

学术前沿专栏:新污染物检测技术

【特约主持人】孟祥周:教育部“新世纪优秀人才支持计划”入选者

【主持人按语】从环境管理角度来看,新污染物一般是指新近发现或被关注,对生态环境或人体健康存在风险,尚未纳入管理或者现有管理措施不足以有效防控其风险的污染物.新污染物是全球环境领域研究热点之一,我国于 2022 年发布实施《新污染物治理行动方案》,全面推进新污染物治理工作,是当前我国生态环境科技创新的重点领域.环境中新污染物种类繁多且总体浓度较低,建立新污染物检测方法是掌握新污染物的环境暴露水平、治理新污染物的重要基础.本专栏围绕新污染物的检测技术进行讨论,其中《PFAS 高分辨率质谱数据的非靶标识别技术综述》聚焦典型新污染物全氟及多氟烷基物质(PFAS),系统梳理了如何基于高分辨质谱检测数据建立 PFAS 的非靶标识别技术,从识别框架、策略、软件等角度展开讨论,可为新污染物筛查提供科技支撑.《基于原位形成低共熔溶剂的酸诱导—分散液液微萃取/高效液相色谱法测定环境水样中烷基酚》以癸酸钠为萃取剂,建立了一种原位检测水样中烷基酚的方法,具有简单、快速、灵敏、低成本和绿色等优点,成功应用于黄河水、湖水、卫河水样中烷基酚的测定,也为开发新污染物原位检测技术提供了借鉴.

PFAS 高分辨率质谱数据的非靶标识别技术综述

孟祥周^{1,2,3},张博暄^{1,3},韩宝苍^{1,3},朱清禾²,杨洁²,张敏超⁴

(1.同济大学 环境科学与工程学院,上海 200092;2.上海市环境科学研究院 生态环境部城市土壤污染控制与修复工程技术中心,上海 200233;3.嘉兴同济环境研究院,浙江 嘉兴 314051;4.上海海关工业品与原材料检测技术中心,上海 201210)

摘要:全氟及多氟烷基物质(per- and polyfluoroalkyl substances, PFAS)使用广泛,在环境中难以降解,且具有生物富集性、迁移性和毒性,引发全球关注.PFAS 种类繁多,基于高分辨率质谱的非靶标分析是发现环境中未知 PFAS 的主要方法,而高效识别技术是其中的难点.系统梳理了 PFAS 非靶标识别技术的框架,总结了不同 PFAS 非靶标识别策略的应用情况与优缺点,比较了不同 PFAS 非靶标识别开源软件的特点,以期对环境未知 PFAS 的精准识别、溯源和管控提供科技支撑.

关键词:全氟及多氟烷基物质(PFAS);高分辨率质谱(HRMS);非靶标识别;疑似筛查;新污染物

中图分类号:X502

文献标志码:A

文章编号:1000-2367(2025)03-0001-11

收稿日期:2024-11-18;**修回日期:**2024-11-23.

基金项目:国家自然科学基金(42177378);国家环境保护城市土壤污染控制与修复工程技术中心开放基金(USCR-202202);嘉兴市公益性研究计划项目(2023AY11053);上海市法医学重点实验室暨司法部司法鉴定重点实验室开放课题(KF202422).

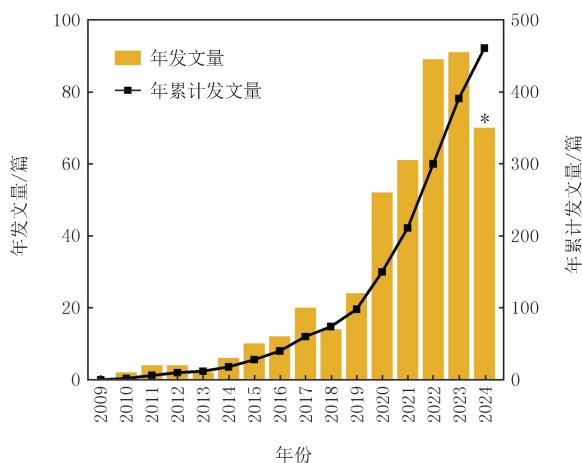
作者简介(通信作者):孟祥周(1979—),男,河南濮阳人,同济大学教授,博士,博士生导师,研究方向为新污染物环境行为及风险管控,E-mail: xzmeng@tongji.edu.cn.

引用本文:孟祥周,张博暄,韩宝苍,等.PFAS 高分辨率质谱数据的非靶标识别技术综述[J].河南师范大学学报(自然科学版),2025,53(3):1-11.(Meng Xiangzhou,Zhang Boxuan,Han Baocang,et al.A review on the nontarget identification technology for PFAS by using high resolution mass spectrometry[J].Journal of Henan Normal University (Natural Science Edition),2025,53(3):1-11.DOI:10.16366/j.cnki.1000-2367.2024.11.18.0003.)

根据经济合作及发展组织(organisation for economic co-operation and development,OECD)的定义^[1],全氟及多氟烷基物质(per- and polyfluoroalkyl substances,PFAS)是含有至少一个全氟甲基或亚甲基碳原子(没有任何 H/Cl/Br/I 原子与其相连)的氟化物质.据初步统计,具有这种结构的物质已超过 700 万种^[2],广泛应用于消防、电镀、食品包装、纺织印染等工业领域^[3].研究表明,PFAS 在全球环境介质中普遍检出,且已经对生态环境和人体健康产生不利影响,作为一类新污染物受到全球重点关注^[4-6].2012 年,PLACE 等^[7]尝试使用非靶标分析方法(nontarget analysis)对水成膜泡沫灭火剂中未知 PFAS 进行识别,鉴定出 10 类 30 种 PFAS,引发了研究者兴趣.

以“PFAS”,“nontarget OR nontargeted OR suspect screening”为主题词,在 Web of Science 数据库进行论文搜索(检索时间截至 2024 年 10 月 31 日),根据内容相关性筛选出 461 篇,其历年发表的论文数量如图 1 所示.自 2019 年开始,每年发表的论文数量呈现快速增长趋势,主要集中在 2022—2024 年.目前,PFAS 非靶标分析方法已经广泛应用于水体^[8]、大气^[9]、土壤^[10]、生物体^[11]、人体组织^[12]、饮用水^[13]、工业生产过程及工业品^[14-15]、污水处理系统^[16]等样品分析中.

