

计算系统毒理学: 形成、发展及应用

李杰, 李柯佳, 张臣, 彭邱鹏, 唐赞*

华东理工大学药学院, 上海市新药设计重点实验室, 上海 200237

* 联系人, E-mail: ytang234@ecust.edu.cn

2014-12-29 收稿, 2015-01-30 接受, 2015-04-27 网络版发表

国家自然科学基金(81373329)、国家高技术研究发展计划(2012AA020308)和中央高校基本科研业务费(WY1113007)资助

摘要 系统生物学的发展为传统毒理学的研究提供了新机遇, 逐步形成了系统毒理学. 计算系统毒理学旨在整合毒理组学实验数据, 构建多水平、多尺度的预测模型, 定量评估化合物的安全性. 目前已经发展了静态网络分析预测、动态网络模拟和有害结局路径等几种研究方法. 虽起步不久, 但在全面理解毒理学机制、发现新的生物标志物及化合物安全性综合评估等方面已表现出良好的应用前景. 本文主要综述了计算系统毒理学的诞生背景、数据资源、研究方法、应用领域及未来展望.

关键词

系统毒理学
组学
数据库
计算方法

纳米材料、绿色农药、新型化妆品、食品添加剂……科研工作者一直致力于“化学改变生活”, 新化学实体(new chemical entity, NCE)不断被批准进入我们的日常生活. 组合化学、点击化学等新技术的应用, 使得化合物数目正以爆炸性速度增长. 截至2014年12月, 美国化学文摘社(Cheical Abstracts Service, CAS)登记的化学品数目已超过9100万种. 食品安全、罕见疾病、生物富集、环境污染、生态失衡等一系列由化学品引发的问题给我们敲响了警钟: 化学研究对效应与安全性的重视已失衡, 新化学实体和已上市分子的风险评估需得到更多的重视. 毒理学研究是化合物风险评估的重要组成部分, 旨在让功能化学分子更安全、更绿色. 20世纪以来, 毒性测试主要是通过整体动物实验, 评估化合物对疾病相关临床体征或病理变化的影响, 如急性毒性、肝损伤等. 但这种方法使用大量动物, 存在时间、成本、效能的局限性以及机制不明确、剂量外推和种属外推不确定性等问题, 也不适用于目前化合物快速增长模式^[1]. 采用计算方法构建毒性预测模型, 可节省大量人力、物力、财力, 已成为化合物早期风险评估有效的研究手段. 但传统的计算方法多采用定量构性关系

(quantitative structure-property relationships, QSPR)分析, 依赖于建模化合物结构的相似性及测试条件的一致性, 适用于简单毒性的分类和回归预测, 已应对不了多机制、慢性毒性的预测.

系统生物学的发展, 为复杂毒性的研究, 提供了新的契机. 系统生物学强调生命过程应该作为统一整体来研究, 而不是各个部分的简单相加^[2]. 化学小分子进入机体以后, 并非作用于单一蛋白、遵循单一机制, 而是作用于细胞、组织、器官不同水平的分子网络体系, 如药物-靶标的多对多关系^[3]. 毒性是机体网络系统对外界化学品应激源免疫、代偿后表现的不良结局. 因此, 应用系统生物学的新思路来指导化合物的安全性研究, 可将西方的还原论和东方的整体论思想有机地结合起来, 进行系统性研究^[4]. 基于此, 毒理学研究不再限于传统动物体内实验和QSPR的简单分析, 基于系统生物学整体思路的系统毒理学应运而生^[5]. 系统毒理学旨在采用人源细胞系、细胞组分进行体外高通量筛选实验, 整合计算预测模型, 直接测试和模拟人体环境, 对化合物人体健康风险进行直接评估. 主要包括毒理组学技术及计算系统毒理学两部分^[6].

引用格式: 李杰, 李柯佳, 张臣, 等. 计算系统毒理学: 形成、发展及应用. 科学通报, 2015, 60: 1751-1760

Li J, Li K J, Zhang C, et al. Computational systems toxicology: Emergence, development and application (in Chinese). Chin Sci Bull, 2015, 60: 1751-1760, doi: 10.1360/N972014-01400

毒理组学技术的发展,得益于后基因组时代高通量分子生物学技术和生物信息学的发展^[7].新一代测序技术、基因表达谱分析、实时定量聚合酶链式反应(PCR)技术促进了基因组学的发展;蛋白质芯片、高通量蛋白分离分析技术极大地扩展了蛋白组学的内涵;新型核磁共振、质谱对代谢变化的精确检测、分析和鉴定,使代谢组学得到进步;体外高通量实验,为测定大批量化合物提供可能;同时工作流等新的生物信息学方法的开发,为大数据的处理和分析提供了有效的手段.针对毒理学研究所面临的挑战和发展机遇,2007年美国国家研究咨询委员会编写了“21世纪毒性测试:愿景与策略(toxicity testing in the 21st century: a vision and a strategy, TT21C)”的研究报告^[8].如图1所示,TT21C包括化学物特征描述、毒性测试、剂量-反应和外推模型、人群暴露数据及风险信息资料等几个方面的研究,强调应用毒理组学、系统生物学和计算毒理学等先进技术,重点开展以“毒性通路”为核心的体外高通量测试方法,为定量系统性毒理学评估提供可靠的依据.TT21C一经提出,便得到各国工业界、学术界的一致响应,纷纷投入巨资启动了多个重大项目:如“毒性预报”ToxCast^[10]、“21世纪毒性测试”Tox21^[11]等.Tox21项目与人类健康风险评估“RISK21”^[12]和“21世纪暴露组学”项目^[13]一起,为使用高新技术全面评估化合物对人体、环境的安全性提供了未来发展的蓝图.

