



基于组学的污染物高通量毒理学筛查与风险评估

闫路¹, 苟潇¹, 夏普¹, 高瑞泽¹, 吉慧敏¹, 史薇^{1,2}, 于红霞¹, 张效伟^{1,2*}

1. 南京大学环境学院, 污染控制与资源化研究国家重点实验室, 南京 210023;

2. 江苏省生态环境保护化学品安全与健康风险研究重点实验室, 南京 210023

* 联系人, E-mail: zhangxw@nju.edu.cn

2022-06-20 收稿, 2022-08-13 修回, 2022-09-15 接受, 2022-10-13 网络版发表

国家重点研发计划(2021YFC3201003)资助

摘要 从大量的环境污染物中筛查出具有高毒性、高风险的化学污染物是环境风险防控的前提和基础。有毒化学污染物经环境暴露进入生命体系后, 可通过不同的分子靶标和生物学通路诱导有害健康结局。现有高通量毒性筛查方法主要依赖少数且孤立的分子靶点, 难以捕获致毒的关键生物学过程, 无法确定化学污染物与疾病之间的直接关系。毒理基因组学利用组学技术所具有的全基因组覆盖、通量高、支持大数据分析等优势, 实现了化学污染物高通量筛查和风险评估方法的革新。本文通过具体研究案例分析, 从4个方面综述了毒理基因组学在化学污染物高通量毒理学筛查与风险评估方法学研究上的技术优势和最新进展。在分子事件上, 毒理基因组学测试覆盖的分子靶点数目多, 且靶点之间的关系可由通路连接。在生物学通路方面, 单一组学、多组学联合以及有害结局路径从多生物学水平上为化学污染物筛查与健康风险评估提供机制信息。在人群疾病方面, 功能基因组学为探究有毒化学污染物暴露下个体疾病风险和易感性的机制提供了新的手段。最后, 针对复合污染, 将组学技术与化学分析结合有助于精准地识别环境介质中构成生态与健康风险的关键化学物质。综上, 未来基于组学的污染物高通量毒理学筛查技术的发展可有力支持传统和新污染物的精准环境风险评估。

关键词 毒理基因组学, 化学污染物, 分子靶点, 生物通路, 人群疾病, 复合污染

环境有毒污染物威胁着人体健康与生物多样性, 筛查与识别出有毒污染物的健康与生态风险是保护人民群众生命健康和生态安全的基础。全球疾病负担(Global Burden of Disease, GBD)通过对195个国家与地区共计80多个疾病负担风险因素的计算发现, 环境因素为全球约60%的死亡负担负责^[1,2]。根据暴露-疾病关联, 每年有900万例死亡(占全球所有死亡人数的16%)归因于空气、水和土壤污染^[3]。由化学污染引发的人类健康威胁通常伴随着生态系统的损害和生物多样性的下降^[4]。欧洲联盟(简称欧盟)最近的两项政策举措, 即“2030年欧盟生物多样性战略”和“零污染行动计划”, 明确承认“污染”是“生物多样性丧失的重要驱动

因素”之一^[5]。因此, 开展化学污染物毒性筛查与风险评估为应对有毒污染物的威胁至关重要。目前, 该工作存在一些方法学的瓶颈问题, 亟待解决。

传统的化学污染物毒性筛查与风险评估方法在从源头到有害健康结局路径上存在4个方面的技术瓶颈问题(图1)。从源头到有害健康结局路径(from source to outcome pathway, STO)描述了有害污染物在不同生物组织水平上引起的生物效应^[6]。首先, 在评估有毒污染物引起生物体(动物或人)内异常生物学信号方面, 传统的靶向测试方案可对化学物质按照毒性终点有针对性地逐一进行毒性测试。美国环境保护署ToxCast计划选取了近500个基因的分子事件作为毒性测试终点^[7]。但

引用格式: 闫路, 苟潇, 夏普, 等. 基于组学的污染物高通量毒理学筛查与风险评估. 科学通报, 2022, 67: 4159–4169

Yan L, Gou X, Xia P, et al. Omics-based high throughput toxicity screening and risk assessment of pollutants (in Chinese). Chin Sci Bull, 2022, 67: 4159–4169, doi: 10.1360/TB-2022-0682

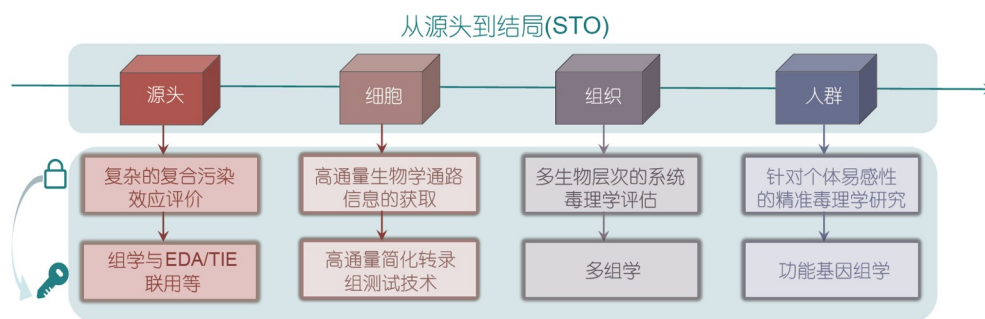


图1 (网络版彩色)化学污染物毒性筛查从源头到有害健康结局面临的四大瓶颈问题。EDA, 效应导向分析; TIE, 毒性鉴别评估

Figure 1 (Color online) Four bottleneck problems for chemical pollutants toxicity screening in the pathway from source to adverse outcomes. EDA: Effect-directed analysis; TIE: Toxicity identification and evaluation

是, 人类基因组已知包含了21306个蛋白质编码基因和21856个非编码基因, 目前靶向测试的分子靶点覆盖的潜在基因靶标仍十分有限, 如何高通量获取污染物的生物学信息需要得到解决。其次, 有毒污染物导致有害健康结局往往是一个复杂的生物过程, 涉及基因、蛋白、代谢等多个生物学水平的扰动。当前以单个分子靶点为检测对象的靶向测试方法(如报告基因、酶活性), 难以捕获有毒污染物所引发的复杂生物学过程, 难以全面评估有毒污染物的生物效应。再次, 个体间基因组差异, 如单核苷酸多态性或表观遗传差异, 可能会影响一个人患病的易感倾向。有毒污染物胁迫对易感性人群形成更大的健康风险, 因此掌握有毒污染物毒性效应的易感性机制才能开展精准的人类健康风险评估^[8,9]。最后, 实际环境污染是有毒污染物复合污染导致的结果, 由于分析技术的限制, 只有极少部分有毒污染物被识别。而一些污染物即使其存在低于其自身的效应阈值和/或分析检测限, 通过相互间作用可能产生毒性, 目前还没有完善的技术手段能够解决复合污染问题^[10]。

