

生物信息技术加速开发旧药新用途*

黄宏斌^{1, 2)} 梁芳²⁾ 熊炜²⁾ 李小玲²⁾ 曾朝阳^{2)**} 李桂源^{2)**}

(¹⁾ 国防科技大学信息系统工程重点实验室, 长沙 410073; (²⁾ 中南大学肿瘤研究所癌变与侵袭原理教育部重点实验室, 长沙 410078)

摘要 传统的技术路线研发新药, 不仅周期很长而且耗资巨大, 开发已获批准药物新的治疗用途, 又称为药物重定位, 比传统的新药研发具有明显的优势。基于芯片的基因表达谱分析, 已常规地广泛用于各种人类疾病的临床研究, 提供了在全基因组水平描述疾病状态的特征信号。同时, 基因芯片也广泛地用于对比药物处理前后细胞基因表达模式的变化, 这也提供了反映药物效应的高质量信号。最近出版的 *Science Translational Medicine* 杂志同时发表了一个研究组的两篇论文, 为我们展示了如何利用生物信息学手段重新解析和比较全基因组基因表达谱数据, 以高效地预测药物的新用途。这两篇论文使用了公共数据库中的 100 种疾病基因表达谱数据, 以及 164 种药物处理前后细胞基因表达谱数据, 通过比较和配对疾病与药物基因表达谱, 得到了一些可以逆转疾病异常表达基因的药物, 其中证实了一些已知的药物-疾病组合, 也预测了一些新的药物-疾病组合。最后通过实验验证了抗溃疡药可用于治疗肺癌, 而抗癫痫药可治疗炎症性肠道疾病, 进一步证实了他们所采用研究策略的正确性。于是, 肺癌和炎症性肠道疾病这两种临床上难治的疾病有了新的候选治疗药物, 我们也有了一种挖掘已有数据快速发现药物新用途的思路和方法。

关键词 生物信息学, 基因表达谱, 药物重定位

学科分类号 Q61

DOI: 10.3724/SP.J.1206.2011.00453

传统的技术路线研发新药, 不仅周期很长而且耗资巨大。有统计表明, 一个新药从研发到上市, 大约需要 15 年, 花费超过 10 亿美元^[1], 其中相当大的一部分研发费用花在了新药的早期筛选和毒性实验中, 大约有 90% 的候选分子因不能通过早期实验和毒性评估而终止。因此, 深入研究已获批准的药物, 充分发掘已有药物的新用途(drug repurposing)^[2], 或者说对药物进行重定位(drug repositioning)^[3], 一直受到生物医药产业的广泛重视。因为批准上市的药物已经历了数期临床试验, 其毒副作用已经通过了严格的评估, 安全性有保障, 因此, 这些药物的新用途更容易获得药品监管部门的批准, 从而可大大降低研发风险, 缩减研发费用, 节省研发时间, 尽早应用于临床, 开拓新的市场。

1 旧药新用的成功案例

旧药新用最有名的例子是辉瑞(Pfizer)公司研发的万艾可(Viagra, 中文又称“伟哥”)。20 世纪 80

年代, 万艾可上市时原本用于治疗冠心病和心绞痛, 后来却偶然地发现它对勃起功能障碍(erection dysfunction, ED)具有显著的疗效。由于这一新用途的发现, 2003 年辉瑞公司的销售额达到 18.8 亿美元, 仅美国就有近 800 万男性服用过万艾可^[4]。

另一个非常有戏剧性的例子是沙利度胺(thalidomide)。该药最早于 1957 年在德国和英国上市, 起初作为镇静剂用于治疗失眠和恶心, 特别是用于缓和孕妇的妊娠反应(因此该药还有一个别名叫“反应停”), 到 1961 年时, 沙利度胺已在 46

* 国家自然科学基金资助项目(30871282, 30871365, 81172189, 81171930), 湖南省自然科学基金资助项目(10JJ7003), 霍英东高校青年教师基金资助项目(121036), 中央高校基本科研业务费专项资金(2011JQ020)和中南大学博士后科学基金资助项目。

** 通讯联系人。Tel: 0731-84805383

曾朝阳。E-mail: zengzhaoyang@xysm.net

李桂源。E-mail: ligy@xysm.net

收稿日期: 2011-10-10, 接受日期: 2011-11-18

个国家销售,在发现其副作用并禁用前,沙利度胺已导致了至少 1.5 万名儿童出现短肢等骨骼发育畸形的严重出生缺陷,这些儿童都是由于母亲在怀孕的头三个月服用了沙利度胺所致。沙利度胺由此臭名昭著,并被广泛禁用,要不是一个偶然的发现,沙利度胺恐怕永无翻身之日。1964 年,马赛大学医院的医生 Sheskin 收治了一例非常严重的麻风结节性红斑(erythema nodosum leprosum, ENL)病人,剧烈的疼痛已使病人几周无法入睡,刚好当时医院里除了沙利度胺,再没有别的镇静剂了,因此 Sheskin 不顾一切地使用了沙利度胺,希望能帮助病人入睡,结果沙利度胺不仅使病人安然入睡,而且还治愈了他的 ENL^[4]。随后 Sheskin 在委内瑞拉开展了一项双盲实验,发现 173 例使用沙利度胺的 ENL 患者有 92% 获得了完全缓解,世界卫生组织随后资助开展了一项针对 4 552 例 ENL 患者的跟踪研究,发现 99% 的 ENL 患者获得了至少两周的完全缓解。至今,沙利度胺仍是 ENL 的主要治疗药物^[4]。随后的研究还发现沙利度胺是肿瘤坏死因子 α (tumor-necrosis factor- α , TNF- α) 的抑制剂^[5],可用于治疗艾滋病人的口腔溃疡和恶病体质(cachexia)^[6]。沙利度胺也被证实可抑制 I 型潜伏 HIV(latent HIV type I) 的活性^[7]。1994 年,波士顿儿童医院的研究者发现沙利度胺还有抑制血管生成的作用^[8],使它成为了抗肿瘤的候选药物。2003 年,沙利度胺的销售收入达到 2.24 亿美元,92% 被用于治疗以多发性骨髓瘤^[9-10]为代表的各种肿瘤^[11-12]。沙利度胺的故事告诉我们,任何一种药物,我们都不敢说对它已经完全地了解了,发现它的新用途尽管不那么容易,但绝非不可能。

如果说万艾可和沙利度胺这两个药物新用途的发现始于偶然,新药巨额的研发成本则迫使生物制药产业主动地去找旧药的新用途,以使现有的药物得到最大化的利用,同时也使制药厂商的利润最大化。比较普遍的研发思路是:首先构建已获批准的药物库(libraries of approved drugs),利用高通量的筛选技术平台^[13]大规模地筛查潜在的药靶,筛选到药物潜在的新用途后,再经过科学的验证体系进行验证,并通过各期临床试验,完成市场前景、知识产权等多方面的评估,最后重新获得药品监管部门的批准,进入新的市场。