TT21C指导下的体外高通量测试会产生大量可分析和预测的毒理组学数据,亟需开发新的计算方法进行大数据整合、挖掘并构建预测模型,以实现动物模型的替代.

随着系统毒理学实验数据的增多,计算系统毒理学逐渐兴起.其主要目的是基于TT21C指导下的高通量毒理组学数据,用计算的手段开发系统性定量评估化合物毒性的方法,为化合物安全性决策的复杂问题提供有效的预测工具.主要研究内容包括:(1)通过对化合物暴露高通量实验数据的挖掘,得到化合物影响复杂生命体和环境的相关知识库;(2)通过构建毒理组学网络模型,并用数学方法表示和模拟中间过程,全面理解外源物致毒的中间机制;(3)发展具有预测功能的毒理学综合模型,以期定量全面评估化合物的安全性^[14].强调“毒性通路”系统的计算思路的出现,使得计算系统毒理学区别于只关注最高效应终点的传统计算毒理学,成为一个新的研究分支.计算系统毒理学发展的初期主要是借鉴计算系统生物学的研究方法,解决21世纪毒理学的预测问题.计算系统毒理学是多学科交叉的新兴学科,涉及化学、生物学、环境科学、医药、数学、计算机等不同学科领域.计算系统毒理学也一定是系统性的,即化合物高通量性、多毒性端点性和通路全面性.鉴于此,计算系统毒理学将有望解决传统毒理学的遗留问题,如高剂量动物测试外推低剂量人体

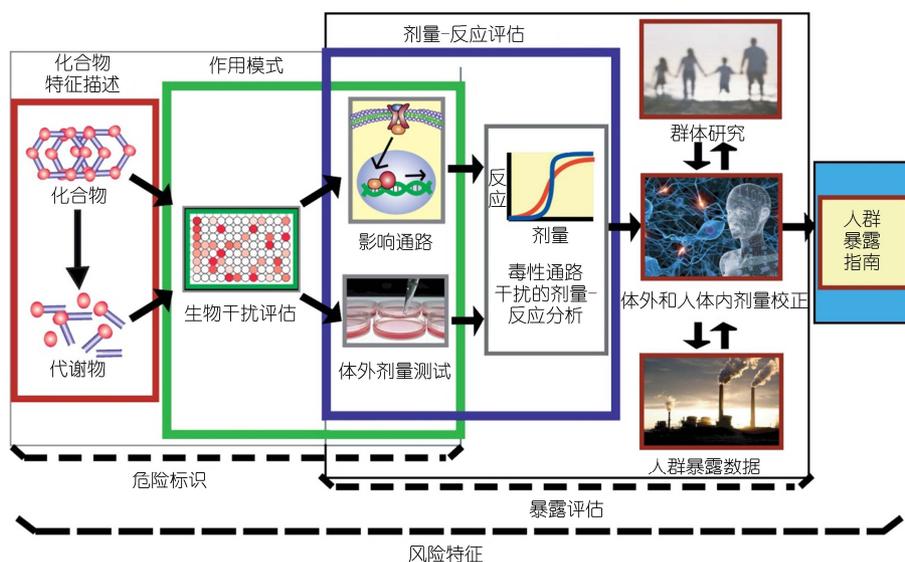


图1 (网络版彩色)21世纪毒性测试指导下的风险评估策略示意图. 图片编译来自文献[9]

Figure 1 (Color online) Risk assessment scheme under the guidance of toxicity testing in the 21st century. Modified after Ref.[9]

效应困难等。同时,在化合物设计早期便可批量预测有害效应,节省了不必要的实验消耗,为寻找绿色可替代化合物提供了新的评价方式。

本文将系统性介绍计算系统毒理学的数据资源、研究方法、应用领域及未来展望,希望为该领域的研究提供一定的参考,提升我国风险决策的预测水平,尽快与国际研究热点接轨。

1 数据资源

构建计算模型的基础是高质量的数据,表S1是相关开源数据库按照类别进行汇总,包含数据库名称、开发者、链接及主要数据内容。计算系统毒理学可利用的数据主要分为2大类:系统生物学数据和暴露相关数据。系统生物学数据的重点是测量机体(尤其是人体)不同解剖学水平如器官、组织、细胞,不同分子水平如基因、转录组、蛋白、代谢的组成及相互作用等。比较有代表性的是京都基因和基因组学百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)^[15]和维基通路(WikiPathways)^[16]2个免费的开源知识数据库。此外,随着系统生物学的发展,数学模型可描述的生命过程也日益增多。CellML^[17]和BioModels^[18]等数据库,使用统一的文件标准,是储存、分享数学模型的平台。更多系统生物学相关数据库可见<http://www.pathguide.org/>^[19]。系统生物学相关数据虽然大都表征正常的生理状态,但可为外界化学品暴露下病理或毒理状态的计算提供参考。

计算系统毒理学最根本的目标是通过毒理学通路分析实现定量化合物风险综合评价,因此需要化学品暴露后机体进行调整、反应所表现的组学实验数据及特定毒性所测得的相关机制组学信息。这类实验数据近几年刚刚兴起。ToxCast计划目前已完成了2000多种农药、工业化学品、食品添加剂、纳米材料等相关化合物的测试^[20,21]。Tox21利用更有效的自动高通量筛选平台,对上万个化合物的潜在毒性关键通路涉及的蛋白组进行测试,是21世纪毒理学里程碑式的工作^[22,23]。Tox21所测数据公开在其化合物检索平台及综合数据库PubChem^[24],ACToR^[25]及CEBS^[26]上。新一代的体外高通量筛选实验数据,不仅可以减少或替代传统毒理学所需的实验动物模型,同时为构建计算模型进行定量评估化合物安全性提供可能。

