法医病理学 DOI：10.13406/j.cnki.cyxb.003687

海水溺死大鼠盲肠微生物群落变化及结合随机森林机器学习算法进行死亡时间推断

徐玉钊，覃小诗，周云超，哈　山，卢江寰，马一新，段智奥，陈建华，邓建强
（海南医科大学，基础医学与生命科学学院，海南省热带法医学司法鉴定工程研究中心和海南省院士工作站/热带法医学，海口　571199）

【摘　要】目的：研究海水中溺死大鼠和CO_2窒息死后入水大鼠盲肠微生物的演替规律，并结合随机森林（random forest，RF）机器学习算法进行死后淹没时间（postmortem submersion interval，PMSI）的推断。**方法：**本研究将70只健康成年SD大鼠随机分为海水溺死组（seawater drowning group，D组）和死后入水组（postmortem submersion group，PS组）。建立动物模型后置于25 ℃恒温气候箱，按照死后经历时间（0、12、24、36、48、72、96 h）对盲肠内容物进行取材，经十六烷基三甲基溴化铵（cetyltrimethylammonium bromide，CTAB）法提取DNA，细菌16S rDNA V3-V4区特异性扩增，高通量测序等得到测序数据，对盲肠微生物群落特征进行多样性分析；结合微生物组测序结果和RF机器学习算法，筛选生物标志物，建立海水溺死和死后入水组大鼠尸体的PMSI预测模型。**结果：**海水溺死组和死后入水组的盲肠微生物群落特征存在差异。结合RF机器学习算法筛选出16种推断PMSI的潜在生物标志物，并分别建立了海水溺死组（D组）和死后入水组（PS组）的PMSI推断模型。海水溺死组预测模型结果：平均绝对误差（mean absolute error，MAE）=13.272 h，R^2=0.798；死后入水组预测模型结果：MAE=9.956 h，R^2=0.793。**结论：**本研究表明，海水溺死和死后入水2种死因的盲肠微生物群落存在规律性差异，证明了结合RF机器学习算法的盲肠内容物微生物群落演替能够成为海水尸体PMSI推断的有效生物学指标。

【关键词】溺死；死后浸没时间；随机森林；微生物组学；死亡时间推断
【中图分类号】R89　　**【文献标志码】**A　　**【收稿日期】**2024-07-26

Changes of cecal microbial community in seawater drowning rats and estimation of postmortem interval in combination with random forest machine learning algorithm

Xu Yuzhao，Qin Xiaoshi，Zhou Yunchao，Ha Shan，Lu Jianghuan，Ma Yixin，Duan Zhiao，Chen Jianhua，Deng Jianqiang
（*Hainan Provincial Tropical Forensic Engineering Research Center and Hainan Provincial Academician Workstation，Tropical Forensic Medicine，School of Basic Medicine and Life Sciences，Hainan Medical University*）

【Abstract】*Objective*：To investigate the succession patterns of cecal microbiota in rats drowned in seawater and those submerged after death by CO_2 asphyxiation，and to infer postmortem submersion interval（PMSI）using the random forest（RF）machine learning algorithm. ***Methods*：**In this study，70 specific pathogen-free healthy adult Sprague-Dawley rats were randomly divided into seawater drowning group（D group）and postmortem submersion group（PS group）. The rats were placed in a constant temperature chamber at 25° C after modeling，and samples of cecal contents were collected at various postmortem time points（0，12，24，36，48，72，and 96 hours）. The CTAB method was used to extract DNA，and the V3-V4 region of bacterial 16S rDNA was specifically amplified；high-throughput sequencing was performed to obtain sequencing data and analyze the diversity of cecal microbiota community. Related biomarkers were obtained based on the results of microbiome sequencing and the RF machine learning algorithm，and then PMSI prediction models were established for the bodies of rats in both the D group and the PS group. ***Results*：**There were differences in the characteristics of cecal microbiota community between the D group and the PS group. A total of 16 potential biomarkers were obtained by the RF machine learning algorithm，and a model for inferring PMSI was established for the D group and the PS group. For the D group，

作者介绍：徐玉钊，Email：1253086902@qq.com，研究方向：病理学与法医病理学方向。
通信作者：邓建强，Email：forensicpatho@163.com。
基金项目：国家自然科学基金资助项目（编号：82060341、81560304）；海南省院士创新平台科研资助项目（编号：YSPTZX202134）；海南省研究生创新课题资助项目（编号：Qhys2023-474）。
优先出版：https://link.cnki.net/urlid/50.1046.r.20241231.1522.002（2025-01-02）

the predictive model had a mean absolute error（MAE）of 13.272 hours and an R^2 value of 0.798，while for the PS group，the predictive model had an MAE value of 8.445 hours and an R^2 value of 0.847. ***Conclusion***：This study shows that there are regular differences in cecal microbiota community between seawater drowning and postmortem submersion，and it proves that the succession of cecal content microbiota based on the RF machine learning algorithm can be used an effective biomarker for inferring the PMSI of bodies in seawater.
【Key words】drowning；postmortem submersion interval；random forest；microbiome；estimation of postmortem interval

微生物群是一个庞大的群落集合体，其中包含单细胞细菌、病毒、真菌、古细菌和真核生物。而微生物组则代表着这些生物体在生物体内所携带的遗传信息及它们之间的相互作用。作为一个动态变化的实体，人类微生物群在人的一生中会经历许多变化，以适应各种环境和个人因素，如吸烟习惯、药物治疗、饮食习惯和医疗保健状况等[1-3]。微生物群落与人类生活紧密相连，它们不仅参与人类的正常生理活动，还在死后尸体的腐败分解过程中以及死亡时间的推断中发挥着重要作用[4]。近年来，对于水中尸体死亡时间推断开展了多个方面的研究，主要包括尸体腐败评分系统[5-6]，生化检测等[7]。然而，这些方法大多数不够精确或客观。近些年，随着新一代测序技术的快速发展，特别是16s rDNA序列技术的发展，使得获得样本的微生物组整体特征信息成为可能，进而实现对微生物群落构成及其动态变化的深度分析，此项技术的发展为全面且系统地探究微生物群落赋予了前所未有的深度与广度[8]。

有研究表明，微生物群落在尸体腐败过程中发挥了重要作用，并且微生物群落及其组成可能发生规律性演替[9]。目前研究发现，尸体微生物的变化在尸体不同部位、不同死因、不同环境下有所不同，其中包括溺死和非溺死尸体间的微生物差异。海南四面环海，溺死案件相对多发，且由于气温高，尸体发现时多腐败，但目前海水尸体内部器官组织的微生物变化的研究报道缺乏。盲肠内微生物较为丰富，且其相对封闭，并便于取材。故本研究拟模拟海南岛年平均气温和湿度的环境[10-11]，构建海水溺死及死后入水动物模型，并运用16s rDNA测序技术对水中尸体盲肠的微生物群落在不同死亡时间的状况进行检验。结合生物信息学方法，探讨溺水与对照组间微生物群落的差异及其死后演替规律。此外，应用随机森林机器学习算法来推断死后淹没时间（postmortem submersion interval，PMSI），为判断海水中腐败尸体死因及推断死亡时间中相关问题进行实验性探索。

1 材料与方法

1.1 试验动物和分组

健康成年SD（sprague dawley）大鼠70只（体质量220~250 g），随机分为海水溺死组（seawater drowning group，以下简称D组）和死后入水组（postmortem submersion group，以下简称PS组）。根据前期预实验结果，D组和PS组大鼠，按照死后经历时间（0、12、24、36、48、72、96 h）分别分为7个实验亚组，每亚组各5只实验动物。

