

· 专题论坛 ·

植物实时荧光定量PCR内参基因的特点及选择

袁伟^{1, 2}, 万红建¹, 杨悦俭^{1*}

¹浙江省农业科学院蔬菜研究所, 杭州 310021; ²南京农业大学园艺学院, 南京 210095

摘要 实时荧光定量PCR(qRT-PCR)具有灵敏度高、特异性强、重复的动态定量范围和高通量等优点, 是进行植物基因表达和转录分析最常用的技术手段之一。选择合适的内参基因是正确运用实时荧光定量PCR分析目标基因表达变化的前提。近年来, 大量研究表明, 内参基因的选择应取决于研究者的实验条件; 随着实验条件的变化, 内参基因的选择也随之变化。因此, 实时荧光定量PCR结果分析的准确性在很大程度上依赖于所选择的内参基因是否适合。该文从内参基因的选择、常用内参基因的特点、新内参基因的挖掘、应用内参基因组合的优点和内参基因的稳定性评价等几方面进行综述, 以期为研究者在实验中选择合适的内参基因提供参考和理论依据。

关键词 基因表达分析, 基因表达的稳定性, qRT-PCR, 内参基因

袁伟, 万红建, 杨悦俭 (2012). 植物实时荧光定量PCR内参基因的特点及选择. 植物学报 47, 427–436.

随着基因组技术的发展越来越迅速, 其为进一步阐明复杂的生物学过程、代谢途径及植物与病原菌的互作机制等提供了强大的动力(Mascia et al., 2010)。实时荧光定量PCR(real time fluorescence quantitative PCR, qRT-PCR)具有灵敏度高、特异性强、重复性好及高通量等特点(Huggett et al., 2005), 已经成为植物科学领域中进行基因组研究的常用技术之一。利用该技术可进行基因表达和转录组分析等(Gachon et al., 2004)。然而, 在其分析过程中, RNA的质量和产量、反转录效率的差异以及其他阻碍因素等都会对目的基因表达分析结果的准确性产生影响(Pfaffl, 2006)。因此需要选择合适的内参基因进行校正和标准化, 从而减小检测样本之间的差异(Quackenbush, 2002; Nolan et al., 2006)。理想的内参基因应该在所有的细胞中和生理状态下都能较稳定地表达, 可以完全用于不同类型的样品。但是随着实验条件的变化, 没有任何一种内参基因的表达是始终稳定的, 即在一种实验条件下合适的内参基因并不一定适用于另一种实验条件(Volkov et al., 2003); 或者不同植物间的内参基因也可能是不同的。因此在具体实验条件下选择表达稳定的内参基因至关重要。由于常用的内参基因是作为构成细胞器骨架的基本组分

(*ACT*、*TUA*、*TUB*等)或参与生物体的基本生化代谢过程(*GAPDH*、*EF-1 α* 、*UBQ*等)(Huggett et al., 2005; Gutierrez et al., 2008), 使得其在某些组织器官中基本可以稳定地表达(胡瑞波等, 2009)。但是qRT-PCR是对基因表达的准确定量分析, 而近年来的研究发现, 这些常用内参基因均存在缺陷, 在不同类型的细胞和组织中、细胞和器官发育的不同阶段以及各种实验条件等情况下, 它们的表达水平通常变异较大(Hu et al., 2009; Die et al., 2010), 已经不能满足定量分析对内参基因的要求。因此, 新内参基因的挖掘已然成为重中之重。目前, 基于基因芯片的表达数据和EST数据库逐渐成为筛选新内参基因的主要渠道。

鉴于内参基因在植物基因组研究特别是qRT-PCR分析中的重要性, 本文从内参基因的选择、常用内参基因的特点、新内参基因的挖掘、应用内参基因组合的优越性和评价内参基因稳定性的方法等几方面进行分析, 以期为针对不同植物在某种实验条件下选择合适的内参基因提供理论依据。

1 植物中内参基因的选择

1.1 内参基因的选择标准

qRT-PCR是目前在mRNA水平上探测基因表达最有

收稿日期: 2011-10-28; 接受日期: 2012-03-15

基金项目: 国家自然科学基金(No.31071800)、浙江省自然科学基金(No.LQ12C150020)和浙江省公益技术研究农业项目(No.2011-C22007)

* 通讯作者。E-mail: hzyyj@yahoo.com.cn

效、精确的方法, 常用于相对定量分析生物体内生理和病理过程中mRNA的变化(Wong and Medrano, 2005)。由于PCR过程中存在各种各样的差异, 对结果的校正和标准化仍是当前面临的最具挑战性的问题(Guénin et al., 2009)。而内参基因的使用被认为是适当的标准方法(Huggett et al., 2005), 因为其能够有效地校正RNA起始量和反转录效率等的差异(Udvardi et al., 2008), 从而获得反映目标基因特异性表达的真正差异。理想的内参基因应满足以下条件:(1) 在不同类型的细胞和组织中及不同的生长发育阶段都稳定表达(Quackenbush, 2002); (2) 表达不受环境、生物或非生物胁迫等内源性和外源性因素的影响(Thellin et al., 1999); (3) 不存在假基因(pseudo gene), 以避免基因组DNA的扩增; (4) 其稳定表达量与目的基因的表达量近似(Dheda et al., 2004)。但是目前尚未发现这样理想化的内参基因, 因为没有任何一种内参基因在不同的植物组织、器官中或不同的实验条件下都恒定表达(Die et al., 2010)。而且, 选择不适当的内参基因往往会导致错误结论(Dheda et al., 2005)。因此选择合适的内参基因已成为应用qRT-PCR进行基因组功能分析的重要前提。