虽然Tox21采用自动高通量筛选平台,1 d可测试相当于1年人力的上千个化合物,但是由于大数据时

代化合物爆炸性增长模式及较昂贵的花费,只能获得有限的具有代表性化合物库的数据。因此,需要其他化学组学相关数据库及传统毒理学数据库来丰富系统毒理学的信息空间。化学组学数据库包括化合物-基因-疾病关联数据库CTD,毒物-靶标数据库T3DB,化合物-蛋白关系数据库Chemprot, STITCH, ChEMBL及BindingDB等。传统的动物体内实验数据库包括:水陆生毒性数据库ECOTOX,猴子及啮齿动物毒性数据库ToxRefDB及多物种体内外数据库SuperToxic^[27]等。本课题组近年来也发展了一些数据库,如含化合物毒性数据的admetSAR^[28]、药物及类药物副作用数据库MetaADEDB^[29]等。以上所述相关数据库详细信息,可见表S1。

2 研究方法

体外高通量筛选实验获得的大量数据,急需开发新的计算方法构建定量的预测模型,为计算系统毒理学的发展提供了契机。计算系统毒理学的关键是应用计算系统生物学的方法解决21世纪毒理学相关问题。如图2所示,这些技术包括把组学数据转化成可计算的静态网络分析、预测,涉及毒性化合物的动态网络模拟,以及有害结局路径的表征和模拟。

2.1 静态网络分析和预测

在计算领域,如何抽象、简化并用数学语言描述化合物对复杂生命系统的作用,是构建预测模型的基础。化合物与细胞的各个组分(如蛋白质、基因、RNA等)之间的相互作用,可以看成多个二部网络:网络中的点表示化合物或组分,点与点之间的连线表示相互作用。网络在数学上可以用邻接矩阵表示,

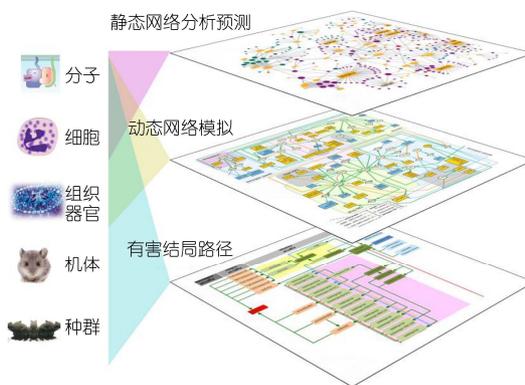


图2 计算系统毒理学研究方法

Figure 2 Computational systems toxicology methodology

因而网络的相关转换和计算,可以转化成矩阵相应的运算.化合物与生物组分相互作用网络的信息,主要来自毒理组学实验相关数据的整合.如Audouze和Grandjean^[30]从ChemProt, Interactome, OMIM和CTD数据库中挖掘整合得到农药二氯二苯基三氯乙烷(DDT)异构体和代谢产物的化合物-蛋白、蛋白-蛋白及蛋白-疾病相关信息,并构建了相互作用网络.构建的毒理机制网络,在一定程度上反应了化合物对机体不同组织、不同分子水平的影响,是生物网络中的一种,因此高质量的毒理组学网络应接近自然界生命网络的拓扑学性质.基本的网络拓扑学性质包括网络中节点的度、聚类系数及核心度^[31], Cytoscape^[32]和Pajek^[33]等软件可以实现网络的可视化及拓扑性质的计算.大部分生物网络具有“无标度”^[34]、“小世界”^[35]、“模块化”^[36]等特点.与其类似,毒理组学抽取的化合物相关网络,也都应具备这些性质.

静态网络的构建及网络拓扑性质的分析,是化合物与机体分子相互作用的知识汇总,为计算预测模型的构建奠定了基础.最基本的预测方法是基于多重网络的打分方法.其前提假设是毒性机制相关测试的阳性率越高则该化合物具有这种毒性的可能性越高.数学评价这种可能性的方法包括简单的计数^[37,38]和线性判别法^[39,40].Hoeng等人^[41]开发了“生物干扰因子”,用于评价化合物不同生物层面的暴露影响.另外一种预测模型是基于化合物的网络拓扑结构相似性,如本课题组^[42]前期开发的网络推理算法(network-based inference, NBI),基于边加权和点加权的网络推理算法首次发现了化合物相关网络中的“弱相互作用”^[43],该算法已成功用于化合物-蛋白^[42,43]、药物-副作用^[29]相互作用网络的预测,同时也构建了化合物-基因-疾病网络^[44]及环境因子-microRNA-疾病网络^[45]的计算系统毒理学框架.Audouze等人^[46]根据蛋白质网络拓扑相似性,预测毒性物质导致疾病的潜在机制蛋白.另外,随机游走也是化合物-靶标相互作用网络很好的预测算法^[47].

静态网络中化合物与生命体分子间的关系,相当于动态生命某一状态的快照,因此不能表征时间尺度.但网络囊括大部分化合物、大部分蛋白和通路信息,可以进行高通量分析.

2.2 动态网络模拟

生命活动是复杂的,仅研究组分及相互作用网

络获得的信息是有限的.系统对外界化学物质压力的反应和调节、细胞间信号转导和组织间物质交换等过程都是动态的.化合物对生命和环境的影响,也具有量效、时效效应.因此,需用数学方程模拟这些动态过程,以弥补静态网络的不足.动态网络的模拟,既包括把静态网络中的相互关系用数学公式表达动力学信息,也含对细胞运动等动态生命过程的描述.事实上,包括基因表达在内的生命过程只有2种状态,即呈二进制变化.因而化合物对机体影响的部分输出结果可简化为0/1的布尔逻辑,大大简化了运算复杂度.构建化合物干扰下动态的多尺度生物模型也是计算系统毒理学的重要研究手段.

