

DOI: 10.13376/j.cblls/2022146

文章编号: 1004-0374(2022)11-1327-09



21世纪生物医学的三个主要发展趋势

吴家睿

(中国科学院分子细胞科学卓越创新中心, 上海 200031)

随着科学技术的发展和社会的进步, 人类的预期寿命在 20 世纪有了显著的增长, 各发达国家陆续进入老龄化社会。中国社会在改革开放的 30 年时间里取得了突出的成就, 在 20 世纪末也进入了老龄化社会。老龄化社会带来了疾病谱的重大转变, 即慢性非传染性疾病(慢性病)显著增加——年龄是慢性病发生的主要风险因素。世界银行在 2015 年一份题为《长寿与繁荣: 东亚和太平洋地区的老龄化社会》的报告中指出, 全球 65 岁以上的老年人中的 36% 目前居住在东北亚地区; 预计到 2030 年, 癌症、心脏病和糖尿病等与高龄相关的慢性病患者将占这个地区全部疾病患者的 85%。

随着 20 世纪中叶 DNA 双螺旋的提出和遗传信息传递的中心法则的建立, 诞生了分子生物学, 开启了在分子水平上研究生命及其活动的生命科学时代。那种依靠经验的传统西方医学在现代生命科学的推动下成功转型, 成为一门依靠实验科学理论和技术进行疾病诊疗的现代生物医学(Biomedicine); 其间不断涌现的“高技术”在抗击疾病的过程中扮演了重要的角色, 尤其是抗生素、疫苗及化学小分子药物的研发和利用, 使人类在全球范围内基本控制了传染病, 甚至消灭了天花等恶性传染病。跨入新世纪, 为了满足不断提升的维护公众健康需求, 以及抗击慢性病日益增大的威胁, 现代生物医学在人类基因组计划的推动下正在进入一个新的转型时期, 其中有三种主要的发展趋势值得我们关注。

1 从简单性思维的分子生物医学转变到复杂性思维的系统生物医学

20 世纪中叶诞生的分子生物学为科研人员提供了这样一种基本的研究范式: 利用生物学实验方法以及物理和化学技术等各种研究手段, 通过在分子层次上揭示单个基因或蛋白质的结构与功能来阐明生物体的生理或病理活动。美国著名的肿瘤生物

学家温伯格 (Weinberg R) 对此有很好的总结: “在 20 世纪, 生物学从传统的描述性科学转变成为一门假设驱动的实验科学。与此紧密联系的是还原论占据了统治地位, 即对复杂生命系统的理解可以通过将其拆解为组成的零部件并逐个地拿出来进行研究”^[1]。

生命科学的进步推动了人类对自身健康和疾病的认识, 使依靠经验的传统西方医学转变成为以分子生物学知识和实验方法为基础的分子生物医学(Molecular Biomedicine)。在生命科学的还原论思维指导下, 广为流行着“一个基因一种疾病”的“分子病”观点, 即疾病意味着某个基因或蛋白质出了问题, 而治疗就是用物理、化学方法去找到并修复这种有问题的分子零件。换句话说, 分子生物医学将复杂的病理现象还原为分子层次的个别生物分子之物理或化学功能异常, 进而以简单化思维方式去理解疾病并给与诊治。

虽然分子生物医学在抗击传染病方面取得了显著的成绩, 但是在抗击肿瘤、代谢性疾病和神经退行性疾病等慢性病方面却面临巨大的挑战, 人们最初认为, 分子生物医学能够在慢性病的“战场”取得同样的胜利, 为此美国政府在 1971 年专门启动了消灭肿瘤的“战争”(War on cancer)。可结果远不是人们所预期的, 被评为是一场失败的“越南战争”。究其失败的根本原因在于, 肿瘤属于复杂疾病, 其发生发展过程涉及到机体众多的内部因素和环境因素, 以及这些因素之间复杂的相互作用。显然, 在简单化思维指导下的分子生物医学难以认识和处理这类复杂疾病。美国肿瘤生物学家温伯格曾经以“一个完整的循环: 从无尽的复杂性变为简单性然后又重回复杂性”为题回顾了这场失败的“肿瘤战争”, 并明确指出: “从事肿瘤研究的科学家见证了这个时候的疯狂转变: 从最初面对无数难以理解的病理现象的困惑, 到树立了还原论必胜的信念, 最

近几年再回到重新面对肿瘤这个疾病无尽的复杂性”^[2]。

世纪之交的“人类基因组计划”通过“组学”(omics)整体研究策略从根本上颠覆了这种“碎片化”的科研范式。英国 *Nature* 杂志曾为此发表了一篇题为“阻止疾病, 现在开始”的社论:“似乎在一夜之间就从一个基因、一个蛋白质、一个分子、一次研究一个, 转变为所有基因、所有蛋白质、所有分子、一次研究所有。一切都按组学的规模进行”^[3]。这种转变不仅仅是将生命体内的研究对象从局部转变为全局, 更重要的是对生命的认知从简单性思维转变为复杂性思维。

这种复杂性思维转变之代表是, 2000 年之初在生命科学领域兴起了一门交叉学科——系统生物学 (Systems Biology), 即通过整合经典的分子细胞生物学、新兴的各种组学, 以及信息科学和数学等非生物学科的研究策略和方法, 对生命复杂系统及其生理病理活动进行系统性和整体性的检测和分析。这门新兴学科很快得到研究人员的接受和重视; *Cell* 杂志在 2011 年 3 月发表了整整一期介绍系统生物学的评论文章, 其中一篇文章的标题就是“系统生物学: 进化成为主流”。