非靶标分析方法一般包括样品前处理、色谱分离和高分辨率质谱分析,而高分辨率质谱数据处理环节是难点.常用的高分辨率质谱主要包括四极杆飞行时间质谱,静电场轨道阱质谱和傅立叶变换离子回旋共振质谱 3 种.质谱数据的采集模式主要包括全扫描(full scan)、数据依赖性采集(data-dependent acquisition,DDA)和数据非依赖性采集(data-independent analysis,DIA)^[17].前人已围绕 PFAS 非靶标分析中的识别技术进行了较多探讨^[18-19],但建立全面的识别方法仍然面临挑战.本文从应用的角度梳理了 PFAS 非靶标识别技术框架,系统综述了非靶标识别策略及其在开源软件中的应用,可为 PFAS 非靶标识别分析方法的发展与实践提供科技支撑.



*表示时间截至2024年10月31日

图1 PFAS非靶标分析研究论文发表数量情况

Fig.1 Number of publications on PFAS nontarget analysis

1 非靶标识别方法框架

非靶标识别技术包括未知化合物的特征提取、特征识别与结构鉴定、结果评估等流程^[20],其框架及对应环节的参数如图 2 所示.框架主体分为 5 部分,Ⅰ:峰提取;Ⅱ:特征提取;Ⅲ:疑似筛查与非靶标筛查;Ⅳ:原始数据核对;Ⅴ:化合物评级.Ⅰ和Ⅱ属于化合物特征提取,部分研究者会根据使用的软件或方法将其合并.Ⅲ为特征识别与结构鉴定,是非靶标识别技术的核心.研究者将 9 种常用的识别策略分为 3 组(图 2),早期研究通常使用其中的 1 组,而目前研究倾向于使用尽可能多的策略来提高识别结果的覆盖面和置信度.Ⅳ和Ⅴ为识别结果评估,即对识别出的未知 PFAS 进行核对和置信度分级(confidence level).SCHYMANSKI 等^[21]依据样品一级质谱(MS)、二级质谱(MS²)、色谱保留时间(retention time,RT)、标准品信息、一级质谱同位素分布特征等将未知化合物的置信度分为 1~5 级.CHARBONNET 等^[22]在分级标准的基础上,结合 PFAS 种类的结构特殊性,细化每个等级的评价标准(图 3).如 GHORBANI 等^[23]通过非靶标方法识别了污染地下水中 88 种 PFAS,其中置信度为 1a 有 22 种,2a 有 20 种,2b 有 7 种,3a 有 37 种,3b 有 2 种.随着识别技术的不断发展,PFAS 置信度评价体系也可能纳入新的指标,如二级质谱碎片质量差异和碳数归一化质量亏损等^[19].

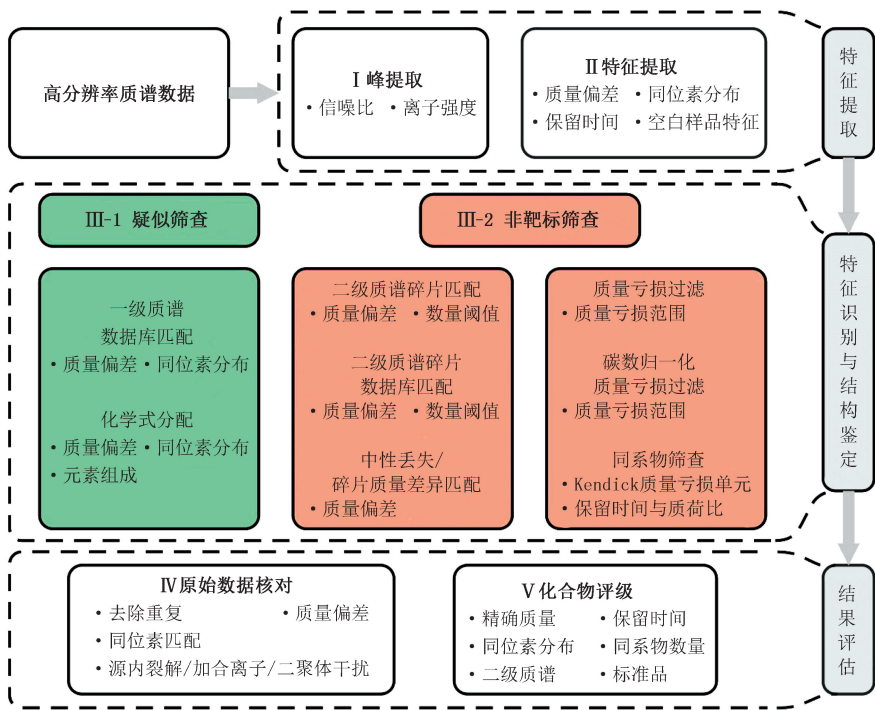


图2 PFAS非靶标识别技术框架

Fig.2 The framework of PFAS nontarget identification technology

2 PFAS 非靶标识别策略

2.1 疑似筛查

2.1.1 一级质谱数据库匹配

数据库匹配是指把检测出的高分辨率一级质谱特征与包含已知化合物的清单或数据库进行比较^[24].仅通过数据库匹配的特征置信度为 5a.常用的数据库包括公开的清单和研究人员自建清单,包括 OECDPFAS($n = 4\ 730$)、NISTPFAS($n = 4\ 948$)、PRIORISKPFAS($n = 4\ 240$)、PFASMASTER($n = 8\ 498$)等,每个清单都包含大量独有的 PFAS^[25].这些数据库一般包括化合物名称、分子式和理论精确质量数.根据分子式可以计算出每个化合物的理论同位素分布,匹配过程中会同时比较同位素的精确质量偏差与丰度偏差.使用化合物通用数据库匹配(如 ChemSpider 和 PubChem)可以涵盖更广泛的 PFAS^[2,26],但识别假阳性概率也随之增加.为了最大限度减少假阳性结果的数量,应根据研究对象(区域污染源特征、PFAS 产品信息等)和分析方法对应的 PFAS 种类和类型(如电离方式,物质电离能力等)调整数据库的内容和范围,以及与其他识别策略进行结合应用.

