

基于高通量测序的数字基因 表达谱技术研究进展

张 渝, 刘玉洁, 郭丹妮, 李惠敏, 秦新民

(广西师范大学 生命科学院, 珍稀濒危动植物生态与环境保护教育部重点实验室, 广西 桂林 541004)

摘 要:数字基因表达谱(digital gene expression profiling, DGE)是一种全面、快速地检测某一物种特定组织在特定状态下基因表达情况的高通量测序技术。通过对基因表达谱的生物信息学搜索、比较和分析,可从中获取基因转录、基因调控、信号转导通路、蛋白质功能等相关信息。该研究主要介绍了数字基因表达谱的原理以及在生物学和医学研究等方面的应用。

关键词:数字基因表达谱;高通量测序技术;差异表达基因

中图分类号:Q 343 **文献标识码:**A **文章编号:**1001-0009(2015)10-0170-07

数字基因表达谱(digital gene expression profiling, DGE)是指通过构建处于某一特定状态下的细胞或组织的非偏性 cDNA 文库,大规模 cDNA 测序,收集 cDNA 序列片段、定性、定量分析其 mRNA 群体组成,从而描绘该特定细胞或组织在特定状态下的基因表达种类和丰度信息^[1]。该技术可以全面、经济、快速地检测某一物种特定组织在特定状态下的基因表达情况,通过生物信息学搜索、查询、比较和分析,人们可从中获取基因的转录水平、样品间差异基因的表达模式、编码蛋白质的功能以及相互联系等信息。

目前研究基因表达谱的技术主要有基因芯片和高通量测序等^[2],与基因芯片相比,数字基因表达谱是一种以新一代高通量测序为基础,对标签直接进行测序的技术,更适用于检测基因低丰度转录以及基因细微变化。因此,已被广泛应用于基础科学研究、医学研究和药物研发等领域。

1 数字基因表达谱的试验流程和原理

数字基因表达谱(DGE)技术的基本原理(图 1)是利用对 mRNAs 经反转录所得 cDNAs 进行双酶切,使得一条 mRNA 得到一个相对应的标签,而后进入高通量测

序和分析流程,经过生物信息均一化后比较不同样本间各种标签条数,找出差异表达标签。

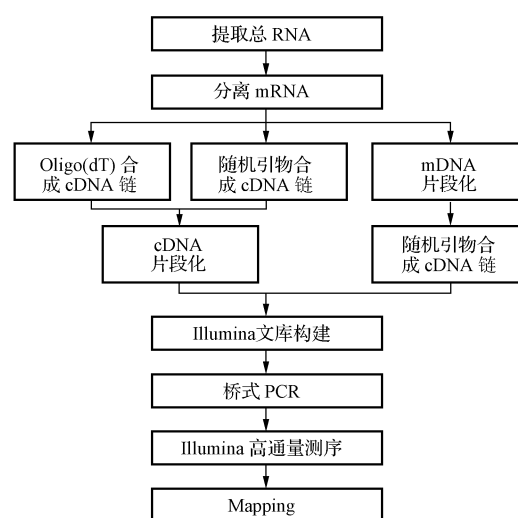


图 1 DGE 技术原理

Fig. 1 The technical principle of digital gene expression profile

DGE 得到的差异表达的标签经过图像识别、碱基识别、数据量产出的统计等处理后,将得到数据进行试验饱和度分析、基因表达量的确定、差异表达基因的检测、差异基因表达模式聚类分析、基因 Ontology 功能显著性富集分析、Pathway 显著性富集分析、蛋白相互作用网络分析、新转录本预测(图 2)。最后,通过样品间的对比,得到基因的差异表达结果。在具有参考基因组的物种和一些没有参考基因组但有单一基因序列集数据库的物种中,可以直接进行 DGE 的应用,而在既没有参考基因组又没有单一基因序列集数据库的物种中,需要先

第一作者简介:张渝(1987-),女,硕士研究生,研究方向为植物基因组学研究。E-mail:1290037351@qq.com.

责任作者:秦新民(1956-),男,教授,现主要从事植物分子生物学等研究工作。E-mail:xmqin@mailbox.gxnu.edu.cn.