毒理基因组学整合基因组学、生物信息学和毒理学的理论与技术, 通过测试不同暴露剂量和时间条件下生命基因组对污染物的反应动态, 已经成为研究环境污染对人体健康影响的重要工具。基因组学具有全基因组覆盖、通量高、支持大数据分析等优势, 为开发创新性的污染物毒性测试与筛查方法提供了历史性机遇。组学技术始于20世纪90年代中期的微阵列技术, 通过微阵列, 可以同时监测数万个基因的表达水平, 从而可以评估由不同化合物诱导或与不同生理条件相关的基因表达谱的改变, 这是传统毒理学测试技术无法实现的^[11]。目前被应用于环境领域更广泛的毒

理基因组学技术主要涉及基因组(genome)、转录组(transcriptome)、蛋白组(proteome)、代谢组(metabolome), 它们能够识别生物体在有毒污染物暴露下从分子到群落多生物学水平的响应变化, 为污染物的风险评估提供了至关重要的基因组学特征^[12]。许多研究已经证实了毒理基因组学的价值^[13,14], 例如, 通过使用毒理基因组学数据构建的内分泌干扰物预测模型, 其测试准确率大于90%, 比基于体外高通量测试数据的预测模型准确率(69%~85%)高^[15,16]。总之, 毒理组学技术有助于有毒污染物筛查与风险评估方法的开发。技术优势包括: (1) 提高测试分子目标的覆盖度; (2) 提供有毒污染物生物效应的机制新知识; (3) 研究目前被视为难以突破的人群易感性及复合污染问题^[6]。

本文将以前污染物毒理学筛查与风险评估遇到的瓶颈问题为出发点, 系统综述高通量组学方法应对问题所开展的研究。包括4个方面的内容: 高通量生物学通路机制信息的获取、多生物层级的系统毒理学评估、针对个体易感性的精准毒理学研究以及复合污染效应评价。

1 高通量生物学通路机制信息的获取

基于组学的生物测试大幅度增加了受试靶点数量。据估计, 人类基因中有3000~5000个可以作为化学物质的分子靶标。传统的毒性方法可通过检测这些分子靶标的激活(蛋白受体结合、酶活性、信号转导等)情况来预测污染物的毒性。以高通量测试计划ToxCast为例, 该项目(始于2007年)涵盖了700余个细胞及以下生物水平的靶点与300多个信号通路^[7]。毒理组学测试通过批量检测受污染物不同暴露剂量(剂量依赖)暴露后生物体内基因、蛋白、代谢物的变化情况来研究污染物的

潜在毒性。测试涉及的基因、蛋白或者代谢物数量可以包括几乎所有该生物体内的基因、蛋白以及代谢物。以人源细胞为受试体的转录组学测试为例,可以实现对近2万个基因的转录本进行测试及分析。Li等人^[17]基于转录组研究双酚A(bisphenol A, BPA)和壬基酚(nonyl phenol, NP)暴露下雌激素(estrogen receptor, ER)促进人类子宫平滑肌瘤细胞增殖的潜在分子靶点和信号通路,试验结果表明, BPA和NP可以引起739个基因表达异常,且与ER信号通路有关,说明酚类环境雌激素可能通过ER信号通路促进子宫平滑肌瘤细胞增殖和细胞周期发展,导致细胞周期调控紊乱。除此之外,基因组、转录组、蛋白组和代谢组可以从不同功能水平增加了数据分析的维度。Wang等人^[18]使用转录组学、蛋白质组学和代谢组学分析了磷酸三苯酯(triphenyl phosphate, TPP)引发的肝毒机制,揭示了TPP肝毒性对不同生物功能的影响。

剂量效应依赖的简化转录组学测试实现了高通量定量分析污染物干扰生物学通路信息。污染物对基因表达的影响,往往表现在生物学通路受到干扰,生物学通路受干扰情况可通过对少量能够代表整个生物学通路的关键基因进行检测^[19]。基于这一原理,简化转录组测试可以获取大量样品转录组表达剂量效应的关系。美国Tox21计划创建了一个TempO-Seq(BioSpyder Technologies, Inc., Carlsbad, CA)平台来测量人类S1500+基因集的基因表达,随后将人类S1500+基因集扩展到大鼠、小鼠以及斑马鱼测试体系中^[20,21]。利用该测试平台,有学者分析并比较了全氟或多氟烷基化合物(per- and polyfluoroalkyl substances, PFAS)在人肝微组织中的生物效力^[22,23]。L1000平台是美国国立卫生研究院NIH的基因组简化实验方案,大幅度加快了药物开发的研发周期。加拿大提出EcoToxChip项目,通过设计EcoTox芯片(包含大约370个基因集)完成高通量的化学品优先排序工作,提高了测试效率并减少了动物的使用量^[24]。在简化转录组测试方面,我国也做了大量的研究。Zhang和Zhao^[25]开发了一种简化转录组图谱的方法,这些基因代表了斑马鱼的神经发生和早期神经元发育,以确定环境化学物质的潜在神经发育毒性。Xia等人^[26]通过构建基于人源细胞系的剂量效应转录组测试方法,识别到与之前103种靶向体外生物测试相似的生物活性图谱,大幅提高了有毒污染物的生物测试通量和效率。Wang等人^[27]结合剂量效应简化斑马鱼转录组(concentration-dependent reduced zebrafish transcrip-

tome, CRZT)和基于表型的高内涵筛选(phenotype-based high content screen, PHCS)评估了5种毒作用模式(mode of action, MoA)的12种有毒污染物的斑马鱼发育毒性。与PHCS比较, CRZT捕捉了生命早期接触有毒污染物时的综合生物活性,提供了一种比PHCS敏感1~3个数量级的识别生物效应方法,并且可以有效地对不同MoAs的有毒污染物分类,显著提高有毒污染物发育毒性的检测能力。

