脂糖凝胶电泳, 对总 RNA 质量进行初步分析。电泳显示有无基因组 DNA 污染, 有无 RNA 降解 (23S rRNA, 16S rRNA 条带); 通过目测 23S rRNA 和 16S rRNA 条带的比例可以初步评价总 RNA 的质量。一般认为 23S rRNA : 16S rRNA ≥ 2 可以初步判定总 RNA 完整性较好^[3]。电泳 RNA 应出现两条清晰的条带 (23S/16S rRNA), 有时会有第三条条带 (5S rRNA), 最上方不可出现 DNA 污染条带, 23S/16S 比例应约为 2: 1。具体评价标准见结果示例分析。若结果较为理想即可进行下一步分析。

以 Agilent Bioanalyzer 进行毛细管电泳 (Capillary Electrophoresis), 并以软件的 RIN (RNA Integrity Number) 分数评估, 10 为 RNA 完整性最好, 0 为最差, 推荐使用 RIN 值在 8.0 以上的 RNA 进行建库和测序^[3]。Agilent bioanalyzer 2100 具体操作步骤简图 (图 1) 如下, 按 Agilent 2100 芯片操作 SOP 依次将胶、染料混合液、marker、Ladder、样品等加入对应的芯片孔内, 随即放入 2100 Bioanalyzer 进行分析。RNA 质量评价标准见结果分析示例。

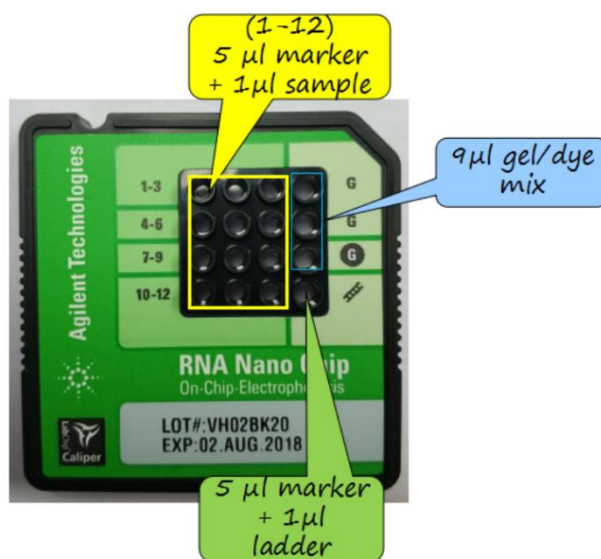


图 1. RNA 上样简图（Agilent 2100 说明书）

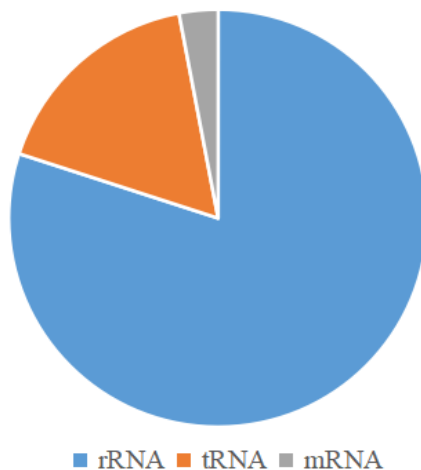
Figure 1. RNA loading diagram

3、 mRNA 的富集。

mRNA 只占原核细胞总 RNA 的 1-5%（图 2），因此建议在转录组测序前对 mRNA 进行富集。虽然 mRNA 的富集不是绝对必要的，但它可以显著增加转录组覆盖率，从而提高所得到的转录组图谱的分辨率^[4]。未经富集的低覆盖率样品可以通过测序更多的 cDNA 来弥补。但是从经济角度出发，这可能会大大增加其实验成本。目前，已知的细菌 mRNA 的分离纯化方法主要有以下几种：rRNA 消减杂交；5'单核苷酸依赖的外切酶处理法；选择性引物扩增法；依赖于双链特异核酸酶的 cDNA 均一化法；大肠杆菌 poly (A) 聚合酶加尾法；与 RNA 结合蛋白 Hfq 免疫共沉淀法^[5]。此处将对 rRNA 消减杂交法进行详细说明，其余方法详细步骤及原理可在附件 1 查看。

rRNA 消减杂交法。目前应用此原理的试剂盒主要有 MICROB Express Ki(Ambion)、RiboMinus bacteria transcriptome isolation Kit (Invitrogen)、Ribo-Zero rRNA removal Kit (Epicentre) 等。本文以 MICROB Express Kit (Ambion) 试剂盒为例进行说明，该试剂盒适用于绝大多数细菌种属（古细菌、支原体等不可用）。其原理为：16S 和 23S rRNA 与总 RNA 样本进行杂交，使用与 rRNA 互补的寡核苷酸探针进行杂交后，带有 16S 和 23S rRNA 的磁珠在磁力的作用下，移至管子的一侧^[6]。上清液中浓缩的 RNA 被转移至新的试管中。洗涤磁珠，以保证磁珠上没有 RNA 的残留。最后用乙醇沉淀 RNA。此过

95 程产生的 RNA 包含 mRNA、tRNA、5S rRNA 和其他小 RNA（图 3）。实验具体操作步
96 骤如下：



97 图 2.细胞内 RNA 分布图

98 Figure 2. Distribution of intracellular RNA

- 100 (1) 将 2-10 μg 总 RNA 添加到 200 μL 结合缓冲液中。
- 101 (2) 加入 4 μL Capture Oligo Mix。
- 102 (3) 加热至 70°C, 10 min。
- 103 (4) 37°C, 持续 15 min。
- 104 (5) 每个样品取出 50 μL Oligo MagBeads 到 1.5 mL 管中。
- 105 (6) 使用磁力试管架捕获 Oligo MagBeads, 并小心地移除和丢弃上清液。
- 106 (7) 用等量的无核酸酶水清洗 Oligo MagBeads。
- 107 (8) 通过与相同体积的缓冲液结合, 以平衡 Oligo MagBeads。
- 108 (9) 重新悬浮 Oligo MagBeads 至相同体积的缓冲液中, 并使缓冲液达到 37°C。
- 109 (10) 同时将洗涤液预热至 37°C
- 110 (11) 在 RNA 预制液中, 加入 50 μL 制备的 Oligo MagBeads, 混匀, 在 37°C 下孵
111 育 15 min。

(12) 捕获 Oligo MagBeads，并将富集的 mRNA 上清液移至收集管。

(13) 用 100 μ L 洗涤液在 37 $^{\circ}$ C 下洗涤，并回收洗涤液，从而回收 Oligo MagBeads 中残留的 mRNA，与前一步骤的上清液合并。

(14) 乙醇沉淀富集的 mRNA。

(15) 将富集的 mRNA 重新悬浮在适当的缓冲液中。

注：MICROB Express Kit（Ambion）试剂盒详细说明书可以在以下网址下载。

<https://www.thermofisher.com/order/catalog/product/AM1905#/AM1905>。

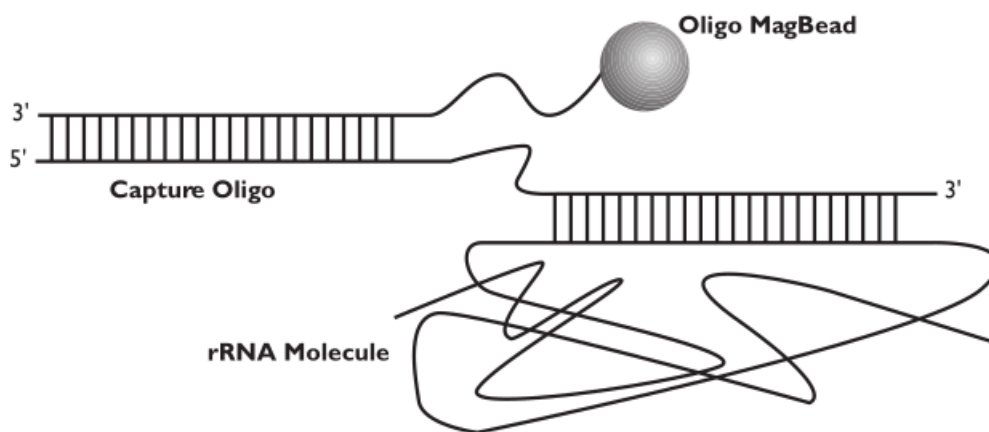


图 3. rRNA 的杂交捕获（MICROB Express Kit 说明书）

Figure 3. Hybridization Capture of rRNA

4、纯化后 mRNA 质量的测定。方法同上文总 RNA 质量检测。

5、片段化处理。

mRNA 纯化之后的文库构建通常有两种思路，一种是首先用 oligo (dT) 引物反转录 mRNA，再进行 cDNA 的片段化；另外一种则是先将 mRNA 打断，再结合随机引物进行反转录。先针对 mRNA 进行打断再进行反转录获得测序序列主要是针对基因本体。而先转录再进行片段化的方法，尤其是结合 oligo (dT) 进行反转录获得的测序结果对转录本 3'端具有比较强的偏好性。所以在 mRNA-Seq 中建议采用先对 mRNA 打断再进行反转录文库构建的方法。以上两种方法具体步骤如下：