基金项目:国家自然科学基金资助项目(31360477);广西教育厅资助项目(2013YB036)。

收稿日期:2015-01-26

采用 RNA-seq 技术构建一个本物种的单一基因序列集数据库,用作 DGE 所得标签的对比参考数据库。

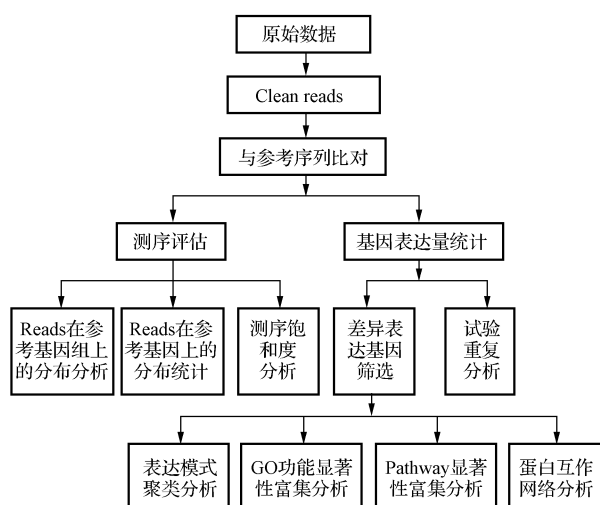


图2 DGE 信息分析流程

Fig. 2 Procedure of bioinformatics analysis for digital gene expression profile

2 数字基因表达谱的特点

DGE 技术与传统基因表达谱芯片相比。有着其独特的优越性:①高通量,一次检验可测出 500 万 reads,几乎涵盖细胞中所有表达的基因;②数字化,测序获得的数字信号具有直接性,避免了噪音和交叉杂交影响。分辨率达到可鉴别序列之间 1 个碱基的差异;③待测序列无需事先设计,具有很好的实时性;④ 2 次独立试验的相关性超过 99%,重复性高;⑤可检测出新转录本,不依赖于已知基因;⑥检测范围可跨越 6 个数量级的检测阈值,从 2 个拷贝到几十万拷贝的不同丰度转录本都能识别。

但 DGE 技术也存在一些不足:如需要较多的样品量,产生的结果数据量庞大,需要具备一定的生物信息学分析基础。

3 数字基因表达谱的应用

DGE 技术因具有低成本、快速、灵敏度高的特点,特别适合于不同样品中的差异表达基因研究,因此已被广泛应用于生命科学和医学等研究领域^[3-5]。

3.1 在动物方面的应用

2009 年,Zoltán 等^[6]用基因表达谱技术研究了被分支杆菌侵染的斑马鱼的转录组变化,从每个样品中得到了上百万的转录标签,发现了被分支杆菌感染后诱导产生的新基因和新的转录机制,更好的阐明了与传染病有关的转录组反应。

Xiao 等^[7]应用数字基因表达谱技术研究猪在感染 N-PRRSV 的 CH1a 株病毒后,体内出现了多种破坏机

制,包括破坏宿主细胞的先天免疫反应、抗凋亡和抗炎能力。受病毒诱导后表达上调的早期炎性相关细胞因子、趋化因子、粘附分子、发炎相关酶蛋白、发炎细胞、抗体以及补体等很可能导致了 N-PRRSV 感染过程中发炎响应机制的发生和发展。

罗非鱼是全世界重要的淡水养殖鱼类。链球菌性的疾病已经严重威胁罗非鱼水产养殖业的发展,但是,到目前为止还没有有效地抑制和控制罗非鱼感染链球菌导致的疾病的方法。Zhong 等^[8]采用 RNA-Seq 技术和数字基因表达谱技术研究罗非鱼感染链球菌前后转录组的变化,774 个 unigenes 表达上调,625 个 unigenes 表达下调。GO、KEEG 分析发现 172 个差异表达的基因与免疫应答有关且涉及 17 个信号通路,TLR 和 NLR 信号通路是识别革兰氏阳性致病菌重要的信号通路。TLR 是跨膜蛋白,它识别 PAMPs,启动宿主天然的免疫应答机制并形成抵抗该病原菌的获得性免疫。成年的斑马鱼感染海鱼分枝杆菌 8 周后 TLR1 和 TLR2 的表达量增加。免疫应答相关的差异表达基因的鉴定和信号通路的研究有助于进一步了解罗非鱼抵抗链球菌类的致病菌免疫应答机制,也为预防该疾病在罗非鱼中的传播提供了思路。

Mu 等^[9]使用 DGE 技术,对遭受嗜水气单胞菌感染的大黄鱼和对照大黄鱼对比分析,发现 1 996 个基因呈现差异表达,其中有 727 个上调表达的基因和 489 个下调表达的基因差异在 1.5 倍以上,而差异变化倍数较大的基因多是炎症反应相关基因。

太平洋牡蛎在夏季由于致病微生物的影响死亡率很高。de Lorgeril 等^[10]研究牡蛎分别在致病性弧菌和无致病性弧菌感染条件下基因的表达变化。发现致病性弧菌感染状态下存活的牡蛎中,上调表达的基因与免疫应答的信号转导、细胞的粘附和通讯、吞噬作用的细胞过程和防御机制、细胞骨架肌动蛋白的重组、细胞的示踪和自噬及抗氧化和抗细胞凋亡有关。结果为牡蛎与病原体之间的相互作用和病发机理的研究开辟了道路。

2012 年,罗庆苗等^[11]利用数字基因表达谱技术探讨小尾寒羊 2 个品种之间繁殖力差异的分子生物学机理。结果表明,处于发情期的 FecB^B FecB^B 型比 FecB⁺ FecB⁺ 型小尾寒羊卵巢组织差异表达基因为 238 条,主要涉及类固醇生物合成与代谢过程、甾醇生物合成与代谢过程、脂质运输与定位、泛素特异的蛋白酶活性、GTP 酶调节活性和 Ras 蛋白信号转导等通路。并初步认为基因 STAR、FLCN、TGFI1、INHA、INHBA 与 Notch 信号通路可能与小尾寒羊的繁殖能力有关。