2 多生物层次的系统毒理学评估

组学测试数据可通过联合毒理学通路为评估有毒污染物的环境风险提供新策略。有害结局路径框架(adverse outcome pathway, AOP)梳理了污染物暴露从分子启动事件到不同生物水平上的毒性信息,可以实现组织、个体水平上的毒性预测。AOP-wiki(有害结局路径数据库, <https://aopwiki.org/>),将有毒污染物暴露引起的分子起始事件(molecular initiating event, MIE)、关键事件(key event, KE)以及最终导致的有害结局(adverse outcome, AO)进行了归纳整理。根据AOP-wiki的记录,目前该数据库包含364条AOP路径,涉及363条分子启动事件(芳香烃受体的激活、DNA烷基化等)、1666条关键事件(细胞死亡、DNA修复受损等)、436条有害结局(繁殖能力受损、种群数量下降等)。基于分子靶标的组学测试可以提供化学污染物的机制信息,这对于预测有毒污染物的毒性(有害结局)、减少对动物试验的依赖至关重要^[28]。随着机制信息数据的增加,如何将AOP数据库中的AOP路径与高通量测试相结合,并进行毒性预测是一个亟须解决的问题^[29]。Martens等人^[30]探索了Wiki-Pathways工具,使组学数据与AOP-Wiki实现了关联,手动匹配关键事件表明67%的关键事件可能与分子途径相关联,但由于AOP-Wiki中目前使用的基因本体尚未能直接与Wiki-Pathways分子本体连接,因此需要手动进行匹配。

多组学技术可从不同生物层次识别与有害结局相关的生物学通路。建立有害结局路径中关键事件与组学测试结果之间的联系是解析有害污染物毒性效应的关键。毒理组学数据的解析离不开生物学数据库的支持,在有毒污染物组学测试中常用的生物学数据库包括GO(gene ontology, <http://geneontology.org/>)数据库,描述生物学中的分子功能、生物过程和细胞成分;KEGG(京都基因与基因组百科全书, <https://www.kegg.jp/kegg/>)数据库,包含核酸分子、蛋白质序列、基因表

达、基因组图谱、代谢通路图等; Hallmark特征基因集 (<http://www.gsea-msigdb.org/gsea/msigdb>), 由50组常用的特征基因组成, 包括细胞凋亡、雌激素效应、信号通路、脂肪酸代谢等. Li等人^[31]使用代谢组和转录组, 并通过KEGG数据库与GO数据库的解析发现, 手性杀虫剂咪唑莫克斯(imazamox, R-IM)干扰了浮萍在光合作用过程中碳固定、谷胱甘肽代谢、磷酸戊糖代谢、玉米蛋白生物合成以及卟啉和叶绿素代谢过程的基因表达. S-IM则主要影响苯丙氨酸代谢、苯丙类化合物生物合成、玉米蛋白生物合成和二次代谢物生物合成. 毒理组学测试数据结果可以在不同生物组织水平和不同分类群中为污染物的毒性筛查提供支持性证据. 在AOP框架下, 组学数据还可以匹配到相应的分子启动事件和关键事件. 以乙酰胆碱酯酶抑制为例, 转录组学数据支持建立以其为分子起始事件的有害结局路径网络. 相关的关键事件包括神经突触中乙酰胆碱的积聚以及肌肉连接内不受控制的兴奋^[32]. 暴露于毒死蜱与二嗪农的秀丽隐杆线虫的微阵列数据揭示, 扰动离子载体活性和脂质代谢是两种农药引发秀丽隐杆线虫死亡的关键生物事件^[33], 该生物机制具有跨物种性质, 在小鼠与斑马鱼的相关研究中也得到了证实^[34,35]. Xia等人^[36]构建了基于AOP网络的短链氯化石蜡(short-chain chlorinated paraffins, SCCPs)浓度依赖性转录组学评估工作流程, 通过人类HepG2细胞和斑马鱼胚胎之间的跨模型比较, 发现模型特异性反应, 例如异种生物代谢, 可能是由于HepG2细胞和斑马鱼胚胎之间存在不同的核受体介导途径. 总之, 高通量组学测试通过识别

与挖掘AOP网络中各生物水平的关键事件, 使全面地构建有毒污染物的机制网络成为可能^[37,38].

从单一组学到多组学联用, 再到有害结局路径网络, 生物学过程复杂度被逐渐提升(图2). 单一组学测试可以通过剂量-效应关系定性和定量地分析有毒污染物扰动下生物学通路的响应. 高通量转录组学与基准浓度模型(benchmark concentration, BMC)的结合可用于识别有毒污染物导致分子变化的阈值^[39], 并提供定量的毒物基因组学信息^[40]. 基于此方法, PFAS类和SCCPs类物质分别在人肝细胞球状模型和斑马鱼胚胎模型下的生物效力得到了评估, 结果发现, 长链PFAS和SCCPs更有可能诱导基因表达的变化, 并且具有较低的转录效力^[23,41]. 多组学联用可以实现多尺度、综合性的生物网络分析. 细胞的调节途径涉及一系列不同的生物分子, 这些分子包含不同的物理化学性质, 并表现出复杂的非线性相互作用, 单组学技术测量特定类型的生物分子, 例如, 转录组学测定的是核糖核酸(RNA), 多组学可以检测到有毒污染物暴露扰动的更多生物信息(RNA和蛋白质等)^[42]. 目前已有一些针对毒理学问题的多组学研究, 特别是双组学研究. 例如, Lee等人^[43]通过蛋白质组学和代谢组学的组合分析发现, 全氟辛烷磺酸暴露对斑马鱼幼鱼神经系统的影响主要是由于轴突变形、神经炎症刺激和钙离子信号传导失调导致的, 而单组学仅能显示, 全氟辛烷磺酸暴露在神经功能、氧化应激和能量代谢的途径中引起扰动. 有害结局路径网络将有毒污染物从分子到个体及以上生物学水平引发的有害结局联系起来, 有助于更加系统地

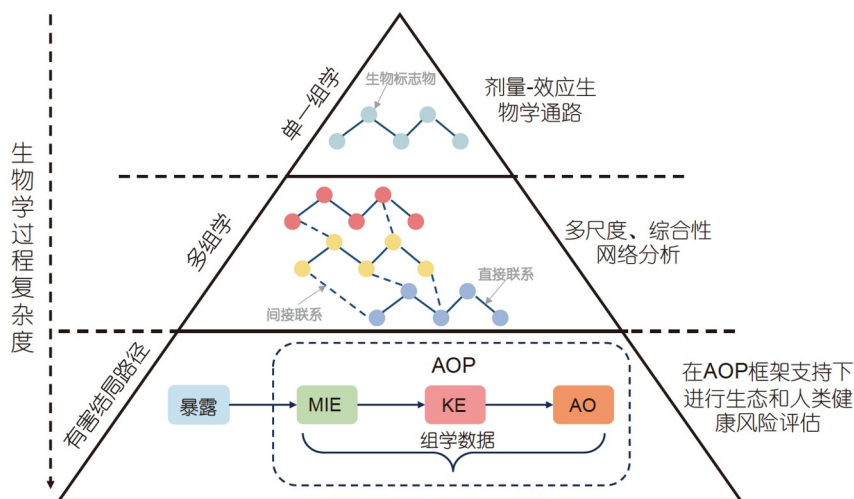


图2 (网络版彩色)单一组学到多组学联用识别化学污染物引起的复杂生物学过程与通路

Figure 2 (Color online) Identification of complex biological processes and pathways induced by chemical pollutants using single/multi-omics

