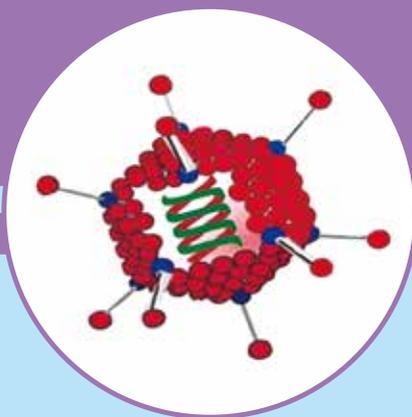


第一章

基因表达系统模型在肿瘤精准 治疗中的临床应用



第一节 概述

个体化治疗，亦称精准治疗，是一种为患者量身定制的新型医疗模式，旨在“在正确的时间为正确的患者提供正确的治疗”。成功的个体化治疗方案制订建立在精准的临床基因组分析基础上。随着人工智能技术的发展和临床基因组表达分析水平的提高，基因组表达谱和系统建模在个体化治疗中的应用日益增多，为肿瘤疾病的靶向治疗开辟了新途径。本章将详细介绍从临床取样、基因组 mRNA 表达水平分析、诊断，到通过系统建模发现特异性基因表达特征（gene expression signature, GES），最终从药物库中筛选敏感药物并应用于临床的整个过程。随着二代测序（next-generation sequencing, NGS）技术的不断发展，与药物发现相关的系统建模对于新一代治疗方法的开发而言将是至关重要的。

科学家经多年研究已发现，癌症并非单一、同质化的疾病，而是一种多因素导致的复杂疾病，这对癌症治疗的有效性构成了巨大挑战^[1]。面对癌症，医师急需引入新一代治疗工具。特别是人工智能技术的发展为个体化用药、提高癌症治疗效果带来了新的突破^[2]。“个体化治疗”或称“精准医疗”有别于传统医学僵化的统一用药方式，它是一种基于每个患者个体特点的新型医疗模式。检测患者的基因组信息是实现个体化治疗的前提。基因组学图谱（profile）整合患者基因组信息，预测疾病对药物的敏感度和药物的个体安全性^[3]。癌症研究的最新进展使医师和科学家可以通过对单核苷酸多态性（SNP）的全基因组关联研究（GWAS）了解某些类型肿瘤对特异性治疗的反应^[4]。

大多数经 FDA 批准的药物都是改变肿瘤细胞表型，这些改变与 mRNA 和蛋白质表达水平相关，而不是针对患者 DNA 中的 SNP 信息进行纠正。因此，SNP 变异导致的 mRNA 变化和蛋白质产物变化，对新一代药物的研发有很大的影响^[5]。研究人员还在肿瘤细胞中发现了其他与 mRNA 或蛋白质表达改变有关的基因组学图谱，如表观遗传学谱、microRNA 谱和非编码 DNA 谱（图 1-1-1）^[6]。为了明确治疗的靶点，科学家和医师们已经深入研究肿瘤细胞的基因组表达图谱，以及这些表达图谱与其他图谱之间的组合图谱，如转录组—GWAS 谱、转录组—表观遗传学谱、转录组—microRNA 谱和全基因组学谱或称为 DNA 元素百科全书（ENCODE）^[7]。更为重要的

是，由于许多患者特定的个体因素会影响药物的吸收和代谢^[8]，因此最佳治疗效果和最小毒性的最佳剂量选择，取决于该患者个体基因表达水平的动态变化^[9]。理论上，在各种基因组表达图谱中，蛋白质组学图谱本应是用来筛选生物标志物和治疗靶点并进行系统建模的最佳参考标准^[10]，但是，现有的研究证据（例如临床样本的纯度和检测灵敏度等问题）已经证明蛋白质组学图谱在临床诊断中的应用受到限制^[11]。在目前的实践中，基于转录组图谱（如微阵列谱和RNA-seq谱）的合理化系统生物学网络模型已成功应用于个体化治疗诊断领域^[12]。

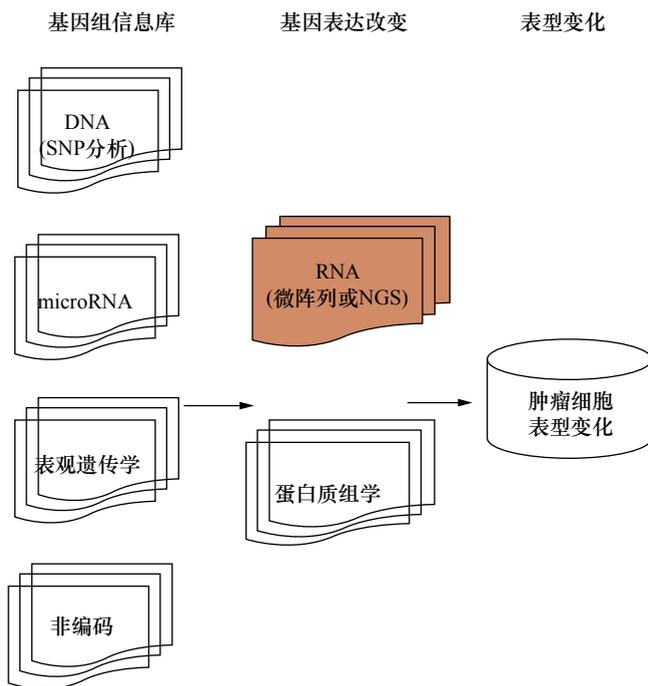


图 1-1-1 三个层次的分析

包括：①基因组信息库（GWAS、microRNA、表观遗传学和非编码基因组）分析；
②包含转录组和蛋白质组学的基因组表达分析；③肿瘤细胞的临床表型变化分析。橙色是本章的重点。

根据系统建模的个体化治疗的工作流程（图 1-1-2），本章将介绍：①临床基因组表达分析技术，这是个体化治疗的基石。②阐述临床基因组诊断的重要性、诊断的精确性对于提高基因组图谱的精度和基因表达水平至关重要，这为后续的系统建模打下了基础。③详细讨论系统模型，包括系统模型的发展和个体化化疗的最新系统模型。④简要阐述与系统建模相关的药物靶向和个体化治疗的多种验证方法。在结论部分，探讨基于系统建模的个体化治疗方法所面临的挑战，并展望其未来的发展方向。

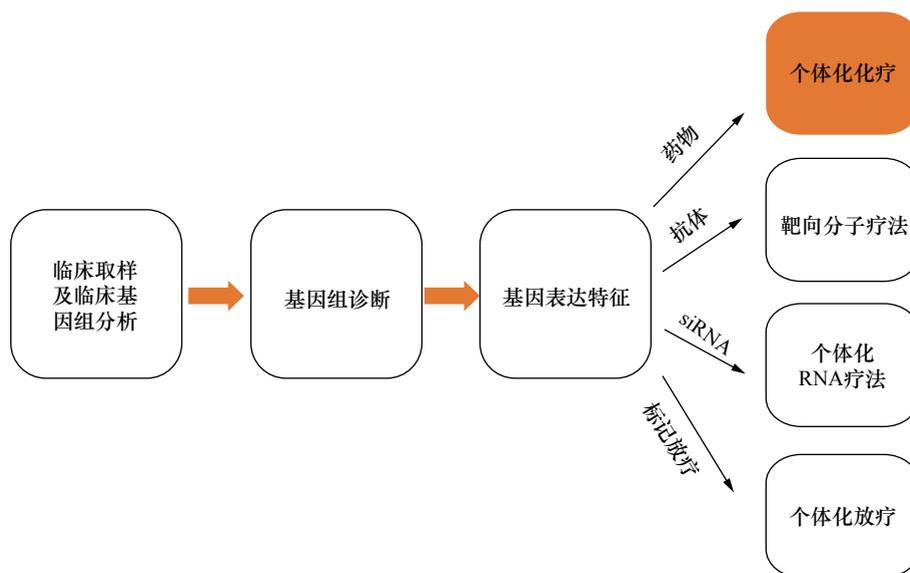


图 1-1-2 个体化治疗的流程图

包括：①临床取样及临床基因组表达分析；②临床基因组诊断；③基因表达特征挖掘；④包括个体化化学治疗、小分子和抗体的靶向分子治疗、个体化RNA治疗和个体化放射治疗的个体化治疗。橙色是本章的重点