2.1.2 化学式分配

质量亏损(mass defect, MD)是指原子或分子的精确质量与标称质量之差.不同元素如碳(12.000 0 Da)、氢(1.007 8 Da)、氧(15.994 9 Da)、氮(14.003 1 Da)和氟(18.998 4 Da)等有着特定的质量亏损,HRMS 可以捕

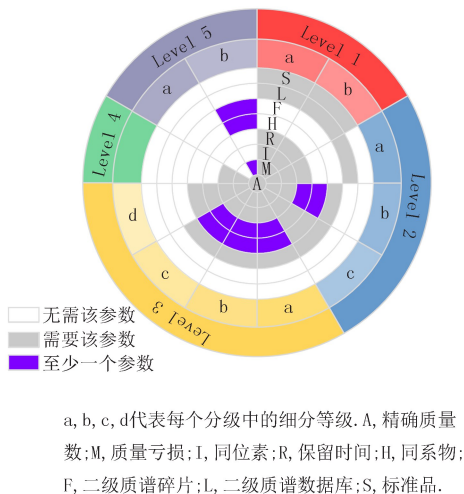


图3 未知PFAS识别置信度的分级

Fig.3 The confidence level of PFAS identified by nontarget technology

捉这样的微小差异,实现将化学式分配给足够精确的质量测量结果.仅确定未知特征分子式的置信度为 4.如全氟已烷磺酸分子离子的精确质量数为 399.943 88,在质量偏差为 5×10^{-6} 时分配的化学式为 41 个, 1×10^{-6} 时为 7 个, 0.2×10^{-6} 时为 4 个^[27].因此,为了缩小化学式分配范围,必须合理限制元素数量与类型,合理运用化学式分配规则^[28].通过比较同位素分布也会大幅减少化学式的可能性,TANG 等^[29]使用 S 同位素分布筛选全氟磺酸类物质,而 JIAO 等^[30]通过特殊同位素分布($[M]^-/[M+2]^- = 3/1$)确定化合物中氯元素的存在.在有限的质量分辨率和精度下,分配的化学式数量随着分子量的增加而呈指数增加,该方法通常与其他方法结合使用以增加识别置信度.

2.2 非靶标筛查

2.2.1 质量亏损过滤

按照质量亏损定义,碳原子不会改变分子的整体 MD,其他原子对分子的 MD 有不同程度的贡献(如 H 和 N 等贡献微小正值,S,O,F,Cl,Br 等贡献微小负值).因此 PFAS 类物质有其特有的 MD 范围(如 $-0.25 \text{ Da} < MD < 0.1 \text{ Da}$,涵盖 OECDPFAS 清单中 92.8% 的 PFAS^[31]),在规定范围内保留的特征置信度为 5b. KOELMEL 等^[32]将 $-0.11 \text{ Da} < MD < 0.12 \text{ Da}$ 范围内(涵盖 PFASMASTER 清单的 90%)的特征标记为潜在 PFAS 特征.使用 MD 过滤方法仅需特征的精确质量数计算即可有效保留氟质量分数大于 60% 的 PFAS,但对于低氟含量的 PFAS 与其他碳氢化合物则无法有效分离^[33].此外,具有高正质量亏损的化合物可能被错误地识别为负质量缺陷,对识别过程造成干扰.

2.2.2 碳数归一化质量亏损过滤

KAUFMANN 等^[34]在 2022 年提出了筛选一级质谱数据 PFAS 特征的新方法.该方法首先估算特征的碳原子数量:

$$C = (I_{M+1}/I_M)/0.011\ 145,$$

式中, C 表示特征碳原子数量, I_{M+1} 和 I_M 分别代表第一同位素峰和单同位素峰的丰度.基于特征碳原子数量,可以用特征质量数除以 C 计算碳数归一化质量(mass over carbon, m/C),用特征质量亏损除以碳数归一化质量亏损(mass defect over carbon, MD/C).通过将 m/C 和 MD/C 绘制为 XY 坐标轴图,研究发现鱼体组织样品中的 PFAS 特征与其他的特征产生了有效分离^[34].ZWEIGLE 等^[33]基于此方法对 49 万种有机化合物进行了 MD/C - m/C 表征,发现含氟量较高的 PFAS [$n(\text{F})/n(\text{C}) > 0.8$, $n(\text{H})/n(\text{F}) < 0.8$,含氟质量分数 $> 55\%$]可以有效地与其他有机化合物及天然有机物分离.结构相关的化合物(如 PFAS 同系物)在 MD/C - m/C 图中有规律的排布,可用于后续同系物筛查.ZWEIGLE 等^[35-36]也开发了系列软件 FindPFAS/PFAScreen 来实现 PFAS 的 MD/C - m/C 筛选流程与可视化,进一步证实将该方法纳入 PFAS 置信度评价体系的可行性.

2.2.3 同系物筛查

由于 PFAS 生产和使用的特点,产品和环境样品中的 PFAS 通常为含有一系列重复单元(如 CF_2 , C_2F_4 , CF_{20})的同系物^[37].Kendrick 质量亏损(Kendrick mass defect, KMD)分析是用于筛查样品中 PFAS 同系物的有效方法^[18].Kendrick 质量(Kendrick mass, KM)是将以碳元素(12.000 0 Da)为基准的质量参考体系转换为以任意重复单元为基准的质量参考体系:

$$KM = M \frac{NM_{RU}}{M_{RU}},$$

式中, KM 为 Kendrick 质量, M 为特征检测质量, NM_{RU} 和 M_{RU} 分别代表重复单元标称质量和精确质量.由于碳元素的质量亏损为零,因此转换后重复单元的质量亏损也为零,特定重复单元同系物的 KMD (由标称 Kendrick 质量与 KM 相减得到)也是相同的,由该分子的剩余部分决定.在 KMD 与质量 XY 轴图中,来自同一化合物类别的同系物出现水平对齐.还可利用不同类别化合物碳链长度与保留时间的相关性对 KMD 识别的特征进一步筛选,这可通过人工或者软件脚本实现^[38].MD 过滤方法也可用于过滤 KMD ,如 MUNOZ 等^[39]保留了 $-0.15 \text{ Da} < KMD < 0.15 \text{ Da}$ 范围内的特征来筛选潜在 PFAS.与 MD 随链长增加而减小不同, KMD 是稳定的,但单一 PFAS 特征无法通过 KMD 分析保留,丰度较低的同系物特征也容易被峰提取算法遗漏.同系物特征和保留时间信息通常作为提高置信度评级的辅助手段.

2.2.4 二级质谱碎片/数据库匹配

二级质谱碎片由质谱碰撞池中母离子的化学键断裂产生,全氟或含氟量较高的化合物通常拥有类似的二级质谱碎片.CHARBONNET 等^[22]将 PFAS 的二级质谱碎片分为两类,其中诊断离子碎片通常为含氟碎片(如 $[\text{C}_n\text{F}_{2n+1}]^-$ 和 $[\text{C}_n\text{F}_{2n+1}\text{O}]^-$ 等),子类离子碎片为不同类别 PFAS 的官能团产生的非氟碎片(如 $[\text{SO}_3]^-$ 和 $[\text{PO}_3]^-$ 等).KOELMEL 等^[32]总结了 PFAS 在质谱碎裂过程可能产生的 777 个碎片用于二级质谱数据匹配.二级质谱数据库包括母离子及其对应的二级质谱碎片的精确质量数,会同时与未知特征进行匹配,匹配特征将分配 3 级及以上置信度.数据库分为公开数据库(如 MassBank^[40])、供应商数据库(如 mz-Cloud^[41])及自建数据库,其中自建数据库包括实验数据库^[42-43]和计算机算法模拟数据库^[44]等.由于较高的置信度分配,二级质谱碎片/数据库匹配已成为 PFAS 非靶标识别的主要方法.