实验前，在清洁环境下饲养实验动物，保证12 h昼夜节律，自由进食、饮水，适应性饲养1周后开展实验。实验中各项操作均符合中华人民共和国国务院颁布的《中国实验动物工作管理》的规定，并通过了海南医科大学实验动物福利与伦理审查（伦理审批号：HYLL-2021-346号）。

1.2 动物模型的制备

实验中海水取自海南岛某海岸附近海水，实验地点为实验室恒温气候箱，整个实验期间的温度控制模拟海南岛年平均气温条件23.1~27.0 ℃，故本研究实验设计将恒温气候箱温度控制在25 ℃，湿度控制在80%[10-11]。动物实验开始前提取现场水样标本，-80 ℃保存备检。

溺死动物模型：实验当天，随机挑选35只大鼠，将它们放入无菌的袋子中，使其完全浸没在水中，并按照将大鼠置于水中1 min，然后在水上呼吸30 s，重复此循环直至大鼠死亡。然后将大鼠放入已提前放置于上述实验条件的气候箱内已盛有前步所取海水的透明塑料箱中。

死后入水动物模型：随机选取35只大鼠置于无菌密封空塑料箱中（箱子的顶部有孔道并通过橡胶管连接于CO_2瓶），打开CO_2瓶的通气阀向泡沫箱中通入CO_2约5 min，直至确定实验动物无活动、无呼吸、瞳孔放大且心跳停止后，再观察2 min，确认死亡。然后将大鼠放入已提前放置于上述实验条件的气候箱内、已盛有海水的透明塑料箱中。

1.3 检材的提取

分别于死后0、12、24、36、48、72、96 h共7个固定时间点分别从2种死因组中随机各取5只大鼠尸体，于无菌环境下进行解剖。提取大鼠尸体盲肠内容物至无菌冻存管中，并浸入液氮速冻，随后取出置于-80 ℃环境中保存备检。实验过程中严格按照无菌操作要求，若发现大鼠出现胸腹部皮肤腐败破溃与胸腹腔相通，则停止对该大鼠尸体的采样工作。

依据以上原则，在实验中，海水溺死组和死后入水组的0、12、24、36、48、72 h这6个时间点样本均能正常采集，海水溺死组的96 h成功采集3具大鼠尸体，死后入水组96 h时间点成功采集2具大鼠尸体，其余大鼠因腐败严重剔除出实验组。以上共计采集65具大鼠尸体的组织样本。

1.4 DNA提取，聚合酶链反应扩增

采用CTAB法提取样品的总基因组DNA。检测DNA浓度和纯度。根据浓度，用无菌水将DNA稀释至1 ng/μL。以上述得到的已稀释的DNA作为模板，使用特异性引物（*341F*：CCTACGGGNGGCWGCAG；*806R*：GGACTACHVGGGTATCTAAT）扩增细菌16S rDNA V3-V4区。所有PCR混合液加入15 μL Phusion® High-FidelityPCRMasterMix（NewEnglandBiolabs）、0.2 μmol/L引物和10 ng基因组DNA模板，在98 ℃下进行1 min的第一次变性，然后在98 ℃（10 s）、50 ℃（30 s）和72 ℃（30 s）下进行30次循环，最后在72 ℃下保持5 min。

PCR产物使用2%浓度的琼脂糖凝胶进行电泳检测。在确认PCR产物合格后，酶标仪测定PCR产物的浓度，根据PCR产物的浓度进行等量混样处理。混匀后再次使用2%的琼脂糖凝胶电泳检测，确认PCR产物的条带。最后，采用通用型DNA纯化回收试剂盒（TianGen）对目的条带进行回收。

1.5 高通量测序

采用NEB Next® Ultra™ Ⅱ FS DNA PCR-free Library Prep Kit构建文库，文库构建完成后经过Qubit和Q-PCR定量。合格的文库使用NovaSeq 6000进行PE 250上机测序。

1.6 测序数据的处理

根据Barcode序列和PCR扩增引物序列从下机数据中将各样本的数据进行拆分，截去Barcode和引物序列后使用FLASH（Version 1.2. 11）对每个样本的reads进行拼接[12]，得到的拼接序列为原始Tags数据（Raw Tags）；使用fastp软件（Version 0.23.1）对拼接得到的Raw Tags经过严格的过滤处理得到高质量的Tags数据（Clean Tags）[13]；将Clean Tags数据进行去除嵌合体序列的处理，Tags序列通过与物种注释数据库（Silva database）进行比对检测嵌合体序列，并最终去除其中的嵌合体序列，得到最终的有效数据（Effective Tags）[14]。

对以上得到的Effective Tags，使用QIIME2（Version QIIME2-202006）软件中的DADA2模块进行降噪，从而获得最终的ASVs（Amplicon Sequence Variants，扩增子序列变异）以及特征表[15]；使用QIIME2软件运用Silva 138.1数据库进行物种注释；使用QIIME2软件进行多序列比对，得到所有ASV序列的系统发生关系；最后对各样本的数据进行均一化处理，以样本中数据量最少的为标准进行均一化处理。后续的Alpha多样性分析和Beta多样性分析都是基于均一化处理后的数据。

1.7 生物信息学分析

1.7.1 Alpha多样性分析 将物种丰度、样本不同分组数据输入至Paleontological Statistics软件中进行α多样性分析，分别计算海水溺死组和死后入水组各时间点各样本的物种多样性：香农-维纳多样性指数（Shannon index）和物种丰富度（Chao-1指数）。

1.7.2 Beta多样性分析 为了衡量不同样本间物种多样性的差异，根据海水溺死组和死后入水组各样本的物种分类信息和物种丰度信息，基于weighted Unifrac距离的计算方法，使用主坐标分析（principal co-ordinates analysis，PCoA）来展示各样本微生物群落的相似性和不同死因（海水溺死和死后入水）引起的微生物群落差异。

1.7.3 生物标志物筛选、PMSI推断模型的建立 使用机器学习中的RF算法基于上述样本数据建立PMSI推断模型，为了筛选生物标志物数据集，使用“randomForest”程序包中内置函数rfcv()探索物种数量与误差之间的关系。并选用相应的生物标志物的OTU作为数据集，运用RF算法对PMSI进行预测，使用Mean Decrease Accuracy和MAE等指标来衡量物种在PMSI推断模型中的重要性。检验水准α=0.05