众所周知，临床样本由混合的多种细胞组成，这种混合细胞群可显著影响基因组诊断和基因组表达分析的准确性。为了建立用于个体化治疗的可靠系统模型，首先必须解决复杂的临床样本问题。单细胞水平的基因组学提供了一种可行的解决方案。根据我们以往的基因组分析经验，以及其他实验室的文献报道^[13, 14]，我们总结了多种临床取样方法，包括：体外临床取样（*in vitro*）（从临床样本中直接获得单细胞进行单细胞基因组分析）、离体临床取样（*ex vivo*）（从临床样本中获得原代细胞进行纯化/扩增培养后再进行基因组分析），以及借助计算机直接进行临床基因组表达分析（*in silico*）（通过使用不同的生物信息学模型直接分析基因组表达水平，进而从给定的组织水平的肿瘤细胞中获得基因组数据）。

一、用于基因组分析的体外临床取样

体外临床取样的样本的基因组表达分析的步骤包括临床肿瘤细胞分离和下游基因组表达分析，其中临床肿瘤细胞分离的技术包括流式细胞分选（FACS）^[15]、磁性细胞分离（MACS）^[16]、激光捕获显微切割（LCM）^[17]和浮力激活细胞分选（BACS）^[18]。

FACS利用细胞表面的特异性生物标志物来分离肿瘤细胞或癌症干细胞（CSC），例如CSC的CD133/CD44^[19]和循环肿瘤细胞（CTC）的EpCAM^[20]。多色FACS通过组合生物标志物选择性地收集同质细胞，可以提高其研究特定类型的肿瘤细胞基因表达谱的能力。

MACS技术和FACS技术相似，也是通过细胞表面生物标志物对CSC和CTC进行选择区分。MACS可以使用多重标记抗体来选择性收集同质细胞，因此也越来越多地被应用于在肿瘤组织水平上揭示特定类型细胞基因表达水平的研究中。

在LCM技术中，LCM可以根据细胞形态变化或对内源性细胞mRNA/蛋白质生物标志物检测区分肿瘤细胞类别，还可以在体内（*in vivo*）环境中特异性选择肿瘤

细胞^[21]。包括免疫组织化学（IHC）和免疫细胞化学（ICC）在内的抗体染色技术，结合DNA/RNA的荧光原位杂交（FISH）染色技术可进一步增强这些生物标志物的细胞特异性。随着LCM技术本身的发展以及不同CSC和肿瘤细胞特异性生物标志物的发现，LCM在临床取样的实际应用中发展迅速，包括：①从固定的细胞发展至活细胞，挑选出来的CSC或肿瘤活细胞可进一步培养以便用于下游的基因组表达分析^[22]；②越来越多的自动化LCM系统被应用于生物标志物的高通量筛选中^[23]。

BACS则是利用与抗体结合的低密度颗粒（微泡）进行细胞分离的方法。浮力激活的微泡由核心的气体和外壳的聚合物、脂质和蛋白质组成。整个技术可以相对缩短时间和减少成本^[24]。

● 二、用于基因组分析的离体临床取样 ●

离体临床样本基因组分析的步骤包括CSC或原代肿瘤细胞培养以及下游基因组分析。1977年，Hamburger和Salmon首先建立了原代肿瘤细胞培养技术，以测定肿瘤患者的药物敏感度。当时是一种在软琼脂上支持人肿瘤干细胞的集落生长的简单的新方法，适用于不同组织病理的各种肿瘤细胞的培养。不同类型肿瘤产生的肿瘤干细胞集落具有不同的生长特性和集落形态，培养的干细胞集落可用于抗癌药物或放射对人肿瘤干细胞影响的临床研究^[25]。之后原代肿瘤细胞培养技术不断地发展完善。1994年，我们也报告了50例用于药物敏感度测定的原代培养肿瘤细胞技术的案例^[26]。现在，我们已经能常规地使用这项技术来增加原代细胞的数量以进行临床基因组分析的研究。随着针对CSC和原代肿瘤细胞的培养技术的不断研发，用于下游基因组分析的临床样本的肿瘤细胞培养系统将在个体化治疗中发挥越来越重要的作用。令人鼓舞的是，来自培养细胞系统的离体药物敏感度测定的结果可以进一步验证个体化化疗的效果，因此造血系统和实体肿瘤CSC成功的离体培养技术案例也越来越多地被报道，这将为个体化化疗带来光明的前景。

● 三、临床基因组的生物信息学直接分析 ●

就目前的肿瘤样本取样技术而言，大多数临床样本在手术切除后就已经直接低温速冻。如前所述，肿瘤组织由混合细胞组成，肿瘤组织的基因组数据是混合了各种细

胞基因组数据的大杂烩。如果组织水平的样本是通过微阵列和RNA-Seq进行检测的,那么用基于基因组学大数据建立的生物信息软件直接分析临床测序数据对于肿瘤细胞基因组的诠释将是非常重要的一步。根据文献报道以及我们的研究结果^[27]可知,可以通过两种组合的生物信息学技术来提高来自异质性细胞临床基因组分析的准确度:第一种是在层次聚类、主成分分析(principal component analysis, PCA)和自组织映射(self-organizing map, SOM)基础上设计的组织水平分析;第二种是在基于肿瘤生物标志物的监督学习(supervised learning)和肿瘤生物标志物随时间的变化基础上设计的分子水平分析。临床基因组生物信息学分析软件的增多,为基因组数据的解释提供了新的工具。

第三节

临床基因组诊断

经过二十多年的基因组分析技术的研发，特别是人类基因组二代测序（NGS）、特异性基因变异识别技术和定量基因组表达技术的发展，导致了基于基因组学的诊断测试的爆发式增长^[28]。这些诊断测试在直接干预治疗方面展现出巨大潜力。例如，一个由英国研究人员领导的国际团队对150例病例进行了NGS检测，发现超过1/5的病例为新确诊的，而传统的筛查方法并未发现这些病例^[29]。此外，一些基因组分析方法已根据标准要求开发了为临床服务的诊断测试方案。例如，临床基因组和外显子测序（clinical genome and exome sequencing, CGES）可用于研究来自患者的数千种CGES测试结果^[30, 31]，其中的工作包括：①为CGES的顺序做标示；②评估“CGES结果”；③解释和传达“CGES结果”；④解释“偶然发现”。然而，由于缺乏足够的临床实践和验证，临床工作者难以区分哪些测试结果对改善实践具有最大价值，导致了基因组诊断测试的重要性被低估。

肿瘤组织分析所面临的主要挑战是识别纯DNA改变和mRNA基因表达的改变。为了解决这些问题，科学家们评估了来自15种肿瘤类型的815对配对的肿瘤样本-正常样本，证实了全外显子组的NGS结果的敏感度>95%，特异性>99.99%。相比之下，仅对肿瘤样本进行测序的方法不能确切地识别肿瘤易感基因的变化，并且在靶点和外显子组分析中分别产生了31%和65%的假阳性结果^[32]。这些数据表明，肿瘤样本-正常样本的配对测序分析对于变异的精确鉴定和诠释至关重要，也是对肿瘤患者进行诊断和治疗的关键步骤。