1.2 内参基因的选择不存在通用性

在植物内参基因研究中, 没有任何一种内参基因在不同的实验条件下表达是恒定的。首先, 不同物种间的同源内参基因表达存在差异, 即在某种植物中表达稳定性高的内参基因在另一种植物中不一定会有较好的表达稳定性, 即不存在通用性。例如在葡萄(*Vitis amurensis*)浆果中表达稳定性较好的GAPDH在小麦(*Triticum aestivum*)中的表达稳定性则最差(Reid et al., 2006; Long et al., 2010); 而在小麦中表达稳定性较好的ACT和UBI在番茄(*Solanum lycopersicum*)中的表达稳定性最差(Long et al., 2010; Mascia et al., 2010)。此外, 同一内参基因家族内的各成员表达稳定性也存在差异, 并不能因为基因家族中的某一个成员表达稳定性好, 适合作内参基因, 就推论出该基因家族其它成员的表达也相对稳定, 即这种情况也不存在通用性。肌动蛋白基因家族作为细胞骨架结构蛋白被广泛用作内参基因, 主要包括 $ACT2/7$ 、 $ACT8$ 、 $ACT11$ 等。Hu等(2009)在对大豆(*Glycine max*)内参基因的研究中发现, $ACT11$ 的表达稳定性明显高于

$ACT2/7$ 。 $Tubulin$ 基因家族同样作为细胞骨架结构蛋白被用作内参基因, 在对大豆(Hu et al., 2009)和桃(*Amygdalus persica*)(Tong et al., 2009)的研究中发现, TUA 的表达稳定性也普遍高于 TUB 。泛素基因家族在细胞的蛋白质降解中发挥重要作用并被广泛用作内参基因, 主要包括多聚泛素(polyubiquitin)和泛素延伸蛋白(ubiquitin extension protein)2类(胡瑞波等, 2009)。Gutierrez等(2008)分析了拟南芥(*Arabidopsis thaliana*)中内参基因的表达, 发现ubiquitin extension protein($UBQ5$)的表达稳定性明显高于同一家族的polyubiquitin($UBQ4$ 、 $UBQ10$ 、 $UBQ11$)。

2 常用内参基因的特点

2.1 常用内参基因的表达稳定性

在植物的qRT-PCR分析中, 内参基因常常是看家基因, 常用的内参基因主要包括肌动蛋白基因(ACT)、18S核糖体RNA(18S rRNA)、3-磷酸甘油醛脱氢酶基因($GAPDH$)、转录延伸因子基因($EF-1\alpha$)、多聚泛素酶基因(UBQ)、 α 微管和 β 微管蛋白基因(TUA 、 TUB)以及亲环蛋白基因(CYP)等(表1)(Dheda et al., 2004)。在众多的内参基因中, 根据其在实验体系内的表达稳定性选择适当的内参基因是进行标准化的关键, 因此越来越多的研究者对植物内参基因的稳定性进行评价与鉴定。目前, 在豌豆(*Pisum sativum*)(Die et al., 2010)、水稻(*Oryza sativa*)(Jain, 2009)、白杨(*Populus tomentosa*)(Brunner et al., 2004)、甘蔗(*Saccharum sinensis*)(Iskandar et al., 2004)、拟南芥(Remans et al., 2008)、马铃薯(*Solanum tuberosum*)(Nicot et al., 2005)、棉花(*Gossypium hirsutum*)(Tu et al., 2007)、大豆(Jian et al., 2008; Hu et al., 2009)、番茄(Mascia et al., 2010)、草类(Hong et al., 2008)、咖啡(*Coffea arabica*)(Cruz et al., 2009)、小麦(Paolacci et al., 2009)、葡萄(Reid et al., 2006)、桃(Tong et al., 2009)和茶树(*Camellia sinensis*)(孙美莲等, 2010)等植物中都鉴定出了在各自不同的实验条件下表达最稳定的内参基因。

选择正确的内参基因很大程度上取决于所研究的组织、细胞及实验条件等。例如不同的组织器官、生长发育阶段、生物或非生物胁迫和激素影响等都会引起内参基因的特异性表达。因此, 每一特定条件下