2 结 果

2.1 海水溺死组和死后入水组大鼠盲肠微生物群落演替

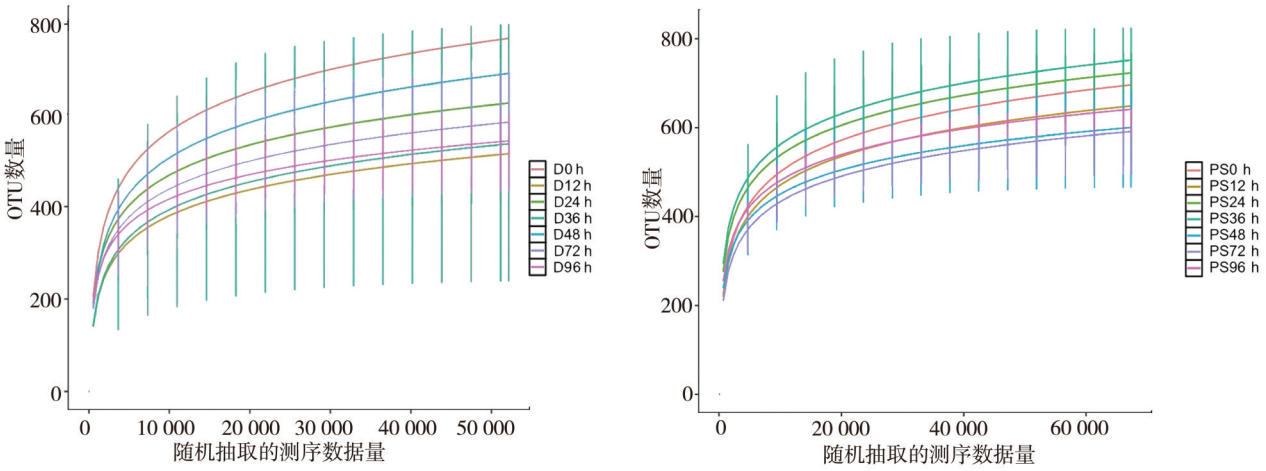
2.1.1 测序结果 海水溺死组和死后入水组共收集到65例盲肠内容物样本，其中，海水溺死组33例，死后入水组32例。通过特异性扩增16S rDNA V3-V4区域，并对产物进行高通量测序，得到原始下机数据后进行数据拼接、过滤及均一化等处理。海水溺死组（图1A）和死后入水组（图1B）的稀释曲线均显示，随着测序深度的增加新的ASV鉴定率逐渐降低，这意味着具有良好的序列覆盖率。

2.1.2 微生物群落构成及分布丰度分析 在门、纲、目、科、属、种6个分类水平对海水溺死组和死后入水组大鼠盲肠微生物分别选择相对丰度最高的前10个物种做柱状图进行分析。

在门分类水平上（图2A、2G），厚壁菌门（Firmicutes）0 h时的相对丰度在海水溺死组和死后入水组中均最高，并随后整体呈下降趋势；拟杆菌门（Bacteroidetes）在0 h时的相对丰度仅次于厚壁菌门，并随PMSI增加，整体呈现上升趋势。尽管厚壁菌门的相对丰度整体呈下降趋势，在死后入水组中，厚壁菌门在0、12、24、36、48、72、96 h时的相对丰度均最高；但在海水溺死组中，随着死亡时间的增加，拟杆菌门（Bacteroidetes）的相对丰度上升至第一位。

在纲分类水平上（图2B、2H），Bacteroidia、Clostridia和Bacilli的相对丰度在海水溺死组和死后入水组中均较高，Bacteroidia的相对丰度在海水溺死组和死后入水组中，随PMSI增加，整体呈现上升趋势。Bacilli的相对丰度在死后入水组的相对丰度远高于海水溺死组，且随着PMSI增加，在海水溺死组中其相对丰度整体呈下降趋势，而在死后入水组中呈升高趋势。随着PMSI增加，Clostridia的相对丰度在海水溺死组中整体呈现先下降后升高的趋势，且在36 h的相对丰度最低；而在死后入水组中，其相对丰度趋于稳定。

在目分类水平上（图2C、2I），海水溺死组和死后入水组大鼠各时间点中，Bacteroidales的相对丰度均保持较高水平，且在12 h组的溺死组大鼠的相对丰度最高，超过了总物种丰度的50%；0 h的溺死组大鼠中，乳杆菌目（Lactobacillales）的



A. D组盲肠菌群稀释曲线　　B. PS组盲肠菌群稀释曲线

图1　大鼠盲肠菌群稀释曲线

相对丰度远远低于其他溺死组以及死后入水组。

在科分类水平上（图2D、2J），Muribaculaceae在海水溺死组中的相对丰度保持较高水平，而其他科的相对丰度则相对较低。在死后入水组中，除了Muribaculaceae以外，毛螺菌科（Lachnospiraceae）和乳杆菌科（Lactobacillaceae）的相对丰度也保持较高水平，且在0 h时，乳杆菌科的相对丰度最高。并且在12 h前后的相对丰度达到高峰，随PMSI延长其相对丰度有下降趋势。在48 h后，海水溺死组和死后入水组的Muribaculaceae、毛螺菌科（Lachnospiraceae）的相对丰度均出现不同程度的下降趋势；乳杆菌科（Lactobacillaceae）在死后入水组中也显示出上述丰度下降的规律，但是在海水溺死组中，乳杆菌科的相对丰度在48 h到72 h间明显增加。

在属分类水平上（图2E、2K），海水溺死组与死后入水组在TOP10物种相对丰度上共有6个相同属，分别为：*Ligilactobacillus*、*Bacteroides*、*Lachnospiraceae_NK4A136_group*、*Lactobacillus*、*Dubosiella*、*Romboutsia*。海水溺死组中，0 h组时，杜波西氏菌属（*Dubosiella*）和罗姆布茨菌属（*Romboutsia*）的相对丰度保持较高水平，随着PMSI的增加，相对丰度呈减少趋势，*Ligilactobacillus*随PMSI的增加，相对丰度逐渐升高，在36 h组相对丰度达到较高水平，随后其相对丰度有下降趋势；0 h组中，阿克曼氏菌（*Akkermansia*）的相对丰度很低，但是随着PMSI的增加，其后各分组中均存在明显的相对丰度。死后入水组中，0 h时，*Lactobacillus*的相对丰度最高，并随着PMSI的增加逐渐下降；*Lachnospiraceae_NK4A136_group*在12 h组的相对丰度达到峰值，随后相对丰度有下降趋势。0 h组的正常未腐败盲肠内容物的微生物群落，海水溺死组，乳酸杆菌（*Lactobacillus*）是非优势种，随着浸没时间的延长，其相对丰度明显增加，而死后入水组中，0 h组时，乳酸杆菌（*Lactobacillus*）是明显的优势属，随着浸没时间的延长，其相对丰度明显减少。