2.2.5 中性丢失/碎片质量差异匹配

中性丢失是指分子中不带电荷的片段在质谱分析过程中丢失的现象,在一级质谱扫描和二级质谱扫描中均会产生.PFAS 物质的中性丢失通常发生在极性的头部基团和非极性的尾部基团^[45].TANG 等^[29]使用 $[\text{CO}_2]$ 和 $[\text{CF}_2\text{O}]$ 作为中性丢失片段对一级质谱数据中全氟羧酸类 PFAS 特征进行搜索识别. $[\text{HF}]$ 是氟调聚和含氢 PFAS 等多氟烷基 PFAS 的常见中性丢失片段,是通过二级质谱数据推断物质结构的重要依据^[46].碎片质量差异源自化合物不同碎片化途径产生的碎片之间的差异,如碎片 $[\text{C}_2\text{F}_5]^-$ 和 $[\text{C}_3\text{F}_7]^-$ 间的 ΔCF_2 差异.ZWEIGLE 等^[35]按照 PFAS 类别总结了对应的碎片质量差异,应用该方法显著提高了非靶标 PFAS 的识别数量.虽然中性丢失和碎片质量差异产生的原理并不相同,但二者都反映为二级质谱中两个碎片之间的中性片段差异,在分析过程中拥有一致的匹配方式,未来可作为置信度评价体系的一个有效指标.

2.3 新兴技术

2.3.1 色谱法-保留时间指数

目前 PFAS 非靶标分析中,液相色谱耦合高分辨率质谱仍是主要方法^[47].因此,借助化合物结构与 RT 的相关性(即保留时间指数 retention time indice,RTI,简记为 I_{RT})可减少假阳性识别结果.AALIZADEH 等^[48]则建立了 I_{RT} 的计算公式:

$$I_{\text{RT}} = \frac{t_{\text{Rx}} - t_{\text{Rmin}}}{t_{\text{Rmax}} - t_{\text{Rmin}}} \times 1\,000 = \alpha \cdot t_{\text{RC}} + C,$$

式中, I_{RT} 是保留时间指数, t_{Rx} 和 t_{RC} 分别表示校准物质和目标化合物的 RT, t_{Rmax} 和 t_{Rmin} 是校准物质实际测量的最大和最小 RT, α 和 C 分别是置信区间 99%对应的斜率和截距.

目前应用 I_{RT} 的研究主要包括:(1)机理模型预测^[49].使用真实溶剂似导体屏蔽模型(conductor-like screening model for realistic solvents,COSMO-RS)计算 PFAS 的辛醇水分配系数(octanol-water partition coefficients, K_{ow}),结合各向同性极化率与实验 I_{RT} 建立多元回归模型,预测未知 PFAS 的 I_{RT} ,但该预测模型不能突破液相条件的限制;(2)机器学习预测.江汉大学丁一^[50]基于 PFAS 分子结构与 RT 之间的构效关系发展了机器学习算法模型,通过筛选多种分子描述符的重要度增加模型的鲁棒性和预测能力.同时,将预测 I_{RT} 与实测 I_{RT} 的误差范围作为条件对识别结果进行过滤,过滤效率高达 52.5%.

2.3.2 离子淌度质谱-碰撞截面积

离子淌度质谱(ion mobility spectrometry,IMS)技术是一种快速气相分离技术,可根据离子在缓冲气体中的大小、形状和电荷状态产生不同的漂移时间,实现离子分离,增加质谱的分析维度^[51].通过漂移时间计算得到的碰撞横截面积(collision cross sections,CCS)是一种分子描述符,可为未知化合物识别提供额外证据^[52].DODDS 等^[51]探究了不同类别 PFAS 的 CCS 与质荷比的关系,系统验证了 CCS 在区分不同子类和异构体 PFAS 方面的稳定性,以及成为未知化合物鉴定指标的潜力.KIRKWOOD-DONELSON 等^[53]使用 IMS 开发了一种基于分子尺寸的 DIA 质谱分析方法,利用 CCS 与碰撞能量的相关性优化了各类 PFAS 的碰撞能量,同时增强了 DIA 数据反卷积的效果.

2.4 识别方法比较

不同 PFAS 识别策略的优缺点如表 1 所示.一级质谱数据经 full scan 得到,包括分子离子及其同位素的精确质量数与离子强度,质荷比范围(m/z)一般在 50~1 000,质量分辨率大于 20 000($m/z=500$),质量偏

差范围为 $(1\sim10)\times10^{-6}$.基于一级质谱数据的识别策略较多,主要是将 PFAS 特征筛选出来,但较少单独使用,对识别置信度的影响也较小.二级质谱数据分为单个母离子的 MS² 数据(经靶向 MS/MS² 或 DDA 得到)和多个母离子的 MS² 数据(经 DIA 得到).二级质谱数据识别策略可以提升识别置信度,常易受数据库容量和二级质谱碎片数量的限制.两种新兴方法的应用效果已经得到证明,但受限于研究数量,仍处于发展阶段,具有较大应用潜力.综合应用识别策略可以弥补各自的劣势,同时应用效果也受到质谱数据质量精确度、碎片丰富度、特征碎裂覆盖率等仪器分析效果的影响,确保高质量质谱数据的获取也是 PFAS 非靶标分析的重要一环.