秦宁^[12]通过 DGE 技术对 2 种猪背最长肌组织 cDNA 文库进行了差异表达基因检测,共筛选到 4 522 个差异

表达基因,其中,上调基因 4 176 个,下调基因 346 个。这些差异表达基因参与了 198 个通路,包括了 RNA 剪接、蛋白质运输、代谢过程、免疫应答、高分子代谢过程、RNA 代谢过程和细胞代谢等。高致病性的生殖和呼吸系统综合征(PRRS)会导致各年龄段猪的高发病率和死亡率。Xiao 等^[13]用 DGE 技术和组织病理学的方法分析猪感染 PRRS 病毒后的肺部基因表达谱与其病理学之间的关系。结果表明异常的宿主对 H-PRRSV 的先天性免疫应答和抗凋亡的诱导有利于 H-PRRSV 的复制和传播。H-PRRSV 的快速繁殖引起促炎性细胞因子和趋化因子异常持续地表达,这些因子能够引起的炎症反应-细胞死亡和抗氧化损伤,最终导致严重的组织损伤和高致病性。该研究能进一步了解 H-PRRSV 的发病机理提纲了分子依据。

3.2 在植物方面的应用

2006 年,Cheung 等^[14]率先将高通量测序技术用于苜蓿表达序列标签测序。随后,Emrich 等^[15]通过 DGE 技术在玉米的基端分生组织中发现了近 400 个玉米特有的转录本,其中 70% 的基因都不存在之前的 cDNA 文库中。

Wang 等^[16]利用 DGE 技术分析了玉米维管束和叶肉细胞中与 C4 光合作用相关基因在发育过程中的表达,从超过 2 700 个表达基因中,确定了几组差异表达基因,这些基因主要参与了光合作用、糖代谢、激素信号转导、氧化还原反应、细胞壁合成等生理过程。

Wu 等^[17]比较了黄瓜的雌性突变株和野生型雌雄同株的基因表达谱,发现差异表达基因参与了植物激素信号转导途径,表明激素及其信号网路在黄瓜花性别决定中起着至关重要的作用。

MiRNAs 可在转录后水平调控基因的表达。Nisha 等^[18]采用 DGE 技术鉴定和描述了土豆 miRNAs 在不同的营养组织和块茎发育的 4 个阶段的特性。总共获得 89 个保守的 miRNAs,其中一些 miRNAs 呈现出在组织或发育阶段特异性的表达。

关于农作物在环境胁迫时基因的应答情况,Nishiyama 等^[19]利用高通量技术研究了拟南芥突变体在盐胁迫下的响应,发现细胞分裂素在胁迫环境下起了重要的调节作用。Min 等^[20]也发现棉花在高温胁迫下酪蛋白激酶 I、糖和生长素是响应高温胁迫的调节子。

坛紫菜是沿海潮间带重要的经济藻类,赖晓娟等^[21]对坛紫菜在高温胁迫下基因表达差异进行研究。通过差异表达基因比较,发现高温胁迫下坛紫菜的基因仅有少部分是上调表达,其中以 HSP 家族、核糖体蛋白 L12、延伸因子 EF-Tu 及部分光合作用相关蛋白为主,而下调表达的基因数量较多。Gene Ontology 功能富集分析结果显示,差异表达的基因涉及核酸代谢、蛋白质翻译和

修饰、糖代谢、脂质代谢、有机酸代谢、萜类等代谢过程。

Qi 等^[22]研究了黄瓜在水涝条件下 24 h 中 5 个时间点的基因表达情况,发现在水涝情况下,许多基因表达下调,差异基因主要与黄瓜的碳代谢、光合作用、活性氧的生成或清除、植物激素的合成有关。结果揭示了黄瓜在水涝情况下应答转录的综合机制。

大豆是重要的农作物,其生长和发育阶段对水分十分敏感。干旱会使大豆的产量和品质受到严重的影响。Chen 等^[23]分析 2 种大豆-耐旱型“金豆 21”和干旱敏感“中豆 33”在干旱和再复水条件下根和叶基因的表达情况。结果发现差异基因主要集中在代谢过程、胁迫应答、植物激素的信号转导、蛋白质加工和植物病原体的相互作用途径。分析表明 7 个差异表达显著的候选基因 *Glyma15g03920*、*Glyma05g02470*、*Glyma15g15010*、*Glyma05g09070*、*Glyma06g35630*、*Glyma08g12590* 和 *Glyma11g16000*,可能决定大豆的耐旱性,这为培育耐旱性的大豆品种奠定了基础。

草甘膦是广谱性除草剂。Jiang 等^[24]用 DGE 技术分析了草甘膦处理过的大豆顶芽的基因表达情况。结果显示草甘膦抑制大豆顶芽莽草酸途径中色氨酸的生物合成,这是草甘膦作用的靶位点。草甘膦可抑制靶位点基因的表达,而对这些靶位点的基因启动子序列分析显示光应答因子是草甘膦诱导途径中的调控因子。结果助于大豆对草甘膦敏感机制进一步的研究。

ATP 合酶 CF1 α 亚蛋白是植物界能量代谢途径中的关键酶,在多细胞的生物过程中起着重要的作用。Shao 等^[25]采用 DGE 技术对甘薯不同器官的 *atpA* 基因的表达差异进行了分析。结果显示,*atpA* 基因在幼叶中的表达量最高,在成熟叶、茎和块根的表达量较低。