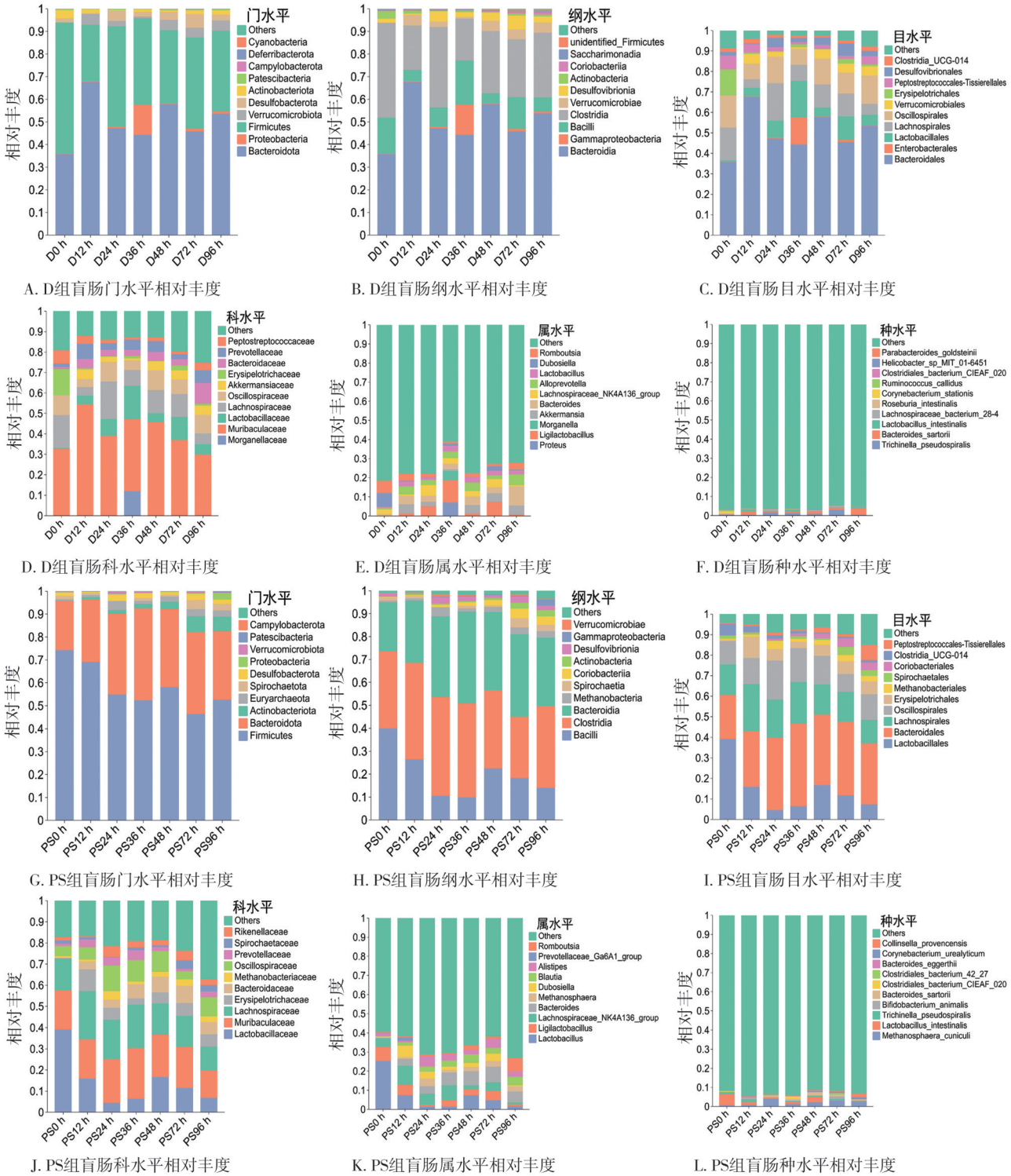
在种的分类水平上（图2F、2L），海水溺死组与死后入水组在TOP10物种相对丰度上有4个相同物种。在0 h时，海水溺死组中，*Corynebacterium_stationis*和*Roseburia_intestinalis*的相对丰度较高，随后其丰度呈明显下降趋势，*Bacteroides_sartorii*和*Trichinella_pseudospiralis*的相对丰度随死亡时间的增加，整体呈上升趋势，且72 h时，*Trichinella_pseudospiralis*的相对丰度达到最高，96 h时*Bacteroides_sartorii*的相对丰度达到最高；死后入水组中，0 h时，*Lactobacillus_intestinalis*的相对丰度最高，并随后呈下降趋势；随着死亡时间的增加，*Methanosphaera_cuniculi*的相对丰度呈逐渐升高趋势并趋于稳定。

2.1.3　盲肠内容物微生物群落Alpha多样性分析结果　基于海水溺死组和死后入水组样本进一步分析盲肠中微生物群落的Alpha多样性，基于OTU水平对每个样本进行Alpha多样性分析，来反映盲肠微生物群落的物种丰度和物种多样性。运用Kruskal-Wallis秩和检验进行组间比较，并绘制Aphla多样性指数箱线图（图3）。

反映群落分布丰富度的指数为Chao-1指数。在海水溺死组中（图3A），死亡后12 h时的物种丰富度最低，且组间有明显差异（$P<0.01$）。说明在0 h后，随着时间的推移，肠道菌群的分解导致微生物菌群的下降。在死后入水组中（图3C），48 h时的物种丰富度最低，且组间有明显差异（$P<0.01$）。

物种多样性指数反映的是物种丰富度和均匀度的综合情况，主要指数有Shannon指数。图中显示，海水溺死组的Shannon指数（图3B）变化不明显，表示0～12 h间，物种多样性存在明显差异（$P<0.01$），其组间多样性没有明显差异；死后入水组的物种多样性（图3D），随着死亡时间的增加，在36 h时达到高峰，随后明显下降（$P<0.01$）。

2.1.4　盲肠内容物微生物群落PCoA分析结果　根据海水溺死组和死后入水组各样本的物种分类信息和物种丰度信息，基于weighted Unifrac距离的计算方法，使用主坐标分析（PCoA）来展示各样本微生物群落的相似性和不同死因（海水溺死和死后入水）引起的微生物群落差异（图4）。在海水溺死组中，在海水溺死组中，导致大鼠腐败尸体微生物群落结构变化的主要因素是第一主成分，贡献值为79.9%；第二主成分则贡献值为9.9%（参见图4A）。总体而言，除个别样本存在差异外，各个时间点分组的样本，在PCoA1的方向上，组间的差异相对较小；在PCoA2的方向上，0 h组和24 h



A. D组盲肠门水平相对丰度　B. D组盲肠纲水平相对丰度　C. D组盲肠目水平相对丰度

D. D组盲肠科水平相对丰度　E. D组盲肠属水平相对丰度　F. D组盲肠种水平相对丰度

G. PS组盲肠门水平相对丰度　H. PS组盲肠纲水平相对丰度　I. PS组盲肠目水平相对丰度

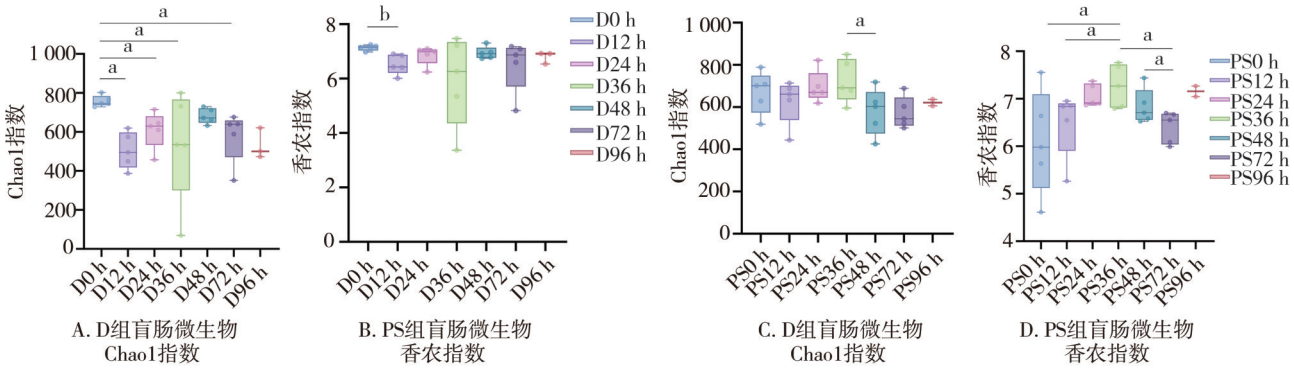
J. PS组盲肠科水平相对丰度　K. PS组盲肠属水平相对丰度　L. PS组盲肠种水平相对丰度

