

[ 文章编号 ] 1007-385X( 2007 )06-0585-04

## 代谢组学——肿瘤学研究的重要技术平台

### Metabolomics—an important technology platform for cancer research

洪毅 综述, 谈治雄, 王红阳\* 审阅( 上海第二军医大学 东方肝胆外科医院, 上海 200438 )

**[ 摘要 ]** 代谢组学( metabolomics )是对某一生物或细胞所有低分子量代谢产物进行定性和定量分析的一门新学科, 与基因组学、转录组学、蛋白质组学等一起构成系统生物学的技术平台。代谢组学分析技术以磁共振和色谱、质谱串联技术为主, 获得的代谢谱数据借助化学计量学工具和模式识别软件进行分析。已报道的肿瘤特异性代谢产物涵盖各类生物大分子的物质代谢, 代谢组学从机体的动态代谢途径寻找肿瘤等疾病特异性代谢产物, 识别人体疾病代谢状态的差异, 从而进行疾病分类和诊断。代谢产物是基因表达的终产物, 分析生物代谢图谱型特征更能够揭示基因和表现型之间的关系, 以达到监测和推断基因功能的目的, 为进一步深入分析肿瘤生物机制提供可能。

**[ 关键词 ]** 代谢组学; 系统生物学; 肿瘤; 代谢产物; 肿瘤标志物

**[ 中图分类号 ]** R730 **[ 文献标识码 ]** A

代谢组学<sup>[1]</sup>是对某一生物或细胞所有低分子量代谢产物进行定性和定量分析的一门新学科。随着人类基因组计划的完成, 生命科学已进入后基因组时代。在基因组学、转录组学、蛋白质组学平台基础上加入代谢组学等新内容, 逐步形成了较完整的系统生物学技术平台, 使人们能更好地从系统生物学水平揭示肿瘤的病因和机制, 发掘新的诊断工具与方法。

恶性肿瘤是一种多因素参与造成机体各系统功能平衡紊乱的复杂疾病, 其代谢作用与正常组织明显不同。通过代谢组学在给定的时间和条件下对某一生物系统所有小分子代谢物质的定量分析, 就可以对生物内源性代谢物的整体情况进行定量或对造成其内因和外因变化的规律进行分析。本文就代谢组学与其它组学的区别和关系、代谢组学研究的技术和分析方法、在肿瘤研究中应用的进展及发展趋势进行综述。

### 1 系统生物学与代谢组学的基本概念

系统生物学<sup>[2]</sup>是在细胞、组织、器官和生物体整体水平多层次、多系统研究其结构和功能各异的各种分子及其相互作用, 用计算生物学方法整合各组学数据来定量描述和预测它们的生物功能、表型和行为。

系统生物学将在基因组序列的基础上完成由生命密码到生命过程的研究, 这是逐步整合的过程, 由生物体内各种分子的鉴别及其相互作用的研究到途径、网络、模块, 最终完成整个生命活动的路线图。基因组学、转录组学、蛋白质组学、代谢组学分别在 DNA、mRNA、蛋白质和代谢产物水平检测和鉴别各种分子并研究其功能。系统生物学使生命科学由描述式的科学转变为定量和预测的科学, 代谢组学是其中新兴发展的重要平台。

转录组学和蛋白质组学还不能完全表达生物体的

全部功能。代谢组学能定量测量生物体基因修饰和受病理生理刺激所产生的与时间相关的多元代谢产物, 目的是通过分析生物体处于不同发育阶段或不同生长环境中的化合物, 跟踪其代谢路径, 以期能描绘出基因、蛋白和代谢产物之间各种复杂的相互作用全图, 系统地对生物体功能进行阐述。代谢组学在疾病诊断方面的应用正逐渐趋于成熟, 机制研究也在进一步深入。