2 生物信息学加速发现旧药新用途

人类基因组计划的完成,特别是公共数据库中

海量的基因表达谱(gene expression profile, GEP)数据^[14-19],结合生物信息学技术^[20],大大地加速了旧药新用途的发现。

2011 年 8 月 17 日出版的 *Science Translational Medicine* 杂志同时发表了美国斯坦福大学医学院 Butte 及其同事的两篇论文^[21-22],在这两篇论文中,他们利用生物信息学技术,分析了公共数据库中已有的各种疾病及各种药物处理细胞后的基因表达谱数据,分别发现并验证了治疗溃疡的药物西咪替丁(cimetidine)可以抑制肺腺癌(lung adenocarcinoma, LA)^[21],而抗癫痫药托吡酯(topiramate)能明显缓解炎症性肠道疾病(inflammatory bowel disease, IBD)的症状^[22]。杂志同期配发了评论文章(Perspective)^[23]和题为“Greening Drug Discovery”的编者按(Editor's Summary)对 Butte 及其同事的工作给予了高度评价。

Butte 等的基本研究思路如图 1 所示:由于基因芯片技术的普及,在公共数据库中已经公开了大量包括各种人类疾病及其相对应正常对照组织的全基因组表达谱数据,通过比较人类疾病与正常状态下的全基因组表达谱,可以获得疾病状态下的差异表达基因,这些差异基因表达谱就构成了可反映疾病病理状态的高质量代表性特征分子信号,或者称为疾病基因表达信号(disease gene expression signature)。同时,基因芯片技术也被广泛用于检测各种药物处理前后人类细胞基因表达谱的变化情况^[24-26],通过比较药物处理前后细胞基因表达谱得到的差异表达基因,则可以代表药物的药效特征信号,或者称为药物基因表达谱(drug gene expression profile)。通过大量的对比和筛查疾病基因表达信号和药物基因表达谱,可以发现一些药物处理后影响的基因表达谱与某些疾病基因表达谱之间存在相关性,根据药物基因表达谱与疾病基因表达信号间相似性的程度和方向,可得到一个“疾病-药物分值”(disease-drug scores),如果某种药物引起一部分基因表达上调而另一部分基因表达下调,而在某种疾病中,与正常对照相比,这些基因也有完全相同的表达变化,也就是说药物的效果和疾病状态高度正相关,这时疾病-药物分值定义为+1;相反,如果“疾病-药物分值”接近-1,表明药物引起的基因表达变化与疾病状态下的基因表达变化完全相反或者高度负相关,该药物则有可能逆转该疾病的表型,或者说这种药物就是我们要寻找的可用于治疗该疾病的候选药物^[21]。

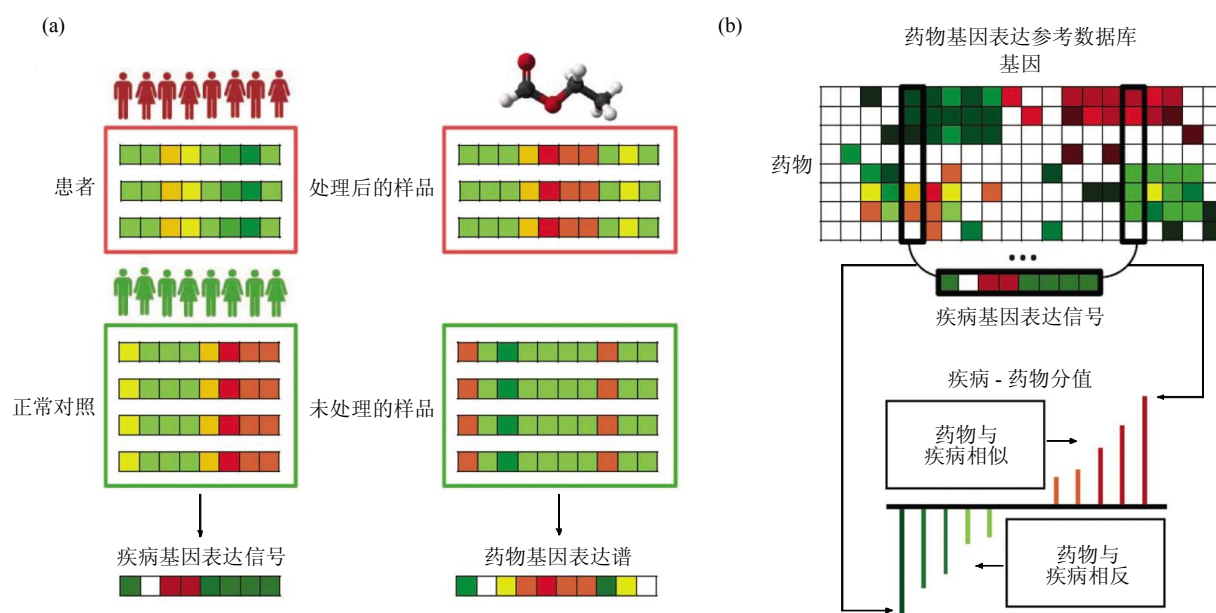


Fig. 1 Bioinformatics analytic strategy for drug repositioning using gene expression profiles^[21]

图1 利用生物信息技术分析基因表达谱数据预测药物新用途的分析策略^[21]

Butte 等从 NCBI 的 GEO 数据库^[15]中获得了涵盖 100 种疾病及它们相应正常对照的 176 组基因表达谱数据^[21-22], 这 176 组数据包含总数达 3 113 份样品的基因芯片结果^[21]. 由于这些基因芯片数据来自不同实验室的不同分析检测平台, 他们首先对这些数据进行了跨平台的分级标准化(rank normalization)^[27-28], 使用 AILUN 软件^[29]将不同平台基因芯片上命名不一的探针统一转换为标准的 NCBI GeneID, 同一芯片上一个基因如果使用了几个探针, 则取不同探针信号的平均值. 标准化后的数据采用 SAM 软件^[30]分别比较每一组数据中疾病与正常对照间的差异表达基因, 从而获得每一种疾病中上调和下调基因的列表, 这个基因列表就是该疾病的基因表达信号^[21].

2006 年, Lamb 及其同事基于“疾病、基因表达改变与药物作用之间必然存在功能上的联系”这一科学假说, 构建了用 164 种有生物活性的小分子化合物, 其中包括大量已被美国食品和药品管理局(FDA)批准的药物, 在不同浓度处理人类培养细胞后, 得到细胞的全基因组表达谱, 并开发了一套软件用于对这些药物影响的基因表达谱进行数据挖掘, 他们将这套系统称为 Connectivity Map^[31], Lamb 等希望利用 Connectivity Map 建立药物、基因与疾病间的联系(disease-gene- drug connections),

加速生物医药的研发. 利用 Connectivity Map, Lamb 等成功地鉴定出了两种所知甚少的天然产物分子可以调控热休克蛋白 90(heat-shock protein 90, HSP90), 抑制前列腺癌细胞中的雄激素受体信号通路(androgen-receptor signaling)^[32], 还发现了雷帕霉素(rapamycin)可以逆转急性淋巴母细胞白血病(acute lymphoblastic leukemia, ALL)对糖皮质激素(glucocorticoid)的耐药性^[33].