表1 常用的植物内参基因及其功能**Table 1** Common reference genes of plants and their function

简称	中文全称	英文全称	功能	适用植物物种及参考文献
ACT	肌动蛋白	Actin	细胞骨架结构蛋白	葡萄(Reid et al., 2006)
18S rRNA	18S核糖体 RNA	18S ribosomal RNA	胞质核糖体小亚基, 翻译	水稻(Kim et al., 2003)
GAPDH	3-磷酸甘油醛脱氢酶	Glyceraldehyde-3-phosphate dehydrogenase	糖酵解、糖异生及光合作用碳固定循环过程中的关键酶	大麦(Faccioli et al., 2007)
EF-1 α	转录延伸因子	Elongation factor 1- α	转录延伸	马铃薯(Nicot et al., 2005)
UBQ	多聚泛素酶	Polyubiquitin	蛋白质修饰、结合、降解	棉花(涂礼莉等, 2007)
TUA	α 微管蛋白	α -tubulin	细胞骨架结构蛋白	番茄(Coker and Davis, 2003)
TUB	β 微管蛋白	β -tubulin	细胞生长, 参与光刺激反应	杨树(Brunner et al., 2004)
CYP	亲环蛋白	Cyclophilin (peptidyl-prolyl cis-trans isomerase)	调节肽基脯氨酸顺反异构酶活性	大豆(Jian et al., 2008)

都有特定的稳定表达的内参基因, 内参基因的选择是否合适需要在不同的实验系统中进行评价与鉴定。Iskandar等(2004)分析了4个内参基因(*ACT*、*TUB*、*GAPDH*、*25S rRNA*)在甘蔗不同组织、器官和不同品种中的表达稳定性, 结果表明, *25S rRNA*的表达水平在所有样品间差异较小, 而在不同组织器官中*GAPDH*的表达稳定性最好。Jian等(2008)分析了10个大豆内参基因(*ACT2/7*、*ACT11*、*TUA*、*TUB*、*EF1 α* 、*UBC2*、*ELF1b*、*CYP2*、*G6PD*、*UBQ10*)在21个不同发育时期和组织器官等样本中的表达稳定性, 结果表明, *ELF1b*和*CYP2*在所有样品中表达都相对最为稳定, 而*ACT2/7*和*TUA*仅在不同发育时期的样品中表达稳定性最好。Mascia等(2010)分析了番茄的8个内参基因(*ACT*、*CYP*、*EF-1 α* 、*GAPDH*、*TUB*、*UBI*、*GAPDH*、*18S rRNA*)在5种不同植物病毒侵染条件下的表达稳定性, 结果表明, *UBI*和*GAPDH*在所有样品中表达都相对稳定, 而*ACT*在侵染样品中表达稳定性最好。Hong等(2008)分析了9个短柄草(*Brachypodium distachyon*)内参基因(*ACT7*、*RCA*、*EF1 α* 、*GAPDH*、*SamDC*、*TUA6*、*UBC18*、*UBi4*、*UBi10*)在21个不同发育时期和组织器官中逆境胁迫(干旱、盐、冷、热)和激素条件下的表达稳定性, 结果表明, *UBC18*在所有样品中表达稳定性最好, *UBi4*和*UBi10*在不同组织器官和激素处理的样品中表达最为稳定, 而在逆境胁迫条件下的样品中*SamDC*的表达稳定性最好。

综上所述, 在不同的实验条件下, 内参基因的表达存在变化; 然而在各自不同的条件下又都有特定的

相对稳定表达的内参基因。因此, 根据具体的样品类型选择适当的内参基因进行标准化是qRT-PCR分析中至关重要的一步。

2.2 常用内参基因的缺陷

长期以来, 人们之所以选择常用内参基因首先是因为它们具有以下功能: (1) 生物体维持生命活动所必需的细胞器骨架的基本组分(*ACT*、*TUA*、*TUB*等); (2) 参与生物体的基本生化代谢过程(*GAPDH*、*EF-1 α* 、*UBQ*等)(具体见表1)。正因为这些特点使得其在某些组织器官中通常可以稳定地表达。其次是因为在以往的基因表达分析中, 研究者主要采用Northern杂交、组织化学染色或RT-PCR等技术, 而这类分析基本上是定性分析或相对定量分析, 由于内参基因表达丰度高, 目标基因表达的上调或下调能够明显地表现出来(胡瑞波等, 2009)。但是qRT-PCR是对基因表达的准确定量分析, 灵敏度较高, 而且有些常用内参基因的表达在不同实验条件下变异较大(Hu et al., 2009; Die et al., 2010), 已经不能满足定量分析对内参基因的要求。最后, 研究者们由于惯性和依赖性一直沿用常用内参基因来进行标准化, 而忽略了其在不同实验条件下的表达是否存在差异。最近, 一些研究结果已经表明, 常用内参基因在某些条件下由于表达不稳定而不适合作为内参基因, 它们已逐渐被新的内参基因所取代。Die等(2010)分析了豌豆的8个常用内参基因和3个新候选内参基因在26个不同组织和不同基因型的豌豆样品中的表达稳定性, 结果表明, *18S rRNA*作为最常用的内参基因表达却最不稳定, 差异最大; 而