系统生物学因其研究生命复杂系统的能力, 很快就被引入到医学领域, 形成了系统生物医学 (Systems Biomedicine)。美国国立卫生研究院 (National Institutes of Health, NIH) 在 2003 年发布的 NIH 路线图, 把采用系统生物学的方法和策略开展慢性病研究列为主要任务^[4]。此外, 欧盟委员会 (European Commission) 也专门成立了一个“系统生物医学行动协调组织” (Coordinating Action Systems Medicine Consortium, CASyM), 并在 2014 年 6 月发布了系统生物医学的研究规划——《CASyM 路线图》。该路线图指出:“系统生物医学就是将系统生物学的方法策略应用到医学概念、研究和实践之中, ……这些活动的开展需要整合不同的学科, 包括数学、计算机科学、数据分析、生物学, 以及临床医学、伦理和社会实践”^[5]。

1.1 生命复杂系统的构成: 从分子到细胞到组织器官的相互作用网络

系统生物学最重要的特性之一就是关注生物体内各种元件之间的相互作用。最初的系统生物学研究主要是针对分子层面的相互作用网络, 如基因转录调控网络、信号转导网络和代谢调控网络等。2005 年 3 月创立的国际上第一个系统生物学的学术

刊物的名字就叫 *Molecular Systems Biology*。中国生物化学与分子生物学会在 2012 年 7 月成立“分子系统生物学专业委员会”也是基于这样的考虑: 由于不同种类的生物分子之间的相互作用是形成生物复杂系统的基础, 所以如何形成分子相互作用网络属于系统生物学的核心科学问题。

系统生物学的引入导致了人们对复杂疾病中有关生物分子作用的全新认识。例如, 传统肿瘤生物学通常是“孤立”地看待基因或蛋白质产生的突变, 认为单个突变可以改变其功能而导致肿瘤的发生发展。不久前, 研究人员利用自动化蛋白相互作用检测新技术系统地分析了肿瘤细胞中数以百万计的蛋白质相互作用, 发现突变会改变蛋白之间的相互作用, 进而形成新的蛋白相互作用网络^[6]。也就是说, 一个氨基酸残基的改变不仅能影响突变蛋白本身的功能, 还可能产生新的相互作用界面而与其他蛋白之间产生新的相互作用, 从而形成基于新的蛋白相互作用网络的功能变异或新功能。

生理或病理活动的复杂性不仅体现在分子层面, 而且还表现在细胞层面和组织器官层面。初期的系统生物学研究技术在细胞层面存在很大的局限性。但随着单细胞 RNA 测序技术的出现, 研究人员能够在细胞层面开展系统生物学方面的研究。例如, 新加坡研究人员利用单细胞 RNA 测序技术, 比较了胚胎发育过程中人类肝脏细胞以及肝癌细胞的单细胞图谱, 并在此基础上发现了一个既可以驱动胎肝发育又可以促进肝癌细胞免疫抑制的肿瘤-胚胎重编程生态系统^[7]。此外, 美国研究人员应用单细胞 RNA 测序技术, 对来自 20 名成人和 8 名儿童胶质母细胞瘤患者的近 25 000 个肿瘤细胞进行了分析, 发现胶质母细胞瘤存在分别具有独特基因表达谱的四种细胞形态, 即神经祖细胞样细胞、少突细胞祖细胞样细胞、星形细胞样细胞和间充质细胞样细胞, 并且这些肿瘤细胞可以在这四种细胞形态中进行转换^[8]。这些发现解释了为什么现有的靶向药物治疗方法难以阻止胶质母细胞瘤的生长, 因为该肿瘤中含有多种类型并可以相互转变的癌细胞。

高分辨空间组学技术的建立和发展则使研究人员更进一步能够对三维空间里的组织器官进行系统性的研究。例如, 耶鲁大学研究人员 2020 年 11 月在 *Cell* 发表论文, 报告了他们发展的一种空间组学技术“DBiT-seq”——将微流控芯片 (microfluidic chip) 和条形码技术 (barcoding) 与单细胞 RNA 测序技术相结合, 可以同时完成组织切片的空间转录组

和蛋白质组的测序，其空间分辨率接近单细胞分辨率^[9]。此外，深圳华大生命科学研究院联合全球多家研究机构组成的时空组学联盟(The Spatio Temporal Omics Consortium, STOC) 2022年5月在*Cell*报告了一项全新的时空组学技术“Stereo-seq”，其分辨率可达500 nm，视野也可达13 × 13 cm²；研究人员利用Stereo-seq技术分析了小鼠第9.5~16.5 d之间的早期胚胎发育过程，获得了单细胞分辨率水平的小鼠器官形成的时空图谱^[10]。

由此可见，当前的系统生物医学不仅有能力揭示基因和蛋白质等生物大分子间的相互作用和功能，而且可以整合生物体不同层次的数据和信息，从而能够更完整地认识人体复杂系统的运行和变化。美国国立肿瘤研究所(National Cancer Institute)最近启动了一个名为“人类肿瘤图谱网络”(The Human Tumor Atlas Network, HTAN)的研究计划，拟从分子、细胞、组织、器官等多个层次开展肿瘤发生发展机制的研究，并将肿瘤的这些多层次生物学数据与患者的临床数据进行整合，形成完整的肿瘤知识网络^[11]。