Butte 等充分利用了 Connectivity Map 提供的这 164 种药物的基因表达谱数据, 对它们同样进行了分级标准化, 然后将药物基因表达谱作为参考数据库(reference database), 将 Connectivity Map 的算法稍加修改, 对前面所述的 100 种疾病信号(disease signature)逐个与药物参考数据库进行比较, 以获得图 1 所述的疾病-药物分值^[21-22]. Butte 等使用的这套分析流程实际上借鉴了 GSEA(gene set enrichment analysis)的原理^[34]. GSEA 能用于评估两种生物学状态下一组有生物学联系的基因集合(gene set)的统计显著性, 克服了以往比较单个基因表达改变的不足, 已广泛地应用于基因芯片数据的分析^[35-36]. Butte 等将每种疾病信号所包含的差异表达基因作为一个 gene set, 所得到的疾病-药物分值就是该疾病的 gene set 在药物基因表达谱参考数据库中得到的富集值(enrichment score). 在该分析

过程中, Butte 等引入了置换(permutation)的策略, 对每一种疾病均随机生成一组与药物信号相同规模的数据, 重新计算模拟的疾病 - 药物分值, 并重复该步骤 100 次, 以计算假发现率(false discovery rate, FDR, 用 q 值表示), 排除偶然因素导致的假阳性. 通过上述方法, 他们比对了 16 400 种疾病 - 药物组合, 其中 2 664 种疾病 - 药物组合有显著的统计学意义($q < 0.05$), 超过半数的药物影响和疾病表型呈负相关(即可供治疗的药物 - 疾病关系). 在 100 种疾病中有 53 种疾病存在至少一种可供治疗的药物 - 疾病关系, 而 164 种药物都与这 53 种疾病中的至少一种存在显著相关^[21].

随后, Butte 等用剩下的 53 种疾病与 164 种药物分别对应的疾病 - 药物分值, 也称为治疗分值(therapeutic scores), 进行了层积聚类(hierarchical clustering)分析, 发现, 具有相似作用机制的药物聚集在一起, 如组蛋白去乙酰化酶(histone deacetylase, HDAC)抑制剂, 包括 trichostatin A、HC Toxin 和 vorinostat 等聚在一起, 也就是说这些药物可能有着相似的可治疗的疾病谱. 同样, 相似的疾病如肺腺癌、胃癌和胸膜间皮瘤等肿瘤聚在一起, 炎症肠道疾病(inflammatory bowel disease, IBD)如溃疡性结肠炎(ulcerative colitis, UC)和克罗恩氏病(Crohn's disease, CD)则集中在聚类树的另一分支中^[21]. 通过上述系统地“疾病 - 基因表达谱 - 药物”分析, Butte 等验证了一些已用于临床的药物和它们所治疗的疾病之间的联系: 比如在预测出的对炎症肠道疾病有最明显治疗分值的几个药物中就包括了皮质类激素氢化波尼松(prednisolone), 治疗分值为-0.216, 也就是说氢化波尼松影响的细胞基因表达信号与溃疡性结肠炎和克罗恩氏病的表达改变存在明显的负相关, 而氢化波尼松正是临床上常用的溃疡性结肠炎和克罗恩氏病的治疗药物. 又如, HDAC 抑制剂与脑瘤、食道癌、肺癌、结肠癌等恶性肿瘤之间有较明显的治疗分值, 而 HDAC 的抑瘤功能也已在实验室得到了验证^[37-38], 这进一步验证了 Butte 等所用的思路和策略的准确性.

3 旧药新用途的实验验证

有趣的是, 临床上常规用于治疗癫痫的药物托吡酯与克罗恩氏病之间得到了比氢化波尼松更明显的治疗分值, 达到了-0.220, 托吡酯与溃疡性结肠炎间的治疗分值同样较高, 而之前从未有过托吡酯可用于炎症肠道疾病, 甚至其他任何肠

道疾病的报道^[21]. 因此, Butte 等在三硝基苯磺酸(trinitrobenzenesulfonic acid, TNBS)诱导的大鼠结肠炎模型中验证了托吡酯对炎症肠道疾病的治疗作用. 实验大鼠被分为 4 组, 每组 12 只: 第一组不做任何处理, 其他 3 组每天根据体重按 100 mg/kg 给予 TNBS 处理以诱导大肠炎, 第二组只进行了 TNBS 处理, 第三组在用 TNBS 诱导的同时用托吡酯处理(每天按 80 mg/kg 体重给药, 实验组), 第四组则在 TNBS 诱导的同时用氢化波尼松处理(每天按 3 mg/kg 体重给药, 阳性对照), 连续用药 7 天, 在第 3 天和第 7 天用微型内窥镜观察大鼠肠内溃疡情况, 实验结束后处死大鼠, 病理切片检测各组大鼠结肠壁黏膜损伤情况. 结果表明托吡酯和氢化波尼松均可明显缓解 TNBS 诱导大鼠的腹泻, 而内窥镜和病理切片结果则显示, 托吡酯甚至比氢化波尼松更有效地治疗 TNBS 诱导的大鼠结肠炎, 明显抑制了 TNBS 诱导的结肠肿胀、溃疡和肠壁黏膜损伤^[21].

Butte 等重新审视了在进行药物 - 疾病预测时用到的基因表达谱数据, 通过比较克罗恩氏病的疾病基因表达信号和托吡酯的基因表达谱, 发现主要是一些免疫应答相关的基因和通路在二者之间存在明显的负相关. 因此, 他们在 TNBS 诱导并用托吡酯治疗的大鼠结肠组织中, 通过定量 PCR 验证了免疫应答相关分子 TRPV1 和 IFI30 的表达改变情况, TNBS 诱导后, 与临床上克罗恩氏病基因表达谱类似, 大鼠结肠组织中 TRPV1 基因表达下调, IFI30 表达上调, 而托吡酯处理能有效地逆转这些基因在疾病中的表达改变, 即 TRPV1 基因表达升高, IFI30 基因表达下降了, 这就从分子机制上进一步解释了托吡酯对免疫性肠病的治疗作用^[21].

相似地, Butte 等发现治疗溃疡的药物西咪替丁与肺腺癌之间存在中等强度的疾病 - 药物分值(-0.088), 尽管在他们的预测结果中还有一些药物对肺腺癌的治疗分值要比西咪替丁高, 但从未有文献报道西咪替丁可用于治疗肺腺癌, 同时西咪替丁也是一种很容易买到、价格比较便宜、且副作用比较小的常规药物, 另外西咪替丁对肺腺癌的治疗分值虽然不如托吡酯与炎症肠道疾病之间那么明显, 但是比广泛应用于肺腺癌治疗的吉非替尼(gefitinib)与肺腺癌间预测出的疾病 - 药物分值(-0.075)还略高一些. 因此, Butte 等决定通过实验来验证西咪替丁对肺腺癌的治疗作用. 他们利用 MTT [3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyl-tetrazolium bromide]

和 TUNEL (terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick end labeling) 等体外实验, 及用重度联合免疫缺陷 (severe combined immunodeficient, SCID) 小鼠构建的移植瘤体内模型, 分别验证了西咪替丁对肺腺癌细胞系 A549 的抑瘤功能, 发现西咪替丁在体内和体外均可抑制 A549 细胞的增殖, 诱导 A549 细胞的凋亡^[21]. 为了验证上述药物 - 疾病预测的特异性, Butte 等还检测了西咪替丁对肾细胞癌的抑瘤能力, 在预测结果中, 西咪替丁与肾细胞癌之间没有相关性, 疾病 - 药物分值为 0 ($P = 1.0$), 在他们构建的 SCID 荷瘤小鼠体内实验模型中, 西咪替丁对肾细胞癌 (ACHN 细胞系) 果然没有抑制作用^[21].