5.1 mRNA 片段化。

建议使用专门试剂 Fragmentation Reagent (BioVendor) 将纯化的 mRNA 片段化。mRNA 片段化完成后，再结合随机引物进行反转录。mRNA 片段化处理方法主要包括碱处理法、金属离子 (Mg^{2+} 、 Zn^{2+}) 溶液处理法、酶 (RNaseIII) 处理法等^[7]。mRNA 片段化之后应立即进行第一条 cDNA 链的合成，因为 mRNA 在该体系下非常容易降解。选择金属离子 (Mg^{2+} 、 Zn^{2+}) 溶液处理法打断时需根据需要的文库大小选择合适的片段化温度和时间。(常用设置条件：150-200 bp, 94°C, 15 min; 200-300bp, 94°C, 10 min; 250-550 bp, 94°C, 5 min)。本文以 Hieff NGS® MaxUp™ II Dual-mode mRNA Library Prep Kit for Illumina® 为例进行详细说明。具体操作步骤如下。试剂盒详细说明书可在此地址下载。<https://www.bio-equip.com/show1equip.asp?equipid=4505343>。

(1) 将 mRNA Capture Beads 从 2-8°C 取出，静置使其温度平衡至室温，约 30 min。

(2) 准备一个 Nuclease free 离心管，取 0.1-4 µg 总 RNA，用 Nuclease free 水将体积补至 50 µL，冰上放置备用。

(3) 颠倒或旋涡振荡混匀磁珠，吸取 50 µL 磁珠悬液加入至 50 µL 总 RNA 样品中，用移液器吹打 6 次，使其充分混匀。

(4) 将磁珠与 RNA 的混合物置于 PCR 仪中，65°C, 5 min; 4°C, hold，使得 RNA 变性。

(5) 室温孵育 5 min，使 mRNA 与磁珠完全结合。

(6) 将样品置于磁力架中，室温静置 5 min，使 mRNA 与总 RNA 分离，小心移除上清。

(7) 将样品从磁力架上取出，用 200 µL Beads Wash Buffer 重悬磁珠，移液器反复吹打 6 次以彻底混匀。将样品置于磁力架中，室温静置 5 min，小心移除上清。

(8) 重复步骤 7，共洗涤两次。

(9) 将样品从磁力架上取出，加入 50 μ L Tris Buffer 重悬磁珠，用移液器反复吹打 6 次以彻底混匀。

(10) 将样品置于 PCR 仪中，80 $^{\circ}$ C，2 min；25 $^{\circ}$ C，hold，将 mRNA 洗脱下来。

(11) 将样品从 PCR 仪中取出，加入 50 μ L Beads Binding Buffer，用移液器反复吹打 6 次以彻底混匀。

(12) 室温放置 5 min，使 mRNA 结合到磁珠上。

(13) 将样品置于磁力架中，室温静置 5 min，小心移除上清。

(14) 将样品从磁力架上取出，用 200 μ L Beads Wash Buffer 重悬磁珠，移液器反复吹打 6 次以彻底混匀，将样品重新放回至磁力架中，室温静置 5 min，吸掉全部上清。

5.2 cDNA 的片段化及文库构建。

先用 oligo (dT) 引物反转录 mRNA，再进行 cDNA 的片段化。首先得到长段的一链 cDNA，之后反转录为双链 cDNA，再利用 DNA 片段化的方法对其进行片段化处理（酶切法或机械法），随后纯化得到 cDNA 文库。酶切法构建 DNA 文库，相较于传统的超声法和转座酶法具有以下优势。首先，无偏好片段化酶：具有稳定的片段化效果，无需复杂的机械片段化过程。酶切产物片段大小同样本类型、Input、DNA 量和 GC 含量等均无相关性，仅与酶切时间有关；其次，建库流程精简。一步完成片段化、末端修复、加 A、纯化等。末端修复加 A (30 min)-接头连接 (15 min)-纯化、分选(30 min)-文库扩增(15 min)-纯化分选(30 min)，常规建库预计耗时约 2 h，PCR-free 建库预计耗时 1h；最后，宽泛的样本投入范围。相比较固定投入量，固定酶切时间的 TN5 建库，酶切法建库适用 500 pg-1 μ g，满足个性化建库。本文以 Hieff NGS[®] OnePot DNA Library Prep Kit for Illumina[®]为例，对酶切法制备 cDNA 文库进行详细说明。试剂盒说明书下载地址为 <https://www.yeasen.com/products/detail/932>。该试剂盒具体步骤如下：

a. DNA 片段化、末端修复、dA 尾添加 (DNA Fragment/End Preparation/dA-Tailing)
该步骤将基因组 DNA 片段化，同时进行末端修复及 dA 尾添加。

(1) 将 Input DNA 、 Smearase[®] Mix 、 10 ddH₂O 解冻后，颠倒混匀，置于冰上备用。

(2) 于冰上配制反应体系。Input DNA X μ L、 Smearase[®] Mix 10 μ L、 补水 ddH₂O 至 60 μ L。

(3) 使用移液器轻轻吹打或低速振荡混匀，并短暂离心将反应液离心至管底。

(4) 将上述 PCR 管置于 PCR 仪，反应程序为热盖，105 $^{\circ}$ C； 4 $^{\circ}$ C， 1 min； 30 $^{\circ}$ C， 3-20 min； 72 $^{\circ}$ C， 20 min； 4 $^{\circ}$ C， Hold。进行 DNA 片段化，末端修复及 dA 尾添加反应。

b. 接头连接(Adapter Ligation)该步骤将 3.1 步骤的产物末端，连接特定的 Illumina[®] 接头。

(1) 根据 Input DNA 量按试剂盒说明稀释 Adapter 至合适浓度。

(2) 将 dA-tailed DNA、 Ligation Enhancer、 Fast T4 DNA Ligase、 DNA Adapter 解冻后颠倒混匀，置于冰上备用。

(3) 于 PCR 管中配制反应体系。具体为： dA-tailed DNA， 60 μ L； Ligation Enhancer， 30 μ L； Fast T4 DNA Ligase， 5 μ L； DNA Adapter 5 μ L。

(4) 使用移液器轻轻吹打或振荡混匀，并短暂离心将反应液收集至管底。

(5) 将 PCR 管置于 PCR 仪中，反应程序为热盖，105 $^{\circ}$ C； 20 $^{\circ}$ C， 15 min； 4 $^{\circ}$ C， Hold。进行接头连接反应。

c. 连接产物磁珠纯化(Post Ligation Clean Up) 该步骤使用磁珠对上一步骤的产物进行纯化或分选。纯化可除去未连接的 Adapter 或 Adapter Dimer 等无效产物。

(1) 将 Hieff NGS[®] DNA Selection Beads 磁珠由冰箱中取出，室温平衡至少 30 min。配制 80% 乙醇。

(2) 涡旋振荡或充分颠倒磁珠以保证充分混匀。

(3) 吸取 60 μ L Hieff NGS[®] DNA Selection Beads (0.6 \times , Beads:DNA=0.6:1) 至 Adapter Ligation 产物中, 室温孵育 5 min。