1.2 生命复杂系统的运行：基于非线性与动力学的控制

在简单性思维指导下，生物体内部的运行关系通常被视为线性的，如许多研究人员认为mRNA表达水平和其翻译产生的蛋白质丰度之间呈现正相关性，前者高则后者高，反之亦然。而越来越多的证据表明，在真实的生命复杂系统中，二者之间存在着复杂的非线性关系。美国研究者不久前对各种人类组织中12 000多个基因的表达水平与相应的蛋白质表达水平进行了定量的比较，发现二者的一致性并不是很高，且“组织特有的蛋白质信息能够解释遗传疾病的表型，而仅仅采用转录组信息则做不到这一点”^[12]。有研究者对mRNA表达水平与蛋白质表达水平的关系进行了系统的总结，认为二者间的数量关系是非线性的，受到细胞内外环境变化、细胞稳态和状态变化，以及mRNA胞内时空分布等各种因素的影响，“转录水平本身在许多情况下不足以用来预测蛋白质表达水平以及解释基因型与表型的关系。因此，获取在不同层次的基因表达水平相关的高质量数据是完全理解生物学过程所必不可少的”^[13]。

研究者发现，生物体内不仅各类生物分子的浓度之间存在着非线性关系，而且这些生物分子的行为和功能也有着复杂的表现形态。例如，比利时鲁

汶大学的研究者发现，催化3-磷酸甘油酸合成丝氨酸的磷酸甘油酸脱氢酶(PHGDH)在原发性乳腺癌细胞中通常表现为高表达，从而促进肿瘤细胞的增殖；但位于高度肿瘤血管化区域的肿瘤细胞则往往表现出较低的PHGDH表达水平，而低表达PHGDH能够促进整合素(integrin) $\alpha\beta3$ 的糖基化，进而导致这类乳腺癌细胞具有较强的转移能力^[14]。也就是说，不同的肿瘤微环境能够使得同一个酶具有不同的表达水平，导致肿瘤细胞形成代谢异质性，进而造成肿瘤细胞的不同行为。

P53蛋白是目前已知最重要的肿瘤抑制因子，在人类50%以上的肿瘤中都发现过p53基因的各种突变。P53能够通过复杂的调控网络影响众多生理活动，如DNA损伤、细胞周期增殖和细胞凋亡等。南京大学王炜教授运用系统生物学的研究方法，分析了p53网络中调控基因表达和信号转导等的动力学机制，发现p53蛋白存在浓度周期性振荡的回路机制，当DNA损伤程度较轻时，p53浓度的周期性变化可诱导短暂的细胞周期阻断，促进DNA修复，并使细胞在完成修复后继续存活，而当DNA损伤较为严重时，持续的p53脉冲则诱发细胞凋亡^[15]。

复杂生理病理过程普遍存在着一种临界现象，即从一个相对稳定状态，经过一个临界期后在很短的时间内快速地进入另一个稳定状态。例如，肿瘤或糖尿病等复杂疾病的发生过程中存在一个临界期，在疾病发生前的临界期为可逆阶段，适当的干预可以转归到“正常状态”；但当病变的进展一旦越过临界期，就迅速到达不可逆的“疾病状态”。显然，临界期就是复杂疾病早期监测和干预的关键时间节点。为了预测这种临界期及其关键驱动因子，中国科学院生物化学与细胞生物学研究所的陈洛南研究员及其合作者建立了基于“动态网络标志物”(dynamical network biomarker, DNB)的临界预测方法，即在复杂生物动态演化或疾病发生发展过程中，存在一个可观测的DNB，它在临界期形成为一个分子之间具有强相关并强震荡的奇异分子网络。DNB不仅可直接用于各种动态生物过程或复杂疾病发生发展的早期诊断，而且可用作复杂疾病发生发展过程的“驱动网络”和关键节点的检测标准^[16]。

1.3 生命复杂系统的研究：定性分析与定量检测的紧密结合

20世纪中叶诞生的分子生物学是一门依靠物理和化学方法的实验科学。那个时代的生命科学研

研究者大多关注定性的研究,以发现新基因或新蛋白质及其结构和功能为主要目标。正如曾担任过美国科学院院长的分子生物学家阿尔伯特(Albert B)所说,“在一个基因克隆占主要地位的时代,许多优秀的科学家在不具备任何定量研究的能力下仍然取得了巨大的成果”^[17]。但阿尔伯特教授同时指出,随着后基因组时代的到来,生命科学研究者的定量研究能力和知识已不再是可有可无的了;“对一种蛋白质机器功能的任何一种真正的认识,不仅需要它在原子精度的静态结构,而且需要有关它的每个反应中间体的动力学和热力学知识”^[17]。而系统生物学正是一门注重定量研究的学科,它不仅注重分子细胞生物学和组学等“湿实验”,而且同样注重生物信息学和计算生物学等“干实验”。成功的系统生物学研究一定是“干实验”与“湿实验”的紧密结合。

传统的肿瘤靶向治疗是直接针对肿瘤细胞中出现突变且功能异常的靶蛋白。但是,美国哥伦比亚大学系统生物学家卡里法纳(Califano A)和其合作者认为,这类突变蛋白通常处于一个调控网络之中,可以找出该网络的关键调控因子来进行靶向治疗。他们采用调控网络的逆向工程思路,建立了一种算法——VIPER (virtual inference of protein activity by enriched regulon analysis),通过对肿瘤细胞的转录组数据分析去寻找肿瘤异常调控网络中的“瓶颈因子”,并提出相应的靶向治疗方案^[18]。为此,哥伦比亚大学计划投入1 500万美元的经费,应用VIPER算法对其医院中的3 000名癌症患者进行分析,并开展相应的临床试验。卡里法纳等人在前期工作的基础上进一步发展了基于网络的新方法——“多组学的主调控因子分析”(multi-omics master-regulator analysis, MOMA),并从不同肿瘤近10 000个样本的多组学数据分析中找到了407个主调控因子^[19]。