动态网络模拟中最常用的方法是常微分方程法^[48].常微分方程法将物质相互作用生成产物的过程用微分方程式表征物质随时间的变化.但常微分方程法不能同时模拟固定容积内正常细胞的随机变化和细胞内的物质转运、扩散及其他不连续的状态变化.因此,引入随机参数对常微分方程法进行扩展,发展了随机微分方程法^[49];引入物质扩散、转运参数来同时模拟时间、空间的变化,发展了偏微分方程法.常微分、随机微分及偏微分方程法是系统生物学模拟中较为成熟的方法.为了能更好地拟合化学反应、信号传导、细胞器官状态等不同尺度的生命过程,一些新的模拟方法被开发,如用于模拟疾病分子机制的基于主体建模(agent-based modeling, ABM)方法^[50]、用于模拟组织、器官状态的元胞机模型(cellular automaton, CA)^[51]方法和用于模拟人体更高水平的生理药代动力学PBPK模型^[52].CellDesigner^[53]等软件包和在线服务为采用不同方法构建模拟模型提供了有效的工具和可视化平台.更多新的模拟算法可参见其他综述^[54,55].

目前,已有的模拟模型可在CellML^[17]和BioModels^[18]等知识库中查询.计算系统毒理学关心的不仅是正常的状态,而是毒理学或化合物暴露下的多尺度生命过程.由于计算系统毒理学的发展要滞后于计算系统生物学和系统毒理学,因此针对化合物暴露的数学模拟模型并不多.随着实验数据的增多和计算方法的完善,更多的毒理学动态模拟模型会被构建.动态网络的模拟相对于静态网络的分析及预测,只能研究相对较小的体系.一般只模拟单个化合物或类似物对机体的暴露影响.随着毒理学机制和关键通路的确证,未来可能用动态模拟的手

段实现定量化合物安全性评估。

2.3 有害结局路径

Adverse outcome pathways(AOP)中文译为“有害结局路径”^[56], 是近年来提出的用于表征化学品是如何通过诱导分子路径变化及其级联反应, 导致最终有害结局的研究框架^[57]。AOP一般只包括1个直接的分子起始事件, 1个最终的有害结局, 中间有细胞、器官、机体、群体等不同生物组织结构层次的多个级联反应步骤。分子起始事件一般是小分子化合物与机体基因、蛋白质等大分子的相互作用。AOP框架是对现有知识的概括和汇总, 不同层次间的关联信息整合了实验测试、计算预测甚至是因果推理的结果, 对化合物有害效应和健康风险评估具有指导意义。图3所示为美国国家环境保护局构建的血管发育毒性有

害结局路径图, 诠释了化合物是如何影响雌激素受体ER, 通过VEGF通路而产生细胞、组织、器官等水平的副作用影响。该AOP是Knudsen课题组^[58]通过分析ToxtCast 1期数据所得, 在胚胎发育毒性的模拟模型中解释了沙利度胺类似物5HPP-33在细胞、组织、器官水平的通路变化^[59]。

虽然目前已确证一些毒性的AOP^[60], 国际经济合作和发展组织(OECD)也出台了相关的标准(<http://www.oecd.org/chemicalsafety/testing/49963554.pdf>), 但是相关研究方法才刚刚起步。不同科研人员根据不同的经验和方法获得的AOP不尽相同, 且大部分都是定性的, 距离构建高质量的定量计算系统毒理学预测模型还有很大距离。相对于其他研究方法, AOP有望构建更高水平的预测模型, 如虚拟器官模型、种群预测模型等。美国国家环境保护局构建的血