(4) 将 PCR 管短暂离心并置于磁力架中分离磁珠和液体, 待溶液澄清后 (约 5 min), 小心移除上清。

(5) 保持 PCR 管始终置于磁力架中, 加入 200 μ L 新鲜配制的 80%乙醇漂洗磁珠, 室温孵育 30 s 后, 小心移除上清。

(6) 重复步骤 5, 总计漂洗两次。

(7) 保持 PCR 管始终置于磁力架中, 开盖空气干燥磁珠至刚刚出现龟裂 (不超过 5 min)。

(8) 将 PCR 管从磁力架中取出, 进行洗脱:

d. 文库扩增 (Library Amplification) 该步骤将对纯化或长度分选后的接头连接产物进行 PCR 扩增富集。

(1) 将 2 \times Super Canace[®] II High-Fidelity Mix、Primer Mix、Adapter Ligated DNA (上一步骤产物) 解冻后颠倒混匀, 置于冰上备用。

(2) 于无菌 PCR 管中配制, 其反应体系为 2 \times Super Canace[®] II High-Fidelity Mix, 25 μ L; Primer Mix, 5 μ L; Adapter Ligated DNA, 20 μ L。

(3) 使用移液器轻轻吹打或振荡混匀, 并短暂离心将反应液收集至管底。

(4) 将 PCR 管置于 PCR 仪中, 反应程序为 98 $^{\circ}$ C, 1 min; 98 $^{\circ}$ C, 10 sec; 60 $^{\circ}$ C, 30 sec; 72 $^{\circ}$ C, 30 sec; 72 $^{\circ}$ C, 5 min; 4 $^{\circ}$ C, Hold。进行 PCR 扩增。

e. 扩增产物磁珠纯化。同上步骤中纯化操作步骤。使用 Hieff NGS[®] DNA Selection Beads (0.9 \times , Beads:DNA=0.9:1) 纯化文库扩增产物。

f. 文库质量控制。步骤同总 RNA 质量检测。

6、cDNA 文库的构建。

为了获得覆盖整个转录组的序列数据，必须从整个 RNA 样本中随机产生小 cDNA 分子。这通常是通过随机六聚体启动的反转录来实现的。由于常规 cDNA 文库构建方法需进行两条链的合成，所以其正向和反向转录组信息相同。这就使得该方法特异性较差。而与常规 cDNA 文库构建方法相比，仅使用第一条链来构建序列库，就可以保持关于转录方向的信息。或者，使用 RNA 片段化、cDNA 片段化方法，通过定向的、特异的、逐步连接的接头序列来保留转录模板链上的信息。此处主要对常规 cDNA 文库构建方法（图 4）进行详细说明，其他方法详见附件 2。

（1）cDNA 第一条链合成。在随机六聚体引物（Random Hexamer Primer）、逆转录酶的作用下，以 mRNA 为模版合成一链 cDNA。具体过程按照 PrimeScript™ 1st strand cDNA Synthesis Kit 说明书进行 PCR 反应。

（2）cDNA 第二链合成并删除 mRNA，产生双链 cDNA。在 cDNA 第一条链混合物中加入 51 μL 超纯水、20 μL 5×SecondStrand Buffer 和 3 μL dNTP mix（10 mM），dNTP mix 中用 dUTP 代替 dTTP，使 cDNA 第二链中仅包含 A/U/T/G。在冰上放置 5 min，然后加入 1 μL RNase H（2U/μL）和 5 μL DNA Pol 聚合酶（10 U/μL）混匀，然后置于 16℃，2.5 h。

（3）双链 cDNA 的末端修复。取纯化 DNA 溶液 35 μL，加入 50 μL T4 DNA 连接酶缓冲液（2×）、4 μL dNTPs mix、T4 DNA 聚合酶、T4 Polynucleotide Kinase、Klenow DNA 聚合酶，在 20℃下反应 30min，进行末端修复。之后用 QIAquick-PCR 纯化试剂盒对 cDNA 进行纯化。

（4）连接 poly（A）端。将前一步纯化的 DNA 中加入 Klenow 缓冲液、dATP 和 Klenow exo，在 37℃下反应 30 min，用 PCR 纯化试剂盒纯化 DNA，最后将样品溶解在 Ethidium Bromide 溶液中。每个连接序列都有一个 Index 序列（6bp），可用于不同的库构建。

（5）添加接头序列。将纯化后的 DNA 加入 T4-DNA 连接酶缓冲液、接头序列寡聚 ligation solution MasterMix 和 T4 DNA 连接酶。然后用 PCR 纯化试剂盒进行纯化。

(6) 回收并纯化产物凝胶进行电泳。将 250~300bp (或自己实验所需的其他大小的目的片段) 的片段回收并放入 Eppendorf 管中。按 QIAquick 凝胶提取试剂盒说明书对凝胶进行纯化回收。

(7) PCR 扩增前, 使用 UNG 酶消化第二链 cDNA, 文库中只含有第一条。用 PCR 方法扩增富集 cDNA 文库。

(8) 使用凝胶回收试剂盒对 PCR 产物进行回收纯化, 待用。具体步骤参照试剂盒说明书。

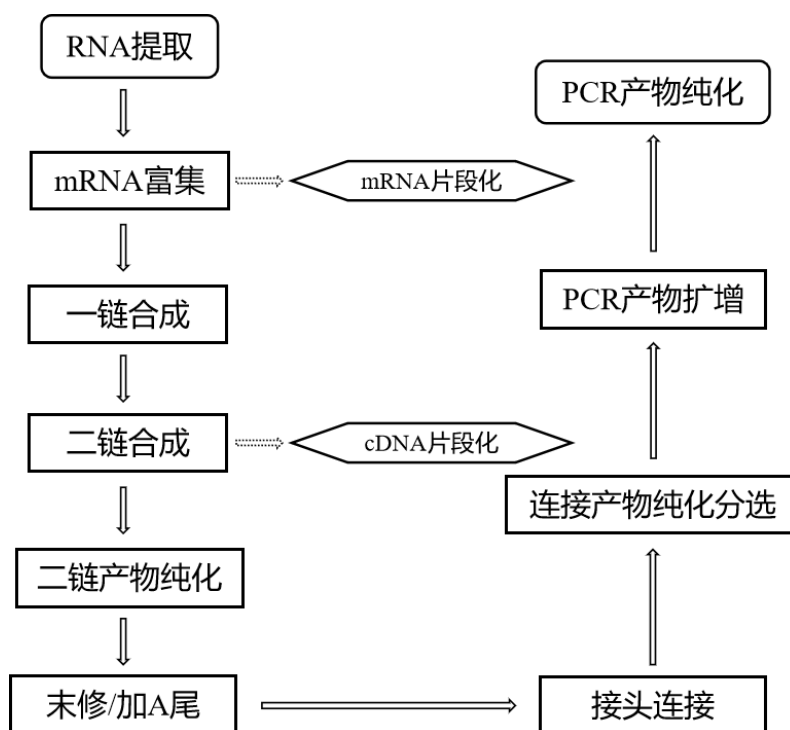


图 4. cDNA 文库构建方法简图

Figure 4. Schematic diagram of cDNA library construction

7、cDNA 文库质量检测。

文库容量、重组率及插入片段的大小是鉴定 cDNA 文库质量的重要指标^[8]。先对第二轮 PCR 产物进行琼脂糖凝胶电泳初步分析, 再利用 Agilent 2100 bioanalyzer 仪进行检测, 检测流程参照其说明书。Bioanalyzer 分析按照 DNA 1000 LabChip 试剂盒所提供的说明进行。每 1 试剂盒含有 25 块芯片和以下试剂: DNA 长度测定梯度标准品 (sizing ladder)、注射器、凝胶基质、染料浓缩液、DNA 相对分子质量标准 (marker) 以及旋

转过滤器。凝胶染料混合物的准备：将 25 μL DNA 染料与凝胶基质混合后， $6000\times g$ 离心，在芯片指定位置分别加入 9 μL 凝胶染料混合物以及 5 μL maker 和 1 μL ladder(1000)，取 7 个 PCR 产物各 1 μL 分别加入 1~7 号样品孔中，最后将芯片涡旋混匀后放入仪器中进行自动检测分析。