4 研究意义和展望

炎性肠道疾病和肺腺癌都是发病率较高且目前临床上比较难治的疾病. 炎性肠道疾病是一类慢性进展性的肠道免疫缺陷疾病, 临床上以溃疡性结肠炎和克罗恩氏病最为常见, 仅北美每年就有超过 100 万人发病, 目前尚无有效的治愈方法, 常用的治疗方案包括长期服用皮质激素类及系统免疫抑制药物以控制原发性炎症, 辅以抗生素治疗控制二次感染, 仅能达到控制综合征, 缓解症状, 防止复发的作用, 且这些药物都存在明显的副作用. 一些靶向治疗药物, 如抗 TNF- α 类药, 则价格高昂且仅对部分患者有效, 未能广泛使用, 手术切除部分感染的肠道也是治疗方案之一, 但是不仅费用高、创伤大, 且手术后容易在其他部位复发^[22]. 肺腺癌则是全世界发病率和死亡率最高的肿瘤之一^[39-41]. 因此这些疾病都需要有更好的治疗方法^[42-43]. Butte 等的研究结果为这两种疾病提供了新的潜在的治疗药物和方案, 尽管他们提供的数据仍然是初步的, 人们还需要进行一系列严格的实验来证实这两种被赋予新用途的药物在人体内的功效. 更重要的是, Butte 等提供了一个通过充分挖掘已有数据资料, 快捷地发现药物新用途的良好范例.

当然, Butte 等使用的基因表达谱数据资料, 反映的是各种疾病及各种药物处理后细胞内蛋白质编码基因 mRNA 水平的状况. 我们知道, 蛋白质才是基因发挥生物学功能的主要执行者, 可以说蛋白质决定表型, 研究蛋白质的表达、结构和功能是研究基因发挥生物学功能的最直接最有说服力的层面. mRNA 的表达水平体现了遗传信息在转录水平传递的状况, 但并不能全面反映蛋白质表达水平

的高低, 因为 mRNA 丰度与蛋白质丰度的相关性并不好, 更重要的是, 蛋白质具有进一步复杂的翻译后剪接和修饰, 而这些过程则几乎无法从 mRNA 水平来判断. 随着蛋白质组学研究理论和技术^[44-53]的发展, 特别是定量蛋白质组学技术的出现^[54-57], 为从蛋白质组的层面全方位地了解疾病的发病机制^[58-69], 寻找疾病新的分子靶标和治疗靶点^[60]提供了有力的技术条件. 可以预计, 随着公共数据库中各种疾病相关的, 以及各种药物处理相关的蛋白质组学数据的快速增长, 不久我们就可以充分利用这些蛋白质组学数据取代基因芯片表达谱数据, 或者作为基因芯片表达谱数据的重要补充, 为新药研发^[67-69]及开发旧药的新用途提供更多的帮助.

另外, 各种蛋白质也不是孤立地在起作用, 疾病的发生以及药物的作用都涉及到多层次、多途径的分子事件和复杂的基因调控网络. 在后基因组时代, 研究疾病的发生和药物的作用机制, 都必然要求在全基因组层面, 从整体、动态、网络的水平全方位地对蛋白质调控网络, 以及蛋白质与蛋白质间的调控规律进行系统深入地解析^[70-88]. 蛋白质相互作用网络体现了蛋白质组学层面的系统描述, 可以使人们在分子水平上更好地理解疾病发生和药物作用的生理机制, 这些机制的阐明对新药的研发及旧药新用途的开发无疑会起到重要的推动作用.

表观遗传修饰, 如 DNA 甲基化、组蛋白乙酰化等的异常^[89-92], 既是疾病发生的重要机制, 也是新药研发的重要靶点. 而以微小 RNA (microRNA)^[93-99] 为主的非编码 RNA (non-coding RNA)^[100-102] 也因为参与了基因表达的调控而日益受到重视. 新一代测序技术的发展^[103-105], 为我们提供了更为丰富的从 DNA 的变异和突变到 mRNA 的不同剪接形式, 从 DNA 甲基化修饰到 micorRNA 的表达水平等各类不同层面的基因表达及调控网络的海量数据^[106-112], 充分利用生物信息技术挖掘这些数据, 将是今后高效地寻找特异性、个体化的分子靶点及创新药物研制的希望所在.

参 考 文 献

- [1] DiMasi J A, Hansen R W, Grabowski H G. The price of innovation: new estimates of drug development costs. *J Health Econ*, 2003, **22**(2): 151-185
- [2] Boguski M S, Mandl K D, Sukhatme V P. Drug discovery. Repurposing with a difference. *Science*, 2009, **324**(5933): 1394-1395
- [3] Chong C R, Sullivan D J Jr. New uses for old drugs. *Nature*, 2007,