系统生物医学通常要面对海量生物分子数据的处理与分析之挑战。但这还不是最难解决的问题;更难的是如何把不同种类生物分子的大数据与生物影像以及健康医疗档案等整合起来指导临床实践。不久前,瑞士科学家提出了一种用来指导肿瘤临床治疗决策的数据整合方案——“TuPro”(tumor profiling);其工作流程大致为:通过不同的分析技术从肿瘤患者样本中获取各种类型的生物分子大数据,然后将这些数据与临床数据进行整合,进而为每位患者生成一份分子研究报告,并提交给多学科

医师小组进行讨论,最终制定出一份具体的治疗方案^[20]。可以说,系统生物医学的大数据整合工作不过是刚刚开始,其临床运用的可行性和可操作性还有待发展和完善。

2 从基于统计研究证据的循证医学转变到关注个体分子特征的精确医学

现代医学的主流是“循证医学”(evidence-based medicine),其诊治方案形成的主要依据是按照各种类型临床研究证据制定的临床指南,以此开展基于统一标准的规范化临床实践活动。临床研究证据的金标准是“随机对照试验”(randomized controlled trial, RCT)。这是一种尽可能排除个体差异对研究结果的统计性影响的临床试验,一方面基于临床试验统计学的要求进行试验设计和招募参试者,另一方面对参试者进行试验组和对照组的随机分配,以减少个体差异可能导致的统计学试验偏倚,从而得到具有普遍意义的统计学规律。

由此可见,基于RCT等各种临床统计学证据的循证医学的主要特征可以说是看“病”而不是看“人”,即患者仅仅是一个“病例”,而不是一个“病人”。循证医学超越了传统医学那种依靠个人经验的医疗实践模式,能够在科学证据的指导下进行更为客观的医疗实践活动。但是,这种排除个体差异的统计学方式同时导致了循证医学在治疗慢性病患者时的实际疗效往往因人而异,因为患者之间广泛存在着由不同的遗传背景和不同的生活环境而产生的个体差异。显然,循证医学的优点——排除了个体差异并具有统计显著性的治疗方案,对具体的患者治疗来说却成了缺点——不够精确!

人类基因组计划在改变循证医学这种“不精确”问题方面同样发挥了重要的作用;正如*Science*周刊在“庆祝基因组”社论中所说:“基因组草图的完成成为一种新的精确医学范式奠定了基础,这种精确医学的目标就是要利用个体独特的基因序列信息去指导治疗和预防疾病的决策”^[21]。事实上,美国研究者对此很早就有清晰的认识:“与人体有关的分子数据正在爆发性的增长,尤其是那些与患者个体相关的分子数据;由此带来了巨大的、尚未被开发的机会,即如何利用这些分子数据改善人类的健康状况”^[22]。基于这样的认识,美国政府在2015年初正式宣布实施“精确医学”(Precision Medicine);此后,包括中国在内的各国迅速跟进,形成了世界范围的精确医学新潮流。

2.1 精确医学的底层逻辑：分子层面的个性与共性之统一

精确医学把主要目标定位在从分子层面认识清楚个体间的遗传差异和表型差异，并相应地把基本任务放在寻找和确定标识个体特征的遗传因子或者表型因子等各种“生物标志物”(biomarker)。例如，欧盟在2014年启动的“创新药物先导项目”(Innovative Medicines Initiative 2, IMI2)中明确指出，精确医学的主要任务就是“生物标志物的发现和验证”。此外，NIH牵头启动的国际癌症基因组项目“癌症基因组图集”(The Cancer Genome Atlas, TCGA)的目标也正是要获取分子层面的信息以进行肿瘤分子分型；该项目目前已经进行了33种不同癌症类型11 000名患者的基因组测序和其他种类生物分子数据的采集与分析^[23]。不久前，英国研究人员报道了肿瘤患者样本规模最大的一项全基因组测序研究；他们通过比较19种癌症类型12 222名患者的全基因组序列，揭示出了58种过去未知的肿瘤基因组序列的突变特征，进而为每种癌症类型确定了常见突变特征与罕见突变特征^[24]。显然，这些分子层面的数据将为未来的个性化精确医疗提供重要的指导信息。

越来越多的研究表明，不同个体在分子层面广泛存在着个体间异质性(inter-heterogeneity)；即使是同卵孪生的双胞胎，也有多项研究发现，二者之间的基因组序列不是完全一样的。更具有挑战性的是细胞之间的内在异质性(intra-heterogeneity)，即个体内同一种组织的同一类型细胞群体的不同细胞也可能存在基因或蛋白质等分子之间的差异。例如，研究人员利用专一结合人体 β 细胞的膜表面蛋白的抗体技术发现，正常成人胰岛组织中的 β 细胞群体中存在4种亚型，这些不同亚型的 β 细胞对葡萄糖的响应有着明显的差别；研究人员还发现，2型糖尿病患者体内的这4种 β 细胞亚型的数量关系发生了明显的改变^[25]。

在肿瘤的发生发展过程中，肿瘤细胞的内在异质性更是扮演了重要角色。中国研究人员曾经通过基因测序等技术分析了一个直径大约为3.5 cm的肝癌组织上基因突变情况，推断出这一肝癌组织拥有上亿个突变，且不同肝癌细胞拥有的突变类型和数量是不一样的^[26]。不久前，一支国际研究团队在*Cell*上发表了目前最大规模的肿瘤细胞间异质性的研究工作——通过分析38种癌症的2 658个肿瘤样本的全基因组测序数据，系统地阐释了肿瘤的异