GH720838(transcription factor IIA) 和 GH720808(histone H3)作为新基因则显示出了较好的表达稳定性。Hu等(2009)分析了大豆7个常用内参基因和7个新候选基因在116个样品(不同的组织器官、生长发育阶段、光周期处理和栽培种)中的表达稳定性, 新基因中的*SKIP16*、*UKN1*和*UKN2*超越常用的内参基因成为表达最稳定的基因, 而常用内参基因*UBQ10*则表达稳定性最差, 不适合作为内参基因。

3 新内参基因的挖掘

常用内参基因由于表达稳定性差已逐渐被新的内参基因所取代。因此, 新内参基因的挖掘已经成为重中之重。目前, 基于基因芯片的表达数据和EST数据库逐渐成为筛选新内参基因的主要渠道。图1显示了通

过这2种方法结合qRT-PCR分析筛选新的内参基因, 并对其稳定性进行评价与鉴定的主要流程。

3.1 基于基因芯片的筛选

基因芯片又称为DNA微阵列(DNA microarray), 随着基因芯片技术越来越成熟, 大量的基因芯片数据为新内参基因的挖掘提供了良好的数据基础和平台。Czechowski等(2005)利用Affymetrix ATH1全基因组基因芯片数据, 对模式植物拟南芥的内参基因进行评价和鉴定, 发现了数以百计的新基因, 其表达稳定性均优于常用内参基因, 筛选出在不同发育时期、逆境(盐分、干旱、冷害)和生物胁迫、激素影响、营养(硫素、碳源、磷)缺乏以及不同光照条件下都稳定表达的22个基因, 并将其与5个常用内参基因(*ACT2*、*TUB6*、*EF-1 α* 、*UBQ10*、*GAPDH*)在不同的实验条

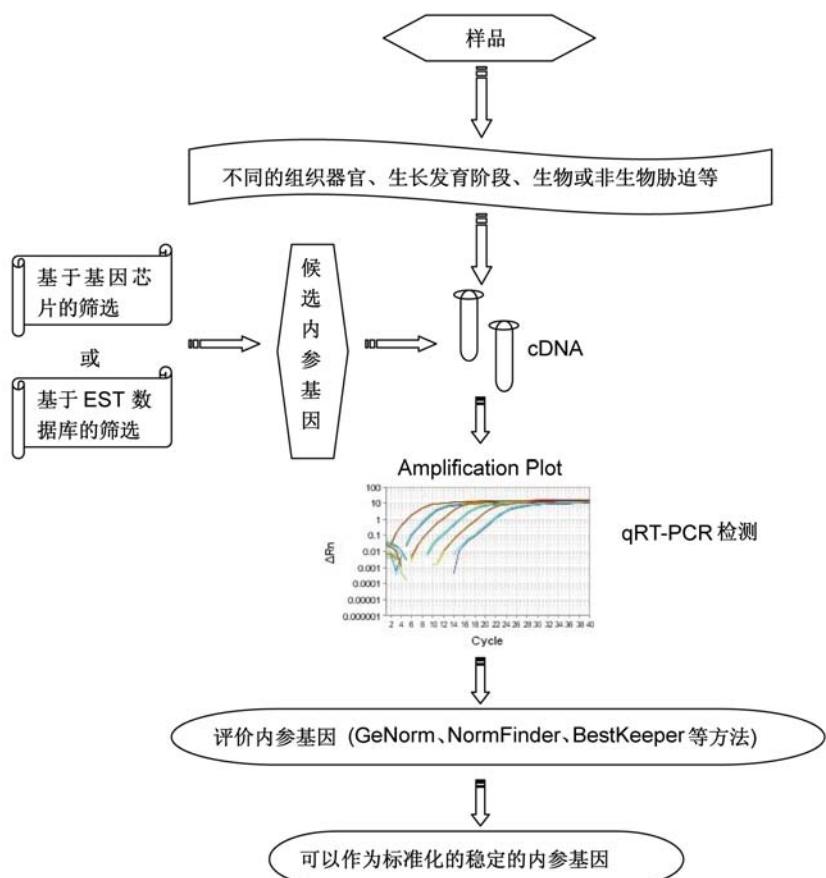


图1 qRT-PCR内参基因稳定性的鉴定

Figure 1 Identification of reference genes for qRT-PCR normalization

件下进行比较。结果表明新基因的表达稳定性普遍优于常用内参基因, 例如*At1g13320(PP2A)*在不同的样品中均显示出较高的表达稳定性, 而在以往研究中常用的内参基因*ACT2*则表现出较差的表达稳定性, 不适合用作内参基因。*Long*等(2010)从小麦Affymetrix基因芯片中选出32个基因作为候选内参基因, 在不同的组织器官、生长发育阶段以及不同的实验条件下通过qRT-PCR技术进行分析, 筛选出表达稳定性较高的15个新基因, 并将其与13个常用内参基因相比较。结果表明, 在所有样品中7个新基因的表达稳定性优于常用内参基因, 可以作为新的内参基因; 而在常用内参基因中只有*EF-1 α* 表现出较高的表达稳定性, 18S rRNA、*TUA*、*ACT*和*GAPDH*的表达稳定性均较差, 不适合作为内参基因。