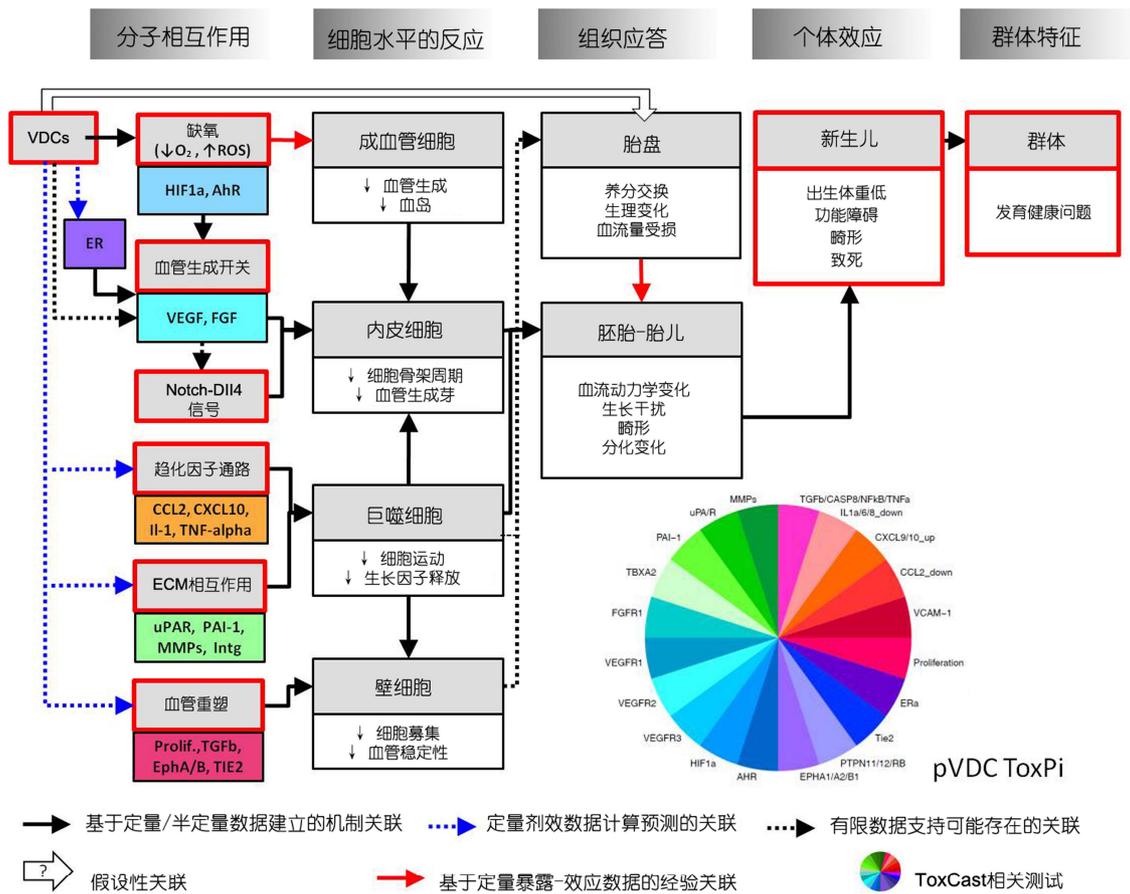


图3 胚胎血管干扰物(VDCs)的有害结局路径。本图编译自 Wiki 有害结局路径 AOP:43(<https://aopkb.org/aopwiki/index.php/Aop:43>)
 Figure 3 Adverse outcome pathways for embryonic vascular disruption (VDCs). Modified after Wiki adverse Outcome pathway (Aop:43) (<https://aopkb.org/aopwiki/index.php/Aop:43>)

管发育毒性AOP就是著名的虚拟胚胎工程的一部分。同时, AOP也可以表征重要步骤对种群差别的影响, 相关的方法开发和评价在环境毒理学中正在进行^[61]。随着计算系统生物学技术的发展, AOP会变得更加特异性、定量性和可预测性, 最终实现定量描述化合物通过分子起始事件产生最终毒性间的详细过程。

3 计算系统毒理学的应用

3.1 全面理解毒理学机制

“21世纪毒性测试”的关键就是对“毒性通路”的全面理解。低成本、高通量的体外测试方法, 全面测得化合物对细胞、组织、器官中基因、蛋白、代谢等不同水平的数据, 通过计算的手段可以找到关键的机制分型。毒理学的研究不再是粗糙的机体最高水平表型的评价, 而是把表型相关的毒性按不同机制进行划分, 从而进行更为精确的毒性综合评估。体外高通量实验, 直接采用人源细胞、细胞系或细胞组分进行实验, 因而对毒理学机制的分析计算, 是人体对化合物毒性的直接反应。

静态网络的分析预测、动态模拟及有害结局路径, 都为全面理解毒理学机制提供了很好的工具。Cheng等人^[44]用网络推理算法对化合物-基因-疾病网络进行预测, 分析了双酚A致病的基因机制网络, 并预测得到双酚A潜在的致病基因, 包括与炎症、神经变性相关的CASP3及肿瘤蛋白TP53等。Qi等人^[62]构建了多巴胺代谢的数学模拟模型, 发现农药导致帕金森病多巴胺神经毒性的不同机制: 鱼藤酮主要抑制转运体VMAT2, 影响多巴胺大部分通路; 百草枯主要抑制酪氨酸羟化酶的活性, 从而影响多巴胺及代谢产物的分布。有害结局路径更是从不同生物组织层次全面展现了毒理学机制, 包括定量实验测得、计算模型预测、经验或假设等信息。

对毒理学机制的全面理解, 弥补了传统动物实验的不足, 也在某些毒性上解决了传统计算毒理学预测准确率低的问题。目前, 化合物皮肤敏感性预测模型, 把不同的作用机制考虑在内, 已经达到可以接受的预测精准度。未来, 类似疾病分型知识库, 根据实验和计算分析得到的毒理学机制分型知识库也会出现。

3.2 发现新的生物标志物

生物标志物是指机体在受到严重损害之前, 体

内一些标志性分子表达水平、修饰上的变化, 或者细胞、体液、个体水平上其他可以定量测试的变化^[63]。这类标志性分子包括基因、蛋白质、代谢物等, 最近microRNA因其特性也被开发为生物标志物^[64]。生物标志物应具有敏感性、特异性、可测性和普遍性, 对于毒性化合物相关损伤的预报和相关疾病的预防具有重要意义。找到准确、易测、便宜又有价值的生物标志物并不容易。计算系统毒理学在多尺度, 对化合物暴露、毒性和疾病进行全面分析、预测和模拟, 可以发现新的重要作用机制来开发生物标志物。Li等人^[45]用网络推理算法预测了187个化合物可能的microRNA标志物, 如mir-155, mir-221, let-7a及mir-126可能作为吸烟引起肺癌、胃癌的标志物。网络打分可以粗略评估化合物对生物分子的影响, 从而发现新机制和潜在标志物。Hoeng等人^[65]综述了外源物代谢网络打分发现吸烟生物标志物的案例。ToxCast高通量筛选实验与毒性的相关性分析, 可获得毒性通路的标志性测试^[38], 测试相关的基因、蛋白等可能成为生物标志物。同时动态网络的模拟中毒性或疾病的关键输入、输出信号, 为标志物的开发提供了相应的研究基础。