质性图谱；研究数据揭示，超过95%的肿瘤里都存在代表肿瘤细胞间异质性的亚克隆扩张(subclonal expansion)，这些具有不同突变特征的亚克隆扩张驱动着肿瘤的演化^[27]。需要强调的是，这种肿瘤细胞间异质性不仅表现在基因组序列的差异上，也表现在基因转录调控和蛋白质表达等各种分子层面上，如在不同微环境下乳腺癌细胞具有不同的代谢异质性，其中高表达磷酸甘油酸脱氢酶的癌细胞具有很强的增殖能力，而低表达该酶的癌细胞则具有较强的转移能力^[14]。

研究者在关注研究“个性”的同时通常也需要关注“共性”，这二者就好像一个钱币的两面，是不可分割的。人类基因组计划最初的目标是建立一个代表全人类的“人类基因组参考图谱”；在2004年该计划的目标达到之后研究者就开始关注个体基因组的检测，如2008年启动的“国际千人基因组计划”。目前国际学术界采用的“人类基因组参考图谱”是用20多个人的基因组序列拼接成的，其中有大约70%的碱基序列是来自同一个人^[28]。不久前，中、美等多个国家的研究者组建了人类泛基因组参考联盟(Human Pangenome Reference Consortium, HPRC)；“人类泛基因组”(human pangenome)的概念，不仅是指一个更高质量和更完整的人类参考基因组，而是指一个更完整的人类基因组变异框架，涵盖包括重复序列以及单核苷酸多态性等整个基因组范围内的变异信息^[28]。换句话说，“人类泛基因组”的提出就是要在分子层面实现个性与共性的整合。

精确医学延续着同样的思路，它并不把其研究工作局限于寻找和鉴定个体之间的分子差异，而是拓展到对不同个体在分子层面的共性研究，其中最具有代表性的就是新的肿瘤类型“泛癌”(PanCancer)概念的提出。为此，在TCGA计划中专门衍生出一个“泛癌图谱计划”(PanCancer Atlas project)，“泛癌图谱计划获得的结果将为下一阶段的工作打下坚实的基础，而后续这类更深入、更广泛和更复杂的工作将有助于实现个体化肿瘤治疗”^[29]。这种“PanCancer”研究可以超越基于病理特征和解剖形态等传统的宏观疾病分类标准，把不同组织/解剖的肿瘤类型视为一个整体。例如，研究者利用TCGA计划获得的RNA转录组数据，对33种肿瘤类型共9 000个样本进行了“增强子表达”(enhancer expression)的共性分析，发现在这些“PanCancer”肿瘤样本中，“基因组整体水平的增强子活性与非整倍体(aneuploidy)正相关，而与基因突变的程度

则没有相关性”^[30]。不久前,荷兰研究者从“泛转移瘤”的角度比较了20多种实体瘤的2520对转移性和原发性肿瘤样本的全基因组序列,发现这些实体转移瘤中的“全基因组扩增”(whole genome duplication, WGD)程度比非转移性瘤的要高很多,前者的WGD平均值达到了55.9%,因此WGD是这些不同类型的转移性肿瘤之共同分子特征^[31]。

最近,美国研究者在*Cell*上发表了关于不同人脑转移瘤(brain metastases)的细胞组成与基因调控模式的研究论文。他们从黑色素瘤和乳腺癌等8种类型的原发性肿瘤的15个患者体内获得了相应的脑转移瘤样本,然后对这些样本进行了单细胞转录组测序等各种分析,一方面发现了不同患者的脑转移肿瘤细胞具有高度的异质性,另一方面也在这些人脑转移肿瘤细胞里鉴定出了8种基因调控模式;此外,他们通过对脑转移肿瘤血液-肿瘤界面的研究,发现可以把这些人脑转移肿瘤细胞分为两种基本类型:一种是增殖型(proliferative),另一种是炎症型(inflammatory)^[32]。这一工作很好地反映了精确医学是如何整合肿瘤的个性与共性研究。

2.2 精确医学的技术路径:理想试验设计与真实世界研究

为了克服RCT研究过程中刻意消除个体差异导致的不精确缺点,精确医学发展出了各种理想化的基于个性差异和共性特征的新型临床研究模式。首先是基于“同病异治”思路的“伞型试验”(umbrella trial),即针对单一疾病采用多种药物治疗并评估其效果^[33],如英国目前正在进行一个“肺癌伞型试验”——National Lung Matrix Trial (NLMT),涉及到具有22个分子标记物的19种非小细胞肺癌患者队列和8种治疗药物。其次是基于“异病同治”思路的“篮型试验”(basket trial),即按照统一的某个分子标志物把不同类型的疾病患者集中在一起,用来评估某一种药物对这些不同类型疾病的治疗效果^[33],如在2018年被美国食品药品监督管理局(Food and Drug Administration, FDA)批准的TRK(tropomyosin receptor kinase)抑制剂Larotrectinib,就是首个依据“篮型试验”结果获批的抗肿瘤药物——不论哪种类型的实体瘤,只要有TRK基因融合突变就可以用此药治疗。

对精确医学而言,理想化的临床研究应该在分子层面实现个性与共性的整合。为此,FDA提出了一种“主方案”(master protocols),不仅同时包括了伞型试验和篮型试验,而且包括了一种“平台试

验”(platform trial),即在同一个研究平台上平行开展在多个不同分子标志物指导下的单臂药物试验^[33]。NIH目前正在开展的“基于分子分析的治疗选择试验”(molecular analysis for therapy choice trial, MATCH)是当前规模最大的一项“主方案”;该方案从6000名肿瘤患者中选出了1000名分别进入到30项治疗单臂试验中;参与这些试验的患者涉及到几乎所有肿瘤类型。