基于基因芯片技术筛选新的内参基因将为新基因的挖掘提供良好的数据基础, 也为内参基因的选择提供了更广阔的空间。但是近年来使用基因芯片数据筛选新基因的研究较少, 且研究成果并没有引起足够的重视, 许多研究者仍然沿用常用基因作为内参, 这成为目前基因表达分析研究的一个弊端。

3.2 基于EST数据库的筛选

表达序列标签(expression sequence tag, EST)数据库作为一种重要的基因组数据库, 已成为新基因发现、基因及重组蛋白表达等研究中一种有力的分子生物学工具。*Coker*和*Davis*(2003)利用番茄EST数据库选择了127个基因, 在不同的组织器官样品中进行分析, 筛选出5个表达稳定性最好的基因(*Dan-J*蛋白基因、翻译控制肿瘤蛋白(*TPCP*)基因、*TUA*、*CYP*和*GAPDH*), 适合作为该实验条件下的内参基因。*Faccioli*等(2007)对大麦(*Hordeum vulgare*)EST数据库中的序列进行分类, 从中寻找在特定cDNA文库中出现频率高的EST序列, 从而筛选出表达稳定性较好的候选基因, 并通过qRT-PCR分析得到5个表达稳定性较高、适合用作标准化的内参基因(*S2*腺苷蛋氨酸脱羧酶基因、*GAPDH*、热激蛋白hsp90基因、翻译控制肿瘤蛋白*TPCP*和一个未知功能基因TC-139056)。

植物学领域中应用EST数据库筛选新内参基因的研究相对较少, 可能是由于目标范围较广, 筛选出的新基因数量较少以及新基因在不同实验条件下表

达水平差异较大等造成的。因此, 仍需寻找更精确有效的方法, 挖掘新的内参基因以代替不适合的常用内参基因, 这还有待于进一步研究。

3.3 新内参基因的表达稳定性评价

基于基因芯片技术和EST数据库筛选可以获得新的内参基因, 但是这些新内参基因的表达稳定性尚不确定, 必须用qRT-PCR等方法进行评价, 并与传统内参基因的表达稳定性进行比较。*Libault*等(2008)通过分析已发表的大豆基因芯片数据, 发现了多达200个表达稳定性较好的基因, 并对其中18个基因在不同发育时期、组织器官、逆境胁迫(创伤、激素、锈病)下的130个样本进行了qRT-PCR分析, 结果表明多数基因的表达稳定性优于传统内参基因*ACT*, 其中肽酶*S16*基因和呼吸爆发氧化酶基因的稳定性最差, 最后筛选出了4个新的内参基因, 分别为ABC转运因子基因、F-box家族基因、金属蛋白酶基因和CDPK蛋白激酶基因。说明通过基因芯片数据筛选到的候选基因必须经过qRT-PCR分析验证, 并与常用内参基因的表达稳定性相比较, 才能确定新的内参基因究竟能不能应用于qRT-PCR结果的校正和标准化。表2列出了目前已开发的一些新的内参基因。qRT-PCR分析证明, 在相应的样品类型中它们的表达稳定性均优于传统的内参基因。

4 应用内参基因组合的优越性

在qRT-PCR分析中, 内参基因的使用被认为是最适当的标准化方法(Nolan et al., 2006), 使用不理想的内参基因会对数据的分析造成偏差(Dheda et al., 2005)。但是使用单一的内参基因进行校准和标准化, 也会对结果的精确性产生影响(Zhu et al., 2008)。*Schmid*等(2003)认为在一组给定的样本或实验条件下, 平行测定2个或2个以上内参基因有助于校正系统偏差。为了消除使用单一内参基因所引起的偏差和波动, *Vandesompele*等(2002)建议在标准化过程中使用至少3个内参基因进行校正。*Hu*等(2009)对大豆内参基因在不同实验条件下的稳定性进行分析, 结果表明: 在不同组织器官中, 内参基因组合(*ACT11*、*UKN1*、*UKN2*)是最适当的; 在不同生长发育阶段, 内参基因组合(*SKIP16*、*TUA5*、*MTP*)为最优组合。同

表2 植物新内参基因的评价**Table 2** Evaluation of new reference genes in plants

植物物种	样品类型	新内参基因	参考文献
拟南芥	发育时期, 组织器官, 营养缺乏(硫、磷、碳源), 非生物胁迫(激素、冷害、甘露醇、盐)	<i>SAND</i> (At2G28390) <i>TIP41</i> (At4G34270) 未知功能基因(At4G26410)	Czechowski et al., 2005
	逆境胁迫(铜、镉)	F-box(At5G15710) <i>SAND</i> (At2G28390) <i>YLS8</i> (At5G08290)	Remans et al., 2008
大麦	组织器官, 高温胁迫	S2腺苷蛋氨酸脱羧酶基因 热激蛋白hsp90基因 未知功能基因(TC139056)	Faccioli et al., 2007
番茄	发育时期, 组织器官	<i>TIP41</i> (At4G34270) <i>SAND</i> (At2G28390) SGN-U346908	Expósito-Rodríguez et al., 2008
大豆	发育时期, 组织器官, 逆境胁迫(创伤、激素、锈病)	ABC转运因子基因 F-box家族基因 金属蛋白酶基因 CDPK蛋白激酶基因	Libault et al., 2008
杨树	形成层	<i>TIP41</i> (At4G34270)	Gutierrez et al., 2008
葡萄	发育时期	<i>SAND</i> (At2G28390)	Lossos et al., 2003