8、若所测 cDNA 文库质量较好，满足建库要求，则可以进行 cDNA 上机测序。

结果与分析

RNA 样本总量及完整性是评判样品质量的关键点，其中 RNA 完整性评估依靠 RIN 值、23S/16S（原核生物）以及 bioanalyzer 检测峰图基线是否平整来综合评判。对于降解样品，实验难以获取完整的转录本信息，这样就会影响数据质量及其完整性。当 RNA 总量较低时，会导致建库成功率低，或数据 Duplication rate 高等问题^[9]。基于此，本文将展示部分较好的结果示例就如何对 RNA、cDNA 质量进行评价作进一步的详细说明。

(1) RNA 质量评价标准——凝胶电泳图。

RNA 质量评价首先要通过凝胶电泳图来对其进行初步分析。主要需查看以下几方面：目的条带是否明亮清晰；泳道内是否有明显的降解弥散区；胶孔处有无蛋白污染；有无 DNA 污染；有无外源物种核酸污染等^[10]。由下文结果示例可以看出，该总 RNA 的凝胶电泳图目的条带明亮清晰，泳道无弥散区，无蛋白和 DNA 的污染。由此得出结论该样品质量较好，如图 5 所示（图 5 为 bioanalyzer 获得）。失败案例如图 6 所示（图 6 为凝胶电泳图）。

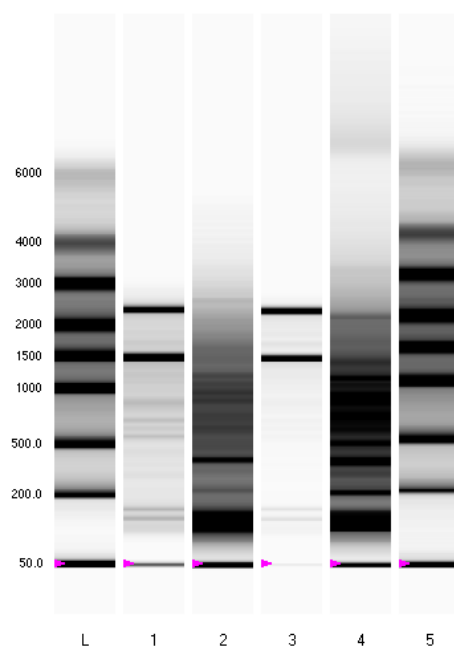


图 5. RNA 质量较好的示例（2017，Han DF 等）

Figure 5. Examples of RNA quality control (good quality) (Bio-Rad)

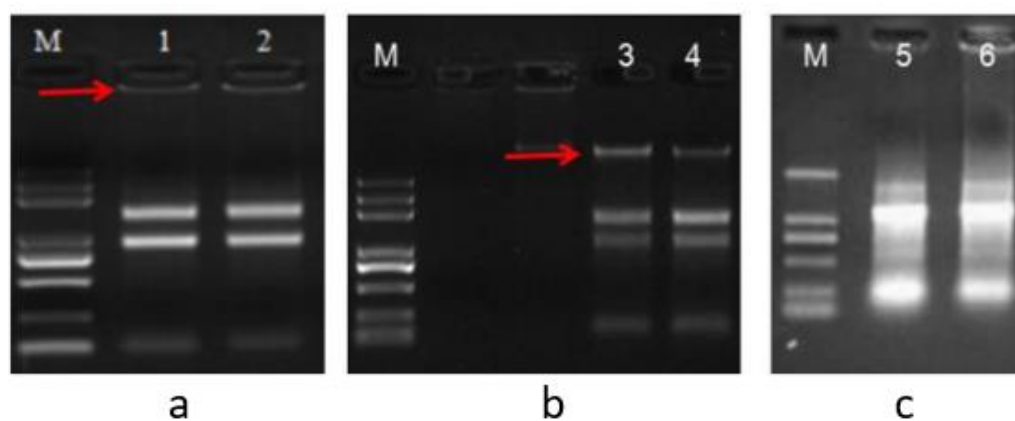


图 6. RNA 质量较差的凝胶电泳图示例

（2020，派森诺生物科技有限公司实例分析）

a. 胶孔处蛋白污染 b. DNA 污染 c. RNA 降解

Figure 6. Examples of RNA quality control (poor quality). a. protein contamination at the sample wells; b. DNA contamination; c. RNA degradation.

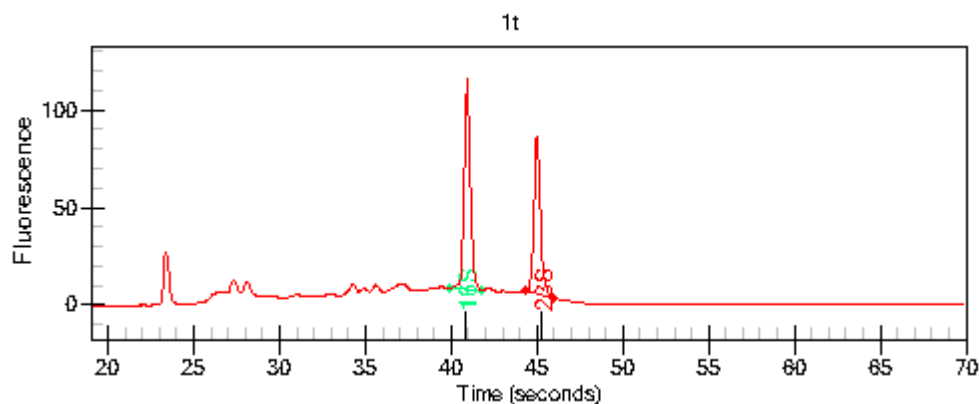
(2) RNA 质量评价标准——电泳图及 RIN 值。

在凝胶电泳检测完成后，需要对其进行 bioanalyzer 分析。图 7 为 bioanalyzer (Bio-Rad) 原核生物典型电泳图。最左侧峰为 Ladder 峰，中间峰为 16S rRNA，最右侧峰为 23S rRNA。整体样品峰图基线较为平整，峰型较好，无拖尾峰、前沿峰、包裹峰及多余的降解峰。通过 mRNA 富集前后对比电泳图，可以清晰地看到 mRNA 富集程度 (即 16S rRNA、23S rRNA 去除程度)。

若样本采用 Agilent bioanalyzer 2100 质检。则需判别其 RIN 值，RIN 值即为 RNA 完整值 (RNA integrity number)，上样完成后，质检报告中会显示其具体数值。RIN 数值范围为 1-10，将生物总 RNA 质量进行分类，其中 1 代表降解最严重的情况，10 代表最完整，建议采用 RIN 数值 8 以上的样品进行 cDNA 文库的构建。通过这种检测方法，有助于解释电泳图，有利于样品之间的相互比较，同时确保实验的可重复性。

(3) cDNA 质量评价标准——凝胶电泳图。

通过 PCR 进行第二链合成后，直接取 5 μ L PCR 产物进行琼脂糖凝胶电泳检测。示例较好的结果如图 8 所示，双链 cDNA 片段弥散分布，大小主要集中在 200~500 bp (实验所需 cDNA 片段大小，片段化条件控制文库 DNA 分子的长度分布具体方法可参考所用试剂盒，比如 TruSeq[®] RNA Sample Preparation v2 Guide 试剂盒)，说明实验所需 mRNA 均成功反转录并得到了有效扩增，其 cDNA 合成质量较高，基本包含全部 cDNA，可满足文库构建的需要，可用于后续筛选目的基因、大规模测序等研究。