随着大数据时代的到来,生物医学大数据也成为了实现精确医学的重要手段。美国国会在2016年通过的《21世纪治疗法案》中提出,日常临床实践中产生的丰富多样的真实世界数据,如电子健康档案和医保数据等构成的“真实世界证据”(real world evidence, RWE)可以作为临床试验证据之外的补充证据。美国FDA在2018年公布了《真实世界证据方案框架》,并明确提出RWE可以作为合成对照臂(synthetic control arm)整合到传统的单臂临床试验。中国近年来也逐渐重视RWE,国家药监局于2020年发布了《真实世界证据支持药物研发与审评的指导原则(试行)》办法;其下属的药品审评中心也在2021年4月发布了《用于产生真实世界证据的真实世界数据指导原则(试行)》。

获得RWE的一个主要途径是真实世界研究(real world study, RWS)。RWS过去通常是一种研究者不对受试对象进行主动干预的观察型研究,如新药上市以后考察其治疗效果和安全性的IV期临床试验。人们认识到,由于RWS源于实际医疗场地或家庭社区等场景,可以避免RCT那样严格受控实验条件带来的局限性,因此RWS也成为了临床试验中新的干预型研究手段,如中国药监局药品审评中心在2021年发布的《以临床价值为导向的抗肿瘤药物临床研发指导原则(征求意见稿)》中,就明确给出了临床研究进入“关键研究阶段”时的3种临床试验设计:(1)随机对照研究;(2)单臂临床试验;(3)真实世界研究。

3 从以治病为中心的临床医学转变到以健康为中心的健康医学

随着当今人类疾病谱从传染病为主转变为慢性病为主,医学的理念和形态正在发生着巨大的变化。首先是医学的关注点从“治疗疾病”转变为“维护健康”。慢性病的发生发展通常都要有一个较长的时间;在出现临床症状之前,会先出现亚健康状态或疾病前期状态等各种过渡态,其高危人群数量往

往比患病人群要大很多；如国内目前糖尿病患者为1亿，但处于糖尿病前期 (prediabetes) 的高危人群可能已达到5亿左右了^[34]。这种疾病演化的“窗口期”给抗击慢性病提供了一个不同于抗击传染病的思路，即将抗击疾病的“关口前移”，加强对人群健康状态的早期监测，并在发现亚健康或疾病前期状态时进行早期干预。

这种“大健康”的思路今天已经上升为中国政府的基本国策，在2016年国家颁布的《“健康中国2030”规划纲要》中，明确提出其指导思想是：把健康摆在优先发展的战略地位；实现从胎儿到生命终点的全程健康服务和健康保障。需要指出的是，“关口前移”并非说要忽略对患者的临床诊断和治疗，而是要把维护健康和临床诊治整合为一体，形成大健康时代的“健康医学”。也就是说，过去的临床医学主要关注患病人群，而今天的健康医学则拓展到所有个体，正如2019年召开的第72届世界卫生大会倡导的主题——全民健康覆盖：不遗漏任何一人。

3.1 健康状态的全过程认识：全人群与全数据

健康医学面临的挑战显然要大于临床医学，不仅要处理临床诊治方面的问题，而且还要解决健康促进和维护方面的问题。此外，对亚健康或者疾病前期的评估也同样充满挑战。例如，国际上通行的高血压诊断标准是 $\geq 140/90$ mmHg；但是，美国心脏协会2017年单方面将高血压的确诊标准调整为 $\geq 130/80$ mmHg，使得美国的高血压患者一下就增加了3千万。这个调整后的新标准导致了各国专家学者的许多争议。由此可见，健康医学不能按照传统的医学研究模式来选择研究对象，而是要采用把所有个体都纳入到研究范围的“全人群”策略。目前NIH正在实施的“全民健康研究项目”(All of Us Research Program)就是这种健康医学“全人群”理念的代表。NIH的负责人特别强调：这个项目不关注疾病 (disease agnostic)， “它不聚焦在某一种疾病、某一种风险因子，或者是某一类人群；反之，它使得研究者可以评估涉及到各种疾病的多种风险因子”^[35]。因此，该项目特别重视参与者的广泛性和多样性，计划招募的全美百万志愿者不限性别、民族和健康状态等，并且要覆盖全美各地区和各阶层，包括过去不受重视的族群 (underrepresented groups)^[35]。

需要指出的是，这种“健康”医学的“全人群”策略是要在一个“长时程”的维度上展开的。首先，要想有效地统计和分析非特定构成的自然人群中各

种常见病的患病率和发病率，不仅需要人群的数量足够大，而且需要连续地对人群进行观察。美国研究者认为，只要美国的自然人群数量达到或超过100万，在5~10年期间内检测到的每种美国人常见病 (如糖尿病、中风、各种类型肿瘤) 的平均发病数量将超过2万例，并将伴随着显著的致死致残率^[36]。其次，对自然人群进行观察和数据收集的时间越长对健康医学就越有价值。例如，在“全民健康研究项目”中，至少要有3年的人群数据才能用于疾病的分类或支撑临床研究，如果是5年的人群数据则用来做这两件事的效果将更好^[37]。