样, 在葡萄(Reid et al., 2006)、马铃薯(Nicot et al., 2005)和白杨(Brunner et al., 2004)中的研究结果也证实, 使用内参基因组合有助于获得更准确的表达分析结果。但是在使用不同的内参基因组合时, 有几点需要注意的基本要求: (1) 内参基因组合中的基因在同一实验条件下其表达丰度应当相近; (2) 应选择退火温度一致的内参基因进行组合; (3) 选择同源基因组合, 在设计引物时应考虑到引物之间的二聚体和发夹结构。总之, 选择不同的内参基因组合, 就必须控制组合内部内参基因的表达差异和扩增的准确性等, 这样才能更好地进行校正和标准化。

5 评价内参基因稳定性的方法

5.1 扩增效率的评价及其软件

在对内参基因的表达稳定性进行评价之前, 必须先通过qRT-PCR来分析内参基因的扩增效率(E)。目前, 主要有2种软件可用于扩增效率的评价。第1种是比利时Biogazelle公司的qBase软件。由于这款软件属于商业化软件, 应用的范围小, 因此在评价扩增效率时大多数人会采用第2种方法——Excel软件的标准曲线法。具体程序是首先运用Excel表格根据qRT-PCR的

结果(Ct 值)计算内参基因的相关系数(R^2), 然后根据公式 $\lg(1+E)=R^2$ 得出扩增效率的值。目前认为只有扩增效率在一定的范围内(90%–105%)时, 才符合进行下一步评价的要求。

5.2 评价内参基因稳定性的方法

在qRT-PCR分析过程中, 评价作为校准和标准化的内参基因有多种分析方法, 包括2种简单数据分析(稳定指数分析、 ΔCt 值分析)和3种基于统计学的软件分析方法(GeNorm、NormFinder、BestKeeper)。其中后者为研究者所常用。

稳定指数分析是一种基于单因素方差分析和线性回归分析的统计学方法, 与植物育种学家评价不同植物品种产量的相对稳定性值的方法类似。通过MSE(均方误差)、MEAN- Ct (qRT-PCR过程中循环数的平均值)、CV(变异系数)和Slope(线性回归方程的斜率)等数值推算出稳定指数。稳定指数越低, 内参基因的表达越稳定。在鉴定植物内参基因的表达稳定性方面, Brunner等(2004)首次使用该分析方法对杨树内参基因进行评价和筛选, 其中UBQ的稳定指数最低, 表明其表达稳定性最高。 ΔCt 值分析是通过比较2个内参基因之间的 ΔCt (qRT-PCR过程中循环数)值来筛选

表达最稳定的内参基因的分析方法。如果在所有的样品中2个内参基因之间的 ΔCt 值没有差异, 就表明这2个基因均稳定表达; 反之则表明至少有1个基因表达不稳定, 需要引入第3个或更多的内参基因来进行两两比较, 直至筛选出 ΔCt 值最小的一对基因继而筛选出表达最稳定的基因。Silver等(2006)在评价人网织红细胞中内参基因的实验中首次提出这种方法, 分析结果表明GAPDH为表达最稳定的基因。

GeNorm程序是Vandesompele等(2002)开发的专门用于分析内参基因稳定性的程序。通过计算内参基因表达稳定性的M值进而分析内参基因在不同样品中的表达稳定性, M值越大稳定性越差; 反之, 稳定性越好。还可以对标准化因子(normalization factor)的配对变异(pairwise variation)V值进行分析, 以确定所需内参基因的最适数目。将 $Vn/Vn+1$ 比值与默认的V值(0.15)进行比较, 如果 $Vn/Vn+1$ 大于0.15, 则有必要引入第n+1个基因; 反之, 则不必引入新的内参基因。GeNorm程序可以用于确定不同实验条件下最适内参基因的数量, 并选择出最优组合, 而不是通常地使用单一内参基因。这将有利于校正系统偏差, 从而获得更可靠的表达分析结果, 对于研究基因表达的细微变化具有极其重要的意义。NormFinder程序是Andersen等(2004)编写的用于选择合适内参基因的程序, 其基于方差分析评价内参基因在不同样品中的表达稳定性。NormFinder的运行原理与GeNorm程序类似, 也是通过计算基因表达稳定性, 然后根据稳定性大小排序, 值越大稳定性越差; 反之, 稳定性越好。但是不足的是, 它只能选择1个最合适的内参基因。而GeNorm程序则能够选出合适的内参基因组合及最适基因数量。BestKeeper程序是Pfaffl等(2004)开发的用于分析基因表达稳定性及基因表达水平的程序, 可以对内参基因和目标基因进行单独分析。BestKeeper程序主要通过比较Ct值的标准偏差和变异系数来选择表达最稳定的基因, 标准偏差和变异系数越小, 稳定性越好; 反之, 稳定性越差。其不同于GeNorm和NormFinder的优点是不但可以分析内参基因表达的稳定性, 而且可以比较目标基因的表达水平。