图2　大鼠盲肠微生物各分类水平相对丰度TOP10

组与其他时间分组存在差异。在死后入水组中，对于腐败尸体的微生物群落结构，第一主成分贡献值为41.1%，第二主成分贡献值为17.2%（图4B）。在PCoA1的方向上，0 h组样本与其他组样本的微生物群落存在一定差异，而其他组间差异较小；在PCoA2的方向上，除个别样本外，各组间的差异较小。

2.2　随机森林分析以及PMSI推断模型的建立

2.2.1　随机森林参数调优　使用R包randomForest对数据进行随机森林模型拟合参数调优图（图5），由图可见无论是单个组别或是所有组别分类误差率随着randomForest函数中决策树数量（treeNum参数）的增大而趋于稳定。拟合模型中决策树数量选择5 000。



注：注：a，组间比较，$P<0.01$；b，组间比较，$P<0.05$

图3　大鼠盲肠微生物 Alpha 多样性分析图

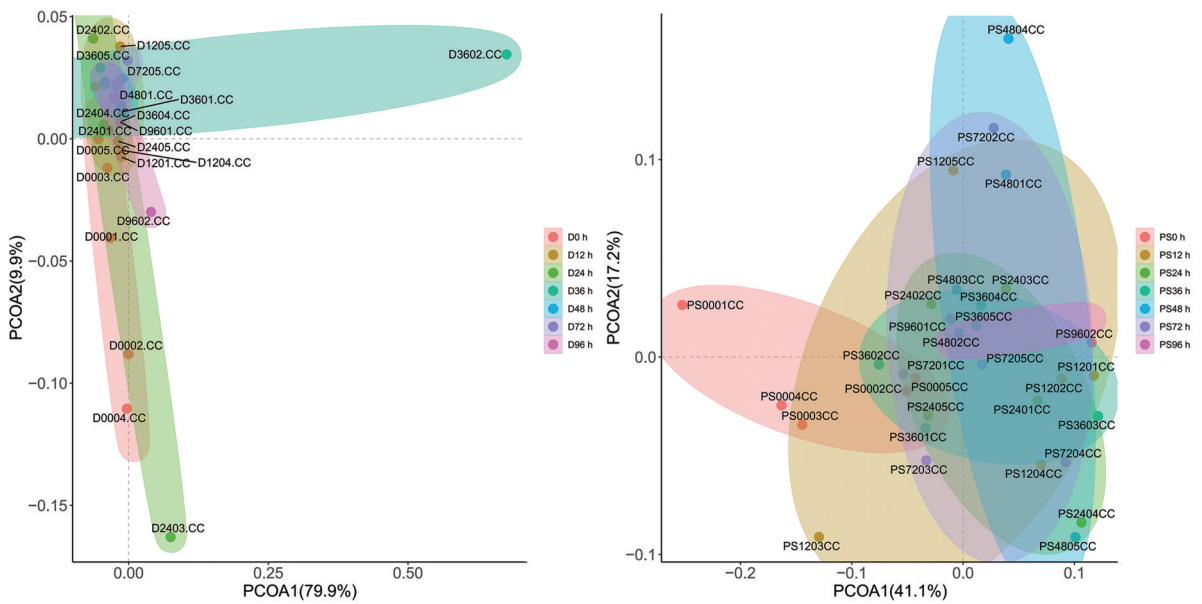


图4　大鼠盲肠微生物主坐标分析图

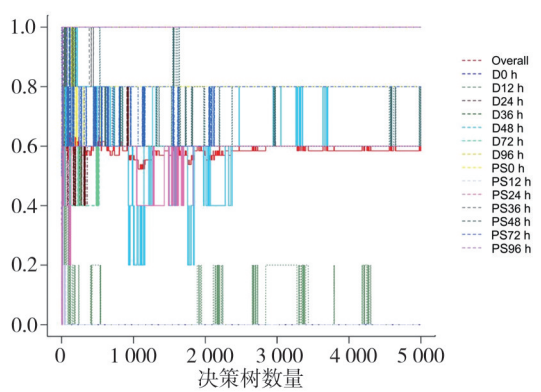
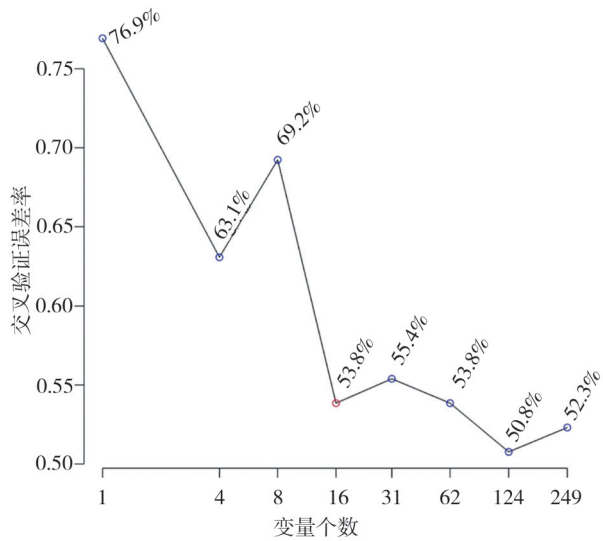


图5　随机森林参数调优图

2.2.2　交叉验证与模型误差率分析　本研究进一步通过 R 包 randomForest 中 rfcv 函数做 10 折交叉验证并重复 1 次获得变量选取与模型分类平均误差率的关系，见图 6。交叉验证的结果显示，随着模型中变量数量升高，模型的误差迅速下降，综合考虑样本量和误差率，以 16 个细菌属来（红圈处对应的 x 轴数值）建立的模型误差最合适。



注：红圈处表明，模型所使用的样本量为 16 个时，所对应的模型误差较小，且合适用来预测分析

图6　模型误差率图

2.2.3 潜在生物标志物的筛选 结合交叉验证结果，选择合适个数的差异特征进行下一步建模。数量选择上可根据上述图6，选择红圈处对应的x轴数值作为相应的特征个数。依据平均精确度减少（Mean Decrease Accuracy）的值从大到小筛选16个细菌为潜在生物标志物，并作图7。16个PMSI推断的潜在标志物见表1。

2.2.4 大鼠RF回归模型的建立 使用python建立RF回归分别对海水溺死组大鼠和死后入水组大鼠进行回归分析，选择上述16个种属的大鼠盲肠微生物数据集中70%的数据作为训练数据集，30%的数据作为验证数据集，并判别正确的大鼠腐败盲肠微生物数据作为测试数据集，最终选择14个属的大鼠盲肠微生物对海水溺死组大鼠的PMSI进行RF回归模型预测；选择16个属的大鼠盲肠微生物对死后入水组大鼠的PMSI进行RF回归模型预测。

结果显示，海水溺死组中：RF回归模型训练集相关系数R^2=0.891，平均绝对误差（mean absolute error，MAE）=6.496；验证集相关系数R^2=0.798，MSE为218.102，MAE=13.272。

死后入水组中：RF回归模型训练集相关系数R^2=0.933，MAE为5.43；验证集相关系数R^2=0.793，MAE=9.956。两组RF预测模型的相应数据如下：

①特征重要性。图8展示了海水溺死组和死后入水组RF预测模型中，各生物标志物的重要性比例。在海水溺死组的预测模型中，Anaerovorax的重要性最高，为24.7%；在死后入水组的预测模型中，Monoglobus的重要性高达38.3%。