显然，“全人群”研究的目标是获得尽可能完整的人群的健康医学大数据，可以称之为“全数据”，即人群的健康数据越完整、越全面，对健康医学就越有价值。目前在健康医学“全数据”方面最成功的是英国的UK Biobank。英国研究者于2006年启动了UK Biobank项目，随后在5年时间里收集了50万名40~69岁英国志愿者的血液、尿液和唾液等生物学样本，以及EHR数据等各种个人信息^[38]。要强调的是，UK Biobank追求的目标正是健康领域的“全数据”。自2012年建成至今，UK Biobank一直在完善其数据的收集工作，如初期对50万名志愿者进行了全外显子测序，2021年则在政府和公司的资助下完成了这些志愿者中20万人的全基因组测序，未来将完成其余30万人的全基因组测序；此外，在2022年1月完成了5万人的器官成像，下一步将扩大到10万人。UK Biobank的数据量预计在2025年将达到40 PB (1 PB = 10^{15} Byte)。由于UK Biobank具有这样的健康医学“全数据”，所以被全世界的研究者用于健康医学领域各种各样问题的研究，目前其拥有包括中国和其他国家在内的全球注册用户28 000多名，基于这些数据已发表了近25 000篇研究论文。例如，研究者不久前利用UK Biobank里27万名欧洲血统参与者的外显子组序列数据，分析了基因变异与表型之间的关联，发现了许多常见疾病的罕见蛋白编码变异^[39]。

3.2 健康状态的全过程管理：全方位的生活方式干预

健康医学的基本形态不同于临床医学，前者是人人参与的“主动健康”，后者则主要是依靠医生的“被动健康”。对临床医学而言，大众把诊治自身疾病的职责交给了医生。但是，健康医学面对维护和管理公众全程健康的需求，依靠医生是远远不够的，需要每一个人的参与。这种“主动健康”理

念已经被纳入 2020 年实施的《中华人民共和国基本医疗卫生与健康促进法》；该法第六十九条清楚地写道：“公民是自己健康的第一责任人，树立和践行对自己健康负责的健康管理理念，主动学习健康知识，提高健康素养，加强健康管理。”

实施健康医学“关口前移”战略的主要举措是早期干预；最常用又最简便的是全方位的生活方式干预，包括营养干预、运动干预、心理干预和睡眠调理等等。例如美国肿瘤协会 (American Cancer Society) 不久前提出了未来 10 年的 10 项预防肿瘤的措施，其中六项都是生活方式干预，包括戒烟、限酒、健康饮食、防晒、加强运动、控制体重^[40]。中国政府在 2019 年 6 月颁布的《国务院关于实施健康中国行动的意见》中也提出了多项具体的生活方式干预行动，如合理膳食行动、全民健身行动、控烟行动和心理健康促进行动等。

饮食对健康的影响从古至今都受到高度重视，中国传统医学很早就提出“药食同源”的观点。美国癌症协会 2016 年 2 月在其网站上推荐了一份用于预防肿瘤的“彩虹食谱”：把蔬果按颜色分成 5 个种类，不同颜色代表不同的植物营养素；人们每天要将不同颜色的蔬果按一定比例搭配进食。此外，许多国家也很重视发布官方的膳食指南，用于指导民众健康饮食，如中国政府于 2017 年 6 月颁布了《国民营养计划 (2017—2030 年)》。

随着健康医学的形成和发展，运动干预近年来越来越受到重视。2018 年第 71 届世界卫生大会通过《2018—2030 年促进身体活动全球行动计划》，希望到 2030 年时将缺乏身体活动现象减少 15%。2020 年世界卫生组织发布了《关于身体活动和久坐行为指南》，建议所有成年人每周至少进行 150~300 min 中等到剧烈的有氧运动。中国国家体育总局在 2017 年也发布了《全民健身指南》，明确指出：“体育活动已经成为增强国民体质、提高健康水平最积极、最有效、最经济的生活方式”。

生活方式干预不仅被用于预防慢性病的发生，而且还正在被视为治疗疾病的重要手段。不久前，美国心脏协会 (American Heart Association) 发布一个声明，建议把运动锻炼作为降低血压和血脂的首选干预措施^[41]。2022 年初，美国纽约市市政府宣布，把生活方式干预纳入大纽约地区的医疗卫生系统，作为针对代谢性疾病等慢性病的一线治疗手段。我国政府和医学界同样很重视生活方式干预的治疗价值，如《中国 2 型糖尿病防治指南 (2017 年版)》

明确规定，运动和饮食等单纯生活方式干预是血糖控制的首选方法，“生活方式干预是 2 型糖尿病的基础治疗措施，应贯穿于糖尿病治疗的始终”^[42]。由此可以看到，健康医学在疾病治疗手段上形成了“三足鼎立”——药物、手术刀、生活方式干预。

4 小结：凡是过往，皆为序章

现代医学在科学技术的推动下正在出现新的发展趋势。从以上的分析可以看到主要表现在三个方面。首先，系统生物学的形成和发展改变了基于还原论的“碎片化”和“简单化”的生命观，使得人们从生命复杂系统的角度去重新认识生命的生理和病理活动。其次，在系统生物学和生物医学大数据的共同推动下，人们从关注疾病发生发展的分子细胞机制转变成关注个体在分子层面的差异和共性。最重要的是，对健康的维护不再局限于临床诊断和治疗，而是对生命从正常到异常到临床的全过程监测，以及从营养到运动到治疗的全方位干预。人类社会迎来了一个全新的“大健康”时代。

致谢：该文得到上海市科学技术委员会基金“支撑主动健康的主要技术对策的研究”(21692113400)和中国科学院先导专项：“多维大数据驱动的中国人人群精准健康研究”，(XDB38020000)的支持。

[参 考 文 献]