2014 年,Lai 等^[26]研究甘薯尿苷二磷酸葡萄糖脱氢酶(UDPGH)基因的表达情况,筛选出了 23 条不同片段长度的 UDPGH contig。通过设计引物扩增编码甘薯 UDPGH 的基因。结果得到 5 种编码 UDPGH 基因(*1bUDPGH1-1bUDPGH5*)。5 种 UDPGH 基因差异表达的分析结果显示,*1bUDPGH1* 在块根的表达量最高,在幼叶和成熟的叶子中的表达量较低,而在须根和茎中的表达量最低。然而 *1bUDPGH2* 和 *1bUDPGH5* 在茎中的表达量最高,根中的表达量较低,叶子中的表达量最低。半定量的 RT-PCR 验证结果与数字基因表达谱的结果一致。这表明 UDPGH 基因的高表达可能有助于甘薯茎和根中细胞比多糖的积累。

类黄酮是植物体内最重要的次生代谢物之一,在植物的生长、发育、抗逆境胁迫等多种生物进程中发挥着重要的作用,同时对人类的健康也具有多种保健与药用功能^[27-28]。韦康等^[29]利用数字基因表达谱技术研究茶树类黄酮合成、调控相关的基因在花瓣、花粉、休眠芽、

萌芽中的表达情况,结果得到了 12 个类黄酮合成关键酶相关基因,为深入分析茶树类黄酮调控机制打下基础。

何首乌合成的化合物 2,3,5,4'-四羟基对二苯基乙炔-2-O- β -D-葡萄糖苷(2,3,5,4'-tetrahydroxy stilbene-2-O- β -d-glucoside, THSG)具有药效作用。但是,详细的生物合成途径的机制仍然不清楚。为了解它的合成机制,Zhao 等^[30]运用 DGE 技术和 RNA-seq 技术研究 4 个不同的何首乌组织转录组的差异。聚类分析和代谢通路分析的结果表明,这些差异基因与 THSG 的生物合成有关。这为进一步了解 THSG 生物合成途径的机制奠定了基础。

Hao 等^[31]对红豆杉(*Taxus mairei*)使用 DGE 技术分析了红豆杉的根、茎、叶 3 种组织的基因差异表达情况,鉴定出一批组织特异性功能相关基因和紫杉烷生物合成途径相关基因。Tang 等^[32]对开花 3、50、70 d 罗汉果(*Siraitia grosvenorii*)果实进行 DGE 分析,从中选出 7 个 CYP450 和 5 个 UDP 基因作为最可能参与罗汉果苷合成的候选基因。

3.3 在医学方面的应用

准确的判定癌症转移的主要位点在医学上一直是个挑战。Kulkarni 等^[33]比较了数字基因表达谱技术与免疫组织化学技术在鉴定癌症转移诊断中的效果,发现数字基因表达谱技术能在癌症转移的鉴定中比免疫组织化学法更高效且使用的组织较少。

宣荣荣等^[34]采用 DGE 技术对脂多糖(LPS)致炎作用的相关基因进行筛选和分析,探讨巨噬细胞经脂多糖(LPS)处理后基因表达谱的变化。结果 LPS 处理后,1 959 个基因出现了差异表达,其中上调 974 个,下调 985 个。GO 分析显示,差异表达基因主要有蛋白结合和催化活性等分子功能,以及调节细胞过程、对环境刺激的反应和免疫系统等生物学功能。验证了 LPS 介导炎症反应的 NF- κ B、丝裂原激活蛋白激酶和 Jak-STAT 3 条主要的信号转导途径,发现了细胞周期、细胞凋亡和细胞外识别等新的途径。

Fan 等^[35]比较了结肠癌细胞和正常的结肠细胞中基因的差异表达,发现在 2 种类型的细胞中新的候选肿瘤标记物 FITM3。用 RT-qPCR 验证在结肠癌细胞中 FITM3 的表达水平提高。标准的组织切片的免疫染色法和组织微阵列的分析得出同样的结论。这些结果表明 FITM3 在早期的结肠癌发展过程起着重要的作用。

慢性阻塞性肺病(COPD)是一种常见的且致残性的肺部疾病,目前几乎没有治疗的方法。Wang 等^[36]研究不吸烟患者 COPD 3 个阶段(即 0、1、2)肺部样品基因表达情况,发现参与细胞外基质的合成或降解、细胞凋亡的基因上调表达;参与抗炎性应答的基因下调表达。在

肺泡巨噬细胞和上皮细胞中尿激酶纤维蛋白溶酶原激活物(urokinase plasminogen activator,PLAU)、尿激酶纤维蛋白溶酶原激活物受体(urokinase plasminogen activator receptor,PLAUR)、血小板反应蛋白(thrombospondin, THBS1)的基因的表达可能与 COPD 的发病机理有关。在其相关通路中的某些基因与转录生长因子 TGF- β 1 和金属蛋白酶的活性有关,这些基因可被 SERPINE2 所抑制。

Nordlund 等^[37]分析了 21 个急性淋巴细胞白血病患者细胞的细胞(ALL)及基因表达谱。4 种不同亚型 B 细胞的差异基因与 TP53 信号通路、干扰素的信号通路和细胞凋亡的信号通路有关。对注释基因非编码链的基因表达研究发现,许多基因具有亚型特异性的表达模式。17 个基因的反链在 BCP ALL 和 T-ALL 亚型中不同,76 个基因的反链区 4 个 BCP 亚型不同。基因重叠区具有可变的聚腺苷酸化和反转录的区域。这些研究为 ALL 细胞 RNA 表达模式提供新的观点。