### 2 代谢组学的分析技术

许多生物组织和体液都可以进行代谢组学研究。以体液为例, 包括唾液、血液、尿液、胆汁、乳液、脑脊液和组织萃取液等, 不同体液的来源、分布及功能各不相同, 因此其包含的代谢产物成分也存在较大差异。代谢组学分析生物体系中所有的代谢产物, 单一的分离、分析技术难以满足这一要求。因此, 越来越多的分离、分析手段及其组合出现在代谢组学的研究中<sup>[3-5]</sup>, 如: 核磁共振( NMR )、薄层色谱( TLC )、紫外或二极管阵列检测器的高效液相色谱( HPLC/UV/PDA )、紫外检测器的毛细管电泳( CE/UV )、激光诱导检测器的毛细管电泳( CE/LIF )、CE-MS、GC-MS、质谱与色谱串联技术( LC-MS )混合型串联傅里叶变换质谱( hKFICR MS )、LC-NMR/MS 等, 其中以 NMR 和色谱与质谱串联技术最为常用。

#### 2.1 核磁共振( NMR )技术

NMR 是当前代谢组学研究中的主要技术之一<sup>[6]</sup>,

**[ 基金项目 ]** 国家自然科学基金资助项目( No. 30400244 )。Supported by the National Natural Science Foundation of China( No. 30400244 )

**[ 作者简介 ]** 洪毅( 1976- ), 男, 福建厦门人, 硕士研究生, 主要从事肿瘤分子生物学研究

\* Corresponding author. E-mail: hywangk@online.sh.cn

它的优势在于能够对样品实现非破坏性、非选择性的分析。Nicholson 研究小组采用魔角旋转 (magic angle spinning, MAS) 技术, 将 NMR 技术广泛地应用于药物毒性分析、基因功能解析<sup>[7]</sup>和疾病的临床诊断中<sup>[8]</sup>。另外, NMR 与 HPLC、MS 联用对多元的数据信息进行分析最具发展潜力, 但缺点是 NMR 的循环周期变长, 同时需要含重氢的流动相。

目前, NMR 主要集中在追踪受疾病或行为模式影响下代谢进程中特定的化学信号, 而不是完整分析代谢途径。

## 2.2 色谱与质谱联用技术

目前, 与质谱联用的色谱技术主要有气相色谱 (GC)、液相色谱 (LC)、毛细管电泳 (CE) 等。LC 可以直接分析不挥发性化合物、极性化合物、热不稳定化合物和大分子化合物 (包括蛋白、多肽、多糖、多聚物等), 分析范围广, 可不衍生化。高性能色谱技术因其良好的分离能力已被广泛用于复杂体系 (如血浆和尿样) 的靶标分析<sup>[9]</sup>。

随着各种新的离子化技术的不断出现, 质谱以其高灵敏度、高通量的特性被广泛地应用于代谢组学研究领域。Castrilloa 等<sup>[5]</sup>应用电喷雾质谱直接进样, 完成了对酵母的代谢组学分析。Fiehn 等<sup>[3]</sup>采用 GC-MS 研究了阿拉伯芥的基因型及其表型的关系。

## 3 代谢组学的数据分析

代谢组学研究产生大量、多维数据, 需要借助专门的数据分析理论与工具软件进行整理、转换、统计和输出。许多化学计量学工具和模式识别技术正不断地应用于代谢组数据分析中, 主要分为两类, 即: 监督学习方法<sup>[11]</sup> (supervised method) 和非监督学习方法<sup>[6]</sup> (unsupervised method)。监督学习方法利用多参数的模型对样品进行识别和分类, 并且在建立模型时有可供学习的样本, 包括传统的显著性分析 (discriminant analysis, DA)、偏最小二乘法 (partial least squares, PLS)、偏最小二乘法 - 显著性分析联合法 (PLS-DA); 人工智能类的规则归纳式学习 (rule induction)、归纳逻辑程序设计 (inductive logic programming, ILP)、进化计算 (evolutionary computation)、人工神经网络 (artificial neural network, ANN) 等。非监督学习方法将得到的分类信息和样品的原始信息进行比较, 找出差别, 包括主成分分析 (principal components analysis, PCA)、层次聚类分析 (hierarchical cluster analysis, HCA) 和自组织图 (self-organizing maps, SOMs)。新的具有良好的重现性组分析的方法仍在不断的被开发出来<sup>[10]</sup>。在代谢组学中比较常用的分析方法主要是 DA、PLS-DA、PCA 和 PCA-DA。