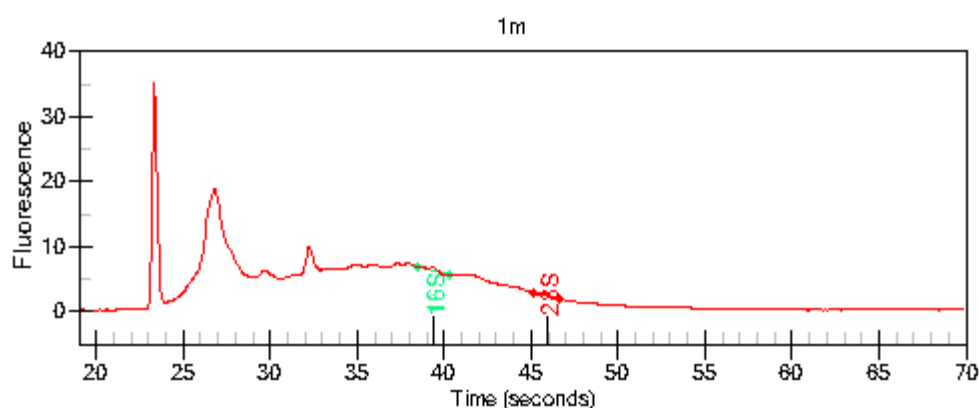


图 7. mRNA 富集前后对比 (2017, Han DF 等)

Figure 7. Comparison of mRNA before and after enrichment (Bio-Rad)

(4) cDNA 质量评价标准——电泳图。

cDNA 质量同样需要进行 bioanalyzer 来进一步确定其质量。图 9 为 bioanalyzer (Bio-Rad) 原核生物 cDNA 典型电泳图。最左侧峰为 1700 bp 内标物，中间峰为 cDNA，最右侧峰为 50 bp 内标物 (内标物可根据实验所需进行设置，包括数量、大小等)。由图可知，该样品整体峰图基线较为平整，峰型较为平滑，没有出现拖尾峰、前沿峰、包裹峰及其他的杂乱峰。cDNA 峰较为明显且峰面积较大。说明文库质量较高，可满足文库构建的需要。

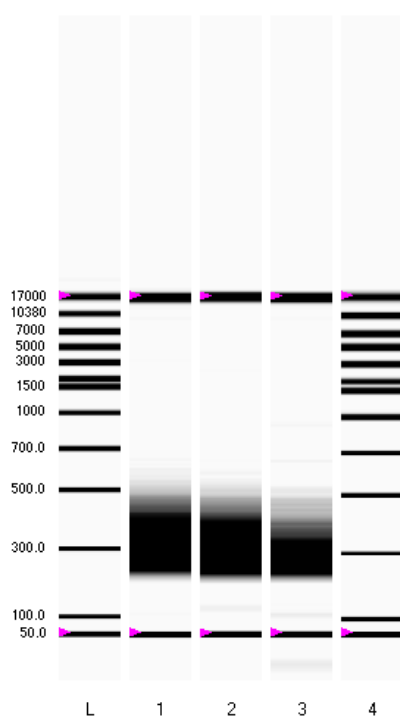


图 8. cDNA 模拟胶图（2017，Han DF 等）

Fig. 8. cDNA simulated gel electrophoresis (Bio-Rad)

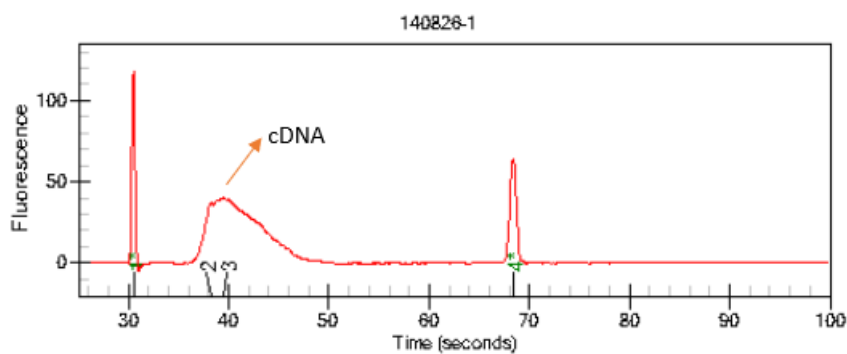


图 9. cDNA 电泳图（2017，Han DF 等）

Figure 9. Electrophoresis of cDNA

致谢

感谢中国农业科学院科技创新工程对本研究的资助。使用本实验方案已发表的文章有：Han, D., Link, H., & Liesack, W. (2017). Response of *Methylocystis* sp. strain SC2 to salt stress: physiology, global transcriptome, and amino acid profiles. *Appl Environ Microbiol*, 83(20), e00866-17. doi:10.1128/aem.00866-17。

参考文献

- [1] Chauhan, A., Sharma, J. N., Modgil, M., & Siddappa, S. (2018). Comparison of various RNA extraction methods, cDNA preparation and isolation of calmodulin gene from a highly melanized isolate of apple leaf blotch fungus *Marssonina coronaria*. J Microbiol Methods, 151, 7-15. doi:10.1016/j.mimet.2018.05.023
- [2] Rodríguez, A., & Vaneechoutte, M. (2019). Comparison of the efficiency of different cell lysis methods and different commercial methods for RNA extraction from *Candida albicans* stored in RNAlater. BMC Microbiol, 19(1), 94. doi:10.1186/s12866-019-1473-z
- [3] Schroeder, A., Mueller, O., Stocker, S., Salowsky R., Leiber M., Gassmann M., Lightfoot S., Menzel W., Granzow M. & Ragg T. (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Molecular Biol 7, 3, doi: 10.1186/1471-2199-7-3
- [4] Kumar, N., Lin, M., Zhao, X., Ott, S., Santana-Cruz, I., Daugherty, S., Dunning Hotopp, J. C. (2016). Efficient enrichment of bacterial mRNA from host-bacteria total RNA samples. Sci Rep, 6, 34850. doi:10.1038/srep34850
- [5]. Kumar, N. , Lin, M. , Zhao, X. , Ott, S. , Santana-Cruz, I. , & Daugherty, S. , et al. (2016). Efficient enrichment of bacterial mRNA from host-bacteria total RNA samples. Scientific Reports, 6, 34850.
- [6] Stewart, F. J., Ottesen, E. A., & DeLong, E. F. (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. ISME J, 4(7), 896-907. doi:10.1038/ismej.2010.18
- [7]. Kumar, R., Ichihashi, Y., Kimura, S., Chitwood, D. H., Headland, L. R., Peng, J., Sinha, N. R. (2012). A high-throughput method for illumina RNA-Seq library preparation. Front Plant Sci, 3, 202. doi:10.3389/fpls.2012.00202
- [8]. Croucher, N. J., & Thomson, N. R. (2010). Studying bacterial transcriptomes using RNA-Seq. Curr Opin Microbiol, 13(5), 619-624. doi:10.1016/j.mib.2010.09.009
- [9] Hayden, C. A. , Wheeler, T. J. , & Jorgensen, R. A. (2005). Evaluating and improving cDNA sequence quality with cQC. Bioinformatics, 21(24), 4414-4415.
- [10] Han, D., Link, H., & Liesack, W. (2017). Response of *Methylocystis* sp. strain SC2 to salt stress: physiology global transcriptome and amino acid profiles. Appl Environ Microbiol, 83(20), e00866-17. doi:10.1128/aem.00866-17

附件 1 其他 mRNA 富集方法原理及步骤

a. 5'单核苷酸依赖的外切酶处理法。

目前已知的应用此原理的试剂盒主要是 mRNA-ONLY Prokaryotic mRNA isolation Kit (Epicentre)。目前该试剂盒已经停产。其具体原理如下：大多数的细菌和古细菌 mRNA 均携带 5'三磷酸 (5'PPP)，类似于真核细胞 RNA 的帽状结构 (如图 10)。经此方法处理过的 RNA 分子，如 rRNAs 和 tRNAs，携带 5'单磷酸 (5'P)。mRNA-ONLY Prokaryotic mRNA isolation Kit (Epicentre) 使用 5' - 3'核酸外切酶降解 5'P RNA 分子，以保持其 mRNA 完整^[11]。文献分析表明，此方法对于革兰氏阳性菌和革兰氏阴性菌均可适用。试剂盒具体说明书可在此网址下载查询。

https://www.massey.ac.nz/massey/learning/departments/centres-research/epicentre/epicentre_home.cfm#studyNav