Bowen 等^[38]通过比较正常人卵巢表面的上皮细胞和卵巢癌表面上皮细胞的基因表达谱发现,与成年干细胞维护相关通路的基因在卵巢表面上皮细胞高表达,而在卵巢的恶性腺瘤中不表达或少量表达。显著性差异表达的 2 000 多个基因中,许多基因与细胞循环通路和癌细胞发育的信号通路有关。人类卵巢表面的上皮细胞是多功能细胞,这可能是卵巢腺瘤的起源。

布氏锥虫是人类致病性疾病的病原体。Veitch 等^[39]用 DGE 技术分析了布氏锥虫在 2 个不同的生命阶段即感染布氏锥虫和未感染布氏锥虫的人类血液中的转录组变化,得到了 1 100 万多个 unique tags,其中 7 360 个基因表达,基因的覆盖率为 81%。分析发现基因组中存在差异表达的基因簇,而且它们的序列是连续的,可能具有相同的多顺反子单元。这些差异表达的基因在转录水平共调控或者在转录的稳定性方面共调控。这些基因能够编码寄生虫主要的表面蛋白。结果为新药的开发和布氏锥虫的防治提供了分子依据。

3.4 其它方面的研究

郑凯^[40]应用 DGE 技术分析了斜卧青霉(*Penicillium deeeumbens*)在诱导和阻遏条件下 2 个数据库的基因表达差异。在诱导条件下共得到差异表达基因 2 396 个,其中显著上调表达的基因 924 个,显著下调表达的基因 1 472 个。显著上调表达的基因是多种木质纤维降解酶类及糖苷水解酶类基因。表达显著下调的基因包括初级代谢相关酶类基因、转运蛋白基因、转录因子相关的基因等。结果为今后利用微生物产生的纤维素酶降解纤维类物质提供了依据。

贵甫^[41]利用 DGE 技术研究草菇基因组内碳水化合物活性酶家族及纤维素酶和果胶酶子家族情况,结果发

现草菇基因组富含降解纤维素、半纤维素和果胶的酶类编码基因,其中含有最高的是植物细胞壁多糖降解相关基因。草菇同、异核体差异表达 CAZymes 编码基因分析表明,草菇异核体 H1521 中显著上调表达的纤维素酶主要是 GH5 和 GH61 家族,半纤维素酶主要是 GH43 家族,果胶酶主要是 GH43、GH88 和 PL4 家族。结果为选育优良草菇栽培种和提高草菇生物转化率提供了理论依据。

棉铃虫和烟青虫都是夜蛾科金源关系很近的物种。棉铃虫是广食性昆虫,烟青虫是单食性昆虫。Li 等^[42]研究这 2 种昆虫 4 个发育阶段基因表达情况,找出了影响宿主对昆虫识别和适应性的差异基因。嗅觉相关基因的转录中,棉铃虫气味结合蛋白和 G-蛋白偶联受体的转录比烟青虫的多。棉铃虫的解毒酶如 P450、羧肽酶和 ATPase 基因的转录比烟青虫的多。这些差异可能与烟青虫狭窄的植物寄主范围有关。解毒酶能够帮助昆虫对事物解毒,提高食物的利用率。这些差异基因为研究宿主植物选择进化提供了重要的信息。

婚飞是性成熟处女蜂王与雄蜂交配过程中的一个重要前奏,在该过程中蜂王体内伴随着一系列重要的生理变化。吴小波等^[43]利用 DGE 技术分析了中华蜜蜂性成熟处女蜂王飞行与未飞行之间的基因表达差异。经 DGE 测序,分别从 2 个样品中获得 5.98 百万和 6.01 百万条 Clean 标签。分析发现有 250 个基因出现了差异表达,其中 133 个基因在飞行蜂王中上调表达,117 个基因在飞行蜂王中下调表达。这些差异基因设计到 348 个功能性类别和 142 个生化途径。这些结果为进一步研究中华蜜蜂蜂王婚飞过程中生理变化的分子机制提供了重要的基因表达信息。

秦秋红^[44]采用 DGE 技术分析了西方蜜蜂经过迷宫视觉学习前后 microRNAs(miRNAs)和 mRNAs 的表达差异。通过小 RNA 测序分析,发现迷宫学习训练组和未训练组中 miRNAs 和 mRNAs 有明显差异。研究结果提示了差异 miRNAs 和 mRNAs 在蜜蜂的学习与记忆过程中很可能都扮演了重要的角色。有助于进一步理解蜜蜂学习与记忆的分子机制。

蝙蝠是唯一能够飞翔的哺乳动物,它的翼由 4 个通过细胞膜连接并延长的足趾(digits II-V)构成。除了 digits II-V,蝙蝠的前翼有一个很短的足趾,在形态学上与后肢的足趾相似。Wang 等^[45]利用蝙蝠前肢足趾形态学上的差异研究足趾的延长和翼的形成的分子机制。其对 2 个鼠耳蝙蝠发育中的足趾基因数字表达谱进行研究,用 2 个大蹄蝙蝠正在发育的足趾的表达基因的 RT-qPCR 对结果进行验证。结果在短足趾和长足趾间差异表达的基因中,有 14 个基因与足趾的延长有关,3 个与足趾的畸形或细胞的增殖有关的基因(*Twist1*,