②模型评估结果。表2、表3分别展示了海水溺死组和死后入水组的模型评估结果，包含训练集和测试集的预测评价指标，通过量化指标来衡量随机森林的预测效果。结果显示：

海水溺死组中：RF回归模型训练集相关系数R^2=0.891，MSE为81.612，MAE为6.496；验证集相关系数R^2=0.798，MSE为218.102，MAE为13.272。

死后入水组中：RF回归模型训练集相关系数R^2=0.933，MSE为51.415，MAE为5.43；验证集相关系数R^2=0.793，MSE为144.892，MAE为9.956。两组RF预测模型的相应数据见下表2、表3。

表1 16个PMSI推断的潜在生物标志物

筛选的16个PMSI推断的潜在生物标志物					
Methanosphaera	*Allobaculum*	*Barnesiella*	*Corynebacterium*	*Adlercreutzia*	*Anaerovorax*
Christensenellaceae_R-7_group	*Veillonella*	*Monoglobus*	*Staphylococcus*	*Jeotgalicoccus*	*Clostridium_sensu_stricto_1*
Candidatus_Saccharimonas	*Nosocomiicoccus*	*Paraprevotella*	*Streptococcus*		

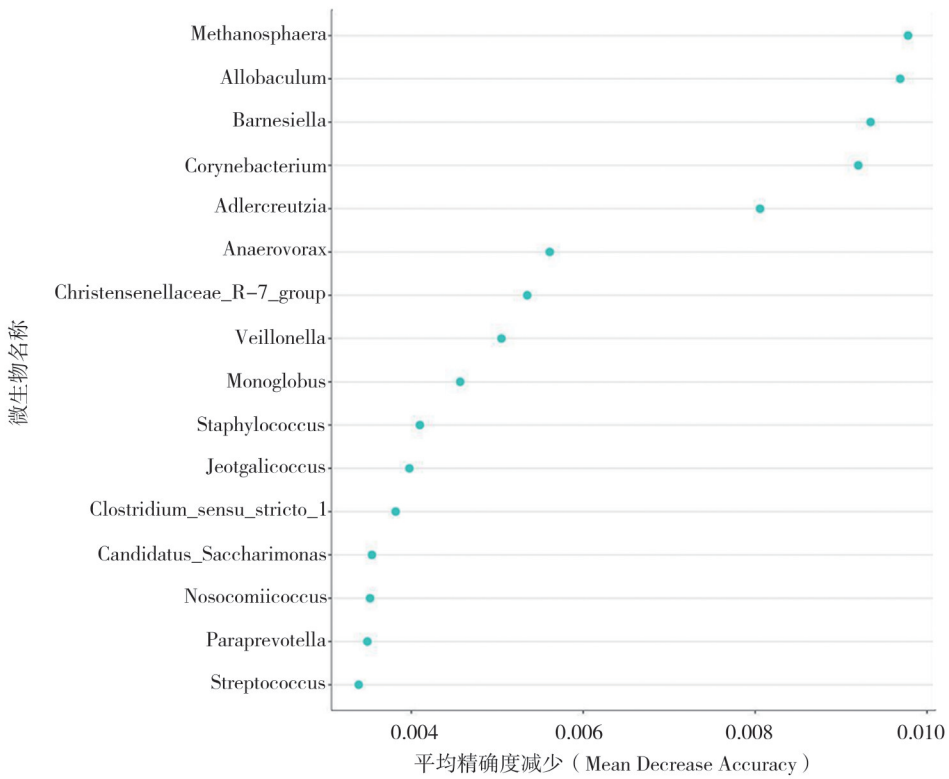
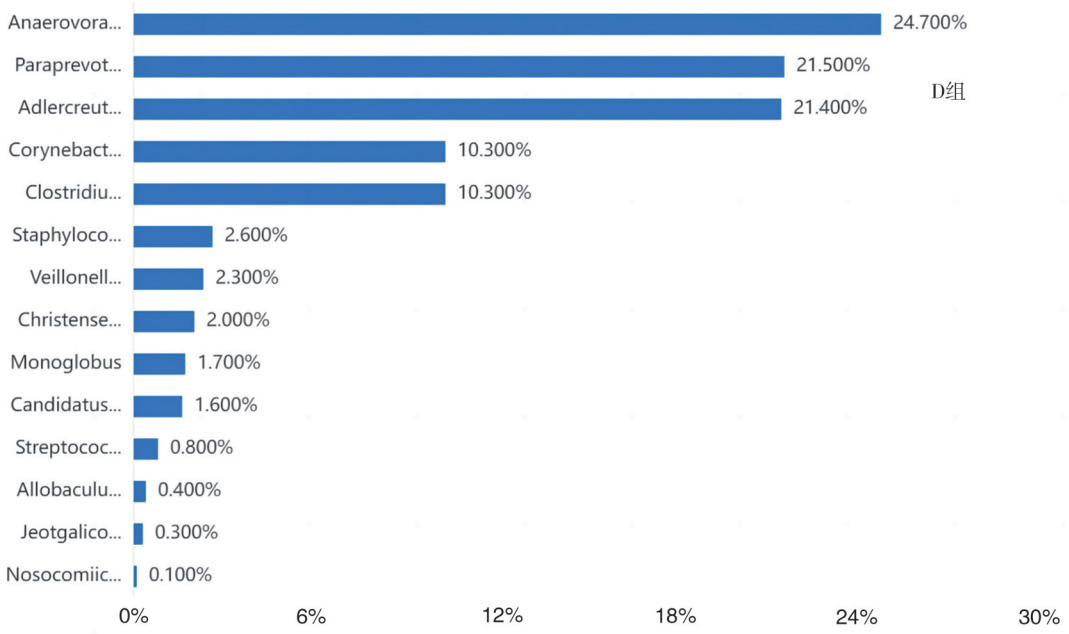
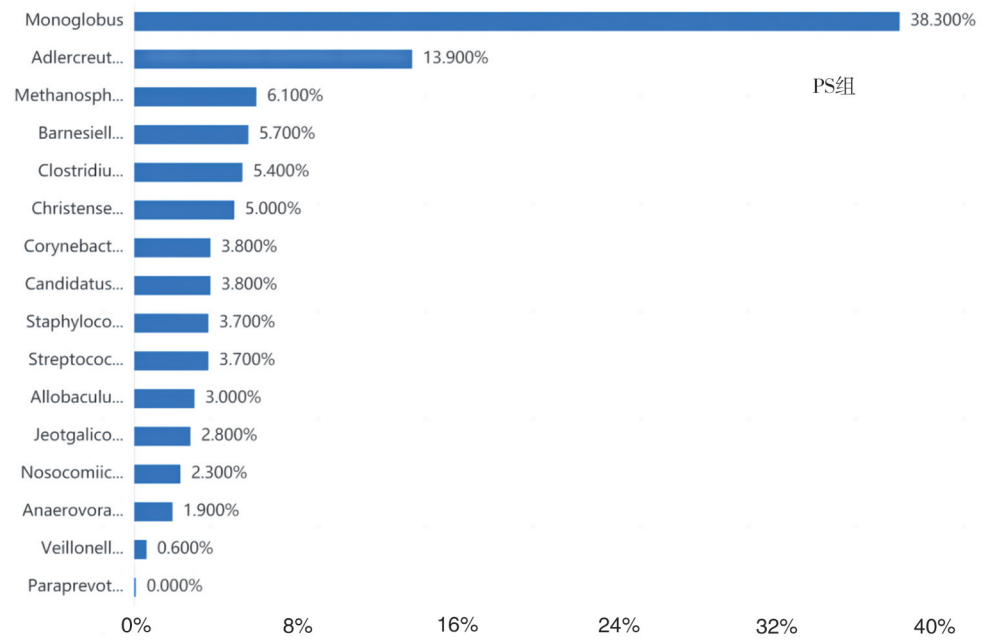