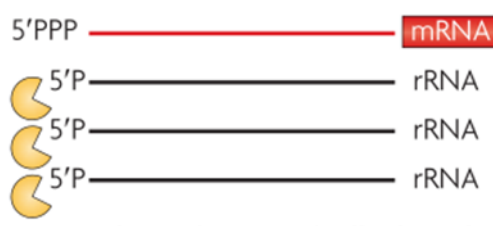


图 10. RNA 降解。(2009, Rotem Sorek 等)

Figure 10. Degradation of processed RNA

b. 选择性引物扩增法。

该方法的具体步骤为：预先设计一系列的不具有 rRNA 偏好性的 NSR 引物，然后使用这些 NSR 引物选择性扩增 rRNA 以外的转录本^[11]。应用此原理的相应试剂盒有 Ovation Prokaryotic RNA-Seq system (NuGEN)，目前该试剂盒已经停产，其升级版本无 rRNA 移除步骤，通过提高其转录效率来富集。该试剂盒具体原理如图 11 所示。

https://www.researchgate.net/publication/51044917_Method_for_improved_Illumina_sequencing_library_preparation_using_NuGEN_Ovation_RNA-Seq_System 该试剂盒详细说明书可以在上面的链接进行下载。

c. 依赖于双链特异核酸酶的 cDNA 均一化法。

这种方法是从真核生物的研究中借鉴而来，在 cDNA 水平上降低 rDNA 和 tDNA 转录物所占的比例，在原核生物中应用的还不多。从标准 cDNA 文库中直接随机测序克隆，对于发现稀有转录物的效率较低，因为中丰度和高丰度的 cDNA 会被重复测序。标准化降低了代表大量转录本的克隆的发生率，显著提高了随机测序和稀有基因发现的效率。目前主要的试剂盒主要有 Trimmer-Direct cDNA Normalization Kit (Evrogen)。该试剂盒专门设计用于规范全长度序列的 SMART Amplified cDNA，完整长度的 cDNA 文库可以在一个克隆步骤中获得每个转录本的完整序列信息。其将 SfiI 酶限制位点整合到 cDNA 中，允许定向克隆标准化 cDNA 文库。标准化是在 cDNA 克隆之前进行的。试剂盒详细说明书可以在此链接下载。具体过程如图 12。

<https://wenku.baidu.com/view/54e5516aa98271fe910ef9f7.html>

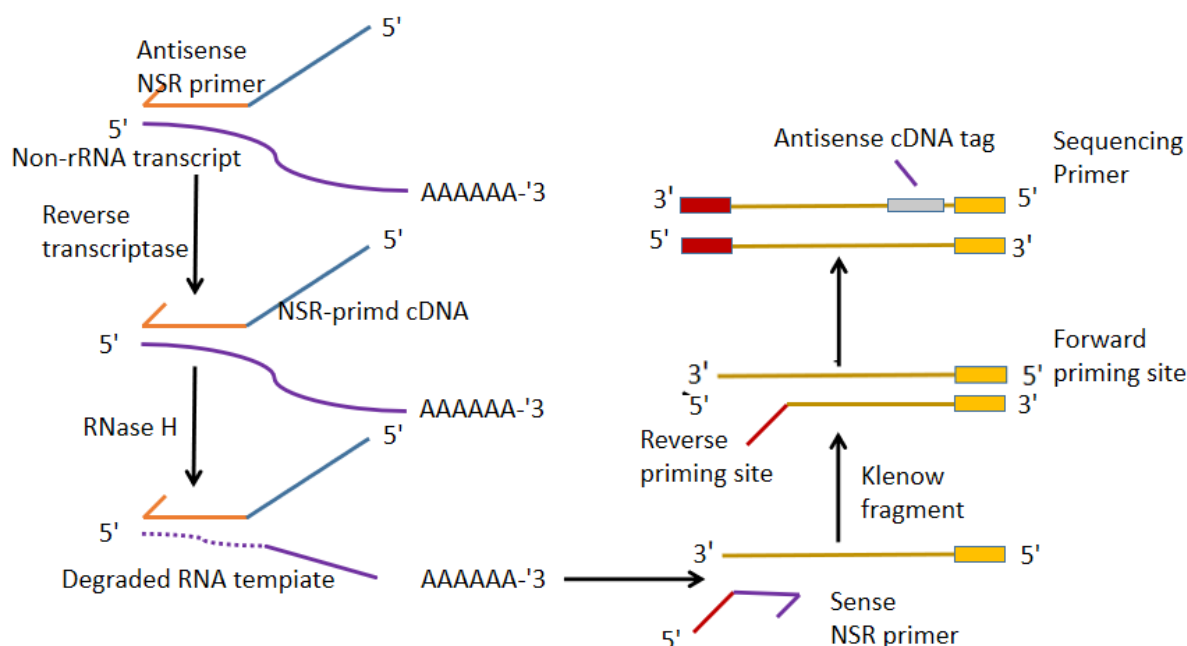


图 11. Ovation Prokaryotic RNA-Seq system (NuGEN) 原理

Figure 11. Principle of Ovation Prokaryotic RNA-Seq system (NuGEN)

d. 大肠杆菌 poly (A) 聚合酶加尾法。

目前应用该原理的试剂盒有 MessageAmp II-bacteria Kit (Ambion)。该试剂盒是一种基于体外转录的线性 RNA 扩增系统。与真核细胞 RNA 的扩增方法相比，主要的区

别在于细菌 mRNA 没有稳定的 poly (A) 尾, 因此它必须经过多聚腺苷酸化才能成为合适的扩增底物^[12-13]。因此, 该程序的第一步是在优化反应中使用 *E. coli* Poly (A) Polymerase (PAP) 进行多聚腺苷酸化, 以确保细菌 RNA 分子的完全和具有代表性的多聚腺苷酸化。接下来, 在带有 T7 启动子的寡核苷酸 (dT) 引物的反应中, 尾部 RNA 被反向转录。但该试剂盒仅适用于大肠杆菌。

该试剂盒使用 ArrayScript™ 逆转录酶, 可产生比野生型酶更高的第一链 cDNA 产量。ArrayScript 可以催化合成几乎全长的 cDNA。为了使 aRNA 产量最大化, 试剂盒中包含 Ambion® MEGAscript® 体外转录试剂。IVT 反应可配置为合成标记的 aRNA (例如生物素、Cy™ 染料或氨基烯丙基) 或未标记的 aRNA, 随后可通过反转录 (例如荧光或放射性 dNTPs) 进行标记。aRNA 可以通过在 IVT 反应中包含标记的核苷酸或在包含标记核苷酸的反应中合成后反向转录 aRNA 来进行标记。因此, 该试剂盒可用于制备 ³³P 标记的杂交 cDNA。具体原理及程序如图 13 所示。

<https://www.thermofisher.com/order/catalog/product/AM1790?SID=srch-srp-AM1790#/AM1790?SID=srch-srp-AM1790> 该试剂盒详细说明书可在以下网址下载。

e. 与 RNA 结合蛋白 Hfq 免疫共沉淀法。

该方法以特定 RNA 为目标 (图 14)。共免疫沉淀 (Co-IP) 用于分离与 Hfq 相关的 RNA, Hfq 是一种介导细菌小 RNA 与其 mRNA 靶点相互作用的蛋白质。Hfq (RNA 相 Qβ 复制的宿主因子) 是大多数 sRNAs 的伴侣, 是转录后的全局调节因子^[14]。Hfq 是一种高度保守的 RNA 结合蛋白, 被认为是细菌基因转录后调控的关键因素。sRNA 通常位于基因间区 (IGR), 长度为 50-500bp, 具有特殊的茎环结构^[15], 这些非编码 RNA 不翻译成蛋白质。它们以 RNA 的形式在维持细菌基因组稳定性、生长代谢和致病机制中起着调节作用。Hfq 作为一种 RNA 伴侣蛋白, 能有效地结合小 RNA, 并辅助其与靶 mRNA 结合。因此, 它被广泛应用于小 RNA 及其靶 mRNA 的研究。其具体实验步骤如下:

(1) 载体构建及 PCR 验证

设计并构建重组载体，加入 3×Flag 标签，重组质粒（Shfq 已敲除成功的菌株和野生型菌株），分别构建带标签的 Hfq 互补株、实验株和带标签的对照株。通过 PCR 鉴定和基因测序证实构建成功。