6 小结

总之, 目前还没有一种理想的内参基因可以完全用于

不同类型的样品, 且其表达恒定不变。内参基因一般都表现为在某些实验条件下表达较稳定, 但在其它实验系统中则是变化的。所以研究目标基因的表达情况时, 研究者需要结合各自的实验条件和样品类型, 根据物种、基因家族等的不同, 选择合适而稳定的内参基因进行校正和标准化, 才能有助于得到可靠的实验结果。

参考文献

- 胡瑞波, 范成明, 傅永福 (2009). 植物实时荧光定量PCR内参基因的选择. 中国农业科技导报 11, 30–36.
- 孙美莲, 王云生, 杨冬青, 韦朝领, 高丽萍, 夏涛, 单育, 骆洋 (2010). 茶树实时荧光定量PCR分析中内参基因的选择. 植物学报 45, 579–587.
- 涂礼莉, 张献龙, 刘迪秋, 金双侠, 曹景林, 朱龙付, 邓锋林, 谭家福, 张存斌 (2007). 棉花纤维发育和体细胞胚发生过程中实时定量PCR内对照基因的筛选. 科学通报 52, 2379–2385.
- Andersen CL, Jensen JL, Ørntoft TF (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64, 5245–5250.
- Brunner AM, Yakovlev IA, Strauss SH (2004). Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol* 4, 14.
- Coker JS, Davis E (2003). Selection of candidate housekeeping controls in tomato plants using EST data. *Biotechniques* 35, 740–742, 744, 746.
- Cruz F, Kalaoun S, Nobile P, Colombo C, Almeida J, Barros LMG, Romano E, Grossi-de-Sá MF, Vaslin M, Alves-Ferreira M (2009). Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. *Mol Breeding* 23, 607–616.
- Czechowski T, Stitt M, Altman T, Udvardi MK, Scheible WR (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* 139, 5–17.
- Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A (2004). Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37, 112–119.
- Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA,

- Johnson MA, Rook GAW, Zumla A** (2005). The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Anal Biochem* **344**, 141–143.
- Die JV, Román B, Nadal S, González-Verdejo CI** (2010). Evaluation of candidate reference genes for expression studies in *Pisum sativum* under different experimental conditions. *Planta* **232**, 145–153.
- Expósito-Rodríguez M, Borges AA, Borges-Pérez A, Pérez JA** (2008). Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol* **8**, 131–142.
- Faccioli P, Ciceri GP, Provero P, Stanca AM, Morcia C, Terzi V** (2007). A combined strategy of “in silico” transcriptome analysis and web search engine optimization allows an agile identification of reference genes suitable for normalization in gene expression studies. *Plant Mol Biol* **63**, 679–688.
- Gachon C, Mingam A, Charrier B** (2004). Real-time PCR: what relevance to plant studies? *J Exp Bot* **55**, 1445–1454.
- Guénin S, Mauriat M, Pelloux J, van Wuytsinkel O, Bellini C, Gutierrez L** (2009). Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot* **60**, 487–493.
- Gutierrez L, Mauriat M, Guénin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerineau F, Bellini C, van Wuytsinkel O** (2008). The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* **6**, 609–618.
- Hong SY, Seo PJ, Yang MS, Xiang FN, Park CM** (2008). Exploring valid reference genes for gene expression studies in *Brachypodium distachyon* by real-time PCR. *BMC Plant Biol* **8**, 112.
- Hu RB, Fan CM, Li HY, Zhang QZ, Fu YF** (2009). Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Mol Biol* **10**, 93.
- Huggett J, Dheda K, Bustin S, Zumla A** (2005). Real-time RT-PCR normalisation: strategies and considerations. *Genes Immun* **6**, 279–284.
- Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, Manners JM** (2004). Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane. *Plant Mol Biol Rep* **22**, 325–337.
- Jain M** (2009). Genome-wide identification of novel internal control genes for normalization of gene expression during various stages of development in rice. *Plant Sci* **176**, 702–706.
- Jian B, Liu B, Bi YR, Hou WS, Wu CX, Han TF** (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol Biol* **9**, 59.
- Kim BR, Nam HY, Kim SU, Kim SI, Chang YJ** (2003). Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotechnol Lett* **25**, 1869–1872.
- Libault M, Thibivilliers S, Bilgin DD, Radwan O, Benitez M, Clough SJ, Stacey G** (2008). Identification of four soybean reference genes for gene expression normalization. *Plant Genome* **1**, 44–54.
- Long XY, Wang JR, Ouellet T, Rocheleau H, Wei YM, Pu ZE, Jiang QT, Lan XJ, Zheng YL** (2010). Genome-wide identification and evaluation of novel internal control genes for Q-PCR based transcript normalization in wheat. *Plant Mol Biol* **74**, 307–311.
- Lossos IS, Czerwinski DK, Wechsler MA, Levy R** (2003). Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. *Leukemia* **17**, 789–795.
- Mascia T, Santovito E, Gallitelli D, Cillo F** (2010). Evaluation of reference genes for quantitative reverse-transcription polymerase chain reaction normalization in infected tomato plants. *Mol Plant Pathol* **11**, 805–816.
- Nicot N, Hausman JF, Hoffmann L, Evers D** (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J Exp Bot* **56**, 2907–2914.
- Nolan T, Hands RE, Bustin SA** (2006). Quantification of mRNA using real-time RT-PCR. *Nat Protoc* **1**, 1559–1582.
- Paolacci AR, Tanzarella OA, Porceddu E, Ciuffi M** (2009). Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol Biol* **10**, 11.
- Pfaffl MW** (2006). Relative quantification. In: Dorak MT, ed. Real-time PCR. New York: International University Line. pp. 63–82.
- Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP** (2004). Determination of stable housekeeping genes, differentially