## 4 代谢组学分析技术在肿瘤研究中的应用

### 4.1 肿瘤特异性代谢产物的鉴定

治愈肿瘤的关键之一是早期诊断, 寻找有诊断价值的特异性肿瘤标志物是各个组学研究的重点之一。美国国家癌症研究所 (The National Cancer Institute) 建立了生物标志物发展评价五阶段标准<sup>[12]</sup>: (1) 临床前探索期, 确定可靠方向; (2) 临床分析和验证期, 分析评价标志物对疾病的检测能力; (3) 回顾性纵向分析设定的指标, 对临床前疾病的检测能力, 制定阳性筛选标准; (4) 前瞻性观察标志物对疾病的检出率及假阳性率; (5) 设计前瞻性临床随机试验, 对普遍人群进行筛查。

代谢组学从机体的动态代谢途径寻找肿瘤特异性代谢产物, 在各阶段均有良好应用前景。特别是在疾病模式分类和病理分型上有显著优势。许国旺等<sup>[9]</sup>采用毛细管电泳方法 (CE), 通过 PCA 法处理数据, 对 68 位癌症患者和 54 位正常人进行分类研究, 识别率达 72%。采用人工神经网络软件对数据进行处理, 肿瘤患者的识别率可达 83%。

已报道的肿瘤特异性代谢产物分为以下几类:

(1) 蛋白质代谢产物 (氨基酸)。肿瘤蛋白质合成分解的特殊性可用于标志物检测。Ippolito 等<sup>[13]</sup>通过 MRS 发现在前列腺神经内分泌癌细胞中 (prostate neuroendocrine cell, PNEC) 相关的代谢标志物为  $\gamma$ -氨基丁酸 (gamma-aminobutyric acid, GABA) 和丙二醇, 两者表达较高且与能量代谢相关。当 GABA、谷氨酰胺和 (或) 甘氨酸的受体刺激激活, 恶性 NE 细胞的增殖速度加快。将 3 种抑制 GABA 通路成员抑制剂的混合物注射 PNEC 裸鼠移植瘤, 移植瘤的生长明显减慢, 这项研究对于神经内分泌型肿瘤的治疗有重要的指导意义。Joseph 等<sup>[14]</sup>采用 GC-MS 技术, 在神经内分泌肿瘤细胞代谢产物中也发现谷氨酸脱羧酶途径加速, 产生的  $\gamma$ -氨基丁酸富集。

(2) 核酸代谢产物 (核苷酸)。肿瘤细胞代谢旺盛, DNA 与 RNA 合成加速, 核苷酸增加, 可提供许多核苷酸代谢指标。Zheng 等<sup>[15-16]</sup>采用反相高效液相色谱法, 对 52 位结肠癌患者和 60 位正常人通过 14 种核苷酸代谢进行分类研究, 识别率达 71%, 而 CEA 检出阳性率只有 29%; 对 26 位乳腺癌患者和 41 位正常人通过 13 种核苷酸代谢进行分类研究, 识别率达 73%。Yang 等<sup>[17]</sup>采用反相高效液相色谱法对 47 位肝癌患者和 70 位正常人通过 15 种核苷酸代谢进行分类研究, 识别率达 83%, 而 AFP 检出阳性率为 73.9%。另外在胶质瘤、卵巢癌检测均有良好应用<sup>[18]</sup>。

(3) 脂类代谢产物。各类恶性肿瘤、炎症时脂质

浓度和定位都会发生变化。Griffin 等<sup>[19]</sup>通过基因治疗的方法处理大鼠脑胶质瘤细胞,增加了细胞凋亡,发现不饱和脂肪(PUFA)浓度升高,研究显示 MRS 可以方便的检测细胞胞质中的脂肪成分,因此检测大鼠脑胶质瘤细胞 PUFA 的浓度可以监测脑胶质瘤基因治疗的效果,为人脑胶质瘤的预后判断打下良好的基础。Millis 等<sup>[20]</sup>采用 NMR 技术,在脑瘤、肉瘤发现细胞膜组分变化,主要是胆碱、磷酸胆碱等。