- 448(7154): 645–646
- [4] Ashburn T T, Thor K B. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov*, 2004, **3**(8): 673–683
 - [5] Sampaio E P, Sarno E N, Galilly R, *et al.* Thalidomide selectively inhibits tumor necrosis factor alpha production by stimulated human monocytes. *J Exp Med*, 1991, **173**(3): 699–703
 - [6] Thompson C. Thalidomide effective for AIDS-related oral ulcers. *Lancet*, 1995, **346**(8985): 1289
 - [7] Jacobson J M, Greenspan J S, Spritzler J, *et al.* Thalidomide for the treatment of oral aphthous ulcers in patients with human immunodeficiency virus infection. National Institute of Allergy and Infectious Diseases AIDS Clinical Trials Group. *N Engl J Med*, 1997, **336**(21): 1487–1493
 - [8] D'Amato R J, Loughnan M S, Flynn E, *et al.* Thalidomide is an inhibitor of angiogenesis. *Proc Natl Acad Sci USA*, 1994, **91**(9): 4082–4085
 - [9] Barlogie B, Tricot G, Anaissie E, *et al.* Thalidomide and hematopoietic-cell transplantation for multiple myeloma. *N Engl J Med*, 2006, **354**(10): 1021–1030
 - [10] Xiong W, Wu X, Starnes S, *et al.* An analysis of the clinical and biologic significance of TP53 loss and the identification of potential novel transcriptional targets of TP53 in multiple myeloma. *Blood*, 2008, **112**(10): 4235–4246
 - [11] 易文渊, 徐波, 李敏, 等. N糖基取代的沙利度胺新衍生物对人鼻咽癌细胞耐药逆转作用研究. *生物化学与生物物理进展*, 2009, **36**(1): 58–64
Yi W Y, Xu B, Li M, *et al.* *Prog Biochem Biophys*, 2009, **36**(1): 58–64
 - [12] Xiong W, Zeng Z Y, Xia J H, *et al.* A susceptibility locus at chromosome 3p21 linked to familial nasopharyngeal carcinoma. *Cancer Res*, 2004, **64**(6): 1972–1974
 - [13] He H, Liu B, Zhang X, *et al.* Development of a high-throughput assay for the HIV-1 integrase disintegration reaction. *Sci China Life Sci*, 2010, **53**(2): 241–247
 - [14] Zeng Z Y, Zhou Y H, Zhang W L, *et al.* Gene expression profiling of nasopharyngeal carcinoma reveals the abnormally regulated Wnt signaling pathway. *Hum Pathol*, 2007, **38**(1): 120–133
 - [15] Barrett T, Suzek T O, Troup D B, *et al.* NCBI GEO: Mining millions of expression profiles—database and tools. *Nucl Acid Res*, 2005, **33**(Database issue): D562–566
 - [16] Zeng Z Y, Zhou Y H, Xiong W, *et al.* Analysis of gene expression identifies candidate molecular markers in nasopharyngeal carcinoma using microdissection and cDNA microarray. *J Cancer Res Clin Oncol*, 2007, **133**(2): 71–81
 - [17] Zhou Y H, Zeng Z Y, Zhang W L, *et al.* Identification of candidate molecular markers of nasopharyngeal carcinoma by microarray analysis of subtracted cDNA libraries constructed by suppression subtractive hybridization. *Eur J Cancer Prev*, 2008, **17**(6): 561–571
 - [18] 周艳宏, 张必成, 曾朝阳, 等. 利用 GenMAPP 筛查鼻咽癌差异表达基因. *生物化学与生物物理进展*, 2005, **32**(12): 1121–1129
Zhou Y H, Zhang B C, Zeng Z Y, *et al.* *Prog Biochem Biophys*, 2005, **32**(12): 1121–1129
 - [19] 邵雪君, 陈子兴, 缪美华, 等. 基于荧光微球液相基因表达阵列的建立. *生物化学与生物物理进展*, 2011, **38**(7): 661–669
Shao X J, Chen Z X, Miao M H, *et al.* *Prog Biochem Biophys*, 2011, **38**(7): 661–669
 - [20] 刘伟, 谢红卫. 基于生物信息学方法发现潜在药物靶标. *生物化学与生物物理进展*, 2011, **38**(1): 11–19
Liu W, Xie H W. *Prog Biochem Biophys*, 2011, **38**(1): 11–19
 - [21] Sirota M, Dudley J T, Kim J, *et al.* Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci Transl Med*, 2011, **3**(96): 96ra77
 - [22] Dudley J T, Sirota M, Shenoy M, *et al.* Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci Transl Med*, 2011, **3**(96): 96ra76
 - [23] Lussier Y A, Chen J L. The emergence of genome-based drug repositioning. *Sci Transl Med*, 2011, **3**(96): 96ps35
 - [24] Zhou M, Liu Z, Zhao Y, *et al.* MicroRNA-125b confers the resistance of breast cancer cells to paclitaxel through suppression of pro-apoptotic Bcl-2 antagonist killer 1 (Bak1) expression. *J Biol Chem*, 2010, **285**(28): 21496–24507
 - [25] Cao J, Fu L, Sun D, *et al.* Taxol inhibits stretch-induced electrophysiological alterations in isolated rat hearts with acute myocardial infarction. *Sci China Life Sci*, 2010, **53**(8): 1009–1014
 - [26] Meng R, Zhou J, Sui M, *et al.* Arsenic trioxide promotes mitochondrial DNA mutation and cell apoptosis in primary APL cells and NB4 cell line. *Sci China Life Sci*, 2010, **53**(1): 87–93
 - [27] Wamat P, Eils R, Brors B. Cross-platform analysis of cancer microarray data improves gene expression based classification of phenotypes. *BMC Bioinformatics*, 2005, **6**: 265
 - [28] 谈效俊, 张永新, 钱敏平, 等. 芯片数据标准化方法比较研究. *生物化学与生物物理进展*, 2007, **34**(6): 625–633
Tan X J, Zhang Y X, Qian M P, *et al.* *Prog Biochem Biophys*, 2007, **34**(6): 625–633
 - [29] Chen R, Li L, Butte A J. AILUN: Reannotating gene expression data automatically. *Nat Methods*, 2007, **4**(11): 879
 - [30] Tusher V G, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA*, 2001, **98**(9): 5116–5121
 - [31] Lamb J, Crawford E D, Peck D, *et al.* The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science*, 2006, **313**(5795): 1929–1935
 - [32] Hieronymus H, Lamb J, Ross K N, *et al.* Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. *Cancer Cell*, 2006, **10**(4): 321–330
 - [33] Wei G, Twomey D, Lamb J, *et al.* Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell*, 2006, **10**(4): 331–342
 - [34] Subramanian A, Tamayo P, Mootha V K, *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*, 2005, **102**(43): 15545–15550
 - [35] Zhang W, Zeng Z, Zhou Y, *et al.* Identification of aberrant cell