同样，转录组数据在个体化治疗的临床应用也需要一个验证过程。用于基因组数据质量控制的“微阵列质量控制（microarray quality control, MAQC）”项目^[33]将在下一节进行讨论，这里只简单展示一些基本证据，如基因组表达谱的特异性和敏感度测试以及其分辨率所面临的挑战。例如，多篇报道表明用上调基因进行敏感度测试，其基因上调的范围在54%~97%，但很少有根据临床试验的标准要求进行特异性测试的报告^[34]。

为了解决基因组表达水平分析的这些问题，我们评估了三组转录组图谱：通过LCM获得心脏肉瘤的肿瘤细胞和与之配对的正常细胞，分别对15个基因进行敏感

度测试，对14个基因进行了特异性测试，结果表明敏感度和特异性分别为87%和60%（图1-3-1）。相比之下，非配对的肿瘤细胞（非小细胞肺癌的活检样本和小细胞肺癌的手术样本，分别使用10~20个基因进行敏感度测试，13~20个基因进行特异性测试）的基因组表达水平表明，敏感度分别为70%和91%，特异性分别为47%和31%。

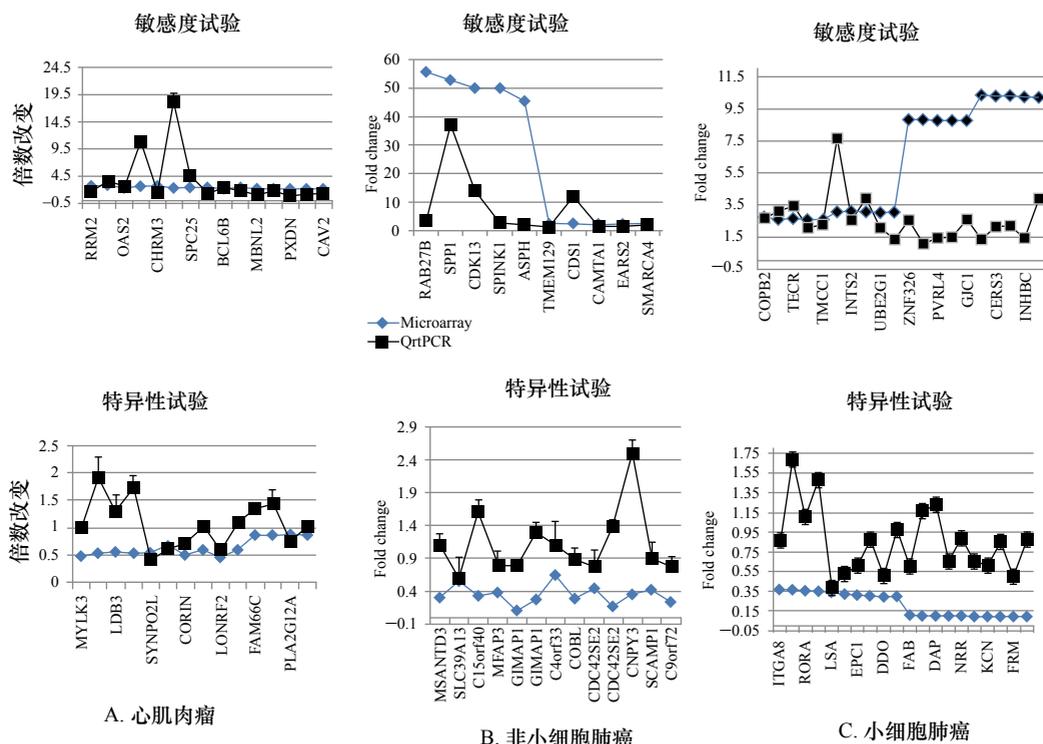


图 1-3-1 敏感度和特异性测试

A. 由 LCM 分离的配对肿瘤细胞-正常细胞的敏感度和特异性；B. 从 NSCLC 活检样本中获得的非配对肿瘤细胞-正常细胞的敏感度和特异性；C. 从 SCLC 手术中获得的非配对肿瘤细胞-正常细胞的敏感度和特异性。深色是 Qrt-PCR 数据，蓝色是转录组数据。

这些数据表明，尽管配对的肿瘤样本-正常样本基因组表达分析的特异性高于非配对的基因组表达分析的特异性，但是配对的肿瘤样本-正常样本基因组分析的敏感度与非配对的基因组分析的敏感度差异并不明显。上述比较研究结果表明，将转录组数据用于临床诊断测试时，需要对基因表达特征（GES）和治疗靶点进行进一步的 Qrt-PCR 验证。

第四节

系统建模和基因表达特征

GWAS是首批预测治疗靶点的模型之一^[35]。如图1-4-1A所示，GWAS旨在预测治疗靶点。对每个样本的DNA进行SNP微阵列或DNA-seq检测后，对配对的病例-对照样本进行统计分析可以确定标记处的等位基因是否可以预测表型。如果重复多次的检测结果显示达到统计学显著性，则可以认为该变异与表型相关。如果基因组某区域中的SNP与肿瘤疾病相关，则可以利用该SNP进一步预测相应的靶向分子疗法，例如小分子和抗体。尽管基因组学的技术创新在不断地扩展，但系统生物学的大规模整合需要具有更强大的潜力来定义治疗靶点以便研究未治愈的肿瘤疾病（图1-4-1B）。本节将深入讨论系统生物学的发展、系统建模的概念，以及系统建模的拓扑分析网络在个体化治疗中的应用。

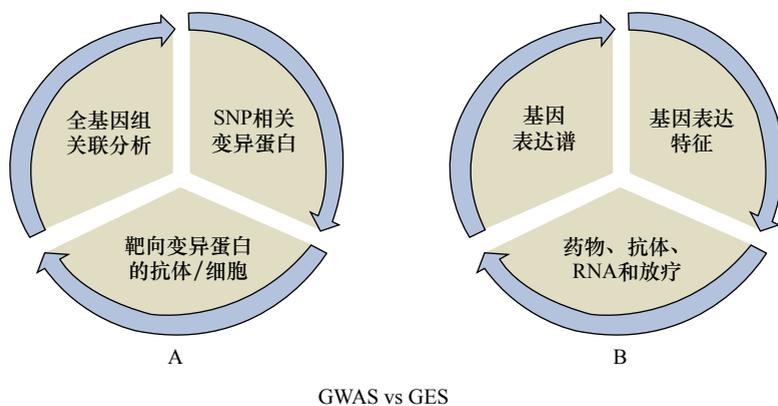


图1-4-1 个体化治疗的流程图

A. GWAS及其用于靶向分子治疗和研究的抗体；B. GES及其药物、抗体和RNA靶向分析

一、系统生物学的发展和系统建模的概念

Ludwig Von Bertalanffy是系统生物学的先驱。随后 Alan Hodgkin 和 Andrew Huxley

在1952年建立了数学模型来解释这种可能性^[36, 37]，他们也因此共同获得了1963年的诺贝尔生理学或医学奖。此后，Mihajlo Mesarovic于1966年在俄亥俄州克利夫兰举行的一次国际研讨会上进一步提出了“系统理论家”学说^[38]。随着基因组学的诞生，自1990年以来，系统生物学已广泛应用于基因组分析中^[39]。

系统生物学是一种复杂生物系统的数学模型。是一种基于生物学的复杂相互作用，涉及到代谢网络或细胞信号网络的整体方法^[40]。生物网络为在生态、进化和生理研究中发现的相互关联性提供了数学分析基础^[41]，对生物网络的研究已经从以前只研究普通对象发展到目前研究人类疾病的网络医学的新领域。肿瘤系统生物学是网络医学的组成部分，它具有研究肿瘤发生和治疗肿瘤的特定目的^[42, 43]。将肿瘤系统生物学应用于精准医疗可以更好地诊断肿瘤并预测个体化治疗的预后。肿瘤系统建模是肿瘤系统生物学跨学科的应用，用以构建系统模型并在精准医疗过程中指导个体化治疗方案的制订^[44]。