(4)糖代谢产物。肿瘤对糖的代谢也会有影响。Tablack 等<sup>[21]</sup>采用 MRS 技术分析了 75 例卵巢癌患者样本(66 例早期浸润癌,9 例基底癌)的代谢物谱,发现两类病理类型的主要区别是嘌呤、嘧啶代谢、糖脂代谢以及能量代谢,通过聚类、主成分分析以及分类预测模型区分两种类型肿瘤的准确度达到 90%。

#### 4.2 结合系统生物学对肿瘤分子机制的分析

代谢产物是基因表达的终产物,通过代谢组学可以监测肿瘤基因转录上调或下调效应的变化,这为肿瘤机制研究提供了新的思路。对不同生理状态的代谢组进行分析,理论上人们就可以全面了解该生物或细胞的生物化学状态,比如因为不同种系动物的基因功能和代谢途径是不同的,通过代谢图谱型特征就可以把不同种系动物区分开<sup>[6,22]</sup>。

肿瘤代谢与多种酶活性密切相关。Chiang 等<sup>[23]</sup>将代谢组与蛋白质组学结合,发现一种新的乙醚脂代谢的关键酶 KIAA1363,作为血小板激活因子和磷脂酶的连接通路,乙醚脂代谢在肿瘤细胞转移和生长中起重要作用,实验显示 KIAA1363 失活阻滞了肿瘤细胞的生长和转移。两种组学的结合使用进一步明确了肿瘤与酶相关代谢通路。

由于代谢组学分析所获得的信息离生物的表现型或生理状态最近,所以从理论上说,代谢组学分析所提供的信息更能够揭示基因和表现型之间的关系,达到监测和推断基因功能的目的。Jacob<sup>[24]</sup>等采用 NMR 技术观测到结肠癌上皮细胞系中核蛋白 P300(转录激活剂)缺失后细胞的磷酸胆碱代谢机制发生改变,造成细胞中甘油磷酸胆碱含量明显增加。代谢组学还能从系统的水平来解释复杂的营养成分对肿瘤防治的机制。Teresa 等<sup>[25]</sup>利用代谢物谱编辑的转录组学分析(metabolomics-edit transcriptomics analysis, META)发现硒和硒蛋氨酸处理肺癌细胞,在糖酵解、磷酸戊糖途径、三羧酸循环、谷氨酰胺代谢以及脂肪酸代谢途径细胞均受硒处理的影响。

通过对基因表达谱的终点——生物代谢产物信息的研究,可以将基因的活动状态比较完整地展现出来,从而系统地研究生命现象及其本质<sup>[22,26]</sup>。对肿瘤等疾病的发生机制、防治诊断措施等的代谢组学研究可能

成为近期研究热点。

## 5 展 望

众所周知,没有任何一个分析技术能够同时对代谢组中的所有化合物进行分析,所以只能通过多次选择不同类型化合物分析来解决这个问题<sup>[3]</sup>。在肿瘤代谢组研究中,样品之间的变异、仪器分析误差等都会给后期分析带来挑战。同时,代谢组学分析会产生出海量的数据<sup>[1]</sup>,尚需建立一个与肿瘤代谢相关的统一标准的通用数据库。

代谢组学研究这几年得到了快速发展<sup>[10,19,23,27]</sup>,目前,代谢组学正日益成为研究的热点。建立基于体液代谢物的异常来诊断恶性肿瘤的方法也正向着实用性方向发展。随着代谢组学定量方法精确度的提高和研究的深入,以及各种组学数据的成功对接,人们可以获得肿瘤基因转录、蛋白以及代谢水平的全景信息,从更高的水平理解生物系统的功能。代谢组学研究必将在揭示基因功能的功能基因组学研究中发挥更大的作用,它能帮助人们更好地了解生物体中各种复杂的相互作用、生物系统对环境 and 基因变化的响应,为人们提供一个了解肿瘤疾病基因的独特途径。