Tmeff2 和 *Enpp2*)。此外,研究还发现 *Tbx4* 和 *Pitx2* 基因与蝙蝠指和后肢足趾形态学上的相似有关,*Acta1*、*Tnnc2*、*Atp2a1*、*Hrc* 和 *Myoz1* 基因与指和后肢足趾功能相似性有关。该研究发现了涉及足趾的延长和翼的形成的候选基因,为了解肢体的发育提供了基础。

4 展望

自 2005 年以来,454 测序技术、Solexa 测序技术和 SOLiD 测序技术等第二代测序技术相继出现。随着高通量测序技术而出现的数字基因表达谱(DGE)技术因在研究基因差异表达方面具有全面、快速等独特的优势,近年来在农业、生物学和医学等领域的研究中得到广泛应用,并取得了显著的成果。但由于高通量测序技术建立的时间还不长,技术方面还存在一些需要改进的地方,特别是测序结果的信息分析比较烦琐,如何从巨量的数据中提取有意义的生物学信息,还有待技术的进一步完善和改进。

相信随着测序技术的发展和新的信息分析方法的建立,DGE 技术将能进一步扩大其应用范围,为动植物的分子育种和人类疾病的防治提供强有力的技术支撑。

参考文献

- [1] Ness R W, Siol M, Barrett S C. De novo sequence assembly and characterization of the floral transcriptome in cross- and self-fertilizing plants [J]. BMC Genomics, 2011, 12: 298-310.
- [2] 滕晓坤, 肖华胜. 基因芯片与高通量 DNA 测序技术前景分析 [J]. 中国科学, 2008, 38(10): 891-899.
- [3] Alagna F, Torchia L, Servili M, et al. Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development [J]. BMC Genomics, 2009, 10: 399-413.
- [4] Severin A J, Woody J L, Bolon Y T, et al. RNA-Seq Atlas of Glycine max: A guide to the soybean transcriptome [J]. BMC Plant Biology, 2010, 10: 160-165.
- [5] Morrissey A S, Morin R D, Delaney A, et al. Next-generation tag sequencing for cancer gene expression profiling [J]. Genome Research, 2009, 19: 1825-1835.
- [6] Zoltán H, Anna Z, Vilmos C Á, et al. Deep sequencing of the zebrafish transcriptome response to mycobacterium infection [J]. Molecular Immunology, 2009, 46: 2918-2930.
- [7] Xiao S Q, Jia J Y, Mo D L, et al. Understanding PRRSV infection in porcine lung based on genome-wide transcriptome response identified by deep sequencing [J]. PLoS ONE, 2010, 5: 1-16.
- [8] Zhong Z, Zhang L L, Ye X, et al. Transcriptome profiling and digital gene expression analysis of Nile tilapia (*Oreochromis niloticus*) infected by *Streptococcus agalactiae* [J]. Mol Biol Rep, 2013, 40: 5657-5668.
- [9] Mu Y N, Ding F, Cui P, et al. Transcriptome and expression profiling analysis revealed changes of multiple signaling pathways involved in immunity in the large yellow croaker during *Aeromonas hydrophila* infection [J]. BMC Genomics, 2010, 11: 506-513.
- [10] de Lorigeril J, Reda Z, Rafael D, et al. Whole transcriptome profiling of successful immune response to vibrio infections in the oyster *Crassostrea gigas* by digital gene expression analysis [J]. PLoS ONE, 2011, 6(8): e23142.
- [11] 罗庆苗, 秦晓玉, 苗向阳. 利用数字基因表达谱对小尾寒羊高繁性状

相关基因的筛选研究[J]. 中国农学通报, 2012, 28(23): 18-25.

[12] 秦宁. 猪骨骼肌数字基因表达谱分析与差异表达基因验证[D]. 长春: 吉林农业大学, 2012.

[13] Xiao S Q, Mo D L, Wang Q W, et al. Aberrant host immune response induced by highly virulent PRRSV identified by digital gene expression tag profiling[J]. BMC Genomics 2010, 11: 544-561.

[14] Cheung F, Haas B J, Goldberg S M, et al. Sequencing *Medicago truncatula* expressed sequenced tag using 454 life sciences technology[J]. BMC Genomics, 2006, 7: 272-282.

[15] Emrich S J, Barbazuk W B, Li L, et al. Gene discovery and annotation using LCM-454 transcriptome sequencing[J]. Genome Res, 2007, 17: 69-73.

[16] Wang L, Li P, Brutnell T P. Exploring plant transcriptomes using ultra high-throughput sequencing[J]. Brief Funct Genomics, 2010, 9: 118-128.

[17] Wu T, Qin Z, Zhou X, et al. Transcriptome profile analysis of floral sex determination in cucumber[J]. J Plant Physiol, 2010, 167: 905-913.

[18] Nisha L, Gopal J, Bhardwaj A R, et al. Identification and characterization of miRNAome in root, stem, leaf and tuber developmental stages of potato (*Solanum tuberosum* L.) by high-throughput sequencing[J]. BMC Plant Biology, 2014, 14: 1-31.

[19] Nishiyama R, Le D T, Watanabe Y, et al. Transcriptome analyses of a salt-tolerant cytokinin-deficient mutant reveal differential regulation of salt stress response by cytokinin deficiency[J]. PLoS One, 2012, 7(2): e32124.