识别有毒污染物的生物效应机制并基于此开展生态和人类健康风险评估。研究提出了针对人类女性生殖毒性的AOP网络,该网络由15个单独的AOP构建,绘制了邻苯二甲酸二酯诱导的人类女性生殖毒性的3种不同途径。第1条途径与雌二醇产生有关,第2条途径与细胞凋亡有关,第3条途径包括属于“卵巢雌二醇减少”和“细胞凋亡”的关键事件,这为更好地理解人类女性生殖功能障碍扰动背后的机制复杂性提供了基础^[44,45]。

3 针对个体易感性的精准毒理学研究

由于人群存在易感性差异,易感人群面临着更大有毒污染物引发的健康风险。易感性是指由遗传因素决定的个体患病风险。例如,存在于人类基因组中的变异序列可能会改变一个人对心脏病和癌症等复杂疾病的倾向,具有这类疾病倾向的人群称为易感人群。单核苷酸多态性数据库(The Single Nucleotide Polymorphism Database)是人类遗传变异的权威且可信的中央存储库(<https://ncbi.nlm.nih.gov/snp/>),据统计,该数据库收录了超过5亿条单核苷酸多态性(single nucleotide polymorphism, SNP)。掌握化学物质毒性效应的遗传易感性机制是开展精准健康风险评估的前提。有研究通过暴露组学和内部多组学之间的潜在动态相互作用揭示了在高度多样化和动态的个人暴露中,农用化学品和真菌占主导地位,提出了依据个人基因组信息进行精准健康评估的设想^[46]。针对个体易感性进行精准毒理学研究是当前遇到的瓶颈问题。

全基因组-环境关联分析研究在基因组学研究不断发展的情况下,被用于鉴定与复杂性状相关联的遗传变异,可建立遗传变异与疾病之间的关联。全基因组关联研究(genome-wide association study, GWAS)是利用全基因组范围内筛选出的高密度分子标记对所研究的群体进行扫描,分析扫描所得的分子标记数据与表型性状之间关联关系的方法。Lind等人^[47]采用GWAS方法,对1016名老年人血液中的甲基化位点进行检测,发现CYP2B6基因的遗传变异程度与血液循环中的滴滴涕(dichlorodiphenyltrichloroethane, DDT)代谢产物(dichlorodiphenyldichloroethylene, *p,p'*-DDE)有关,并阐述了该遗传变异介导的心血管疾病机制。Traglia等人^[48]的一项孕妇和胎儿遗传学研究阐明了多溴联苯醚(polybrominated diphenyl ethers, PBDE)和多氯联苯(polychlorinated biphenyl, PCB)暴露影响母体和胎儿代谢功能障碍的遗传分子机制,全基因组关联分析确定

了与异物和脂质代谢相关的基因CYP2B6、PRKCDBP、SUMF1和NDUFS4是有机卤素暴露下母体和胎儿患代谢性疾病(肥胖、2型糖尿病)的遗传决定性因素。为了探究多环芳烃(polycyclic aromatic hydrocarbons, PAHs)造成的遗传损伤与SNPs的关系,Dai等人^[49]利用GWAS研究了1557名高PAHs暴露人群(焦炉工人)血液的基因变异情况,阐明了与DNA损伤修复相关基因ANRIL的SNPs可能是PAHs促进肺癌发生的关键事件。Zeng等人^[50]为了研究呼吸系统症状(咳嗽、呼吸困难、咯痰)的不同职业暴露(粉尘、气体、烟雾、农药和溶剂)与SNPs之间的相互作用关系,通过基因分型和Meta分析确定了TMPRSS9、SERPINH1、TOX3和ARHGAP18为呼吸系统疾病的易感基因。

功能基因组学技术通过建立有毒污染物与基因、有害结局的直接关联,揭示污染物有毒效应的分子机制,并与分子流行病学相结合识别有毒污染物暴露的疾病易感性差异影响因素(图3)。Huang等人^[51]使用CRISPR/Cas9基因敲除技术小鼠研究了CPRC5A基因与苯并[a]芘(Benzo[a]pyrene, BaP)暴露下引发的特发性肺纤维化(idiopathic pulmonary fibrosis, IPF)存在潜在联系,发现CPRC5A低表达引发的维生素代谢失衡和有丝分裂融合2(mitofusion2, Mfn2)调节的动力学紊乱是导致BaP引发肺纤维化的重要机制,表明GPCR5A基因可作为污染物损伤和肺纤维化的新治疗靶点。为了明晰BaP引发神经毒性的机制,Yang等人^[52]利用CRISPR/Cas9质粒转染的PC-12细胞发现了基因SNAP-25的显著下调能够降低多巴胺受体表达,验证了BaP可通过促进多巴胺水平上升而引发大鼠大脑皮层的超微结构病理变化。Tian等人^[53]开发了一种人体细胞CRISPR筛查与流行病学分析相结合的方法,通过对人肺癌细胞A549进行CRISPR筛选,将HTR4基因的过表达与慢性阻塞性肺病的发生联系起来。

4 复合污染

在不断变化的环境中追踪复杂的有毒污染物。在实际环境中,我们接触到的是化学物质混合物,而其中只有一小部分被识别出来,大量化学物质是未知的^[10]。结合生物分析的污染物毒性筛查可用于识别复合污染中的风险驱动因素,并分离可能被基质效应(由于样品提取物不纯导致的)掩盖的生物活性化学物质^[54]。一项关于化学物质污染负担的研究报告指出,研究检测到的290种化学物质仅解释了生物效应的一小部分

(<8%)^[55]. 在对源水和自来水进行全面的生物效应评估和关键生物效应贡献识别方面, Guo等人^[56]通过基于剂量效应转录组的效应导向分析发现长江中下游水体污染严重, 其雌激素效应最为普遍, 并通过报告基因法确定了己烯雌酚是该效应的主要贡献者.