表 1 不同识别策略比较分析
Tab. 1 Comparison of different nontarget identification strategies

数据类型	识别方法	优点	缺点
一级质谱(MS)	数据库匹配	可快速锁定候选物质范围	单独使用候选结果较多
	化学式分配	特定化学式物质识别效果好	结果受化学式分配参数限制,分配结果多
二级质谱(MS ²)	质量亏损过滤	有效过滤非含氟特征	无法有效过滤低含氟特征
	碳数归一化质量亏损过滤	非含氟特征过滤效果优良,可估算氟含量	离子强度较低的特征过滤误差大
	同系物筛查	提高置信度,获取分组信息	单一物质无法有效识别
	二级质谱碎片/数据库匹配	提高置信度,获取结构信息	受输入碎片或输入库限制
	中性丢失/碎片质量差异匹配	可识别未知含氟组分	需预先识别二级碎片
色谱	保留时间指数	基于机理的合理预测,减少假阳性结果	受限于模型应用域
离子淌度质谱	碰撞横截面积	增加识别维度,增强异构体识别	依赖仪器与对应的数据库

3 非靶标识别开源软件

一些高分辨率质谱供应商(如 Agilent, Thermo Fisher 和 Waters 等)针对各自仪器的数据格式开发了相应的非靶标识别软件^[54].这些软件配有可视化界面并涵盖大部分非靶标分析流程,在软件运行速度上具有优势,成为非靶标分析的主要工具.但是,商业软件存在识别算法不透明的缺点(如特征峰提取,峰对齐等算法),用户无法灵活调整数据解析方式,造成识别结果不理想,新功能的集成也比较滞后^[55].近来,研究者专门针对 PFAS 非靶标识别开发了新方法及开源软件,各具特点.开源软件在部署环境和功能模块设置上更具灵活性,尽管需要使用者掌握 R 语言、Python、Java 等典型编程语言,但已经较大程度降低了对编程知识的依赖.

3.1 Fluoromatch

Fluoromatch 软件由耶鲁大学 KOELMEL 领导的团队于 2020 年开发^[56],发布后历经数次更新,当前版本为 5.4(时间截至 2024 年 10 月 31 日).该软件是首个开源的自动化 PFAS 高分辨率质谱数据分析软件,实现了从特征峰挑选,空白过滤,到 PFAS 特征注释与分组的完整工作流.其突出特点是包含大量由标准品产生或计算机预测生成的 PFAS 碎片信息,初级版本中(Fluoromatch 1.0)包含约 7 000 个 PFAS 的母离子与对应的碎片信息.该软件通过将样品特征信息(精确质量数、保留时间、二级质谱碎片)与软件数据库匹配,增加对样品 PFAS 特征注释的覆盖率.研发者还提供了支持多条件交叉过滤的可视化软件 FluoroMatch visualizer^[57],方便使用者进行人工筛查与导出数据.

3.2 EnviMass

EnviMass 由 Martin Loos 开发,是一种基于 R 编码环境的自动化高分辨率质谱数据挖掘工具,包含 Web 浏览器中的图形用户界面^[58].该软件最初用于从 LC-HRMS 数据中检测水生系统中已知和未知微污染物的变化趋势和泄漏信息,在 4.4 版本增加了针对 PFAS 分析的脚本,可进行靶向和疑似筛查、源内片段识别、同系物筛查、二级质谱碎片匹配、质量亏损过滤.其中源内片段识别会检查每个特征是否具有其他共洗脱离子,并与自建的源内片段库进行筛选匹配,从而排除因源内裂解和加合离子导致的假阳性识别.

3.3 FindPFAS/PFAScreen

图宾根大学 ZWEIGLE 团队^[35]于 2022 年开发了基于 Python 的开源软件 FindPFAS,该软件通过搜索二级质谱原始数据中碎片质量的特定差异(如 ΔCF_2 、 ΔCF_3 等)来发现相关联的 PFAS 类别.FindPFAS 同时提供二级质谱碎片匹配和前体离子同系物筛查等功能.在 FindPFAS 功能的基础上,该团队又开发了 PFAScreen 软件^[36],该软件主要增加了一级质谱的碳数归一化质量亏损(MD/C)过滤功能,并增强了色谱图、质谱图和离子共洗脱等可视化功能.

3.4 APP-ID

南京大学韦斯团队在 2024 年开发了一个自动 PFAS 识别平台(APP-ID)^[59],该平台突出亮点是对常规 PFAS 非靶标识别流程的特征提取和特征识别环节进行改进.这些改进包括使用自主开发的 Flink 算法构建 PFAS 分子网络,开发一级质谱精确搜索模块和转换搜索模块,开发机器学习预测 PFAS 结构指纹模块,构建候选结构和预测指纹相似性的排序评分系统,综合提高对疑似 PFAS 特征的提取覆盖率和未知 PFAS 的识别准确率.

3.5 SWATH-F/IonDecon

在 3.1~3.4 节所列出的软件主要适用于 DDA 数据,且功能已经较为完善,而针对 DIA 数据的 PFAS 非靶标识别软件仍在不断开发中.南京大学韦斯团队针对 SWATH(sequential window acquisition of all theoretical fragment-ion spectra)数据(一种 DIA 数据)开发了基于 Python 的 SWATH-F^[60],该脚本可以从反卷积与峰提取后的质谱数据中筛查 PFAS 同系物,并自动化注释其结构.耶鲁大学 KOELMEL 团队^[61]则开发了解卷积软件 IonDecon,完成了 DIA 数据到 DDA 数据的转化,并集成 Fluoromatch 软件用于后续 PFAS 识别.

3.6 应用实践

表 2 总结了几款软件的基本信息与特点.已有研究比较了不同软件(包括商业软件与开源软件)在 PFAS 非靶标识别效果的差异,不同软件之间识别结果的一致性约为 70%,其差异主要源自峰提取算法和数据库容量^[54,62].软件的迭代更新可不断提高 PFAS 非靶标识别的数量与准确性,也会进一步减少识别结果的差异.Fluoromatch 软件更新频率较高,目前使用该软件进行 PFAS 非靶标识别的研究在逐步增加.数据采集模式也是影响识别效果的一个主要因素,PARTINGTON 等^[63]研究表明一级质谱扫描、DIA 和 DDA 适用于不同识别目的,需要在扫描精度、鉴定置信度和识别数量上进行取舍.因此,研究者可根据实验数据情况和预期功能来选择软件进行数据分析.

4 结论与展望

PFAS 的广泛使用及其在环境中的普遍检出对生态环境与人体健康形成潜在危害,而非靶标 PFAS 的高效识别是进行 PFAS 分析和管控的重要基础.本文系统总结了利用高分辨率质谱数据开展 PFAS 非靶标识别的技术框架、识别策略应用情况及优缺点,全面梳理了 PFAS 非靶标识别开源软件的特点及应用实践.对于 PFAS 非靶标识别技术的研发,将来可从以下方面展开:

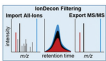
(1)扩大非靶标分析涵盖化学空间.样品的前处理方法、色谱和质谱条件等分析过程均可影响所涵盖的 PFAS 种类,但目前尚需构建简易通用工作流程,用于提取样品中绝大部分 PFAS 组分.同时,也需要开展与其他分析方法(总可氧化前体、红外光谱、核磁共振氟谱等)的联合使用研究,用来更好解释样品中有机氟的组成和来源.