◦ 二、系统建模和 GES ◦

建立系统模型后，“基因表达特征（GES）”（单个或组合的基因表达改变）即为系统模型中用于诊断、预后或预测治疗反应的特异性靶点^[45, 46]。这个科学术语已越来越多地被用于肿瘤生物学、预测肿瘤疾病的进展并根据患者的个体差异选择最适合的治疗方案。本章将重点介绍如何将系统建模应用于临床上的患者治疗^[47]。

网络是细胞内生物组成成分和分子成分，以及它们在系统建模中作为纽带的直接或间接相互作用的形象化体现^[48]。不同类型的功能细胞代表了不同类型的细胞内分子和生物网络。有许多不同类型的网络可用于系统建模，例如有向或无向、布尔网络（Boolean network）和斯坦纳树（Steiner tree）^[49]。为了深度覆盖这些数据结构和算法，生物信息学使来自许多不同研究的数据整合到单个框架中成为可能。例如，可以直接从时间序列数据或扰动数据生成网络。网络拓扑可作为直接将数据表变为 mRNA 表达/蛋白质数量变化的“模拟工程师”，它们可以定义在治疗反应系统模型中经过验证的特异性 GES^[50]。目前，一些软件工具有助于分析 GES：例如 Cytoscape 平台和公共链接：Pajek（<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>）、GraphViz（<http://www.graphviz.org/>）和 yEd（<http://www.yworks.com/products/yed/>）^[51~53]。

三、用GES确定个体化治疗的药物靶点

网络的拓扑结构使我们能够通过GES找到肿瘤细胞的关键靶点。网络拓扑结构包括节点的一般属性和特定属性、边的属性、网络整体属性（全局拓扑属性）和网络内的模块（图1-4-2A）^[54, 55]，介绍如下：

（1）节点的属性：包括连接度（connectivity degree, CD, 即每个节点的连接数）、中介中心度（betweenness centrality, BC, 即在所有可能相连的节点对间通过一个节点的最短路径的数量）、邻近中心度（closeness centrality, 从一个节点到所有其他节点的平均最短路径）和特征向量中心度（eigenvector centrality, 一种更复杂的中心度量, 用于评估与紧密连接的节点间的邻近度）。

（2）边的属性：包括节点间关系的类型（通过磷酸化、结合、一对节点之间的基因调控来激活或抑制）和边的方向性（上游或下游）。

（3）网络的全局拓扑属性：包括连接性分布（connectivity distribution, 显示节点及其连接的直方图）、路径长度（path length, Floyd-Warshall算法或Dijkstra算法）和聚类系数（clustering coefficient, 通过测量整个网络中平均每个节点与其邻居的连接性来计算交互的局部密度）。

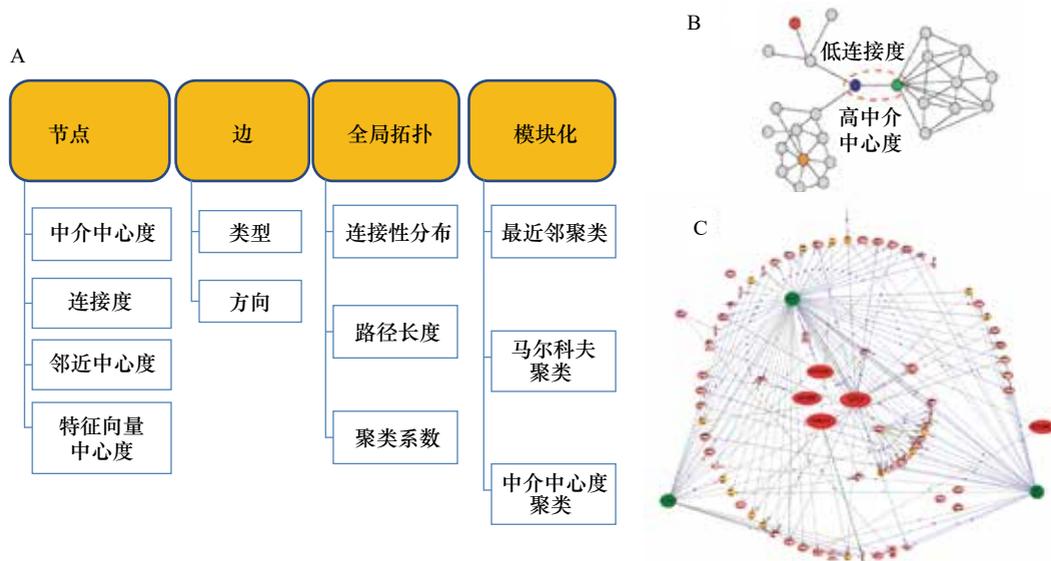


图1-4-2 GES定量网络图

A. 整个定量网络分析；B. 具有较高“中介中心度”和较低“连接度”的GES；C. 具有较高“中介中心度”和较低“连接度”的GES举例。红色是具有较高的“中介中心度”和较低“连接度”的I型靶点，绿色是具有较低的“中介中心度”和较高“连接度”的II型靶点

(4) 模块 (或网络群集): 包括无监督的群集算法, 例如最近邻居群集、马尔可夫群集和基于中介中心度的群集 (通过高的中介中心度和低的连接度来分隔群集)。

在了解了上述网络结构和拓扑分析参数之后, 在临床上, 应选择对生物网络中的肿瘤细胞具有最佳治疗效果且对正常细胞毒性最小的GES作为靶点。尽管许多参数都可以用来评估网络中节点的重要性, 例如, 具有高中介中心度 (BC) 和高连接度 (CD) 的节点都可以用作有效的药物靶点, 但具有高CD值的靶点也可能由于在全系统中具有广泛的影响而对正常细胞具有毒性, 因此我们寻找的作为GES的治疗靶点应具有较高的中介中心度以获得治疗的有效性, 且具有较低的连接度以获得较低的毒性, 从而降低对整个系统的影响 (图 1-4-2B)^[56]。结果如图 1-4-2C所示, 网络拓扑类型为 I 型的GES靶点基因可用于进一步的个体化治疗分析^[57]。

第五节

个体化治疗

GES可用于识别肿瘤和药物的靶向标志物及其剂量依赖性细胞反应。如前文所述，GES还可预示应回避的毒理学通路。挖掘GES之后，可以通过与GES相关联的药物库筛选经FDA批准的药物。

一、从GES到药物靶点

在2007年之前，微阵列技术的临床应用受限于测定数据的误差。例如，在不同微阵列平台上杂交的同一RNA样品可能会产生相互矛盾的结果。为了使基因组数据更好地应用于包括靶向药物筛选在内的临床实践，需要统一基因组数据，于是FDA发起了名为“微阵列质量控制（MAQC）”的项目，来自51个学术机构和行业合作伙伴的137名研究人员参加了该项目，旨在系统地解决不同实验室和不同微阵列平台之间微阵列检测技术的可重复性问题^[58]。最终结果显示，在不同测试地点使用同一平台检测差异表达基因可获得平均89%的重叠率，使用不同的微阵列平台则有74%的重叠率^[59]。经过十多年的努力，尤其是将二代测序技术引入这一新领域之后，基于准确基因组数据的系统建模和GES挖掘为肿瘤疾病个体化治疗的药物筛选和新药研究作出了巨大的贡献。目前，几个常用的公共链接在证明药物靶点方面趋于完善：例如Genecards（<http://www.genecards.org/>）、Drug-bank（<http://www.drugbank.ca/>）以及一些商业软件工具，例如GeneGo和Pathway Studio^[60, 61]。