3.3 化合物安全性综合评估

传统的计算毒理学研究方法主要是构建QSPR预测模型。分类模型可根据正负类毒性的信息空间, 构建数学模型预测未知化合物的毒性简单分类; 回归模型也仅可粗略预测单一毒性值。传统的计算毒理学方法, 已不能解决复杂机制毒性或慢性毒性的预测问题^[66]。计算系统毒理学通过高通量实验数据分析, 把复杂毒性按照机制进行描述和分类, 克服了传统计算毒理学的关键难题。计算系统毒理学最终的目标是构建多水平、多机制、多尺度的数学预测模型, 以定量描述毒性的关键机制、通路等信息, 实现化合物安全性综合评估。

致癌性、生殖毒性、发育毒性及肝毒性等, 都是传统计算毒理学的预测难点。计算系统毒理学的研究方法, 把化合物对机体多尺度的分子、通路机制考虑在内, 对复杂毒性的预测取得了突破性进展。Nicole等人^[37]对比ToxCast I期高通量筛选数据和ToxRefDB数据库, 分析得到致癌性通路的关键基因测试集, 通过对关键测试集实验结果的打分排序, 构建了啮齿动物致癌性预测模型, 准确地把33个化合物按照致

癌性分成4类。类似方法得到的胚胎血管发育干扰物预测模型,对大鼠发育毒性的预测AUC高达90%^[38]。Martin等人^[40]用线性判别法对关键通路相关测试进行打分,构建了大鼠生殖毒性预测模型,在外部验证集的准确率为76%。计算系统毒理学目前还在发展的初期,体外高通量筛选数据的分析方法、计算预测模型的构建方法等都还处于摸索阶段。相对于传统计算毒理学对化合物特征及顶端毒性的描述,计算系统毒理学已开始结合高通量组学数据定量描述化合物对毒性通路的影响。

对于某些复杂毒性,仅通过定量毒性通路进行化合物安全性评估是不够的。哈姆纳健康科学研究所已商业化的“药物引起肝毒性”预测软件DILTsymTM^[67]采用常微分方程法方法,定量模拟了肝毒性的分子、细胞、组织、器官、个体等多个尺度的重要步骤,包括谷胱甘肽的消耗和分解模拟、线粒体功能障碍毒性机制模型、肝细胞循环和凋亡模型、免疫反应模型、PBPK动力学模型及临床终点指标模型等。通过参照大鼠、小鼠和人体的体外高通量数据,对定量模型的关键参数进行近似,实现了体外实验结合计算系统毒理学模型对体内测得的临床终点指标的时效、量效准确推测。由此可见,计算系统毒理学多尺度模型可对生命过程进行多尺度定量描述,基于毒性相关的临床终点指标集,实现对化合物安全性的定量综合评价。

4 挑战及展望

21世纪的毒理学研究把体外高通量筛选数据与计算方法相结合,以期替代动物体内实验,把计算科

学放在了重要的位置。通过静态网络分析、预测,动态网络模拟和有害结局路径的理解,计算系统毒理学为化合物安全性评价及动物模型替代提供了强有力的手段。虽然计算系统毒理学具有广阔的应用前景,但是目前仍处在初级阶段,尚存在不少挑战,比如:(1)用于替代动物体内实验的体外高通量组学实验技术仍有待提高,如蛋白质在体内多为复合物,涉及蛋白的通路测量会产生很大误差^[68];(2)体外高通量筛选实验条件、方法的不同导致数据质量也参差不齐,如何将这些数据标准统一化,是保证建模质量的关键问题^[25];(3)建模方法有待完善以提高预测准确率,如综合评估化合物对生命复杂网络的影响^[41],考虑更多时毒、量毒关系以更接近机体真实状态,AOP框架中大量的关联信息还需定量数据确证。

展望未来,计算系统毒理学将吸取研究小体系网络的经验,动态模拟更高水平的网络体系,实现虚拟器官和虚拟群体的定量评估。随着对分子、细胞、组织、器官甚至群体等多尺度的定量模型方法的开发,最终会实现化合物安全性的定量综合评估。基于体外高通量实验的分析,毒性机制分型的相关知识数据库也会出现。计算方法开发与实验技术向来是相辅相成的,可靠的预测模型依赖于高质量的测试数据,计算预测结果又可反过来指导实验的设计。计算系统毒理学在化学、系统生物学中是关键领域,不应满足于把相关领域的研究方法应用于毒理学的研究,而应开发独特的计算系统毒理学方法,实现化合物安全性综合评估的同时,也为系统生物学等其他领域提供新的研究方法。

参考文献

- 1 Cheng F X, Li W H, Liu G X, et al. *In silico* ADMET prediction: Recent advances, current challenges and future trends. *Curr Top Med Chem*, 2013, 13: 1273–1289
- 2 Kitano H. Systems biology: A brief overview. *Science*, 2002, 295: 1662–1664
- 3 Yildirim M A, Goh K I, Cusick M E, et al. Drug-target network. *Nat Biotechnol*, 2007, 25: 1119–1126
- 4 Looijen R C. Holism and Reductionism in Biology and Ecology: The Mutual Dependence of Higher and Lower Level Research Programmes. Dordrecht: Kluwer Academic Publishers, 2000
- 5 Waters M D, Fostel J M. Toxicogenomics and systems toxicology: Aims and prospects. *Nat Rev Genet*, 2004, 5: 936–948
- 6 Plant N J. An introduction to systems toxicology. *Toxicol Res*, 2015, 4: 9–22
- 7 Hartung T, van Vliet E, Jaworska J, et al. Systems toxicology. *Altex*, 2012, 29: 119–128
- 8 Krewski D, Acosta Jr D, Andersen M, et al. Toxicity testing in the 21st century: A vision and a strategy. *J Toxicol Env Health Pt B Crit Rev*, 2010, 13: 51–138
- 9 Krewski D, Westphal M, Al-Zoughool M, et al. New directions in toxicity testing. *Annu Rev Public Health*, 2011, 32: 161–178