（2） 生长曲线绘制

实验菌株分别接种于含相应抗生素的培养基中，培养至对数后期。1:40 稀释至新鲜培养基中，每隔 2 h 测其 Deco 值，连续监测 46 h 并绘制生长曲线，每株菌重复 3 次。

（3） 细胞裂解液上清制备

离心收集细胞沉淀物，弃上清液，用预先冰浴冷却的 PBS 重悬细胞，4℃、12000 r/min 离心 5min，弃上清，洗涤 2 次；加入适量 RIP 洗涤缓冲液、蛋白酶抑制剂混合液、RNAs 抑制剂，充分混匀，避免产生气泡；超声波裂解细胞，离心，上清液将用于免疫沉淀，-80℃贮存。

（4） 磁珠的准备

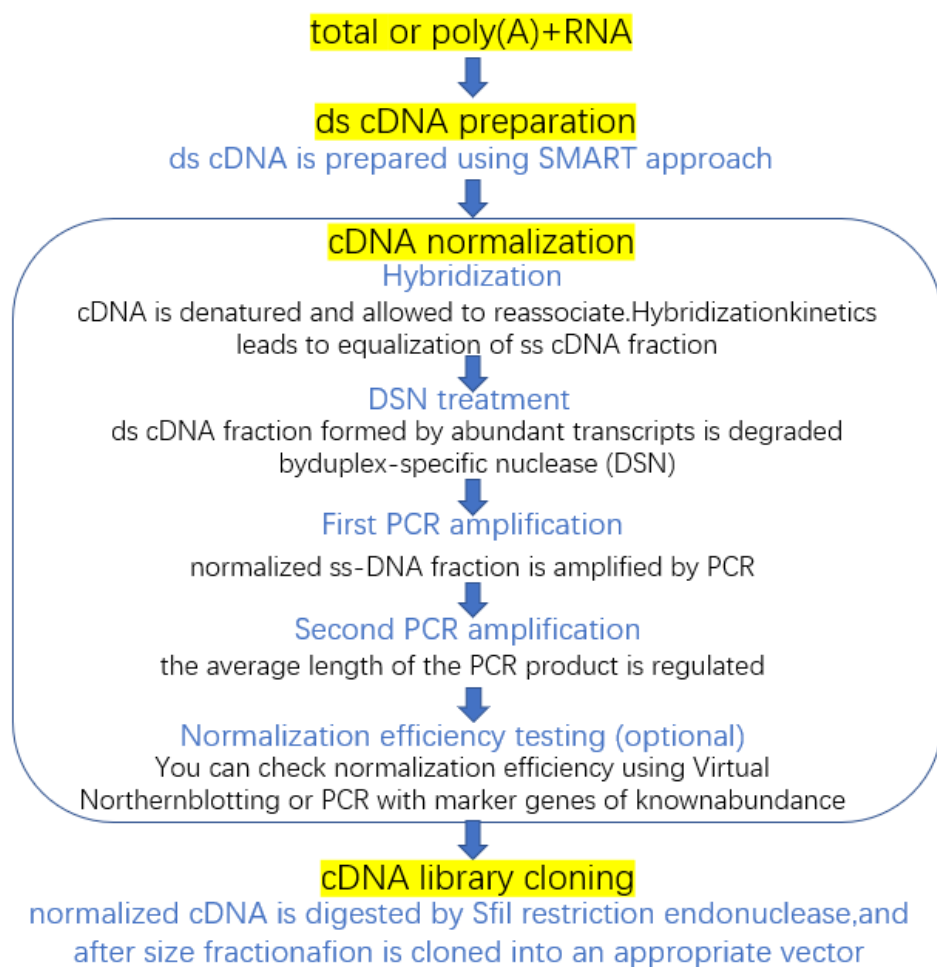
在清洗过程中，吸管头须无 RNase 污染。重悬磁珠，每管加 50 μL 磁珠悬液，再加 0.5 mL RIP 洗涤缓冲液，涡旋洗涤后将管置于磁力分离器上，当磁珠聚集后弃上清，共 2 次洗涤；每管加 100 μL RIP 洗涤缓冲液，重悬磁珠后加 5 μg 抗体，室温旋转孵育 30 min；将每管短暂离心后置于磁力架上，弃上清，每管加 0.5 mL RIP 洗涤缓冲液，涡旋后置于磁力分离器上，弃上清液；重复洗涤 1 次；置于冰上。

（5） RIP 裂解

制备适量含 RIP 洗涤缓冲液、0.5 mol/L EDTA、RNase 抑制剂的 RIP 反应液。磁珠放磁力架上，弃上清液，每管添加 900 μL RIP 反应液；将第 I 部分的 RIP 裂解上清液迅速解冻，4℃下 12000 r/min 离心 10 min，将 100 μL 上清液加入磁珠中，使每管 RIP 反应液达到 1.0 mL；从第一部分的 RIP 裂解上清液中取出 100μL 提取 RNA，此部分为总 RNA（或 input RNA）；4℃旋转孵育 3 h 或过夜，短暂离心，置于磁力分离器上，留存上清，提取 RNA。此部分 RNA 为未结合 RNA（Unbound RNA）；加入 0.5 mL 预冷的 RIP 洗涤缓冲液，涡旋后置于磁力分离器上，留存上清液；重复此步 5 次，共 6 次，

457 对获取的上清液进行标记；第 6 次重悬后，取磁珠悬液 50 μ L 用于 Western 印迹，验证
458 磁珠上的 Hfq 蛋白。

459 (6) RNA 的纯化用蛋白酶 K 溶液重悬磁珠，55°C 消化 10 min；再用 TRizol 法抽
460 提 RNA，用乙醇沉淀，此部分 RNA 为结合 RNA (Bound RNA)。