- cycle regulation in Epstein-Barr virus- associated nasopharyngeal carcinoma by cDNA microarray and gene set enrichment analysis. *Acta Biochim Biophys Sin*, 2009, **41**(5): 414-428
- [36] 黄 波, 李文婷, 李 雯, 等. 基因集合的参数统计分析方法. *生物化学与生物物理进展*, 2009, **36**(11): 1415-1422
- Hwang B, Li W T, Li W, *et al.* *Prog Biochem Biophys*, 2009, **36**(11): 1415-1422
- [37] 蒋智文, 刘新光, 周中军. 组蛋白修饰调节机制的研究进展. *生物化学与生物物理进展*, 2009, **36**(10): 1252-1259
- Jiang Z W, Liu X G, Zhou Z J. *Prog Biochem Biophys*, 2009, **36**(10): 1252-1259
- [38] 王秀莉, 冯云鹏, 赵 静, 等. 可逆的组蛋白乙酰化调控 p16 的表达. *生物化学与生物物理进展*, 2010, **37**(6): 600-606
- Wang X L, Feng Y P, Zhao J, *et al.* *Prog Biochem Biophys*, 2010, **37**(6): 600-606
- [39] 欧阳卿, 胡瑞成, 戴爱国, 等. 吸烟对肺组织 D4-GDI 表达的影响及其与慢性阻塞性肺疾病相关性的蛋白质组学研究. *生物化学与生物物理进展*, 2010, **37**(11): 1204-1211
- Ouyang Q, Hu R C, Dai A G, *et al.* *Prog Biochem Biophys*, 2010, **37**(11): 1204-1211
- [40] 赵艾君, 郑毛根, 王国臣, 等. RNA 干扰 ABCE1 基因后可增加肺癌 95-D/NCI-H446 细胞的 E- 钙黏附蛋白表达并减低细胞侵袭力. *生物化学与生物物理进展*, 2010, **37**(8): 891-896
- Zhao A J, Zheng M G, Wang G C, *et al.* *Prog Biochem Biophys*, 2010, **37**(8): 891-896
- [41] 高书颖, 李恩民, 孟令英, 等. 肺癌细胞中调控 ezrin 基因基本转录活性的顺式作用元件及转录因子的鉴定. *生物化学与生物物理进展*, 2009, **36**(3): 288-296
- Gao S Y, Li E M, Meng L Y, *et al.* *Prog Biochem Biophys*, 2009, **36**(3): 288-296
- [42] 郑毛根, 高 英, 黄 波, 等. RNA 干扰 ABCE1 基因后可抑制肺癌 95-D/NCI-H446 细胞的增殖和诱导细胞凋亡. *生物化学与生物物理进展*, 2009, **36**(11): 1475-1482
- Zheng M G, Gao Y, Huang B, *et al.* *Prog Biochem Biophys*, 2009, **36**(11): 1475-1482
- [43] Yue W, Zhao X, Zhang L, *et al.* Overexpression of cyclin Y in non-small cell lung cancer is associated with cancer cell proliferation. *Sci China Life Sci*, 2010, **53**(4): 511-516
- [44] He F. At a glance: Proteomics in China. *Sci China Life Sci*, 2011, **54**(1): 1-2
- [45] Gao X, Zhang X, Zheng J, *et al.* Proteomics in China: Ready for prime time. *Sci China Life Sci*, 2010, **53**(1): 22-33
- [46] Li G, Yang X, Zeng H, *et al.* Stable isotope labelled mass spectrometry for quantification of the relative abundances for expressed proteins induced by PeaT1. *Sci China Life Sci*, 2010, **53**(12): 1410-1417
- [47] Mi W, Liu X, Jia W, *et al.* Toward a high resolution 2-DE profile of the normal human liver proteome using ultra-zoom gels. *Sci China Life Sci*, 2011, **54**(1): 25-33
- [48] Zhou N, Wang N, Xu B, *et al.* Whole-cell matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for rapid identification of bacteria cultured in liquid media. *Sci China Life Sci*, 2011, **54**(1): 48-53
- [49] 李 宁, 吴松锋, 朱云平, 等. 鸟枪法蛋白质鉴定质量控制方法研究进展. *生物化学与生物物理进展*, 2009, **36**(6): 668-675
- Li N, Wu S F, Zhu Y P, *et al.* *Prog Biochem Biophys*, 2009, **36**(6): 668-675
- [50] 孙瑞祥, 董梦秋, 迟 浩, 等. 基于电子捕获裂解 / 电子转运裂解串联质谱技术的蛋白质组学研究. *生物化学与生物物理进展*, 2010, **37**(1): 94-102
- Sun R X, Dong M Q, Chi H, *et al.* *Prog Biochem Biophys*, 2010, **37**(1): 94-102
- [51] 章 杰, 刘琼明, 许丹科, 等. 无细胞系统原位制备蛋白质芯片及其应用. *生物化学与生物物理进展*, 2009, **36**(4): 391-397
- Zhang J, Liu Q M, Xu D K, *et al.* *Prog Biochem Biophys*, 2009, **36**(4): 391-397
- [52] 蔡潭溪, 刘平生, 杨福全, 等. 脂质组学研究进展. *生物化学与生物物理进展*, 2010, **37**(2): 121-128
- Cai T X, Liu P S, Yang F Q, *et al.* *Prog Biochem Biophys*, 2010, **37**(2): 121-128
- [53] 孙汉昌, 张纪阳, 刘 辉, 等. 串联质谱图谱从头测序算法研究进展. *生物化学与生物物理进展*, 2010, **37**(12): 1278-1288
- Sun H C, Zhang J Y, Liu H, *et al.* *Prog Biochem Biophys*, 2010, **37**(12): 1278-1288
- [54] 张 伟, 张纪阳, 刘 辉, 等. 蛋白质质谱分析的无标记定量算法研究进展. *生物化学与生物物理进展*, 2011, **38**(6): 506-518
- Zhang W, Zhang J Y, Liu H, *et al.* *Prog Biochem Biophys*, 2011, **38**(6): 506-518
- [55] 熊 霞, 沈剑英, 李建军, 等. 大鼠背根神经节细胞质膜的双水相法纯化及其蛋白质组学研究. *生物化学与生物物理进展*, 2009, **36**(11): 1458-1468
- Xiong X, Shen J Y, Li J J, *et al.* *Prog Biochem Biophys*, 2009, **36**(11): 1458-1468
- [56] 张志强, 李茂玉, 张桂英, 等. 激光捕获显微切割技术结合 18O 标记定量蛋白质组技术在胃癌标志物筛查中的应用研究. *生物化学与生物物理进展*, 2009, **36**(3): 311-322
- Zhang Z Q, Li M Y, Zhang G Y, *et al.* *Prog Biochem Biophys*, 2009, **36**(3): 311-322
- [57] 米 薇, 王 晶. 基于电感耦合等离子体质谱的蛋白质定量技术研究进展. *生物化学与生物物理进展*, 2010, **37**(2): 224-229
- Mi W, Wang J. *Prog Biochem Biophys*, 2010, **37**(2): 224-229
- [58] Li G, Xiao Z, Liu J, *et al.* Cancer: A proteomic disease. *Sci China Life Sci*, 2011, **54**(5): 403-408
- [59] Wang L, Chen G. Current advances in the application of proteomics in apoptosis research. *Sci China Life Sci*, 2011, **54**(3): 209-219
- [60] 叶 桦, 刘 伟, 虞朝辉, 等. 非酒精性脂肪性肝病蛋白质组学研究进展. *生物化学与生物物理进展*, 2011, **38**(6): 487-498
- Ye H, Liu W, Yu C H, *et al.* *Prog Biochem Biophys*, 2011, **38**(6): 487-498
- [61] Cao J, Shen C, Zhang J, *et al.* Comparison of alternative extraction methods for secretome profiling in human hepatocellular carcinoma cells. *Sci China Life Sci*, 2011, **54**(1): 34-38
- [62] Gao J, Gao G, Zhang Y, *et al.* Proteomic analysis of human epithelial ovarian cancer xenografts in immunodeficient mice