(2)开发准确高效智能的分析软件.目前商用软件和开源软件在峰提取算法、分析速度、识别置信度等方面均存在优化的空间,需要开发更多数据“无损”过滤方法来进一步降低识别假阳性率和人工干扰.人工智能和机器学习可辅助提高 PFAS 特征的识别覆盖率和准确率.基于 PFAS 的数量和类型,可靠的自动评级、自动分组与数据可视化功能也将大幅减少人工处理时间.

(3)拓展 PFAS 非靶标数据的应用.目前纳入详细讨论分析的 PFAS 非靶标数据主要用于未知 PFAS 识别,环境归趋与溯源.可以将其他学科如毒理学、医学等学科的数据与 PFAS 非靶标数据结合,实现对跨学科

问题的多角度研究,如与定量构效关系预测和效应导向分析结合可实现“有毒”PFAS 的准确鉴定,与代谢组学非靶标数据的结合可扩大对未知 PFAS 在化学暴露组中重要性的认识。

表 2 PFAS 非靶标识别开源软件特征

Tab. 2 Characteristics of open-source nontarget identification software for PFAS			
软件名称	峰提取算法	数据采集模式	软件主要特点
Fluoromatch 	MZmine	DDA	庞大的 PFAS 分类数据库和碎片库
Envimass 	Envimass	DDA/DIA	同系物筛查、源内片段识别
FindPFAS/SPFAS Screen 	OpenMS	DDA	同系物筛查、碎片差异分析、碳数归一化质量亏损过滤
APP ID 	MS-DIAL	DDA	PFAS 分子网络识别、一级质谱精确搜索、二级质谱结构预测
SWATH F 	CorrDec/MS-DIAL	SWATH	SWATH 数据解卷积后的同系物筛查与识别
IonDecon 	MZmine/IonDecon	DIA	将 DIA 数据解卷积为 DDA 数据

参 考 文 献

[1] WANG Z Y, BUSER A M, COUSINS I T, et al. A new OECD definition for per- and polyfluoroalkyl substances[J]. Environmental Science & Technology, 2021, 55(23): 15575-15578.

[2] SCHYMANSKI E L, ZHANG J, THIESSEN P A, et al. Per- and polyfluoroalkyl substances (PFAS) in PubChem: 7 million and growing [J]. Environmental Science & Technology, 2023, 57(44): 16918-16928.

[3] GLÜGE J, SCHERINGER M, COUSINS I T, et al. An overview of the uses of per- and polyfluoroalkyl substances (PFAS) [J]. Environmental Science Processes & Impacts, 2020, 22(12): 2345-2373.

[4] COUSINS I T, JOHANSSON J H, SALTER M E, et al. Outside the safe operating space of a new planetary boundary for per- and polyfluoroalkyl substances (PFAS) [J]. Environmental Science & Technology, 2022, 56(16): 11172-11179.

[5] EVICH M G, DAVIS M J B, MCCORD J P, et al. Per- and polyfluoroalkyl substances in the environment [J]. Science, 2022, 375(6580): eabg9065.

[6] ACKERMAN GRUNFELD D, GILBERT D, HOU J, et al. Underestimated burden of per- and polyfluoroalkyl substances in global surface waters and groundwaters [J]. Nature Geoscience, 2024, 17: 340-346.

[7] PLACE B J, FIELD J A. Identification of Novel Fluorochemicals in Aqueous Film-Forming Foams Used by the US Military [J]. Environmental Science & Technology, 2012, 46(13): 7120-7127.

[8] WANG Q, RUAN Y F, YUEN C N T, et al. Tracing per- and polyfluoroalkyl substances (PFASs) in the aquatic environment: target analysis and beyond [J]. TrAC Trends in Analytical Chemistry, 2023, 169: 117351.

[9] YU N Y, GUO H W, YANG J P, et al. Non-target and suspect screening of per- and polyfluoroalkyl substances in airborne particulate matter in China [J]. Environmental Science & Technology, 2018, 52(15): 8205-8214.

[10] ZWEIGLE J, BUGSEL B, RÖHLER K, et al. PFAS-contaminated soil site in Germany: nontarget screening before and after direct TOP assay by Kendrick mass defect and FindPFAS [J]. Environmental Science & Technology, 2023, 57(16): 6647-6655.

[11] SPAAN K M, VAN NOORDENBURG C, PLASSMANN M M, et al. Fluorine mass balance and suspect screening in marine mammals from the Northern Hemisphere [J]. Environmental Science & Technology, 2020, 54(7): 4046-4058.

[12] LI Y Q, YU N Y, DU L T, et al. Transplacental transfer of per- and polyfluoroalkyl substances identified in paired maternal and cord sera using suspect and nontarget screening [J]. Environmental Science & Technology, 2020, 54(6): 3407-3416.

[13] JIAO E M, LARSSON P, WANG Q, et al. Further insight into extractable (organo) fluorine mass balance analysis of tap water from