尽管第一代基因组研究和二代测序都可用于临床的抗癌药物的发现，但它们在很大程度上仍不是最佳选择，例如还存在临床应答与病理应答方面的问题、联合化疗与单一药物的选择问题^[62-65]。然而，GES数据在改善肿瘤疾病的临床管理和增进我们对肿瘤生物学机制的理解方面仍具有巨大潜力。尽管FDA已经批准了一些用于临床的表达谱分析平台，但要想使这些特征（signature）具有更强的临床价值，仍然需要增加大量的验证和病例报告。

二、药物靶点的确认

一旦从药物库平台筛选出了推荐的药物，就会有一套流程来确认该药物是对肿瘤细胞普遍有作用，还是对特定类型的肿瘤具有特异性作用。根据 Chibon F 等的报告^[65]，当推荐的抗癌药物从药物库中筛选出后，该药物在被用于每个个体化治疗前都必须进行独立验证。Chibon 为每个个体化治疗设置了三个步骤来验证抗癌药物：特征鉴定、特征验证和临床验证，并通过被称为“自上而下法”“自下而上法”和“候选基因法”的三个模块来进行。

(1) “自上而下法”是一种监督法，通过鉴定与转移和生存率相关的表达谱来建立预测模型。René Bernards 的团队是最早使用这种方法的小组之一，于 2002 年报告了对乳腺癌预后特征的鉴定^[66]。他们报告了一种三段式方法，即鉴定所有样本中的基因（5000 个）、与转移相关的特征基因（231 个）以及作为最佳预测因子的最终特征基因（70 个）。“自上而下法”的第二个例子是套细胞淋巴瘤，其中 28/48 个基因（58%）可作为淋巴瘤细胞增殖特征基因^[67]。

(2) “自下而上法”不同于“自上而下法”的监督法，它是一种生物学假说。通过这种方法鉴定出 677 个与细胞周期、细胞运动和细胞外基质重塑有关的基因^[68]。

(3) “候选基因法”是一种代表性基因法，通过 PCR 和免疫组织化学技术在福尔马林固定、石蜡包埋的肿瘤（FFPE）样本中检测少量具有代表性的基因^[69]。

用以上三种方法完成特征鉴定之后，方可进入临床实践阶段，从而可以更有效地验证下游的个体化治疗。

尽管 Chibon 的“三个步骤和三个方法”可以被病理学家和临床肿瘤学家应用于临床领域，但是用于指导临床医师采用推荐的抗肿瘤药物进行临床个体化治疗仍存在争议。目前，针对个体化治疗的预测抗肿瘤药物的验证被越来越多地报道^[70]。根据已发布的临床应用，本文总结了大多数验证方法，如图 1-5-1 和表 1-5-1 所示。

表 1-5-1 验证 GES 药物的敏感度和治疗靶点

方法	技术	优点	缺点
间接确认	Qrt-PCR 或 ICC/IHC 染色	简单便宜	间接数据
计算机验证	生物信息学模块	生物信息学支持	模拟数据
离体验证	原代肿瘤细胞或 CSC 培养	不同药物的客观应答	技术困难
体内验证	人肿瘤细胞的异种移植	不同药物的客观应答	成本很高，受时间限制
体外验证	肿瘤细胞系培养	不同药物简单易用	非个体化反应

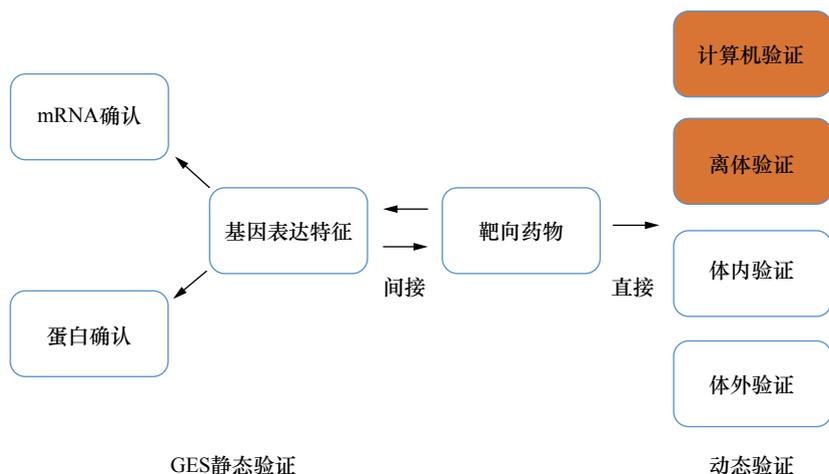


图 1-5-1 个体化治疗的验证流程图

包括：①通过 Qrt-PCR 和蛋白质水平进行 GES 确认的间接方法；
②计算机验证、离体验证、体内验证和体外验证的直接方法。橙色主要用于实验室

1. 间接方法

鉴定可作为药物靶点的 GES 的技术至关重要，例如 PCR、免疫组织化学和蛋白质组学^[71]。基于上述敏感度和特异性的结果，我们常规地将 Qrt-PCR 测定作为 GES 和药物挖掘后的必要验证步骤。此外，在 2007 年我们还报道了蛋白质组学作为应用于发现生物标志物和治疗靶点的候选方法。遗憾的是，如前所述，蛋白质组学与基因组数据（25/750 个基因）的重叠率非常低^[57, 72]。

2. 直接方法

直接的验证分析也可预测药物的功效，直接验证包括计算机验证（in silico）、离体验证（ex vivo）、体内验证（in vivo）和体外验证（in vitro）。

（1）计算机验证：与药物发现有关的网络已有大量报道。Zheng J 等首先建立了一个基于 python 的网络模型，使用评分分析系统来研究肿瘤细胞系的增殖情况，然后使用该模型对网络中的药物进行评分^[73]。2015 年，我们进一步报告了该模型用于患有三阴性乳腺癌（TNBC）的患者基因组分析的案例。令人兴奋的是，在通过 python 模型进行预测和验证-模拟分析后，发现并确认了两组适合患者的药物，该患者在 3 个月的推荐药物治疗后获得了完全缓解，如图 1-5-2A 和 1-5-2B 所示^[74]。一些实验室进行了类似的药物挖掘和验证工作，例如，在 2013 年，Riddick G 等利用计算机模拟筛选结合胶质母细胞瘤 CSC 验证的方法建立了用于药物应答的预测模型。通过计算机筛选，他们报告了 185 种化合物，其中，7 种化合物的应答在胶质瘤细胞系中得到了验证，21 种化合物的应答在 3 种胶质母细胞瘤干细胞中得到了验证^[75]。