- exposed to chronic psychological stress. *Sci China Life Sci*, 2011, **54**(2): 112-120
- [63] 刘迎福, 肖志强, 张鹏飞, 等. 人原发性肺腺癌转移相关分子的定量蛋白质组学研究. *生物化学与生物物理进展*, 2009, **36**(4): 448-457
- Liu Y F, Xiao Z Q, Zhang P F, *et al.* *Prog Biochem Biophys*, 2009, **36**(4): 448-457
- [64] 李美香, 肖志强, 彭芳, 等. LCM 纯化的鼻咽癌间质和正常鼻咽间质的定量蛋白质组学研究. *生物化学与生物物理进展*, 2009, **36**(9): 1122-1133
- Li M X, Xiao Z Q, Peng F, *et al.* *Prog Biochem Biophys*, 2009, **36**(9): 1122-1133
- [65] 朱益民, 徐其, 董来荣, 等. 应用蛋白质组学技术筛选胃癌 SGC-7901 细胞中的 let-7a 功能相关蛋白. *生物化学与生物物理进展*, 2011, **38**(5): 441-448
- Zhu Y M, Xu Q, Dong L R, *et al.* *Prog Biochem Biophys*, 2011, **38**(5): 441-448
- [66] Shao C, Wang Y, Gao Y. Applications of urinary proteomics in biomarker discovery. *Sci China Life Sci*, 2011, **54**(5): 409-417
- [67] Huang L, Xiao P, Guo L, *et al.* Molecular pharmacognosy. *Sci China Life Sci*, 2010, **53**(6): 643-652
- [68] Li F, Wen H, Zhang Y, *et al.* Purification and characterization of a novel immunomodulatory protein from the medicinal mushroom *Trametes versicolor*. *Sci China Life Sci*, 2011, **54**(4): 379-385
- [69] 曹长姝, 沈伟哉, 李药兰, 等. 中药臭灵丹中 3, 5-二羟基-6, 7, 3', 4'-四甲氧基黄酮对人鼻咽癌 CNE 细胞凋亡的影响及机制. *生物化学与生物物理进展*, 2011, **38**(3): 254-261
- Cao C S, Shen W Z, Li Y L, *et al.* *Prog Biochem Biophys*, 2011, **38**(3): 254-261
- [70] 赵研, 陈丽娜, 张良才, 等. 基于功能一致性和网络拓扑属性预测冠心病致病基因. *生物化学与生物物理进展*, 2009, **36**(6): 781-786
- Zhao Y, Chen L N, Zhang L C, *et al.* *Prog Biochem Biophys*, 2009, **36**(6): 781-786
- [71] 王宏, 曲晓莉, 赵研, 等. 基于表达及网络拓扑结构挖掘动脉粥样硬化风险疾病基因. *生物化学与生物物理进展*, 2010, **37**(8): 916-922
- Wang H, Qu X L, Zhao Y, *et al.* *Prog Biochem Biophys*, 2010, **37**(8): 916-922
- [72] 罗湘建, 曹亚. 肿瘤能量代谢机制研究进展. *生物化学与生物物理进展*, 2011, **38**(7): 585-592
- Luo X J, Cao Y. *Prog Biochem Biophys*, 2011, **38**(7): 585-592
- [73] 韦晓兰. 细胞黏附分子互作网络的构建与分析. *生物化学与生物物理进展*, 2011, **38**(4): 347-352
- Wei X L. *Prog Biochem Biophys*, 2011, **38**(4): 347-352
- [74] Yao C, Zhang M, Zou J, *et al.* Functional modules with disease discrimination abilities for various cancers. *Sci China Life Sci*, 2011, **54**(2): 189-193
- [75] 向波, 王理, 易梅, 等. 鼻咽癌下调新基因 NOR1 相互作用蛋白的筛选和鉴定. *生物化学与生物物理进展*, 2009, **36**(6): 709-714
- Xiang B, Wang L, Yi M, *et al.* *Prog Biochem Biophys*, 2009, **36**(6): 709-714
- [76] 蒯艳, 钱志明, 柯亚. 铁代谢蛋白在肾脏的表达与功能. *生物化学与生物物理进展*, 2011, **38**(2): 113-118
- Lin Y, Qian Z M, Ke Y. *Prog Biochem Biophys*, 2011, **38**(2): 113-118
- [77] Zhao M, Chen Y, Qu D, *et al.* A database of transporter substrates linking metabolic pathways and transporter systems on a genome scale *via* their shared substrates. *Sci China Life Sci*, 2011, **54**(1): 60-64
- [78] 刘齐军, 王正华, 刘万霖, 等. 一种基于基因表达模型识别酵母细胞周期条件特异调控子网的方法. *生物化学与生物物理进展*, 2010, **37**(4): 402-415
- Liu Q J, Wang Z H, Liu W L, *et al.* *Prog Biochem Biophys*, 2010, **37**(4): 402-415
- [79] 刘中扬, 李栋, 朱云平, 等. 蛋白质相互作用网络进化分析研究进展. *生物化学与生物物理进展*, 2009, **36**(1): 13-24
- Liu Z Y, Li D, Zhu Y P, *et al.* *Prog Biochem Biophys*, 2009, **36**(1): 13-24
- [80] 孙汉昌, 李栋, 王建, 等. PNmerger: 一个整合生物学通路和蛋白质相互作用网络的 Cytoscape 插件. *生物化学与生物物理进展*, 2009, **36**(12): 1613-1616
- Sun H C, Li D, Wang J, *et al.* *Prog Biochem Biophys*, 2009, **36**(12): 1613-1616
- [81] 陈丽娜, 王倩, 尚玉奎, 等. 人类蛋白质结构互作网络——结构域对网络拓扑与蛋白质功能的影响. *生物化学与生物物理进展*, 2010, **37**(5): 517-526
- Chen L N, Wang Q, Shang Y K, *et al.* *Prog Biochem Biophys*, 2010, **37**(5): 517-526
- [82] 李非, 伯晓晨, 李鹏, 等. 一种融合表达谱相关性信息的激活子网辨识算法. *生物化学与生物物理进展*, 2010, **37**(2): 208-217
- Li F, Bo X C, Li P, *et al.* *Prog Biochem Biophys*, 2010, **37**(2): 208-217
- [83] 欧阳玉梅. 结构域相互作用数据库的产生、发展与应用. *生物化学与生物物理进展*, 2009, **36**(3): 280-287
- Ouyang Y M. *Prog Biochem Biophys*, 2009, **36**(3): 280-287
- [84] Gong X, Liu B, Chang S, *et al.* A holistic molecular docking approach for predicting protein-protein complex structure. *Sci China Life Sci*, 2010, **53**(9): 1152-1161
- [85] 周婷婷, 容健锋, 陈振冲, 等. MetaGen: 从 KEGG 建模代谢网络的新工具. *生物化学与生物物理进展*, 2010, **37**(1): 63-68
- Zhou T T, Yung K F, Chan C CK, *et al.* *Prog Biochem Biophys*, 2010, **37**(1): 63-68
- [86] 江丽华, 李亦学, 刘琪. 综合 ChIP-chip 数据、基因敲除数据和表达谱数据重构基因调控网络. *生物化学与生物物理进展*, 2010, **37**(9): 996-1005
- Jiang L H, Li Y X, Liu Q. *Prog Biochem Biophys*, 2010, **37**(9): 996-1005
- [87] 涂馨, 闵凌峰, 陈琼, 等. 氧化铁磁性纳米颗粒介导缺氧诱导因子 1ashRNA 质粒转染逆转 A549/CDDP 细胞对顺铂耐药的实验研究. *生物化学与生物物理进展*, 2010, **37**(10): 1090-1100
- Tu X, Min L F, Chen Q, *et al.* *Prog Biochem Biophys*, 2010, **37**(10): 1090-1100