图7 大鼠盲肠菌群重要变量图



A. D组特征标志物比例图



B. PS组特征标志物比例图

图 8　特征标志物重要性比例图

表 2　海水溺死组预测模型评估结果

海水溺死组模型	MSE	RMSE	MAE	MAPE	R^2
训练集	81.612	9.034	6.496	24.076	0.891
测试集	218.102	14.768	13.272	51.050	0.798

表 3　死后入水组预测模型评估结果

死后入水组模型	MSE	RMSE	MAE	MAPE	R^2
训练集	51.415	7.170	5.430	30.661	0.933
测试集	144.892	12.037	9.956	34.569	0.793

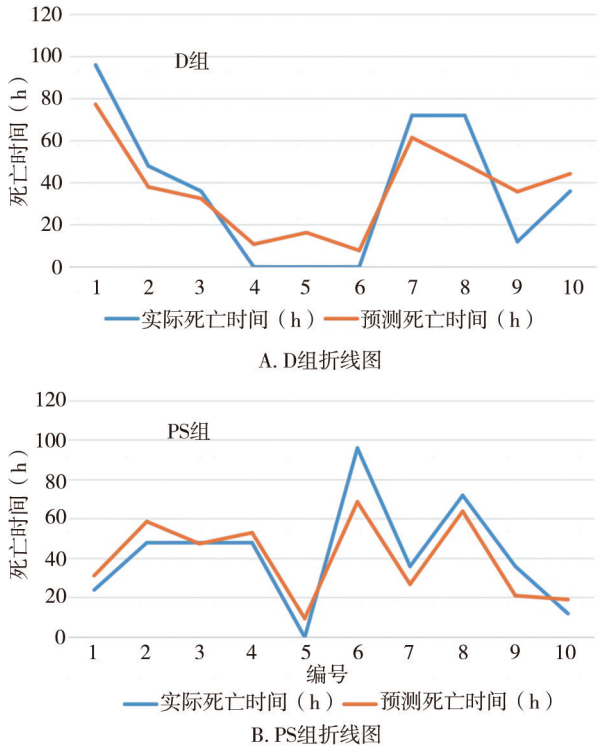
③测试集预测结果。表 4、表 5 分别展示了海水溺死组（D 组）和死后入水组（PS 组）的随机森林模型对测试数据的预测情况。并根据此测试集的数据，做出以下实际 PMSI（真实值）和预测的 PMSI（预测值）的折线图（图 9），蓝色折线为实际 PMSI，橙色折线为预测的 PMSI，图中展示了随机森林模型对测试数据的预测情况。从图中可看出，死后入水组的预测模型的预测值与真实值更相近。

表 4　海水溺死组预测结果表

D组预测测试集结果(h)	实际死亡时间(h)
77.28	96
38.04	48
32.52	36
10.80	0
16.32	0
7.80	0
61.44	72
48.96	72
35.76	12
44.28	36

表 5　死后入水组预测结果表

PS组预测测试集结果(h)	实际死亡时间(h)
31.30	24
58.72	48
47.48	48
53.08	48
9.54	0
68.70	96
26.92	36
63.94	72
21.16	36
19.12	12



A. D组折线图

B. PS组折线图

图 9　测试数据 PMSI 预测与实际 PMSI 比较折线图

3　讨　论

3.1　大鼠盲肠内容物微生物群落分析

法医微生物学是法医调查中评估水中尸体 PMSI 的一种新兴工具[16-17]。之前有研究检验了小鼠、猪和人类遗体的皮肤、骨骼、尸体相关土壤和腹部等的微生物群落，这些研究中的微生物群落会随着周围的生物和环境条件（例如天气、昆虫可及性和尸体位置）而波动[18-21]。盲肠靠近肠道的下段，相对封闭不易受外界环境影响，微生物丰富，且靠近肛门，便于取材，目前关于死后微生物群落在海水尸体中的盲肠微生物群落演替和死亡时间推断仍然缺乏。考虑到上述环境因素等的影响，以及物种的组成与相对丰度变化会导致微生物群落演替。本研究以盲肠样本为例，对海水溺死大鼠和死后入水大鼠进行细菌群落演替分析。

研究发现，在门水平上，海水溺死组与死后入水组中厚壁菌门（Firmicutes）、拟杆菌门（Bacteroidetes）和放线菌门（Actinobacteria）占主导地位，这与以往研究的几篇报道相吻合[22]。2013 年，Hyde ER 等[16]在腐败的口腔中发现，细菌群落主要以厚壁菌门和变形菌门为主；放线菌通过分解作用和腐殖质的形成，在难降解生物成分的再循环中起着至关重要的作用，2015 年，Cobaugh KL 等[23]和 Damann FE 等[24]的研究证明，一组丝状放线菌是重要的多糖、蛋白质和脂质的分解者，通过产生胞外酶，并在坟墓土壤和尸体腐烂晚期的骨头中发现。

在科水平上，Muribaculaceae 在海水溺死组和死后入水组中的相对丰度均保持较高水平，Muribaculaceae 含量显著性升高，该菌属为新确立的菌属名，曾被称为 S24-7 家族，在鼠肠道中含量较高，属于拟杆菌门，Muribaculaceae 的变化主要与各种饮食治疗、宿主条件或啮齿动物定殖过程有关[25]。此外，毛螺菌科（Lachnospiraceae）和乳杆菌科（Lactobacillaceae）的相对丰度也保持较高水平。在厚壁菌门中，毛螺菌科、乳杆菌科水解淀粉和其他糖以产生丁酸盐和其他短链脂肪酸。毛螺菌科是一种潜在的有益菌，参与多种碳水化合物的代谢，比如水果中果胶的代谢，发酵导致乙酸和丁酸的产生，为宿主提供能量的主要来源[26]。食用富含纤维的植物性饮食与毛螺菌科（Lachnospiraceae）、Roseburia

和 Prevotella 的丰度增加可能与短链脂肪酸产量增加相关。本研究显示出的48h后，两组别的 Muribaculaceae、毛螺菌科（Lachnospiraceae）的相对丰度均出现不同程度的下降趋势，这种现象也出现于淡水和陆地环境中人体和小鼠尸体腐败过程[27-29]。有研究表明，梭菌科和梭菌属是陆地和水生环境中尸体分解的关键物种[16-17,30]，在本研究中同样可观察到其较高的相对丰度。同时，也可观察到海水溺死组乳杆菌科（lactobacillus）相对丰度在48~72 h间略升高，与上述陆地尸体情况相反。

在属的分类水平上，海水溺死组和死后入水组的物种相对丰度存在显著差异，并且阿克曼氏菌（*Akkermansia*）和 *Alloprevotella* 在海水溺死组中的相对丰度明显高于死后入水组，这2种菌主要常见于肠道内，其中阿克曼氏菌又称为嗜粘蛋白-阿克曼氏菌，这种菌是在2004年从人类粪便中发现的，是一种革兰氏阴性菌[31]。随着PMSI的延长，物种组成和相对物种丰度的变化使得微生物群落发生演替。在本研究中，海水溺死组和死后入水组的乳杆菌属（*Lactobacillus*）在0 h时差异明显，表明水源性细菌进入胃肠道，对盲肠微生物群落产生影响。