[20] Min L, Li Y, Hu Q, et al. Sugar and auxin signaling pathways respond to high temperature stress during anther development as revealed by transcript profiling analysis in cotton[J]. Plant Physiology, 2014, 164: 1293-1308.

[21] 赖晓娟, 严小军, 杨悦, 等. 高温胁迫下坛紫菜的数字基因表达谱研究[J]. 海洋学报, 2014, 36(6): 104-111.

[22] Qi X H, Xu X W, Lin X J, et al. Identification of differentially expressed genes in cucumber (*Cucumis sativus* L.) root under waterlogging stress by digital gene expression profile[J]. Genomics, 2012, 99: 160-168.

[23] Chen L M, Zhou X A, Li W B, et al. Genome-wide transcriptional analysis of two soybean genotypes under dehydration and rehydration conditions[J]. BMC Genomics, 2013, 14: 687-705.

[24] Jiang L X, Jin L G, Guo Y, et al. Glyphosate effects on the gene expression of the apical bud in soybean (*Glycine max*) [J]. Biochemical and Biophysical Research Communications, 2013, 437: 544-549.

[25] Shao H H, Cao Q H, Tao X, et al. Cloning and characterization of ATP synthase CF1 α gene from sweet potato[J]. African Journal of Biotechnology, 2011, 10: 19035-19042.

[26] Lai X J, Gu Y H, Tao X, et al. Cloning and characterization of uridine diphosphate glucose dehydrogenase gene from *Ipomoea batatas* [J]. Russian Journal of Plant Physiology, 2014, 61: 298-308.

[27] Mahadevan S, Park Y. Multifaceted therapeutic benefits of *Ginkgo biloba* L: chemistry, efficacy, safety, and uses[J]. Journal of Food Science, 2008, 73: 14-19.

[28] Heiss C, Keen C L, Kelm M. Flavanols and cardiovascular disease prevention [J]. European Heart Journal, 2010, 31: 2583-2592.

[29] 韦康, 王丽鸳, 成浩, 等. 基于数字基因表达谱对茶树类黄酮合成、调

控相关基因表达的研究[J]. 茶叶科学, 2014, 34(2): 195-203.

[30] Zhao W, Xia W, Li J, et al. Transcriptome profiling and digital gene expression analysis of *Fallopia multiflora* to discover putative genes involved in the biosynthesis of 2,3,5,4'-tetrahydroxy stilbene-2-O- β -d-glucoside[J]. Gene, 2014, 547: 126-135.

[31] Hao D C, Ge G B, Xiao P G, et al. The first insight into the tissue specific taxus transcriptome via illumina second generation sequencing[J]. PLoS One, 2011, 6: e21220.

[32] Tang Q, Ma X J, Mo C M, et al. An efficient approach to finding *Siraitia grosvenorii* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis[J]. BMC Genomics, 2011, 12: 343-355.

[33] Kulkarni A, Pillai R, Ezekiel A M, et al. Comparison of histopathology to gene expression profiling for the diagnosis of metastatic cancer[J]. Diagn Pathol, 2012, 7: 110-119.

[34] 宣荣荣, 陈海敏. 脂多糖活化巨噬细胞的数字基因表达谱分析[J]. 中国药理学与毒理学杂志, 2013, 27(6): 971-977.

[35] Fan J F, Peng Z P, Zhou C Z, et al. Gene-expression profiling in chinese patients with colon cancer by coupling experimental and bioinformatic genome-wide gene-expression analyses[J]. Cancer, 2008, 113: 266-275.

[36] Wang I M, Sergey Stepaniants S, Boie Y, et al. Gene expression profiling in patients with chronic obstructive pulmonary disease and lung cancer[J]. American Journal of Respiratory and Critical Care Medicine, 2008, 177: 402-411.

[37] Nordlund J, Kiialainen A, Karlberg O, et al. Digital gene expression profiling of primary acute lymphoblastic leukemia cells[J]. Leukemia, 2012, 26: 1218-1227.

[38] Bowen N J, Walker L D, Matyunina L V, et al. Gene expression profiling supports the hypothesis that human ovarian surface epithelia are multipotent and capable of serving as ovarian cancer initiating cells[J]. BMC Medical Genomics, 2009, 2: 71-84.

[39] Veitch N J, Johnson P C D, Trivedi U, et al. Digital gene expression analysis of two life cycle stages of the human-infective parasite, *Trypanosoma brucei* gambiense reveals differentially expressed clusters of co-regulated genes [J]. BMC Genomics, 2010, 11: 124-137.

[40] 郑凯. 斜卧青霉基因表达谱的差异分析研究[D]. 济南: 山东大学, 2009.

[41] 贵甫. 草菇数字基因表达谱揭示同核体与异核体 CAZymes 表达差异 [D]. 福州: 福建农业大学, 2012.

[42] Li H C, Zhang H, Guan R, et al. Identification of differential expression genes associated with host selection and adaptation between two sibling insect species by transcriptional profile analysis[J]. BMC Genomics, 2013, 14: 1-12.

[43] 吴小波, 王子龙, 张飞, 等. 婚飞行为影响中华蜜蜂性成熟处女蜂王的基因表达[J]. 昆虫学报, 2013, 56(5): 486-493.