- [88] 张红彦, 刘学琦, 孙玉娜, 等. 热休克因子结合蛋白(HSBP1)的表达、纯化、结晶与初步晶体学研究. 生物化学与生物物理进展, 2010, **37**(4): 441-444
Zhang H Y, Liu X Q, Sun Y N, *et al.* Prog Biochem Biophys, 2010, **37**(4): 441-444
- [89] 凡时财, 张学工. DNA 甲基化的生物信息学研究进展. 生物化学与生物物理进展, 2009, **36**(2): 143-150
Fan S C, Zhang X G. Prog Biochem Biophys, 2009, **36**(2): 143-150
- [90] 张祖萍, 武明花, 唐海林, 等. 5-Aza-CdR 对胶质瘤细胞生长及 LRRc4 基因异常甲基化的影响. 生物化学与生物物理进展, 2009, **36**(7): 904-909
Zhang Z P, Wu M H, Tang H L, *et al.* Prog Biochem Biophys, 2009, **36**(7): 904-909
- [91] 杜霞, 袁洪, 邢晓为. DNA 甲基化与原发高血压的研究进展. 生物化学与生物物理进展, 2010, **37**(4): 364-369
Du X, Yuan H, Xing X W. Prog Biochem Biophys, 2010, **37**(4): 364-369
- [92] 于建宁, 李少华, 王丹秋, 等. 5-脱氧杂氮胞苷抑制小鼠附植前的胚胎发育. 生物化学与生物物理进展, 2009, **36**(2): 228-237
Yu J N, Li S H, Wang D Q, *et al.* Prog Biochem Biophys, 2009, **36**(2): 228-237
- [93] 丛肖静, 谭建军, 刘明, 等. HIV-1 跨膜蛋白 gp41 与 N-取代吡咯衍生物的结合模式研究. 生物化学与生物物理进展, 2010, **37**(8): 904-915
Cong X J, Tan J J, Liu M, *et al.* Prog Biochem Biophys, 2010, **37**(8): 904-915
- [94] Yang C, Wei W. The miRNA expression profile of the uveal melanoma. Sci China Life Sci, 2011, **54**(4): 351-358
- [95] 郭志云, 茆灿泉, 熊莉丽, 等. p53 直接调控 microRNA 及其靶基因的高通量筛选. 生物化学与生物物理进展, 2009, **36**(9): 1154-1164
Guo Z Y, Mao C Q, Xiong L L, *et al.* Prog Biochem Biophys, 2009, **36**(9): 1154-1164
- [96] 马宁, 高旭. 转基因动物在 microRNA 研究中的应用. 生物化学与生物物理进展, 2009, **36**(9): 1095-1100
Ma N, Gao X. Prog Biochem Biophys, 2009, **36**(9): 1095-1100
- [97] 汪进业, 蔡荣, 罗本燕. MicroRNA 与神经系统重大疾病. 生物化学与生物物理进展, 2009, **36**(1): 25-32
Wang J Y, Cai R, Luo B Y. Prog Biochem Biophys, 2009, **36**(1): 25-32
- [98] 严丽梅, 吴建勇, 于潇华, 等. 内含子源性 microRNA 对内皮型一氧化氮合酶表达及血管内皮细胞增殖的作用. 生物化学与生物物理进展, 2010, **37**(7): 747-753
Yan L M, Wu J Y, Yu X H, *et al.* Prog Biochem Biophys, 2010, **37**(7): 747-753
- [99] 金幼芳, 徐根明, 李艳, 等. 定量 PCR 检测 microRNA 表达的数据归一化新技术. 生物化学与生物物理进展, 2011, **38**(5): 473-481
Jin Y F, Xu G M, Li Y, *et al.* Prog Biochem Biophys, 2011, **38**(5): 473-481
- [100] Zheng L, Qu L. Computational RNomics: Structure identification and functional prediction of non-coding RNAs in silico. Sci China Life Sci, 2010, **53**(5): 548-562
- [101] Jia Q, Wu H, Zhou X, *et al.* A "GC-rich" method for mammalian gene expression: A dominant role of non-coding DNA GC content in regulation of mammalian gene expression. Sci China Life Sci, 2010, **53**(1): 94-100
- [102] 黄文涛, 郭向前, 戴甲培, 等. MicroRNA, lncRNA 与神经退行性疾病. 生物化学与生物物理进展, 2010, **37**(8): 826-833
Huang W T, Guo X Q, Dai J P, *et al.* Prog Biochem Biophys, 2010, **37**(8): 826-833
- [103] Wu J, Xiao J, Zhang R, *et al.* DNA sequencing leads to genomics progress in China. Sci China Life Sci, 2011, **54**(3): 290-292
- [104] Zhou X, Ren L, Li Y, *et al.* The next-generation sequencing technology: A technology review and future perspective. Sci China Life Sci, 2010, **53**(1): 44-57
- [105] 王曦, 汪小我, 王立坤, 等. 新一代高通量 RNA 测序数据的处理与分析. 生物化学与生物物理进展, 2010, **37**(8): 834-846
Wang X, Wang X W, Wang L K, *et al.* Prog Biochem Biophys, 2010, **37**(8): 834-846
- [106] 曾朝阳, 熊炜, 沈守荣, 等. 利用动态等位基因特异性杂交技术进行单核苷酸多态高通量分型. 生物化学与生物物理进展, 2002, **29**(5): 806-810
Zeng Z Y, Xiong W, Shen S R, *et al.* Prog Biochem Biophys, 2002, **29**(5): 806-810
- [107] 熊炜, 曾朝阳, 肖炳焱, 等. 一个新硝基还原酶基因 NOR1 编码区单核苷酸多态及与鼻咽癌的关联分析. 生物化学与生物物理进展, 2003, **30**(3): 401-405
Xiong W, Zeng Z Y, Xiao B Y, *et al.* Prog Biochem Biophys, 2003, **30**(3): 401-405
- [108] 熊炜, 曾朝阳, 沈守荣, 等. 泛肽相关蛋白基因 UBAP1 单核苷酸多态及与鼻咽癌的相关研究. 生物化学与生物物理进展, 2002, **29**(5): 766-770
Xiong W, Zeng Z Y, Shen S R, *et al.* Prog Biochem Biophys, 2002, **29**(5): 766-770
- [109] 熊炜, 曾朝阳, 李小玲, 等. NGX6 基因单核苷酸多态及与鼻咽癌的相关性. 生物化学与生物物理学报, 2002, **34**(4): 512-515
Xiong W, Zeng Z Y, Li X L, *et al.* Acta Biochim Biophys Sin, 2002, **34**(4): 512-515
- [110] Hao D, Xiao P, Chen S. Phenotype prediction of nonsynonymous single nucleotide polymorphisms in human phase II drug/xenobiotic metabolizing enzymes: perspectives on molecular evolution. Sci China Life Sci, 2010, **53**(10): 1252-1262
- [111] Rajasekaran R, Sethumadhavan R. Exploring the structural and functional effect of pRB by significant nsSNP in the coding region of RB1 gene causing retinoblastoma. Sci China Life Sci, 2010, **53**(2): 234-240
- [112] 张惠丹, 王晓楠, 周喆, 等. 基于微流控芯片的温度梯度毛细管电泳检测大肠癌石蜡标本中 K-ras 基因突变. 生物化学与生物物理进展, 2010, **37**(7): 794-800
Zhang H D, Wang X N, Zhou Z, *et al.* Prog Biochem Biophys, 2010, **37**(7): 794-800

Bioinformatics Accelerates Drug Repositioning*

HUANG Hong-Bin^{1,2)}, LIANG Fang²⁾, XIONG Wei²⁾, LI Xiao-Ling²⁾, ZENG Zhao-Yang^{2)**}, LI Gui-Yuan^{2)**}

⁽¹⁾ Key Laboratory of Information System Engineering, National University of Defense Technology, Changsha 410073, China;

⁽²⁾ Key Laboratory of Carcinogenesis and Cancer Invasion of Ministry of Education, Cancer Research Institute, Central South University, Changsha 410078, China)

Abstract Traditional approaches to drug discovery are generally regarded as protracted and costly. The application of established drug compounds to new therapeutic indications, known as drug repositioning, offers several advantages over traditional drug development. Gene expression microarrays are regularly and broadly applied in clinical studies of human diseases, providing genome-wide characterization of a disease state. Microarrays are also widely used to discover gene expression patterns that signify pharmacologic perturbation, allowing for the development of high-quality signatures of drug effect. A pair of papers from one group recently published in *Science Translational Medicine* provided a concrete example of how to using bioinformatics approach to reinterpret and compare genome-wide gene expression data, that allows us to effectively hypothesize which drugs from one disease-indication can be repurposed for another disease. They examined publicly available gene expression data and determined the genes affected in 100 diseases and 164 drugs. By pairing drugs that correct abnormal gene expression in diseases, they confirmed known effective drug-disease pairs and predicted new indications for already approved agents. Experimental validation that an antiulcer drug and an antiepileptic can be reused for lung cancer and inflammatory bowel disease reinforced the promise of this approach. These two drugs are therefore good candidates for repositioning to treat lung cancer and inflammatory bowel disease that in need of better therapies, and we now have a way to mine available data for fast routes to new disease therapies.

Key words bioinformatics, gene expression profile, drug repositioning

DOI: 10.3724/SP.J.1206.2011.00453

*This work was supported by grants from The National Natural Science Foundation of China (30871282, 30871365, 81172189, 81171930), The Hunan Province Natural Sciences Foundation of China (10JJ7003), The Fok Ying Tong Education Foundation (121036), The Fundamental Research Funds for the Central Universities (2011JQ020) and The Postdoctoral Science Foundation of Central South University

**Corresponding author. Tel: 86-731-84805383

ZENG Zhao-Yang. E-mail: zengzhaoyang@xysm.net

LI Gui-Yuan. E-mail: ligy@xysm.net

Received: October 10, 2011 Accepted: November 18, 2011