- Shanghai, China[J]. *Environmental Science & Technology*, 2023, 57(38): 14330-14339.
- [14] CHEN Y J, WANG R D, SHIH Y L, et al. Emerging perfluorobutane sulfonamido derivatives as a new trend of surfactants used in the semiconductor industry[J]. *Environmental Science & Technology*, 2024, 58(3): 1648-1658.
- [15] HARRIS K J, MUNOZ G, WOO V, et al. Targeted and suspect screening of per- and polyfluoroalkyl substances in cosmetics and personal care products[J]. *Environmental Science & Technology*, 2022, 56(20): 14594-14604.
- [16] LIU T, HU L X, HAN Y, et al. Non-target discovery and risk prediction of per- and polyfluoroalkyl substances (PFAS) and transformation products in wastewater treatment systems[J]. *Journal of Hazardous Materials*, 2024, 476: 135081.
- [17] GUO J, HUAN T. Comparison of full-scan, data-dependent, and data-independent acquisition modes in liquid chromatography-mass spectrometry based untargeted metabolomics[J]. *Analytical Chemistry*, 2020, 92(12): 8072-8080.
- [18] LIU Y N, DAGOSTINO L A, QU G B, et al. High-resolution mass spectrometry (HRMS) methods for nontarget discovery and characterization of poly- and per-fluoroalkyl substances (PFASs) in environmental and human samples[J]. *TrAC Trends in Analytical Chemistry*, 2019, 121: 115420.
- [19] BUGSEL B, ZWEIGLE J, ZWIENER C. Nontarget screening strategies for PFAS prioritization and identification by high resolution mass spectrometry: a review[J]. *Trends in Environmental Analytical Chemistry*, 2023, 40: e00216.
- [20] 钱慧敏, 刘艳娜, 姚林林, 等. 非靶标技术在新污染物识别中的应用[J]. *环境化学*, 2024, 43(2): 363-376.
- QIAN H M, LIU Y N, YAO L L, et al. Recent advances in nontarget discovery of emerging pollutants in the environment[J]. *Environmental Chemistry*, 2024, 43(2): 363-376.
- [21] SCHYMANSKI E L, JEON J, GULDE R, et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence[J]. *Environmental Science & Technology*, 2014, 48(4): 2097-2098.
- [22] CHARBONNET J A, MCDONOUGH C A, XIAO F, et al. Communicating confidence of per- and polyfluoroalkyl substance identification via high-resolution mass spectrometry[J]. *Environmental Science & Technology Letters*, 2022, 9(6): 473-481.
- [23] GHORBANI GORJI S, GÓMEZ RAMOS M J, DEWAPRIYA P, et al. New PFASs identified in AFFF impacted groundwater by passive sampling and nontarget analysis[J]. *Environmental Science & Technology*, 2024, 58(3): 1690-1699.
- [24] PLACE B J, ULRICH E M, CHALLIS J K, et al. An introduction to the benchmarking and publications for non-targeted analysis working group[J]. *Analytical Chemistry*, 2021, 93(49): 16289-16296.
- [25] JOERSS H, MENDER F. The complex 'PFAS world'—How recent discoveries and novel screening tools reinforce existing concerns[J]. *Current Opinion in Green and Sustainable Chemistry*, 2023, 40: 100775.
- [26] LI X R, CUI D N, NG B, et al. Non-targeted analysis for the screening and semi-quantitative estimates of per- and polyfluoroalkyl substances in water samples from South Florida environments[J]. *Journal of Hazardous Materials*, 2023, 452: 131224.
- [27] YOUNG R B, PICA N E, SHARIFAN H, et al. PFAS analysis with ultrahigh resolution 21T FT-ICR MS: suspect and nontargeted screening with unrivaled mass resolving power and accuracy[J]. *Environmental Science & Technology*, 2022, 56(4): 2455-2465.
- [28] KIND T, FIEHN O. Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry[J]. *BMC Bioinformatics*, 2007, 8: 105.
- [29] TANG C M, LIANG Y T, WANG K, et al. Comprehensive characterization of per- and polyfluoroalkyl substances in wastewater by liquid chromatography-mass spectrometry and screening algorithms[J]. *NPJ Clean Water*, 2023, 6: 6.
- [30] JIAO Z Y, YU N Y, MAO J D, et al. The occurrence, tissue distribution, and PBT potential of per- and polyfluoroalkyl substances in the freshwater organisms from the Yangtze River via nontarget analysis[J]. *Journal of Hazardous Materials*, 2023, 458: 131868.
- [31] MOHAMMED TAHA H, AALIZADEH R, ALYGIZAKIS N, et al. The NORMAN suspect list exchange (NORMAN-SLE): facilitating European and worldwide collaboration on suspect screening in high resolution mass spectrometry[J]. *Environmental Sciences Europe*, 2022, 34(1): 104.
- [32] KOELMEL J P, STELBEN P, MCDONOUGH C A, et al. FluoroMatch 2.0—making automated and comprehensive non-targeted PFAS annotation a reality[J]. *Analytical and Bioanalytical Chemistry*, 2022, 414(3): 1201-1215.
- [33] ZWEIGLE J, BUGSEL B, ZWIENER C. Efficient PFAS prioritization in non-target HRMS data: systematic evaluation of the novel MD/C-m/C approach[J]. *Analytical and Bioanalytical Chemistry*, 2023, 415(10): 1791-1801.
- [34] KAUFMANN A, BUTCHER P, MADEN K, et al. Simplifying nontargeted analysis of PFAS in complex food matrixes[J]. *Journal of AOAC International*, 2022, 105(5): 1280-1287.
- [35] ZWEIGLE J, BUGSEL B, ZWIENER C. FindPFAS: non-target screening for PFAS—Comprehensive data mining for MS² fragment mass differences[J]. *Analytical Chemistry*, 2022, 94(30): 10788-10796.
- [36] ZWEIGLE J, BUGSEL B, FABREGAT-PALAU J, et al. PFAScreen—an open-source tool for automated PFAS feature prioritization in non-target HRMS data[J]. *Analytical and Bioanalytical Chemistry*, 2024, 416(2): 349-362.
- [37] BARZEN-HANSON K A, ROBERTS S C, CHOYKE S, et al. Discovery of 40 classes of per- and polyfluoroalkyl substances in historical aqueous film-forming foams (AFFFs) and AFFF-impacted groundwater[J]. *Environmental Science & Technology*, 2017, 51(4):

2047-2057.