（2）离体验证：如前所述，在我们实验室中工作超过 20 年的 Hamburger 和 Salmon

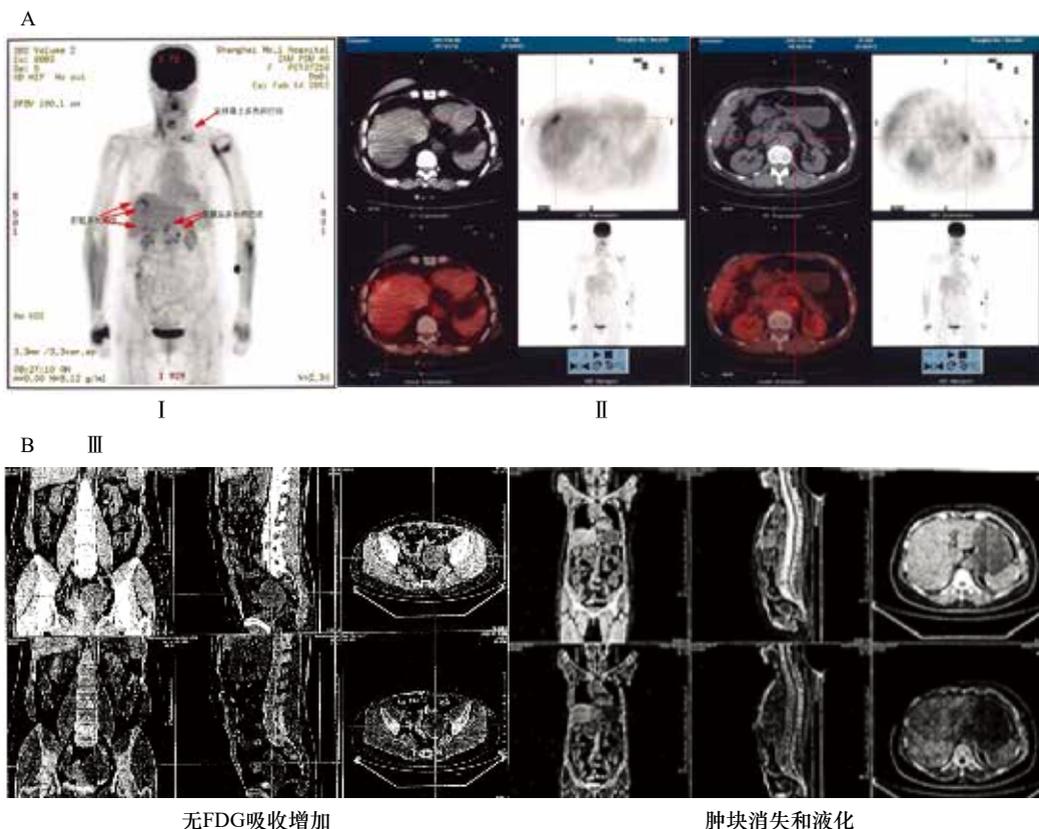


图1-5-2 正电子发射断层扫描/计算机断层扫描 (PET/CT) 的结果

A. 显示了个体化治疗之前的多个转移病灶；B. 显示了个体化治疗5个月后反应良好。

建立了包括原代肿瘤细胞在内的临床细胞培养方法以进行药物敏感度测定。至今，已有好几种原代肿瘤细胞培养的方法及其技术可应用于个体化治疗的验证系统^[76]。因此，尽管技术非常复杂，但离体模块提供了一个很好的系统来验证用于个体化治疗的推荐药物。

(3) 体内验证：人体肿瘤细胞的异种移植也是验证药物的良好模型。一些科学家开发了人类肿瘤细胞的异种移植以验证药物。例如，科学家选择性分离了原发性肿瘤的转移肿瘤细胞亚群进行基因表达谱分析。通过这种方式，他们发现了一种乳腺癌转移和侵袭的特异性GES，称为人类入侵特征 (human invasion signature, HIS)。通过独立的生物学重复试验，他们证实了这些功能基因在转移性肿瘤细胞中被上调。他们还通过患者来源的乳腺肿瘤证明了体内侵袭所必需的特定基因的功能。最后，他们使用统计分析表明该特征可以有效预测乳腺癌转移的风险^[77]。虽然该模型可能很有价值，但体内模块验证药物的成本很高。

(4) 体外验证：特定药物的敏感度可能会严重影响治疗方案的决策。一些科学家开发了肿瘤细胞系作为模型来验证患者的应答。例如，结直肠癌中基于耐药性

的评分系统可以将患者对三种一线抗癌药物（5-FU、奥沙利铂和伊立替康）的应答进行排名^[78]。

这些耐药性细胞系的基因表达模式为开发特异性耐药性GES提供了坚实的基础。由于肿瘤细胞系的遗传背景/突变有助于对被测药物产生耐药性，因此特殊的基因表达模式表明不同药物的耐药性机制不同。尽管已报道肿瘤细胞系可用于药物筛选，但该模型可能仅能在肿瘤细胞系中并且在遗传背景/突变的有限范围内对治疗靶点进行分类。

个体化治疗或精准治疗的问世，使医师能够根据患者个体自身的基因组数据进行个体化治疗。全面合理的个人基因组信息分析对于此项新颖工具的成功应用至关重要。在这里，我们列出一些需要进一步解决的问题：

一、用于基因组分析的临床取样需要改进

用于基因组分析的临床取样，有几个方面需要改进。

(1) 利用体外临床取样进行基因组表达分析需要在细胞表面、细胞质蛋白、mRNA 或 DNA 上有更多特异性的生物标志物。

(2) 利用离体临床取样进行基因组表达分析需要开发用于原代肿瘤细胞和 CSC 培养的标准培养技术和培养系统。我们已经在长达 25 年的时间里大量报道了来自肺、肝、结肠、脑、胰腺和肾脏的包括原代肿瘤细胞、CSC 和干细胞在内的临床细胞培养技术。不同的样本需要不同的培养条件，例如化疗前和化疗后收集的原代细胞培养方法就有所不同，“一刀切”的培养系统无法适应个体化治疗，因此需要更多有经验的专家来进行临床取样。幸运的是，德克萨斯大学 MD 安德森癌症中心 (UT MD Anderson Cancer Center) 最近报告了一些简化个体化治疗方案的设想。在不久的将来，我们相信用于基因组表达分析的离体临床细胞培养技术将越来越多地用于个体化治疗。

(3) 利用计算机软件直接进行临床基因组表达分析则需要更多操作性强的能获取全临床生物信息的使用工具。

大多数临床样本在手术切除后直接在肿瘤组织水平被低温冷冻，如何更好地通过这些临床样本获得更多的生物学信息需要我们去研究。

● 二、临床基因组诊断需要常规化的特异性和敏感度测试 ●

在直接的治疗干预中，基于基因组学的诊断测试已经得到了很大的发展。许多病例在临床基因组学分析的指导下治疗效果有了很大改善。在此基础上，科学家和医师们将进一步开发临床基因组诊断测试方法，为患者提供常规的个体化治疗选择。如上所述，转录组水平分析的基因组数据用于临床所面临的挑战是如何提高肿瘤组织中肿瘤细胞的纯度。根据当前的数据，作为常规临床测试，应为每个样本的敏感度和特异性建立标准。

● 三、系统建模和GES需要更多的改进 ●

肿瘤研究的最新发展使医师和科学家能够使用不同的基因组数据，例如GWAS、微阵列和RNA-Seq的转录组谱、表观遗传学谱、microRNA谱和非编码DNA谱。此外，ENCODE也已出现在基因组分析中。为了将来建立合理的个体化治疗网络，基于转录组的系统生物学还应包括基因组档案，例如GWAS、表观遗传学谱或microRNA谱。定量网络评分系统应涵盖所有基因组学图谱，包括转录组、GWAS、表观遗传学和microRNA。

● 四、药物靶向疗法和个体化疗法需要增加其他的可选项 ●

挖掘GES之后，通过关联的药物库首先选择经FDA批准的药物用于临床靶向治疗。如图1-1-2所示，其他三种个体化疗法（用抗体进行的靶向分子疗法的抗体、个体化RNA疗法和个体化放射疗法）尚未广泛用于临床患者。临床医师应在不同分子水平（RNA水平、抗体抗蛋白水平和针对分子的药物水平）上研究个体化疗法并开展精准治疗的临床应用。