效应导向分析(effect-directed analysis, EDA)按分子量、疏水性等物化特性通过色谱或物理法分馏, 将分离的组分进行包括组学在内的生物测试, 每个生物活性组分被进一步分馏, 直到鉴定出一个或多个生物活性亚组分, 这些亚组分含有解释大多数观察到的效应物质, 从而识别出关键致毒物质(图4). Xiao等人^[57]基于EDA, 使用荧光蛋白的方法检测环境污染物胁迫下RTL-W1细胞中脱乙基酶(ethoxyresorufin-*O*-deethyl-

lase, EROD)的活性, 识别了三峡大坝沉积物激活AhR的关键致毒组分为4~5环的PAHs、苯并噻唑等, 并发现其关键致毒物质存在区域差异性. Kim等人^[58]通过检测H4IIE-luc荧光信号, 识别出韩国蔚山湾沉积物中激活AhR的关键致毒组分为苯并[j]荧蒽等7种新PAHs, 共可解释环境样品0.07%~16%的AhR激活效应, 并发现研究区域新PAHs的AhR风险强于传统PAHs物质BaP. Qi等人^[59]基于EDA, 使用荧光蛋白、钼酸铵法等方法, 检测10 d内摇蚊5种酶的效应, 综合评价珠江沉积物的毒性, 发现氯氰菊酯、烯酰吗啉、小珠螨和苯基氯的氧化应激效应是沉积物引发摇蚊死亡的重要原因. Yue等人^[60]基于EDA, 使用酵母雌激素筛选测试(吸光光度法), 发现加拿大某地废水雌激素活性的关键致毒物质

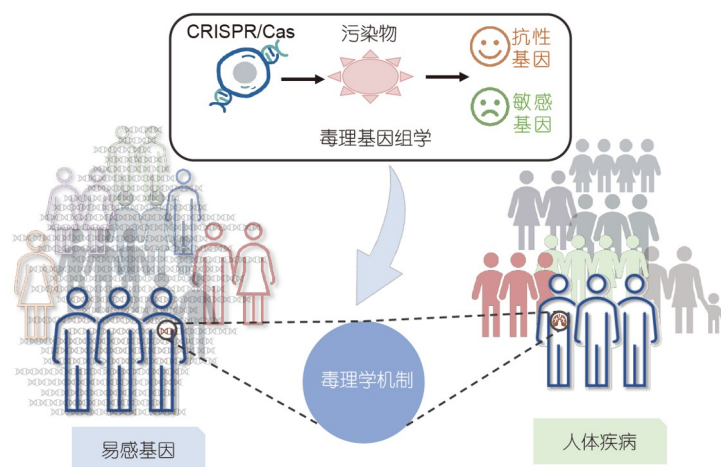


图3 (网络版彩色)毒理基因组策略识别有毒污染物-基因-疾病的关联

Figure 3 (Color online) Identification of the connection among toxic pollutants, genes and diseases by functional toxicogenomic approaches



图4 (网络版彩色)高通量组学联合效应导向性分析识别复合污染中关键致毒化学物质

Figure 4 (Color online) Identification of key chemical toxicants in complex mixtures by the combination of high-throughput omics and effect-directed analysis

结构为双键当量在6~10之间的化合物。沈艳红^[61]对长江流域水厂的进水与出水使用人体细胞MCF7进行剂量效应转录组的测试,发现水源水对激素、DNA损伤与增殖的干扰效力较强,而水厂处理能有效降低水样品的DNA损伤与增殖干扰效力。结合EDA,研究发现三苯基氧化膦、邻苯二甲酸二(2-乙基己)酯等是长江流域水厂的雌激素受体活性的主要贡献者。但是基于传统靶向生物活性测试方法的EDA存在一些不足之处,如致毒通路覆盖度低不能对样品进行综合活性评估,所识别的高活性物质类型具有偏向性,以及难以应用于致毒模式未知或复杂的混合物样品中。

剂量依赖的组学测试可建立更多化学物质结构与分子、生物学通路的机制关联,提供新的分子毒性鉴别评估(molecular toxicity identification and evaluation, mTIE)策略。组学测试通过探明生物整体基因网络对混合物中关键致毒物质结构的响应关系,从而评估化学复合污染整体效应和阐释化学复合污染潜在机理,并鉴别关键致毒物质类型。Biales等人^[62]基于TIE,使用微阵列检测伊丽沙白河沉积物暴露下的基因变化情况,发现PAHs是研究区域沉积物影响钩虾(*Ampelisca abdita*)基因表达水平变化的主要致毒物质,该方法能显著地将稀释10%~40%的沉积物样品与空白对照的毒性效应区分开来,证明了检测基因水平的变化比传统急性毒性方法识别有毒污染物的灵敏度高出至少1个数量级。Antczak等人^[63]基于mTIE,使用微阵列检测有毒污染物胁迫下大型溞(*Daphnia magna*)基因表达变化,并基于基因表达变化数据结合机器学习,成功地将不同结构的化学物质进行区分。方文迪^[64]使用剂量效应转录组方法,检测某化工园区对HepG2细胞的全生物学通路干扰效力,发现该复合污染水体主要通过DNA损伤致毒。她结合比较毒理基因数据库(Comparative Toxicogenomics Database, CTD)的有毒污染物结构效应关系,预测了高风险物质结构为杂环、氨基、醛基,最后

通过与工业园区物质名录比对,筛选出了70种高风险物质。

5 结论

随着化学品种类和使用量的不断增加,污染物毒性筛查与健康风险评估对保护人体健康以及环境安全是非常重要的。我们综述了组学技术在4个方面的优势,可用于评估大规模有毒污染物的需求,主要包括:(1)可以实现高通量获取生物学通路机制的信息;(2)系统获取有害污染物在不同生物层次上的毒性信息并进行机制解释;(3)构建遗传变异-环境暴露-人群疾病之间的直接关系,实现精准的人体健康评估;(4)从分子水平评价环境样品的总体毒性以及识别出关键致毒物质(表1)。同时,毒理组学测试所产生的大量生物学信息数据尚未被充分利用,如何解析这些数据是非常大的挑战。未来随着组学数据分析工具不断地被开发,尤其是针对多组学数据的网络分析工具可为污染物毒性筛查与健康风险评估提供更高效的手段^[65]。此外,现有的毒理组学测试还存在一些技术不足之处,例如,功能基因组学中使用的CRISPR/Cas9全基因敲除技术会存在脱靶现象,测试对象局限于模式物种等。