- 10 Dix D J, Houck K A, Martin M T, et al. The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicol Sci*, 2007, 95: 5–12
- 11 Andersen M E, Krewski D. Toxicity testing in the 21st century: Bringing the vision to life. *Toxicol Sci*, 2009, 107: 324–330
- 12 Pastoor T P, Bachman A N, Bell D R, et al. A 21st century roadmap for human health risk assessment. *Crit Rev Toxicol*, 2014, 44: 1–5
- 13 Lioy P J, Smith K R. A discussion of exposure science in the 21st century: A vision and a strategy. *Environ Health Perspect*, 2013, 121: 405
- 14 Sturla S J, Boobis A R, FitzGerald R E, et al. Systems toxicology: From basic research to risk assessment. *Chem Res Toxicol*, 2014, 27: 314–329
- 15 Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 2000, 28: 27–30
- 16 Pico A R, Kelder T, van Iersel M P, et al. WikiPathways: Pathway editing for the people. *PLoS Biol*, 2008, 6: e184
- 17 Lloyd C M, Lawson J R, Hunter P J, et al. The CellML model repository. *Bioinformatics*, 2008, 24: 2122–2123
- 18 Le Novère N, Bornstein B, Broicher A, et al. BioModels Database: A free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*, 2006, 34: 689–691
- 19 Bader G D, Cary M P, Sander C. Pathguide: A pathway resource list. *Nucleic Acids Res*, 2006, 34: 504–506
- 20 Knudsen T B, Houck K A, Sipes N S, et al. Activity profiles of 309 ToxCast™ chemicals evaluated across 292 biochemical targets. *Toxicology*, 2011, 282: 1–15
- 21 Kavlock R, Chandler K, Houck K, et al. Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chem Res Toxicol*, 2012, 25: 1287–1302
- 22 Schmidt C W. TOX 21: New dimensions of toxicity testing. *Environ Health Perspect*, 2009, 117: 348–353
- 23 Tice R R, Austin C P, Kavlock R J, et al. Improving the human hazard characterization of chemicals: A Tox21 update. *Environ Health Perspect*, 2013, 121: 756–765
- 24 Wang Y, Suzek T, Zhang J, et al. PubChem BioAssay: 2014 update. *Nucleic Acids Res*, 2014, 42: 1075–1082
- 25 Judson R S, Martin M T, Egeghy P, et al. Aggregating data for computational toxicology applications: The US environmental protection agency (EPA) aggregated computational toxicology resource (ACToR) system. *Int J Mol Sci*, 2012, 13: 1805–1831
- 26 Waters M, Stasiewicz S, Merrick B A, et al. CEBS-chemical effects in biological systems: A public data repository integrating study design and toxicity data with microarray and proteomics data. *Nucleic Acids Res*, 2008, 36: 892–900
- 27 Schmidt U, Struck S, Gruening B, et al. SuperToxic: A comprehensive database of toxic compounds. *Nucleic Acids Res*, 2009, 37: 295–299
- 28 Cheng F X, Li W H, Zhou Y D, et al. AdmetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties. *J Chem Inf Model*, 2012, 52: 3099–3105
- 29 Cheng F X, Li W H, Wang X C, et al. Adverse drug events: Database construction and in silico prediction. *J Chem Inf Model*, 2013, 53: 744–752
- 30 Audouze K, Grandjean P. Application of computational systems biology to explore environmental toxicity hazards. *Environ Health Perspect*, 2011, 119: 1754–1759
- 31 Seebacher J, Gavin A C. SnapShot: Proteinprotein interaction networks. *Cell*, 2011, 144: 1000
- 32 Shannon P, Markiel A, Ozier O, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res*, 2003, 13: 2498–2504
- 33 Batagelj V, Mrvar A. Pajek-program for large network analysis. *Connections*, 1998, 21: 47–57
- 34 Barabási A L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286: 509–512
- 35 Watts D J, Strogatz S H. Collective dynamics of “small-world” networks. *Nature*, 1998, 393: 440–442
- 36 Milo R, Shen-Orr S, Itzkovitz S, et al. Network motifs: Simple building blocks of complex networks. *Science*, 2002, 298: 824–827
- 37 Kleinstreuer N C, Dix D J, Houck K A, et al. *In vitro* perturbations of targets in cancer hallmark processes predict rodent chemical carcinogenesis. *Toxicol Sci*, 2013, 131: 40–55
- 38 Kleinstreuer N C, Judson R S, Reif D M, et al. Environmental impact on vascular development predicted by high-throughput screening. *Environ Health Perspect*, 2011, 119: 1596–1603
- 39 Sipes N S, Martin M T, Reif D M, et al. Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicol Sci*, 2011, 124: 109–127
- 40 Martin M T, Knudsen T B, Reif D M, et al. Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biol Reprod*, 2011, 85: 327–339
- 41 Hoeng J, Deehan R, Pratt D, et al. A network-based approach to quantifying the impact of biologically active substances. *Drug Discov Today*, 2012, 17: 413–418