461

462 图 12. Evrogen TRIMMER DIRECT 使用的规范化过程概述

463 Figure 12. Overview of the normalization procedure using Evrogen TRIMMER DIRECT.

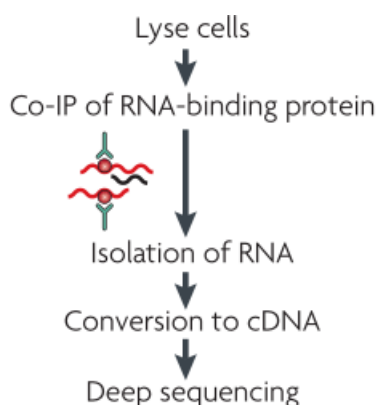
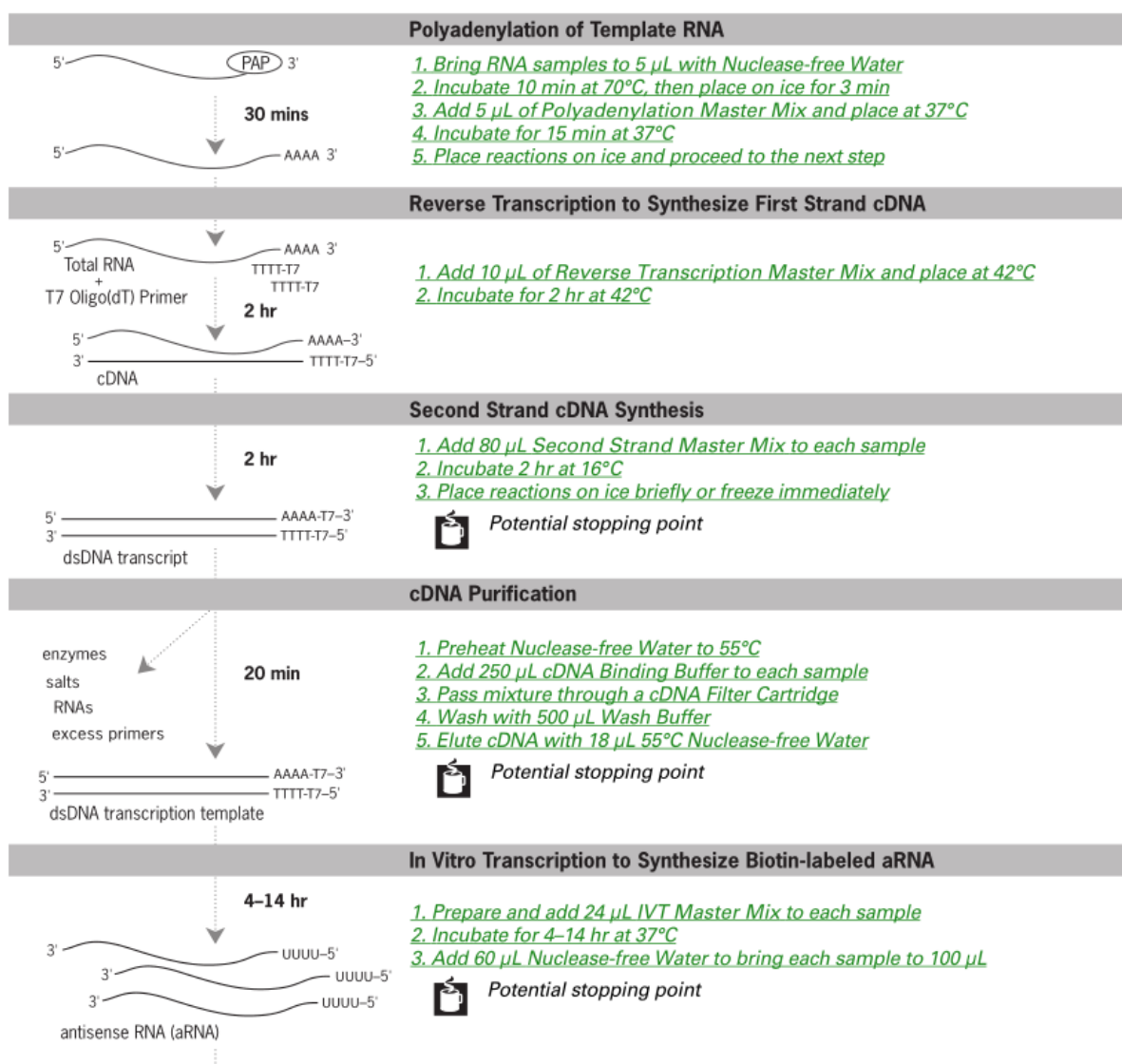


图 14. 捕捉与特定蛋白质相互作用的 RNAs (2009, Rotem Sorek 等)

Figure 14. Capture of RNAs that interacts with a specific protein



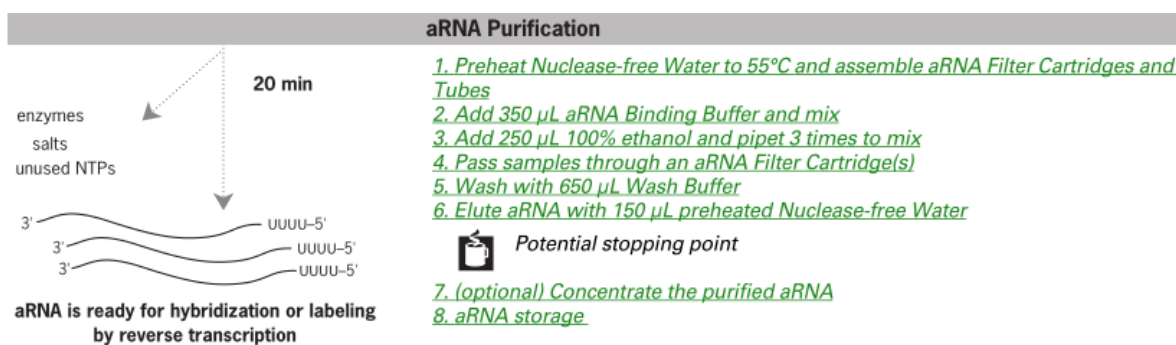


图 13. Message Amp II-bacteria Kit 过程概述（Ambion 试剂盒说明书）

Figure 13. Overview of Message Amp II bacterium Kit （Ambion）

附件 2 其他 cDNA 构建原理及方法

a. 仅使用第一条 cDNA 链来构建序列库。

在最初的 RNA-Seq 方案中，经过广泛的 DNase 处理，RNA 通常通过随机六聚体启动的反转录转化为 cDNA，然后进行第二条 DNA 链的合成^[16]。然而，使用双链 cDNA 构建测序文库会导致正向链和反向链上的信号水平相等，从而丢失有关转录方向的信息。保持 RNA 序列数据中方向性信号的一个简单方法是仅从第一链 cDNA 构建 Illumina 文库。其具体的操作方法和上文相同，只是不进行第二条链的合成，直接以第一条 cDNA 链来构建序列库。目前 cDNA 第一链合成一般使用随机引物六聚体反转录或 BeyoRT™ II cDNA 第一链 cDNA 合成试剂盒（RNase H）。该试剂盒 20 µL 体系反转录 1 µg 总 RNA，引物使用 Random Hexamer Primer，操作均按使用说明进行。其具体过程如下。

（<https://beyotime.com/d7168.htm> 试剂盒详细说明书下载地址）。

（1）设置反转录反应。

（2）轻轻混匀（用移液器轻轻吹打混匀或用涡旋混合器在最低速度轻轻混匀），随后离心沉淀液体。

（3）如果使用 Oligo(dT)₁₈ 或基因特异性引物，42°C 孵育 60 min。如果使用 random hexamer（随机六聚体）作为引物，先在 25°C 孵育 10 min，随后在 42°C 孵育 60 min。注意：对于 GC 含量较高或二级形成现象结构比较严重的模板 RNA，可以 50°C 孵育 60 min，

以充分利用本产品中的反转录酶在 50°C 时仍有良好活性这一特点, 在较高温度进行反转录可以有效减少二级结构的干扰。

(4) 80°C 孵育 10 min 以失活 BeyoRT™ II M-MLV 反转录酶 (RNase H) 并终止反转录反应。说明: 对于 5kb 以上的长片断 cDNA 不推荐采用加热的方法失活反转录酶, 该方法易导致部分长片断 DNA 被剪切, 此时可考虑酚氯仿抽提或柱纯化方法。

(5) 反转录产物可以直接用于后续的 PCR 反应等, 也可以 -20°C 冻存以备以后使用。用于后续 PCR 反应时, 如果 PCR 的反应体系为 20 和 50 μL, 则推荐相应地使用 0.8 μL 和 2 μL 反转录产物。

(6) 其他试剂盒方法构建文库方法, 具体过程参照器具具体说明书。比如 EpiNext DNA Library Preparation Kit (Illumina), Collibri™ PS DNA Library Prep Kit for Illumina Systems, with UD indexes (Set A, 1-24) 等。

附件 1、2 参考文献

- [10].Sorek, R., & Cossart, P. (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet*, 11(1), 9-16. doi:10.1038/nrg2695
- [11].Amara, R. R. & Vijaya, S. (1997) Specific polyadenylation and purification of total messenger RNA from *Escherichia coli*. *Nucleic Acids Res.* 25, 3465–3470. doi:10.1093/nar/25.17.3465
- [12]. Volker F. Daniel P. Khodursky A. (2001). Isolation of *Escherichia coli* mRNA and comparison of expression using mRNA and total RNA on DNA Microarrays. *Anal. Biochem.* 290, 205–213. doi:10.1006/abio.2000.4982.
- [13].Sittka, A., Lucchini, S., Papenfort, K., Sharma, C. M., Rolle, K., Binnewies, T. T., .Vogel, J. (2008). Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet*, 4(8), e1000163. doi:10.1371/journal.pgen.1000163
- [14] Assis, N. G., Ribeiro, R. A., da Silva, L. G., Vicente, A. M., Hug, I., & Marques, M. V. (2019). Identification of Hfq-binding RNAs in *Caulobacter crescentus*. *RNA Biol*, 16(6), 719-726. doi:10.1080/15476286.2019.1593091
- [15] Tim, C. , Matthew, B. , Adrian, T. , Chinmay, P. , Böhme Ulrike, & Barrell, B. G. , et al. (2008). Artemis and act: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* (23), 2672-2676.