（张逸飞 陆静 李彪如）

参考文献

- [1] Rodríguez-Antona C, Taron M. Pharmacogenomic biomarkers for personalized cancer treatment [J]. *J Intern Med*, 2015, 277(2): 201-217.
- [2] Schilsky RL. Opinion: Personalized medicine in oncology: the future is now [J]. *Nature Reviews Drug Discovery*, 2010, 9: 363-366.
- [3] Hudson TJ. Genome variation and personalized cancer medicine [J]. *J Intern Med*, 2013, 274(5): 440-450.
- [4] Yoshida T, Ono H, Kuchiba A, et al. Genome-wide germline analyses on cancer susceptibility and GeMDBJ database: Gastric cancer as an example [J]. *Cancer Sci*, 2010, 101(7): 1582-1589.
- [5] Quetglas IM, Moeini A, Pinyol R, et al. Integration of genomic information in the clinical management of HCC [J]. *Best Pract Res Clin Gastroenterol*, 2014, 28(5): 831-842.
- [6] Peterson SM, Thompson JA, Ufkin ML, et al. Common features of microRNA target prediction tools [J]. *Front Genet*, 2014, 5: 23.
- [7] Pazin MJ. Using the ENCODE Resource for Functional Annotation of Genetic Variants [J]. *Cold Spring Harb Protoc*, 2015, 6: 522-536.
- [8] Ben-Horin S, Chowers Y. Review article: loss of response to anti-TNF treatments in Crohn's disease [J]. *Aliment Pharmacol Ther*, 2011, 33(9): 987-995.
- [9] Wu P, Walker BA, Brewer D, et al. A gene expression-based predictor for myeloma patients at high risk of developing bone disease on bisphosphonate treatment [J]. *Clin Cancer Res*, 2011, 17(19): 6347-6355.
- [10] Polotskaia A, Xiao G, Reynoso K, et al. Proteome-wide analysis of mutant p53 targets in breast cancer identifies new levels of gain-of-function that influence PARP, PCNA, and MCM4 [J]. *Proc Natl Acad Sci USA*, 2015, 112(11): E1220-E1229.
- [11] Calvo K R, Liotta L A, Petricoin E F. Clinical Proteomics: From Biomarker Discovery and Cell Signaling Profiles to Individualized Personal Therapy [J]. *Bioscience Reports*, 2005, 25(1-2): 107-125.
- [12] Kim J, Vasu V T, Mishra R, et al. Bioinformatics-driven discovery of rational combination for overcoming EGFR-mutant lung cancer resistance to EGFR therapy [J]. *Bioinformatics*, 2014, 30(17): 2393-2398.
- [13] Li B. A strategy to identify genomic expression at single-cell level or a small number of cells [J]. *Journal of Biotechnology*, 2005, 8(1): 71-81.
- [14] Li B. Clinical Genomic Analysis and Diagnosis-Genomic Analysis Ex Vivo, in Vitro and in Silico [J]. *Clinical Medicine and Diagnostics*, 2012, 2(4): 37-44.
- [15] Ormerod M. *Flow Cytometry: A practical approach* [M]. Oxford: Oxford University Press, 2000.
- [16] Bacon K, Lavoie A, Rao B M, et al. Past, Present, and Future of Affinity-based Cell Separation Technologies [J]. *Acta Biomater*, 2020, 112: 29-51.

- [17] Emmert-Buck M R, Bonner R F, Smith P D, et al. Laser capture microdissection [J]. *Science*, 1996, 274(5289): 998-1001.
- [18] Liou Y, Wang Y, Lee C, et al. Buoyancy-activated cell sorting using targeted biotinylated albumin microbubbles [J]. *PLoS One*, 2015, 10(5): e0125036.
- [19] Pustovalova M, Blokhina T, Alhaddad L, et al. CD44⁺ and CD133⁺ Non-Small Cell Lung Cancer Cells Exhibit DNA Damage Response Pathways and Dormant Polyploid Giant Cancer Cell Enrichment Relating to Their p53 Status [J]. *Int J Mol Sci*, 2022, 23(9): 4922.
- [20] Magbanua M J, Park J W. Isolation of circulating tumor cells by immunomagnetic enrichment and fluorescence-activated cell sorting (IE/FACS) for molecular profiling [J]. *Methods*, 2013, 64(2): 114-118.
- [21] Niyaz Y, Stich M, Sigm. Isolation of circulating tumor cells by immunomagnetic enrichment and fluorescence-activated cell sorting (IE/FACS) for molecular profiling [J]. *Methods*, 20 Mol Med, 2005, 114: 1-24.
- [22] Steen J, Morrison J A, Kulesa P M. Multi-position photoactivation and multi-time acquisition for large-scale cell tracing in avian embryos [J]. *Cold Spring Harb Protoc*, 2010, 10: 101-110.
- [23] Vandewoestyne M, Van Hoofstat D, Van Nieuwerburgh F, et al. Automatic detection of spermatozoa for laser capture microdissection [J]. *Int J Legal Med*, 2009, 2: 169-175.
- [24] Shi G, Liu Y. Laboratory-Scale Production of Sterile Targeted Microbubbles [J]. *Methods Mol Biol*, 2022, 2394: 591-599.
- [25] Hamburger A W, Salmon S E. Primary bioassay of human tumor stem cells [J]. *Science*, 1977, 97(4302): 461-463.
- [26] Li B R, Tong S Q, Zhang X H, et al. A new experimental and clinical approach of combining usage of highly active tumor-infiltrating lymphocytes and highly sensitive antitumor drugs for the advanced malignant tumor [J]. *Chin Med J (Engl)*, 1994, 107(11): 803-807.
- [27] Lähdesmaki H, Shmulevich L, Dun mire V, et al. In silico microdissection of microarray data from heterogeneous cell populations [J]. *BMC Bioinformatics*, 2005, 6: 54-58.
- [28] Roundtable on Translating Genomic-Based Research for Health, Board on Health Sciences Policy, Institute of Medicine. Refining Processes for the Co-Development of Genome-Based Therapeutics and Companion Diagnostic Tests: Workshop Summary [R]. Washington (DC): National Academies Press (US), 2014.
- [29] Sharma V P, Fenwick A L, Brockop M S, et al. Mutations in TCF12, encoding a basic helix-loop-helix partner of TWIST1, are a frequent cause of coronal craniosynostosis [J]. *Nature Genetics*, 2013, 45(3): 304-307.
- [30] Trubetskoy V, Rodriguez A, Dave U, et al. Consensus Genotype for Exome Sequencing (CGES): improving the quality of exome variant genotypes [J]. *Bioinformatics*, 2015, 31(2): 187-193.
- [31] Liu Z, Zhang L, Ren C, et al. Whole genome and exome sequencing identify NDUFV2 mutations as a new cause of progressive cavitating leukoencephalopathy [J]. *J Med Genet*, 2022, 59(4): 351-357.
- [32] Jones S, Anagnostou V, Lytle K, et al. Personalized genomic analyses for cancer mutation discovery and interpretation [J]. *Sci Transl Med*, 2015, 7(283): 283.