- [1] Weinberg R. Point: hypotheses first. *Nature*, 2010, 464: 678
- [2] Weinberg R. Coming full circle—from endless complexity to simplicity and back again. *Cell*, 2014, 157: 267-71
- [3] Editorial. To thwart disease, apply now. *Nature*, 2008, 453: 823
- [4] National Institutes of Health. The NIH roadmap[EB/OL]. <http://nihroadmap.nih.gov>
- [5] The CASyM Consortium. The CASyM roadmap: implementation of systems medicine across Europe[EB/OL]. <http://www.casym.eu/publications>
- [6] Mo X, Niu Q, Ivanov AA, et al. Systematic discovery of mutation-directed neo-protein-protein interactions in cancer. *Cell*, 2022, 185: 1974-85
- [7] Sharma A, Seow JW, Dutert CA, et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell*, 2020, 183: 377-94
- [8] Neftel C, Laffy J, Filbin MG, et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell*, 2019, 178: 835-49.e21
- [9] Liu Y, Yang M, Deng Y, et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell*, 2020, 183: 1665-81

- [10] Chen A, Liao S, Cheng M, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 2022, 185: 1777-92
- [11] Rozenblatt-Rosen O, Regev A, Oberdoerffer P, et al. The human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. *Cell*, 2020, 181: 236-49
- [12] Jiang L, Wang M, Lin S, et al. A quantitative proteome map of the human body. *Cell*, 2020, 183: 269-83.e19
- [13] Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. *Cell*, 2016, 165: 535-50
- [14] Rossi M, Altea-Manzano P, Demicco M, et al. PHGDH heterogeneity potentiates cancer cell dissemination and metastasis. *Nature*, 2022, 605: 747-53
- [15] Zhang XP, Liu F, Wang W. Two-phase dynamics of p53 in the DNA damage response. *Proc Natl Acad Sci USA*, 2011, 108: 8990-95
- [16] Chen L, Liu R, Liu ZP, et al. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep*, 2012, 2: 342
- [17] Alberts B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, 1998, 92: 291-4
- [18] Califano A, Alvarez JM. The recurrent architecture of tumour initiation, progression and drug sensitivity. *Nat Rev Cancer*, 2017, 17: 116-30
- [19] Paull EO, Aytes A, Jones SJ, et al. A modular master regulator landscape controls cancer transcriptional identity. *Cell*, 2021, 184: 334-51
- [20] Irmisch A, Bonilla X, Chevrier S, et al. The Tumor Profiler Study: integrated, multi-omic, functional tumor profiling for clinical decision support. *Cancer Cell*, 2021, 39: 288-93
- [21] Fraser CM. A genome to celebrate. *Science*, 2021, 371: 545
- [22] National Research Council. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease[EB/OL]. <http://www.nap.edu/catalog/13284/>
- [23] Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. *Cell*, 2018, 173: 283-5
- [24] Degasperi A, Zou X, Amarante TD, et al. Substitution mutational signatures in whole-genome-sequenced cancers in the UK population. *Science*, 2022, 376: science.abl9283
- [25] Dorrell C, Schug J, Canaday PS, et al. Human islets contain four distinct subtypes of β cells. *Nat Commun*, 2016, 7: 11756
- [26] Ling S, Hu Z, Yang Z, et al. Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. *Proc Natl Acad Sci USA*, 2015, 112: E6496-505
- [27] Dentre SC, Leshchiner I, Haase K, et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell*, 2021, 184: 2239-54
- [28] Wang T, Antonacci-Fulton L, Howe K, et al. The Human Pangenome Project: a global resource to map genomic diversity. *Nature*, 2022, 604: 437-45
- [29] Ding L, Bailey MH, Porta-Pardo E, et al. Perspective on oncogenic processes at the end of the beginning of cancer genomics. *Cell*, 2018, 173: 305-20
- [30] Chen H, Li C, Peng X, et al. A pan-cancer analysis of enhancer expression in nearly 9000 patient samples. *Cell*, 2018, 173: 386-99
- [31] Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature*, 2019, 575: 210-6
- [32] Gonzalez H, Mei W, Robles I, et al. Cellular architecture of human brain metastases. *Cell*, 2022, 185: 729-45
- [33] Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. *N Engl J Med*, 2017, 377: 62-70
- [34] Xu Y, Wang L, He J, et al. Prevalence and control of diabetes in Chinese adults. *JAMA*, 2013, 310: 948-58
- [35] NIH-Wide Strategic Plan for Fiscal Years 2021–2025[EB/OL]. <https://www.nih.gov/sites/default/files/about-nih/strategic-plan-fy2021-2025-508.pdf>
- [36] Precision Medicine Initiative (PMI) Working Group. The precision medicine initiative cohort program – building a research foundation for 21st century medicine[EB/OL]. <https://www.nih.gov/sites/default/files/research-training/initiatives/pmi/pmi-working-group-report-20150917-2.pdf>
- [37] The All of Us Research Program Investigators. The “All of Us” Research Program. *N Engl J Med*, 2019, 381: 668-76
- [38] Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 2018, 562: 203-9
- [39] Wang Q, Dhindsa RS, Carss K, et al. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature*, 2021, 597: 527-32
- [40] Siegel RL, Jemal A, Wender RC, et al. An assessment of progress in cancer control. *CA Cancer J Clin*, 2018, 68: 329-39
- [41] Barone GB, Hivert MF, Jerome GJ, et al. Physical activity as a critical component of first-line treatment for elevated blood pressure or cholesterol: who, what, and how?: A scientific statement from the American Heart Association. *Hypertension*, 2021, 78: e26-e37
- [42] 中华医学会糖尿病学分会中国2型糖尿病防治指南(2017年版). *中华糖尿病杂志*, 2018, 10: 4-67