组学技术具有全基因组覆盖、支持大数据分析等优势,可以克服传统污染物毒性筛查与风险评估方法在从源头到有害健康结局路径上存在的技术瓶颈问题。在分子事件上,毒理组学测试可以实现对整个基因组分子事件的检测,且通过生物信息学分析分子事件之间的联系。在生物学通路方面,多组学联合从不同生物层级对有毒污染物毒性机制信息进行系统分析。在人群疾病方面,功能基因组学助力搭建有毒污染物-基因-疾病之间的直接关系,这是开展精准健康风险评估的前提。在复合污染方面,联合转录组学的化学分析可以从整体效应出发评估样品的毒性效应并识别关键致毒物质。

表 1 高通量组学技术在不同生物学水平上评估化学污染物效应的应用以及不足

Table 1 The application of high-throughput omics at different biological levels induced by chemical pollutants and the corresponding shortcomings

方向名称	研究特色	优势	不足
有害健康结局的分子事件获取	高通量获取生物学通路机制信息	通量高	信息多, 成本高, 技术要求高
关键事件的系统生物通路分析	可以系统获取有害污染物在不同生物层次上的毒性信息	机制信息解释性和逻辑性强	对研究者的分析能力要求高
基于毒理学的人群疾病遗传特征筛查	构建遗传变异-环境暴露-人群疾病之间的直接关系	精准	脱靶效应
基于分子效应的复合污染分析	突破了化学分析仪器检测限的限制以及单一生物测试覆盖度低的问题	从分子水平评价真实环境的毒性以及识别出关键致毒物质	缺少参考化学物质数据库; 测试平台局限于模式物种

参考文献

- Vermeulen R, Schymanski E L, Barabási A L, et al. The exposome and health: Where chemistry meets biology. *Science*, 2020, 367: 392–396
- Factors G R. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet*, 2017, 390: 1345–1422
- Landrigan P J, Fuller R, Acosta N J R, et al. The Lancet Commission on pollution and health. *Lancet*, 2018, 391: 462–512
- Brack W, Barcelo Culleres D, Boxall A B A, et al. One planet: One health. A call to support the initiative on a global science-policy body on chemicals and waste. *Environ Sci Eur*, 2022, 34: 21
- Groh K, Vom Berg C, Schirmer K, et al. Anthropogenic chemicals as underestimated drivers of biodiversity loss: Scientific and societal implications. *Environ Sci Technol*, 2022, 56: 707–710
- Zhang X, Xia P, Wang P, et al. Omics advances in ecotoxicology. *Environ Sci Technol*, 2018, 52: 3842–3851
- Richard A M, Judson R S, Houck K A, et al. ToxCast chemical landscape: Paving the road to 21st century toxicology. *Chem Res Toxicol*, 2016, 29: 1225–1251
- Tian M M, Xia P, Zhang X W. Applications of functional genomics in uncovering the toxicity mechanisms of environmental chemicals (in Chinese). *Asian J Ecotoxicol*, 2022, 17: 1–17 [田明明, 夏普, 张效伟. 功能基因组学在环境化学品毒性机制研究中的应用. *生态毒理学报*, 2022, 17: 1–17]
- McHale C M, Zhang L, Hubbard A E, et al. Toxicogenomic profiling of chemically exposed humans in risk assessment. *Mutat Res Rev Mutat Res*, 2010, 705: 172–183
- Escher B I, Stapleton H M, Schymanski E L. Tracking complex mixtures of chemicals in our changing environment. *Science*, 2020, 367: 388–392
- Chen M, Zhang M, Borlak J, et al. A decade of toxicogenomic research and its contribution to toxicological science. *Toxicol Sci*, 2012, 130: 217–228
- Liu Z, Huang R, Roberts R, et al. Toxicogenomics: A 2020 vision. *Trends Pharmacol Sci*, 2019, 40: 92–103
- Pettit S, des Etages S A, Mylecraine L, et al. Current and future applications of toxicogenomics: Results summary of a survey from the HESI Genomics State of Science Subcommittee. *Environ Health Perspect*, 2010, 118: 992–997
- Chepelev N L, Moffat I D, Labib S, et al. Integrating toxicogenomics into human health risk assessment: Lessons learned from the benzo[a]pyrene case study. *Crit Rev Toxicol*, 2015, 45: 44–52
- Sakhteman A, Failli M, Kublbeck J, et al. A toxicogenomic data space for system-level understanding and prediction of EDC-induced toxicity. *Environ Int*, 2021, 156: 106751
- Mansouri K, Abdelaziz A, Rybacka A, et al. CERAPP: Collaborative estrogen receptor activity prediction project. *Environ Health Perspect*, 2016, 124: 1023–1033
- Li Z, Yin H, Shen Y, et al. The influence of phenolic environmental estrogen on the transcriptome of uterine leiomyoma cells: A whole transcriptome profiling-based analysis. *Ecotoxicol Environ Saf*, 2021, 211: 111945
- Wang X, Li F, Liu J, et al. Transcriptomic, proteomic and metabolomic profiling unravel the mechanisms of hepatotoxicity pathway induced by triphenyl phosphate (TPP). *Ecotoxicol Environ Saf*, 2020, 205: 111126
- 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: 1929–1935
- Yeakley J M, Shepard P J, Goyena D E, et al. A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling. *PLoS One*, 2017, 12: 178302
- Balik-Meisner M R, Mav D, Phadke D P, et al. Development of a zebrafish S1500+ sentinel gene set for high-throughput transcriptomics. *Zebrafish*, 2019, 16: 331–347
- Subramanian A, Narayan R, Corsello S M, et al. A next generation connectivity Map: L1000 Platform and the first 1000000 profiles. *Cell*, 2017, 171: 1437–1452.e17
- Reardon A J F, Rowan-Carroll A, Ferguson S S, et al. Potency ranking of per- and polyfluoroalkyl substances using high-throughput transcriptomic analysis of human liver spheroids. *Toxicol Sci*, 2021, 184: 154–169
- Basu N, Crump D, Head J, et al. EcoToxChip: A next-generation toxicogenomics tool for chemical prioritization and environmental management. *Environ Toxicol Chem*, 2019, 38: 279–288
- Zhang K, Zhao Y. Reduced zebrafish transcriptome atlas toward understanding environmental neurotoxicants. *Environ Sci Technol*, 2018, 52: 7120–7130
- Xia P, Zhang X, Zhang H, et al. Benchmarking water quality from wastewater to drinking waters using reduced transcriptome of human cells. *Environ Sci Technol*, 2017, 51: 9318–9326
- Wang P, Wang Z, Xia P, et al. Concentration-dependent transcriptome of zebrafish embryo for environmental chemical assessment. *Chemosphere*,