此外,代谢组学分析系统的硬件、软件技术都将进一步提高,朝着整合化、自动化和高通量的方向发展,使人们能够以更快的速度对代谢组学研究实现自动分析和可视化。

## [ 参 考 文 献 ]

- [1] Nicholson JK, Connelly J, Lindon JC, *et al.* Metabonomics: a platform for studying drug toxicity and gene function [J]. *Nat Rev Drug Disc*, 2002, 1(2): 153-161.
- [2] Kitano H. Systems biology: a brief overview [J]. *Science*, 2002, 295(5560): 1662-1664.
- [3] Baggett BR, Cooper JD, Hogan ET, *et al.* Profiling isoflavonoids found in legume root extracts using capillary electrophoresis [J]. *Electrophoresis*, 2002, 23(11): 1642-1651.
- [4] Sumner LW, Mendes P, Dixon RA. Plant metabolomics: large-scale phytochemistry in the functional genomics era [J]. *Phytochemistry*, 2003, 62(6): 817-836.
- [5] Castrillo JI, Hayes A, Mohammed S, *et al.* An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry [J]. *Phytochemistry*, 2003, 62(6): 929-937.
- [6] Griffin JL, Williams HJ, Sang E, *et al.* Metabolic profiling of genetic disorders: a multi-tissue 1H NMR spectroscopic and pattern recognition study into dystrophic tissue [J]. *Anal Biochem*, 2001, 293(1): 16-21.
- [7] Coen M, Ruepp SU, Lindon JC, *et al.* Integrated application of transcriptomics and metabonomics yields new insight into the toxicity due to paracetamol in the mouse [J]. *J Pharm Biomed Anal*, 2004, 35(1): 93-105.

- [ 8 ] Griffin JL, Shockcor JP. Metabolic profiles of cancer cells [ J ]. Nat Rev Cancer, 2004, 4( 7 ): 551-561.
- [ 9 ] Xu G, Liebich HM, Lehmann R, et al. Capillary electrophoresis of urinary normal and modified nucleosides of cancer patients [ J ]. J Methods Mol Biol, 2001, 162: 459-474.
- [ 10 ] Gika HG, Theodoridis GA, Wingate JE, et al. Within-day reproducibility of an HPLC-MS-based method for metabolomic analysis: application to human urine [ J ]. Proteome Res, 2007, 6( 8 ): 3291-303.
- [ 11 ] Beckwith-Hall BM, Brindle JT, Barton RH, et al. Application of orthogonal signal correction to minimise the effects of physical and biological variation in high resolution 1H NMR spectra of biofluids [ J ]. Analyst, 2002, 127( 10 ): 1283-1288.
- [ 12 ] Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer [ J ]. J Natl Cancer Inst, 2001, 93( 14 ): 1054-1061.
- [ 13 ] Ippolito JE, Merritt ME, Bäckhed F, et al. Linkage between cellular communications, energy utilization, and proliferation in metastatic neuroendocrine cancers [ J ]. Proc Natl Acad Sci USA, 2006, 103( 33 ): 12505-12510.
- [ 14 ] Ippolito JE, Xu J, Jain S, et al. An integrated functional genomics and metabolomics approach for defining poor prognosis in human neuroendocrine cancers [ J ]. Proc Natl Acad Sci USA, 2005, 102( 28 ): 9901-9906.
- [ 15 ] Zheng YF, Yang J, Zhao XJ, et al. Urinary nucleosides as biological markers for patients with colorectal cancer [ J ]. World J Gastroenterol, 2005, 11( 25 ): 3871-3876.
- [ 16 ] Zheng, YF, Kong HW, Xiong JH, et al. Clinical significance and prognostic value of urinary nucleosides in breast cancer patients [ J ]. Clin Biochem, 2005, 38( 1 ): 24-30.
- [ 17 ] Yang J, Xu G, Zheng YF, et al. Diagnosis of liver cancer using HPLC-based metabolomics avoiding false-positive result from hepatitis and hepatocirrhosis diseases [ J ]. J Chromatogr B Analyt Technol Biomed Life Sci, 2004, 813( 1-2 ): 59-65.
- [ 18 ] Lehtimäki KK, Valonen PK, Griffin JL, et al. Metabolite changes in BT4C rat gliomas undergoing ganciclovir-thymidine kinase gene therapy-induced programmed cell death as studied by 1H NMR spectroscopy *in vivo*, *ex vivo*, and *in vitro* [ J ]. J Biol Chem, 2003, 278( 46 ): 45915-45923.
- [ 19 ] Griffin JL, Kauppinen RA. A metabolomics perspective of human brain tumours [ J ]. FEBS J, 2007, 274( 5 ): 1132-1139.
- [ 20 ] Millis K, Weybright P, Campbell L N, et al. Classification of human liposarcoma and lipoma using *ex vivo* proton NMR spectroscopy [ J ]. Magn Reson Med, 1999, 41( 2 ): 257-267.
- [ 21 ] Denkert C, Budczies J, Kind T, et al. Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors [ J ]. Cancer Res, 2006, 66( 22 ): 10795-10804.
- [ 22 ] Holmes E, Nicholson JK, Tranter G. Metabonomic classification of genetic variations in toxicological and metabolic responses using probabilistic neural networks [ J ]. Chem Res Toxicol, 2001, 14( 2 ): 182-191.
- [ 23 ] Chiang KP, Niessen S, Saghatelian A, et al. An enzyme that regulates ether lipid signaling pathways in cancer annotated by multidimensional profiling [ J ]. Chem Biol, 2006, 13( 10 ): 1041-1050.
- [ 24 ] Bundy JG, Iyer NG, Gentile MS, et al. Metabolic consequences of p300 gene deletion in human colon cancer cell [ J ]. Cancer Res, 2006, 66( 15 ): 7606-7614.
- [ 25 ] Fan TW, Higashi RM, Lane AN, et al. Integrating metabolomics and transcriptomics for probing SE anticancer mechanisms [ J ]. Drug Metab Rev, 2006, 38( 4 ): 707-732.
- [ 26 ] Soga T, Ueno Y, Naraoka H, et al. Simultaneous determination of anionic intermediates for bacillus subtilis metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry [ J ]. Anal Chem, 2002, 74( 10 ): 2233-2239.
- [ 27 ] Huhman DV, Sumner LW. Metabolioc profiling of saponin glycosides in Medicago sativa and Medicago truncatula using HPLC coupled to an electrospray iontrap mass spectrometer [ J ]. Phytochemistry, 2002, 59( 3 ): 347-360.