- [33] Shi L, Campbell G, Jones W D, et al. The Micro-Array Quality Control (MAQC)- II study of common practices for the development and validation of microarray-based predictive models [J]. *Nat Biotechnol*, 2010, 28(8): 827-838.
- [34] Cabanski C R, Qi Y, Yin X, et al. SWISS MADE: Standardized Within Class Sum of Squares to evaluate methodologies and dataset elements [J]. *PLoS One*, 2010, 5(3): e9905.
- [35] Yang Z, Paschou P, Drineas P. Reconstructing SNP allele and genotype frequencies from GWAS summary statistics [J]. *Sci Rep*, 2022, 12(1): 8242.
- [36] Drack M, Ludwig von Bertalanffy's organismic view on the theory of evolution [J]. *J Exp Zool B Mol Dev Evol*, 2015, 324(2): 77-90.
- [37] Raju T N. The Nobel chronicles. 1963: Sir Alan Lloyd Hodgkin(1914-98), Sir Andrew Fielding Huxley (b 1917), and Sir John Carew Eccles (1903-97) [J]. *Lancet*, 1999, 354(9174): 263.
- [38] Rosen R. A Means Toward a New Holism [J]. *Science*, 1968, 161(3836): 34-35.
- [39] Yu D, Kim M, Xiao G, et al. Review of biological network data and its applications [J]. *Genomics Inform*, 2013, 11(4): 200-210.
- [40] Liu G, Qin Y, Li Z, et al. Development of highly efficient, low-cost lignocellulolytic enzyme systems in the post-genomic era [J]. *Biotechnol Adv*, 2013, 31(6): 962-975.
- [41] Kueffer C, Py Li Z, et al. Development of highly efficient, low-cost lignocellulolytic enzyme systems in the post-genomic era [J]. *Biotechnol Adv*, 2013, 31(6): 962-975.
- [42] Costantini S, Colonna G, Castello G. A holistic approach to study the effects of natural antioxidants on inflammation and liver cancer [J]. *Cancer Treat Res*, 2014, 159: 311-323.
- [43] Pathmanathan S, Grozavu I, Lyakisheva A, et al. Drugging the undruggable proteins in cancer: A systems biology approach [J]. *Curr Opin Chem Biol*, 2022, 66: 102079.
- [44] Novosyadlyy R, Leroith D. Insulin-like growth factors and insulin: at the crossroad between tumor development and longevity [J]. *J Gerontol A Biol Sci Med Sci*, 2012, 67(6): 640-651.
- [45] Fumagalli D, Blanchet-Cohen A, Brown D, et al. Transfer of clinically relevant gene expression signatures in breast cancer: from Affymetrix microarray to Illumina RNA-Sequencing technology [J]. *BMC Genomics*, 2014, 15: 1008.
- [46] Latha NR, Rajan A, Nadhan R, et al. Gene expression signatures: A tool for analysis of breast cancer prognosis and therapy [J]. *Crit Rev Oncol Hematol*, 2020, 151: 102964.
- [47] Shah N, Lankerovich M, Lee H, et al. Exploration of the gene fusion landscape of glioblastoma using transcriptome sequencing and copy number data [J]. *BMC Genomics*, 2013, 14: 818.
- [48] Proulx S R, Promislow D E L, Phillips P C. Network thinking in ecology and evolution [J]. *Trends in Ecology and Evolution*, 2005, 20(6): 345-353.
- [49] Ballerstein K, Haus U U, Lindquist J A, et al. Discrete, qualitative models of interaction networks [J]. *Front Biosci (Schol Ed)*, 2013, 5: 149-166.
- [50] Janjić V, Pržulj N. The topology of the growing human interactome data [J]. *J Integr Bioinform*, 2014, 11(2): 238.
- [51] Sanders E, Diehl S. Analysis and interpretation of transcriptomic data obtained from extended Warburg effect genes in patients with clear cell renal cell carcinoma [J]. *Oncoscience*, 2015, 2(2): 151-186.

- [52] Zhao J H. Pedigree-drawing with R and graphviz [J]. *Bioinformatics*, 2006, 22(8): 1013-1014.
- [53] Helaers R, Bareke E, De Meulder B, et al. gViz, a novel tool for the visualization of co-expression networks [J]. *BMC Res Notes*, 2011, 4: 452.
- [54] Olesen J M, Bascompte J. The modularity of pollination networks [J]. *PNAS*, 2007, 104(50): 19891-19896.
- [55] Telesford Q K, Simpson S L, Burdette J H, et al. The brain as a complex system: using network science as a tool for understanding the brain [J]. *Brain Connect*, 2011, 1(4): 295-308.
- [56] Hu G, Zhou J, Yan W, et al. The topology and dynamics of protein complexes: insights from intra-molecular network theory [J]. *Curr Protein Pepl Sci*, 2013, 14(2): 121-132.
- [57] Li B, Senzer N, Rao D, et al. Bioinformatics Approach to Individual Cancer Target Identification, 11th Annual Meeting of the American Society of Gene Therapy [J], 2008, 8: 45-46.
- [58] MAQC Consortium, Shi L, Reid L H, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements [J]. *Nat Biotechnol*, 2006, 24(9): 1151-1161.
- [59] Patterson T A I, Lobenhofer E K, Fulmer-Smentek S B, et al. Performance comparison of one-color and two-color platforms within the Micro-Array Quality Control (MAQC) project [J]. *Nat Biotechnol*, 2006, 24(9): 1140-1150.
- [60] Wishart D S, Knox C, Guo A C, et al. Drug Bank: a knowledgebase for drugs, drug actions and drug targets [J]. *Nucleic Acids Res*, 2008, 36 (Database issue): D901-D906.
- [61] Harel A, Dalah I, Pietrokovski S, et al. Omics data management and annotation [J]. *Methods Mol Biol*, 2011, 719: 71-96.
- [62] Sotirioul C, Piccart M J. Opinion: Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? [J]. *Nature Reviews Cancer*, 2007, 7: 545-553.
- [63] Ulahannan D, Kovac M B, Mulholland P J, et al. Technical and implementation issues in using next-generation sequencing of cancers in clinical practice [J]. *Br J Cancer*, 2013, 109(4): 827-835.
- [64] De M G, Pasello G, Dono M, et al. The storm of NGS in NSCLC diagnostic-therapeutic pathway: How to sun the real clinical practice [J]. *Crit Rev Oncol Hematol*, 2022, 169: 103561.
- [65] Chibon F. Cancer gene expression signatures-The rise and fall? [J]. *European Journal of Cancer*, 2013, 49(8): 2000-2009.
- [66] van't Veer L J, Dai H, van de Vijver M J, et al. Gene expression profiling predicts clinical outcome of breast cancer[J]. *Nature*, 2002, 415: 530-536.
- [67] Rosenwald A, Wright G, Wiestner A, et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma [J]. *Cancer Cell* 2003, 3(2): 185-197.
- [68] Ramaswamy S, Ross K N, Lander E S, et al. A molecular signature of metastasis in primary solid tumors [J]. *Nat Genet*, 2003, 33(1): 49-54.
- [69] Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer [J]. *N Engl J Med*, 2004, 351(27): 2817-2826.

- [70] Vari S, Pilotto S, Maugeri-Sacca A. Advances towards the design and development of personalized non-small-cell lung cancer drug therapy [J]. *Expert Opin Drug Discov*, 2013, 8(11): 1381-1397.
- [71] Lossos I S, Czerwinski D K, Alizadeh A A, et al. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes [J]. *N Engl J Med*, 2004, 350(18): 1828-1837.
- [72] Tan D S, Thomas G V, Garrett M D, et al. Biomarker-driven early clinical trials in oncology: a paradigm shift in drug development [J]. *Cancer J*, 2009, 15(5): 406-420.
- [73] Zheng J, Zhang D, Przytycki P F, et al. SimBoolNet-a Cytoscape plugin for dynamic simulation of signaling networks [J]. *Bioinformatics*, 2010, 26(1): 141-142.
- [74] Hu H, Zhang Q, Li S, et al. A Therapeutic Targeting Identification from Microarray Data and Quantitative Network Analysis [J]. *The Open Access Journal of Science and Technology*, 2015, 3: 1-10.
- [75] Riddick G, Song H, Holbeck S L, et al. An in silico screen links gene expression signatures to drug response in glioblastoma stem cells [J]. *Pharmacogenomics J*, 2014, 61: 10.
- [76] Dairkec S H, Ji Y G, Ben Y. A molecular 'signature' of primary breast cancer cultures; patterns resembling tumor tissue [J]. *BMC Genomics*, 2004, 5: 47.
- [77] Patsialou A, Wang Y R, Lin J. Selective gene-expression profiling of migratory tumor cells in vivo predicts clinical outcome in breast cancer patients [J]. *Breast Cancer Res*, 2012, 14(5): R139.
- [78] Zheng Y, Zhou J, Tong Y. Gene signatures of drug resistance predict patient survival in colorectal cancer [J]. *The Pharmacogenomics Journal*, 2015, 15, 135-143.