- [38] WANG Y, YU N Y, ZHU X B, et al. Suspect and nontarget screening of per- and polyfluoroalkyl substances in wastewater from a fluorocarbon chemical manufacturing park[J]. *Environmental Science & Technology*, 2018, 52(19): 11007-11016.
- [39] MUNOZ G, MICHAUD A M, LIU M, et al. Target and nontarget screening of PFAS in biosolids, composts, and other organic waste products for land application in France[J]. *Environmental Science & Technology*, 2022, 56(10): 6056-6068.
- [40] HORAI H, ARITA M, KANAYA S, et al. MassBank; a public repository for sharing mass spectral data for life sciences[J]. *Journal of Mass Spectrometry*, 2010, 45(7): 703-714.
- [41] VINAIXA M, SCHYMANSKI E L, NEUMANN S, et al. Mass spectral databases for LC/MS- and GC/MS-based metabolomics; state of the field and future prospects[J]. *TrAC Trends in Analytical Chemistry*, 2016, 78: 23-35.
- [42] BOATMAN A K, CHAPPEL J R, POLERA M E, et al. Assessing per- and polyfluoroalkyl substances in fish fillet using non-targeted analyses[J]. *Environmental Science & Technology*, 2024, 58(32): 14486-14495.
- [43] FU Y, JI Y Y, TIAN Y W, et al. Unveiling priority emerging PFAS in Taihu Lake using integrated nontarget screening, target analysis, and risk characterization[J]. *Environmental Science & Technology*, 2024, 58(42): 18980-18991.
- [44] GETZINGER G J, HIGGINS C P, FERGUSON P L. Structure database and in silico spectral library for comprehensive suspect screening of per- and polyfluoroalkyl substances (PFASs) in environmental media by high-resolution mass spectrometry[J]. *Analytical Chemistry*, 2021, 93(5): 2820-2827.
- [45] CHEN Y J, YANG J S, LIN A Y C. Comprehensive nontargeted analysis of fluorosurfactant byproducts and reaction products in wastewater from semiconductor manufacturing[J]. *Sustainable Environment Research*, 2024, 34(1): 14.
- [46] STRYNAR M, MCCORD J, NEWTON S, et al. Practical application guide for the discovery of novel PFAS in environmental samples using high resolution mass spectrometry[J]. *Journal of Exposure Science & Environmental Epidemiology*, 2023, 33(4): 575-588.
- [47] MANZ K E, FEERICK A, BRAUN J M, et al. Non-targeted analysis (NTA) and suspect screening analysis (SSA): a review of examining the chemical exposome[J]. *Journal of Exposure Science & Environmental Epidemiology*, 2023, 33(4): 524-536.
- [48] AALIZADEH R, ALYGIZAKIS N A, SCHYMANSKI E L, et al. Development and application of liquid chromatographic retention time indices in HRMS-based suspect and nontarget screening[J]. *Analytical Chemistry*, 2021, 93(33): 11601-11611.
- [49] GUARDIAN M G E, ANTLE J P, VEXELMAN P A, et al. Resolving unknown isomers of emerging per- and polyfluoroalkyl substances (PFASs) in environmental samples using COSMO-RS-derived retention factor and mass fragmentation patterns[J]. *Journal of Hazardous Materials*, 2021, 402: 123478.
- [50] 丁一. 基于机器学习实现全氟及多氟化合物非靶标快速筛查[D]. 武汉: 江汉大学, 2023.
- DING Y. Rapid non-target screening of Per- and Polyfluoroalkyl Substances based on machine learning algorithm[D]. Wuhan: Jiangnan University, 2023.
- [51] DODDS J N, HOPKINS Z R, KNAPPE D R U, et al. Rapid characterization of per- and polyfluoroalkyl substances (PFAS) by ion mobility spectrometry-mass spectrometry (IMS-MS)[J]. *Analytical Chemistry*, 2020, 92(6): 4427-4435.
- [52] SONG X C, CANELLAS E, DREOLIN N, et al. Application of ion mobility spectrometry and the derived collision cross section in the analysis of environmental organic micropollutants[J]. *Environmental Science & Technology*, 2023, 57(51): 21485-21502.
- [53] KIRKWOOD-DONELSON K I, DODDS J N, SCHNETZER A, et al. Uncovering per- and polyfluoroalkyl substances (PFAS) with nontargeted ion mobility spectrometry-mass spectrometry analyses[J]. *Science Advances*, 2023, 9(43): ead7048.
- [54] JACOB P, WANG R, CHING C, et al. Evaluation, optimization, and application of three independent suspect screening workflows for the characterization of PFASs in water[J]. *Environmental Science: Processes & Impacts*, 2021, 23(10): 1554-1565.
- [55] SADIA M, BOUDGUIYER Y, HELMUS R, et al. A stochastic approach for parameter optimization of feature detection algorithms for non-target screening in mass spectrometry[J]. *Analytical and Bioanalytical Chemistry*, 2024. DOI: 10.1007/s00216-024-05425-3.
- [56] KOELMEL J P, PAIGE M K, ARISTIZABAL-HENAO J J, et al. Toward comprehensive per- and polyfluoroalkyl substances annotation using FluoroMatch software and intelligent high-resolution tandem mass spectrometry acquisition[J]. *Analytical Chemistry*, 2020, 92(16): 11186-11194.
- [57] KOELMEL J P, STELBEN P, GODRI D, et al. Interactive software for visualization of nontargeted mass spectrometry data—FluoroMatch visualizer[J]. *Exposome*, 2022, 2(1): osac006.
- [58] MARTIN LOOS. *EnviMass* version 3.5(2019)[CP].
- [59] WANG X B, YU N Y, JIAO Z Y, et al. Machine learning-enhanced molecular network reveals global exposure to hundreds of unknown PFAS[J]. *Science Advances*, 2024, 10(21): eadn1039.
- [60] LI L H, GAO R J, WANG X B, et al. SWATH-F: a novel nontarget strategy based on the SWATH-MS deconvolution method assisting in annotating PFAS homologues in multisample studies[J]. *Analytical Chemistry*, 2023, 95(39): 14551-14557.
- [61] KOELMEL J P, KUMMER M, CHEVALLIER O, et al. Expanding per- and polyfluoroalkyl substances coverage in nontargeted analysis using data-independent analysis and IonDecon[J]. *Journal of the American Society for Mass Spectrometry*, 2023, 34(11): 2525-2537.

[62] NASON S L,KOELMEL J,ZUVERZA-MENA N,et al.Software comparison for nontargeted analysis of PFAS in AFFF-contaminated soil[J].Journal of the American Society for Mass Spectrometry,2021,32(4):840-846.

[63] PARTINGTON J M,RANA S,SZABO D,et al.Comparison of high-resolution mass spectrometry acquisition methods for the simultaneous quantification and identification of per- and polyfluoroalkyl substances(PFAS)[J].Analytical and Bioanalytical Chemistry,2024,416(4):895-912.

A review on the nontarget identification technology for PFAS by using high resolution mass spectrometry

Meng Xiangzhou^{1,2,3}, Zhang Boxuan^{1,3}, Han Baocang^{1,3}, Zhu Qinghe², Yang Jie², Zhang Minchao⁴

(1. College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China; 2. Ministry of Ecology and Environment Engineering Center for Urban Soil Contamination Control and Remediation, Shanghai Academy of Environmental Sciences, Shanghai 200233, China; 3. Jiaxing Tongji Environmental Research Institute, Jiaxing 314051, China; 4. Technical Center for Industrial Products and Raw Materials Inspection and Testing of Shanghai Customs District, Shanghai 201210, China)

Abstract: Per- and polyfluoroalkyl substances(PFAS) have been widely used in various commercial products, which raised global concerns because of their persistence, bioaccumulation, mobility, and toxicity. Large unknown PFAS in the environment are generally identified by using nontarget analysis with high-resolution mass spectrometry, however, the development of fast and efficient identification technology is difficult and urgent. This review aims to draw the framework of nontarget identification technology for PFAS, to summarize the application, advantages and disadvantages of different nontarget identification strategies, and to compare the open-source nontarget identification software. The results are helpful for the precise identification, source exploration and control of novel PFAS in the environment.

Keywords: per- and polyfluoroalkyl substances(PFAS); high-resolution mass spectrometry(HRMS); nontarget identification; suspect screening; new pollutants

[责任编辑 赵晓华 刘洋]