[44] 秦秋红. 东方蜜蜂与西方蜜蜂学习记忆比较及蜜蜂学习记忆相关分子机理分析[D]. 南昌: 江西农业大学, 2013.

[45] Wang Z, Dong D, Ru B H, et al. Digital gene expression tag profiling of bat digits provides robust candidates contributing to wing formation [J]. BMC Genomics, 2010, 11: 619-630.

Research Progress on Digital Gene Expression Profiling Based on High-throughput Sequencing

ZHANG Yu, LIU Yujie, GUO Dan-ni, LI Hui-min, QIN Xin-min

(Laboratory of Ecology of Rare and Endangered Species and Environmental Protection, Ministry of Education, College of Life Science, Guangxi Normal University, Guilin, Guangxi 541004)

DOI:10.11937/bfyy.201510044

盐碱胁迫对植物种子萌发的影响及 生理生化机制研究进展

代明龙, 王平, 孙吉康, 周 韬

(中南林业科技大学 生命科学与技术学院, 湖南 长沙 410004)

摘 要:土壤盐渍化是影响农业生产和生态环境的重要非生物因素,也是目前制约着我国农业增产的两大土壤因素之一,土壤盐渍化由于其高盐且高碱环境对包括种子萌发在内的植物生长发育具有重要影响。现总结了盐胁迫和碱胁迫对种子萌发的影响规律,并初步分析了盐碱胁迫对种子萌发影响的生理生化机制,用以指导生产和进一步开展相关研究提供参考。

关键词:盐胁迫;碱胁迫;发芽率;发芽势

中图分类号:S 604⁺.1 **文献标识码:**A **文章编号:**1001-0009(2015)10-0176-04

随着全球环境的不断恶化,土壤盐碱化已成为全球性的环境问题,日益威胁着人类赖以生存的有限土地资源。据联合国教科文组织(UNESCO)和粮农组织(FAO)的不完全统计^[1],世界盐渍土面积约 9.5 亿 hm^2 ,占世界陆地面积的 7.6%。我国盐渍土面积大,约 $1 \times 10^8 \text{ hm}^2$,且分布范围广,类型多^[2]。通常土壤含盐量在 0.2%~0.5%即不利于植物的生长^[3],盐渍土由于其高盐环境且常常伴随着高 pH 值,对植物的生长发育产生较强的毒害作用^[4],耐盐性低的植物很难生存,土壤盐渍化已成为世界范围内影响农作物产量的最关键因素^[5]。植物种子萌发和幼苗生长是植物生活史中的 2 个关键阶段^[6],同时也是对盐碱胁迫较为敏感的时期,常常将植物种子萌发期的耐盐性作为该品种的耐盐性^[7]。因此,盐碱胁迫对种子萌发的影响引起全世界的广泛关注。目前,已经在盐生植物和农作物中开展了大量研究^[8],现在前人研究基础上,总结了盐碱胁迫对种子萌发的影

响规律,并初步分析了盐碱环境对种子萌发的影响机理,为耐盐植物品种的培育,盐碱地环境的改良以及荒漠植被的恢复与重建提供理论依据,以发挥盐碱地巨大的经济效益和生态效益。

土壤盐分过多对植物造成的危害称为盐害,也称盐胁迫^[9],自然界造成盐胁迫的盐分主要是 NaCl 、 Na_2SO_4 、 Na_2CO_3 、 NaHCO_3 。 NaCl 、 Na_2SO_4 等中性盐造成的胁迫称为盐胁迫, Na_2CO_3 、 NaHCO_3 等碱性盐造成的胁迫称为碱胁迫。不同种类植物,同一植物的不同品种对盐分的敏感程度不同,耐盐性存在差异^[10-11];不同的盐由于毒害机理不同以及同一种盐的浓度不同对植物的危害程度不同^[12]。评价种子发芽情况的常用指标有发芽率、发芽势、发芽指数、活力指数等^[13]。发芽势是指种子发芽初期在规定的时间内能正常发芽的种子数占种子总数的比例,体现了种子的发芽速度和整齐度,能在一定程度上反映种子抵御逆境胁迫的能力。发芽指数与活力指数是表征种子活力的指标,能较全面的反映种子的萌发率及幼芽的生长状况等。

1 盐碱胁迫对种子萌发的影响

1.1 中性盐胁迫对种子萌发的影响

通常,低浓度的中性盐对种子萌发无显著影响,高浓度的中性盐抑制种子萌发,且随着盐浓度的升高,抑

第一作者简介:代明龙(1989-),男,硕士研究生,研究方向为生物化学与分子生物学。E-mail:446058997@qq.com.

责任作者:王平(1964-),男,博士,教授,现主要从事生物技术方面的教学与科研工作。E-mail:csfuwp@163.com.

基金项目:国家林业公益性行业科研专项资助项目(201204606)。

收稿日期:2015-01-22

Abstract: Digital gene expression profiling (DGE) is a high-throughput sequencing technologies, which can comprehensively and quickly detect gene expression in a specific tissue under certain conditions of a species. The related information about gene transcription, gene regulation, signal transduction pathway and protein function can be obtained by the search, comparison and analysis on bioinformatics of gene expression profiling. This paper mainly described the principles of digital gene expression profiling and applications in biological, medicine and so on.

Keywords: digital gene expression profiling; high-throughput sequencing; differentially expressed gene