- regulated target genes and sample integrity: BestKeeper-excel-based tool using pair-wise correlations. *Biotechnol Lett* **26**, 509–515.
- Quackenbush J** (2002). Microarray data normalization and transformation. *Nat Genet* **32**, 496–501.
- Reid KE, Olsson N, Schlosser J, Peng F, Lund ST** (2006). An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biol* **6**, 27.
- Remans T, Smeets K, Opdenakker K, Mathijssen D, Vangronsveld J, Cuypers A** (2008). Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta* **227**, 1343–1349.
- Schmid H, Cohen CD, Henger A, Irrgang S, Schlöndorff D, Kretzler M** (2003). Validation of endogenous controls for gene expression analysis in microdissected human renal biopsies. *Kidney Int* **64**, 356–360.
- Silver N, Best S, Jiang J, Thein SL** (2006). Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol* **7**, 33.
- Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E** (1999). Housekeeping genes as internal standards: use and limits. *J Biotechnol* **75**, 291–295.
- Tong ZG, Gao ZH, Wang F, Zhou J, Zhang Z** (2009). Selection of reliable reference genes for gene expression studies in peach using real time PCR. *BMC Mol Biol* **10**, 71.
- Tu LL, Zhang XL, Liu DQ, Jin SX, Cao JL, Zhu LF, Deng FL, Tan JF, Zhang CB** (2007). Suitable internal control genes for qRT-PCR normalization in cotton fiber development and somatic embryogenesis. *Chin Sci Bull* **52**, 3110–3117.
- Udvardi MK, Czechowski T, Scheible WR** (2008). Eleven golden rules of quantitative RT-PCR. *Plant Cell* **20**, 1736–1737.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F** (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, 0034.1–0034.11.
- Volkov RA, Panchuk II, Schöffl F** (2003). Heat-stress dependency and developmental modulation of gene expression: the potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. *J Exp Bot* **54**, 2343–2349.
- Wong ML, Medrano JF** (2005). Real-time PCR for mRNA quantitation. *Biotechniques* **39**, 75–85.
- Zhu J, He FH, Song SH, Wang J, Yu J** (2008). How many human genes can be defined as housekeeping with current expression data? *BMC Genomics* **9**, 172.

Characterization and Selection of Reference Genes for Real-time Quantitative RT-PCR of Plants

Wei Yuan^{1,2}, Hongjian Wan¹, Yuejian Yang^{1*}

¹Institute of Vegetables, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China

²College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China

Abstract Real-time quantitative RT-PCR (qRT-PCR) is one of the most common technologies used for gene expression and transcriptome analysis, with its high sensitivity, specificity, good reproducibility, wide dynamic quantification range and high-throughput capacity. Selecting the appropriate reference genes is the first step in analyzing the expression of genes of interest. Selecting appropriate reference genes depends on experimental conditions, and selection of reference genes changes after the experimental conditions. Therefore, the accuracy of results from qRT-PCR analysis largely depends on the reference genes used. In this paper, we give a comprehensive summary of reference genes for qRT-PCR, including their selection, characteristics of traditional reference genes, mining new reference genes, the advantage of combining different reference genes, and how to assess stable reference gene expression. The results provide a theoretical foundation for selecting appropriate reference genes for qRT-PCR of plants.

Key words gene expression evaluation, gene expression stability, qRT-PCR, reference genes

Yuan W, Wan HJ, Yang YJ (2012). Characterization and selection of reference genes for real-time quantitative RT-PCR of plants. *Chin Bull Bot* 47, 427–436.

* Author for correspondence. E-mail: hzyj@yahoo.com.cn

(责任编辑: 刘慧君)