- 2020, 245: 125632
- 28 Svingen T, Villeneuve D L, Knapen D, et al. A pragmatic approach to adverse outcome pathway development and evaluation. *Toxicol Sci*, 2021, 184: 183–190
 - 29 FitzGerald R E. Adverse outcome pathway bridge building from research to regulation. *Chem Res Toxicol*, 2020, 33: 849–851
 - 30 Martens M, Verbruggen T, Nyman P, et al. Introducing WikiPathways as a data-source to support adverse outcome pathways for regulatory risk assessment of chemicals and nanomaterials. *Front Genet*, 2018, 9: 661
 - 31 Li R, Luo C, Qiu J, et al. Metabolomic and transcriptomic investigation of the mechanism involved in enantioselective toxicity of imazamox in *Lemna minor*. *J Hazard Mater*, 2022, 425: 127818
 - 32 Russom C L, LaLone C A, Villeneuve D L, et al. Development of an adverse outcome pathway for acetylcholinesterase inhibition leading to acute mortality. *Environ Toxicol Chem*, 2014, 33: 2157–2169
 - 33 Vinuela A, Snoek L B, Riksen J A G, et al. Genome-wide gene expression analysis in response to organophosphorus pesticide chlorpyrifos and diazinon in *C. elegans*. *PLoS One*, 2010, 5: 12145
 - 34 Garcia-Reyero N, Escalon L, Prats E, et al. Targeted gene expression in zebrafish exposed to chlorpyrifos-oxon confirms phenotype-specific mechanisms leading to adverse outcomes. *Bull Environ Contam Toxicol*, 2016, 96: 707–713
 - 35 Moreira E G, Yu X, Robinson J F, et al. Toxicogenomic profiling in maternal and fetal rodent brains following gestational exposure to chlorpyrifos. *Toxicol Appl Pharmacol*, 2010, 245: 310–325
 - 36 Xia P, Peng Y, Fang W, et al. Cross-model comparison of transcriptomic dose-response of short-chain chlorinated paraffins. *Environ Sci Technol*, 2021, 55: 8149–8158
 - 37 Brockmeier E K, Hodges G, Hutchinson T H, et al. The role of omics in the application of adverse outcome pathways for chemical risk assessment. *Toxicol Sci*, 2017, 158: 252–262
 - 38 Guan R, Li N, Wang W, et al. The adverse outcome pathway (AOP) of estrogen interference effect induced by triphenyl phosphate (TPP): Integrated multi-omics and molecular dynamics approaches. *Ecotoxicol Environ Saf*, 2022, 234: 113387
 - 39 Slob W, Moerbeek M, Rauniomaa E, et al. A statistical evaluation of toxicity study designs for the estimation of the benchmark dose in continuous endpoints. *Toxicol Sci*, 2005, 84: 167–185
 - 40 Paul Friedman K, Gagne M, Loo L H, et al. Utility of *in vitro* bioactivity as a lower bound estimate of *in vivo* adverse effect levels and in risk-based prioritization. *Toxicol Sci*, 2020, 173: 202–225
 - 41 Peng Y, Fang W, Yan L, et al. Early life stage bioactivity assessment of short-chain chlorinated paraffins at environmentally relevant concentrations by concentration-dependent transcriptomic analysis of zebrafish embryos. *Environ Sci Technol*, 2020, 54: 996–1004
 - 42 Canzler S, Schor J, Busch W, et al. Prospects and challenges of multi-omics data integration in toxicology. *Arch Toxicol*, 2020, 94: 371–388
 - 43 Lee H, Sung E J, Seo S, et al. Integrated multi-omics analysis reveals the underlying molecular mechanism for developmental neurotoxicity of perfluorooctanesulfonic acid in zebrafish. *Environ Int*, 2021, 157: 106802
 - 44 Pogrmic-Majkic K, Samardzija Nenadov D, Tesic B, et al. Mapping DEHP to the adverse outcome pathway network for human female reproductive toxicity. *Arch Toxicol*, 2022, 96: 2799–2813
 - 45 Tsamou M, Roggen E L. Building a network of adverse outcome pathways (AOPS) incorporating the Tau-driven AOP toward memory loss (AOP429). *J Alzheimers Dis Rep*, 2022, 6: 271–296
 - 46 Gao P, Shen X, Zhang X, et al. Precision environmental health monitoring by longitudinal exposome and multi-omics profiling. *Genome Res*, 2022, 32: 1199–1214
 - 47 Lind L, Ng E, Ingelsson E, et al. Genetic and methylation variation in the CYP2B6 gene is related to circulating p,p'-DDE levels in a population-based sample. *Environ Int*, 2017, 98: 212–218
 - 48 Traglia M, Croen L A, Lyall K, et al. Independent maternal and fetal genetic effects on midgestational circulating levels of environmental pollutants. *G3-Genes Genomes Genet*, 2017, 7: 1287–1299
 - 49 Dai X, Deng S, Wang T, et al. Associations between 25 lung cancer risk-related SNPs and polycyclic aromatic hydrocarbon-induced genetic damage in coke oven workers. *Cancer Epidemiol Biomarkers Prev*, 2014, 23: 986–996
 - 50 Zeng X, Vonk J M, van der Plaats D A, et al. Genome-wide interaction study of gene-by-occupational exposures on respiratory symptoms. *Environ Int*, 2019, 122: 263–269
 - 51 Huang Z, Wang S, Liu Y, et al. GPRC5A reduction contributes to pollutant benzo[a]pyrene injury via aggravating murine fibrosis, leading to poor prognosis of IIP patients. *Sci Total Environ*, 2020, 739: 139923
 - 52 Yang K, Jiang X, Cheng S, et al. Synaptic dopamine release is positively regulated by SNAP-25 that involves in benzo[a]pyrene-induced neurotoxicity. *Chemosphere*, 2019, 237: 124378
 - 53 Tian M, Xia P, Yan L, et al. Toxicological mechanism of individual susceptibility to formaldehyde-induced respiratory effects. *Environ Sci Technol*, 2022, 56: 6511–6524