- 42 Cheng F X, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol*, 2012, 8: e1002503
- 43 Cheng F X, Zhou Y D, Li W H, et al. Prediction of chemical-protein interactions network with weighted network-based inference method. *PLoS One*, 2012, 7: e41064
- 44 Cheng F X, Li W H, Zhou Y D, et al. Prediction of human genes and diseases targeted by xenobiotics using predictive toxicogenic-derived models (PTDMs). *Mol Biosyst*, 2013, 9: 1316–1325
- 45 Li J, Wu Z R, Cheng F X, et al. Computational prediction of microRNA networks incorporating environmental toxicity and disease etiology. *Sci Rep*, 2014, 4: 5576
- 46 Audouze K, Juncker A S, Roque F J, et al. Deciphering diseases and biological targets for environmental chemicals using toxicogenomics networks. *PLoS Comput Biol*, 2010, 6: e1000788
- 47 Chen X, Liu M X, Yan G Y. Drug-target interaction prediction by random walk on the heterogeneous network. *Mol Biosyst*, 2012, 8: 1970–1978
- 48 Jones D S, Plank M, Sleeman B D. *Differential Equations and Mathematical Biology*. London: CRC Press, 2011
- 49 Meng T C, Somani S, Dhar P. Modeling and simulation of biological systems with stochasticity. *In Silico Biol*, 2004, 4: 293–309
- 50 Macal C M, North M J. Tutorial on agent-based modelling and simulation. *J Simul*, 2010, 4: 151–162
- 51 Moreira J, Deutsch A. Cellular automaton models of tumor development: A critical review. *Adv Complex Syst*, 2002, 5: 247–267
- 52 Nestorov I. Whole body pharmacokinetic models. *Clin Pharmacokinet*, 2003, 42: 883–908
- 53 Funahashi A, Matsuoka Y, Jouraku A, et al. CellDesigner 3.5: A versatile modeling tool for biochemical networks. *Proc IEEE*, 2008, 96: 1254–1265
- 54 Hoban S, Bertorelle G, Gaggiotti O E. Computer simulations: Tools for population and evolutionary genetics. *Nat Rev Genet*, 2012, 13: 110–122
- 55 Dada J O, Mendes P. Multi-scale modelling and simulation in systems biology. *Integr Biol*, 2011, 3: 86–96
- 56 Peng S Q. The definition of AOP in risk assessment and considering in its Chinese translation (in Chinese). *Chin J Pharmacol Toxicol*, 27(Suppl 1): 305 [彭双清. 危险度评定中 AOP 的概念以及关于其中文译名的思考. *中国药理学与毒理学杂志*, 2013, 27(增刊): 305]
- 57 Ankley G T, Bennett R S, Erickson R J, et al. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem*, 2010, 29: 730–741
- 58 Knudsen T B, Kleinstreuer N C. Disruption of embryonic vascular development in predictive toxicology. *Birth Defects Res C*, 2011, 93: 312–323
- 59 Kleinstreuer N, Dix D, Rountree M, et al. A computational model predicting disruption of blood vessel development. *PLoS Comput Biol*, 2013, 9: e1002996
- 60 Vinken M. The adverse outcome pathway concept: A pragmatic tool in toxicology. *Toxicology*, 2013, 312: 158–165
- 61 Kramer V J, Etterson M A, Hecker M, et al. Adverse outcome pathways and ecological risk assessment: Bridging to population-level effects. *Environ Toxicol Chem*, 2011, 30: 64–76
- 62 Qi Z, Miller G W, Voit E O. Rotenone and paraquat perturb dopamine metabolism: A computational analysis of pesticide toxicity. *Toxicology*, 2014, 315: 92–101
- 63 Silins I, Högberg J. Combined toxic exposures and human health: Biomarkers of exposure and effect. *Inter J Env Res Public Health*, 2011, 8: 629–647
- 64 Schöler N, Langer C, Döhner H, et al. Serum microRNAs as a novel class of biomarkers: A comprehensive review of the literature. *Exp Hematol*, 2010, 38: 1126–1130
- 65 Hoeng J, Talikka M, Martin F, et al. Case study: The role of mechanistic network models in systems toxicology. *Drug Discov Today*, 2014, 19: 183–192
- 66 Allen T E H, Goodman J, Gutsell S, et al. Defining molecular initiating events in the adverse outcome pathway framework for risk assessment. *Chem Res Toxicol*, 2014, 27: 2100–2112
- 67 Bhattacharya S, Shoda L K, Zhang Q, et al. Modeling drug- and chemical-induced hepatotoxicity with systems biology approaches. *Front Physiol*, 2012, 3: 462
- 68 Basketter D A, Clewell H, Kimber I, et al. A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing. t4 report. *Altex*, 2011, 29: 3–91

Computational systems toxicology: Emergence, development and application

LI Jie, LI KeJia, ZHANG Chen, PENG QiuPeng & TANG Yun

Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China

With the development of systems biology, there are new opportunities for the transformation of classical toxicology. Computational systems toxicology aims at building multi-level and multi-scale predictive models to quantitatively assess chemical safety, combining toxicogenomic experimental data. Many methods have been developed for computational systems toxicology; examples are methods that employ static network analysis and prediction, dynamic network simulation and adverse outcome pathways. Although in its early stage, computational systems toxicology has been applied to the overall understanding mechanism of toxicology, allowing the discovery of new biomarkers and the comprehensive assessment of chemical safety. This review mainly focuses on related data sources, the research status, applications, challenges and perspectives.

systems toxicology, omics, database, computational methods

doi: 10.1360/N972014-01400

补充材料

表 S1 计算系统毒理学有关开源数据库汇总

本文以上补充材料见网络版 csb.scichina.com. 补充材料为作者提供的原始数据, 作者对其学术质量和内容负责.