[ 收稿日期 ] 2007 - 05 - 10

[ 修回日期 ] 2007 - 11 - 04

[ 本文编辑 ] 王莹

· 读者 · 作者 · 编者 ·

## 文稿中数字用法的要求

本刊严格执行国家标准《出版物上数字用法的规定》，文稿中凡是可以使用阿拉伯数字且很得体的地方，均应使用阿拉伯数字。（1）公历世纪、年代、年、月、日和时、时刻必须使用阿拉伯数字，如 20 世纪 90 年代、2006 - 02 - 15、5 h、30 min、30 s、14:36:08 等；年份不能用简称，“1998 年”不能写作“98 年”。（2）物理量量值必须使用阿拉伯数字。（3）非物理量量词前面数字一般也应使用阿拉伯数字，如 3 支、5 根等。（4）数值范围的表达要求：5 万至 10 万应写成 5 万 ~ 10 万，不能写成 5 ~ 10 万； $3 \times 10^9$  至  $5 \times 10^9$  应写成  $3 \times 10^9 \sim 5 \times 10^9$ ，或  $(3 \sim 5) \times 10^9$ ，不能写成  $3 \sim 5 \times 10^9$ ；60% 至 70% 不能写成 60 ~ 70%，应写成 60% ~ 70%； $25.5 \pm 0.5 \text{ mg}$  应写成  $(25.5 \pm 0.5) \text{ mg}$ 。（5）带单位的量值相乘时，每个数值后单位不能省略，如  $4 \text{ mm} \times 2 \text{ mm} \times 3 \text{ mm}$ ，不能写成  $4 \times 2 \times 3 \text{ mm}$  或  $4 \times 2 \times 3 \text{ mm}^3$ 。

（本刊编辑部）