- 54 Brack W, Ait-Aissa S, Burgess R M, et al. Effect-directed analysis supporting monitoring of aquatic environments—An in-depth overview. *Sci Total Environ*, 2016, 544: 1073–1118
- 55 Neale P A, Braun G, Brack W, et al. Assessing the mixture effects in *in vitro* bioassays of chemicals occurring in small agricultural streams during rain events. *Environ Sci Technol*, 2020, 54: 8280–8290
- 56 Guo J, Shen Y, Zhang X, et al. Effect-directed analysis based on the reduced human transcriptome (RHT) to identify organic contaminants in source and tap waters along the Yangtze River. *Environ Sci Technol*, 2022, 56: 7840–7852
- 57 Xiao H, Krauss M, Floehr T, et al. Effect-directed analysis of Aryl hydrocarbon receptor agonists in sediments from the Three Gorges Reservoir, China. *Environ Sci Technol*, 2016, 50: 11319–11328
- 58 Kim J, Hong S, Cha J, et al. Newly identified AHR-active compounds in the sediments of an industrial area using effect-directed analysis. *Environ Sci Technol*, 2019, 53: 10043–10052
- 59 Qi H, Li H, Wei Y, et al. Effect-directed analysis of toxicants in sediment with combined passive dosing and *in vivo* toxicity testing. *Environ Sci Technol*, 2017, 51: 6414–6421
- 60 Yue S, Ramsay B A, Brown R S, et al. Identification of estrogenic compounds in oil sands process waters by effect directed analysis. *Environ Sci Technol*, 2015, 49: 570–577
- 61 Shen Y H. Identification of the active chemicals in water from treatment plants based on molecular EDA (in Chinese). Master Dissertation. Nanjing: Nanjing University, 2019 [沈艳红. 基于分子EDA的水厂水中关键活性物质鉴别技术研究. 硕士学位论文. 南京: 南京大学, 2019]
- 62 Biales A D, Kostich M, Burgess R M, et al. Linkage of genomic biomarkers to whole organism end points in a toxicity identification evaluation (TIE). *Environ Sci Technol*, 2013, 47: 1306–1312
- 63 Antczak P, Jo H J, Woo S, et al. Molecular toxicity identification evaluation (mTIE) approach predicts chemical exposure in *Daphnia magna*. *Environ Sci Technol*, 2013, 47: 11747–11756
- 64 Fang W D. Development of reduced transcriptome approach for evaluation and diagnosis of environmental toxic chemicals (in Chinese). Doctor Dissertation. Nanjing: Nanjing University, 2020 [方文迪. 基于简化转录组的环境化学物质毒性评估与诊断. 博士学位论文. 南京: 南京大学, 2020]
- 65 Li J, Chen F, Liang H, et al. MoNET: An R package for multi-omic network analysis. *Bioinformatics*, 2021, 38: 1165–1167

Summary for “基于组学的污染物高通量毒理学筛查与风险评估”

Omics-based high throughput toxicity screening and risk assessment of pollutants

Lu Yan¹, Xiao Gou¹, Pu Xia¹, Ruize Gao¹, Huimin Ji¹, Wei Shi^{1,2}, Hongxia Yu¹ & Xiaowei Zhang^{1,2*}

¹ State Key Laboratory of Pollution Control & Resource Reuse, School of Environment, Nanjing University, Nanjing 210023, China;

² Jiangsu Province Ecology and Environment Protection Key Laboratory of Chemical Safety and Health Risk, Nanjing 210023, China

* Corresponding author, E-mail: zhangxw@nju.edu.cn

Screening priority pollutants with high toxicity and risk potential from a large number of environmental chemical pollutants is the prerequisite for prevention and control of environmental risks. After entering the living system, toxic pollutants can induce adverse outcomes through a series of disturbed molecular targets and biological pathways. Existing high-throughput toxicity screening methods rely on few and isolated cellular or molecular endpoints, making it difficult to capture the key biological processes that cause toxicity and to determine the direct relationship between chemical pollutants and disease. Meanwhile, the awareness is increasing that we are exposed to a true cocktail of chemicals, among which only a fraction has been identified through analytical methods. Integrated bioanalytical approaches, such as effect-directed analysis and toxicity identification evaluation (TIE), have been used to identify risk drivers in complex mixture pollutants. However, it is still difficult to elucidate the molecular toxicological drivers of chemical stress in the environment.

Toxicogenomics take advantage of genome-wide coverage, high throughput, and big data capacity that are associated with omics technologies to achieve technology innovation in screening and risk assessment of chemical pollutants. We reviewed the technical advantages and the recent progress of toxicogenomic methodology in high-throughput toxicological screening and risk assessment of chemical pollutants from four aspects with specific research cases. To identify the molecular events initiated by toxic pollutants, toxicogenomics could uncover a large number of molecular events and pathways potentially disturbed at the early development stage of toxicity. Dose-dependent transcriptomes for high-throughput chemical testing have been developed to quantitatively assess responses induced by chemical pollutants at the molecular level, which improve the sensitivity of identifying bioactivity by 1–3 orders of magnitude.

To uncover the biological pathways of adverse health outcome by chemical pollutants, systemic toxicological assessment by omics strategies provide mechanistic information for chemical pollution screening and health risk assessment at multiple biological levels. Through biological pathway enrichment analysis, such as the gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) pathway, the molecular mechanistic information of toxicity could be captured. Moreover, multi-omics integration allows for multi-scale, comprehensive biological network analysis, which provide a systematic understanding of toxicity pathways or adverse outcome pathways (AOPs) to guide ecological and human health risk assessments.

For human diseases diagnostics, functional genomics provide a new tool to explore mechanisms of individual disease risk and susceptibility under toxic pollutants exposure. Variations of individual susceptibility to toxicant-induced disease are largely due to genetic susceptibility across different populations. The direct relationship among pollutants, genes and adverse outcomes could be established by combining functional genomics and molecular epidemiological approaches.

Finally, for risk management of complex mixture pollutants, it is important to assess the overall toxicity effects of the pool of pollutants and identify the key toxic pollutant. The combination of high-throughput bioassays and chemical analysis can be used to identify the highly bioactive fraction and to screen the key chemical pollutant that pose ecological and health risks. Recently, progresses have been made in integrating transcriptomic techniques with molecular toxicity identification and evaluation (mTIE) to assess the overall effects of mixtures and distinguish the key pollutants in inducing the corresponding biological responses. Particularly, dose-response transcriptome could be used to quantitatively evaluate the overall toxic effects and to guide the key toxicant identification by high mechanistic resolution.

In summary, the development of omics-based high-throughput toxicity screening technologies will strongly support accurate environmental risk assessment of conventional and new pollutants in the future.

toxicogenomics, chemical pollutants, molecular targets, biological pathways, human diseases, mixture pollutants

doi: [10.1360/TB-2022-0682](https://doi.org/10.1360/TB-2022-0682)