在 Alpha 多样性分析中，12~36 h 间，死后入水组微生物的多样性差异显著（$P<0.05$）。结合上述微生物群落物种丰度分析，且0 h时海水溺死组的物种丰富度明显高于死后入水组，同样表明海水溺死组的盲肠微生物群落受到了水源性细菌的影响。

3.2 随机森林预测 PMSI 的相关分析

RF是一种机器学习技术，该模型通过随机特征选择和未修剪分类树集合进行分类或回归[32]。它基于独立决策树，通过多数投票得出结果，旨在最大化样本间的区分度[33-34]。作为有监督学习算法，RF通过构建分类标签分布模型，利用预测器特征对测试实例分类，其中预测器特征已知但类标签未知。RF具有快速训练、易调整、抗噪声和防止过拟合的特点[35]。

以往学者曾利用类或门级分类数据集用机器学习方法预测 PMI 或 PMSI。2013 年，Metcalf JL 等[36]利用皮肤微生物数据集建立 RF 模型，在 34 d 内估计 PMI 误差为 2.52~3.30 d；2016 年，Johnson HR 等[37]结合微生物样本，用 k 近邻回归器预测 PMI 平均误差约 3 d；Wallace JR 等[38]研究猪尸体表面细菌群落演替，建立 PMSI 推断模型解释 77.2% 变异；2022 年，Zhang FY 等[39]基于肠道微生物群落建立的模型能够解释 93.42% 的变异，进一步证明了利用水中尸体的菌群微生物群落结合机器学习建立死亡时间推断模型的可行性。

在本研究中，首先运用随机森林分析，筛选出了16种可能用于PMSI推断的细菌作为生物标志物，并基于这些生物标志物来建立海水中的大鼠尸体PMSI推断模型，并分为海水溺死组和死后入水组。首先，基于以下原则对数据进行筛选，本研究剔除了分解过程中相对丰度低、相对丰度方差较小的OTU，这些OTU在本研究中被认为是噪声和低贡献特征。因此，经过筛选和剔除，14种细菌用于海水溺死组PMSI模型的建立，16种细菌用于死后入水组PMSI模型的建立。在所得到的结果中，海水溺死组预测模型所得到的 MAE=13.272 h，R^2=0.798；死后入水组预测模型所得到的 MAE=9.956 h，R^2=0.793。本研究拓展了水中尸体肠道微生物推断死亡时间的场景，为更全面了解水中尸体的腐败过程提供了实验基础。当前的研究仍存在一定的局限性，目前所得数据仅使用了相应生物标志物进行RF机器学习算法对PMSI推断的预测，后续仍需要对模型进行优化，以及采用更大或更优的微生物群落进行分析。

目前，在开发用于预测PMI或PMSI的可推广微生物模型方面，最大的障碍是尸体相关数据集的规模限制和采样频率低。因此，在未来的工作中，需要完善和建立更长的时间体系来研究微生物群落的变化对微生物演替的影响[40-41]。此外，本研究中的盲肠内容物微生物群落仅反映了死后内部的微生物演替。外界环境（水体环境、空气、不同温度）的影响以及与宿主之间复杂的相互作用是今后研究的重点。目前的实验研究工作为今后对水中尸体的微生物相关调查奠定了初步的实验基础。然而，如何将本研究结果运用到实际的水中尸体法医调查中，还需要进一步的研究。

3.3 16s rDNA 扩增子测序现存局限性

目前，大多数使用微生物组推断死亡时间的研究常采用16s rDNA扩增子测序分析。此外，基于宏基因组测序的新NGS技术也可应用于微生物物种鉴定。然而，此方法成本更高，需要更复杂的仪器和复杂操作才能获得可信的结果。但是，采用传统的扩增子测序方法，也存在以下不足：鉴定微生物

物种时需要将序列信息与数据库中的参考基因组进行对比，然而现有数据库的差异可能会影响已识别微生物的数量和类型，且未知微生物可能无法识别；其分类能力只能精确到属水平，虽然本研究尝试对已分析出的物种进行种水平的Top10的相对丰度分析（图3F和图3I），但是对种水平的分类能力有限；单样本研究通常可以解决和识别个体之间的差异以及法医学问题，但无法探索多元数据间的变量。尽管收集的微生物组样本数量正在增加，但用于采样的方法和报告数据的方式的多样性和差异性也妨碍分析信息的完整性和作为一个同质实体的客观性[42]。

4 小 结

本研究基于海南岛的高温高湿环境，研究在腐败过程中，大鼠盲肠微生物群落演替对不同死因的影响，为水中尸体腐败的死因诊断研究提供相应实验基础。本研究分析了海水中不同死因的大鼠盲肠微生物群落演替，发现与以往的淡水中尸体和陆上尸体的研究有相互印证的部分，同时也存在差异，海水溺死组和死后入水组的乳杆菌属（*Lactobacillus*）的差异明显，具有作为PMSI推断的生物标志物的潜能。此外，本研究对微生物菌群数据进行随机森林分析，并筛选出甲烷球形菌属（*Methanosphaera*）、异杆菌属（*Allobaculum*）、巴氏杆菌（*Barnesiella*）等16类菌株，这些菌株具有作为生物标志物的潜能，有助于PMSI预测模型的建立。基于上述实验基础和相应潜在生物标志物，本研究利用RF机器学习算法成功建立了海水溺死组和死后入水组的PMSI推断模型，来评估2种不同死因的数据集推断PMSI的效果。本研究表明了海水溺死和死后入水2种死因的盲肠微生物群落存在规律性差异，同时为进一步研究利用基于微生物群落数据的机器学习模型来估计PMSI提供了方向和实验基础。

参 考 文 献

[1] Dahl WJ, Rivero Mendoza D, Lambert JM. Diet, nutrients and the microbiome[J]. Prog Mol Biol Transl Sci, 2020, 171: 237-263.

[2] Lynch SV, Pedersen O. The human intestinal microbiome in health and disease[J]. N Engl J Med, 2016, 375(24): 2369-2379.

[3] Robinson JM, Pasternak Z, Mason CE, et al. Forensic applications of microbiomics: a review[J]. Front Microbiol, 2020, 11: 608101.

[4] Metcalf JL. Estimating the postmortem interval using microbes: Knowledge gaps and a path to technology adoption[J]. Forensic Sci Int Genet, 2019, 38: 211-218.

[5] Megyesi MS, Nawrocki SP, Haskell NH. Using accumulated degree-days to estimate the postmortem interval from decomposed human remains[J]. J Forensic Sci, 2005, 50(3): 618-626.

[6] van Daalen MA, de Kat DS, Oude Grotebevelsborg BF, et al. An aquatic decomposition scoring method to potentially predict the postmortem submersion interval of bodies recovered from the north sea[J]. J Forensic Sci, 2017, 62(2): 369-373.

[7] Pérez-Cárceles MD, del Pozo S, Sibón A, et al. Serum biochemical markers in drowning: diagnostic efficacy of Strontium and other trace elements[J]. Forensic Sci Int, 2012, 214(1/2/3): 159-166.

[8] Kuiper I. Microbial forensics: next-generation sequencing as catalyst: The use of new sequencing technologies to analyze whole microbial communities could become a powerful tool for forensic and criminal investigations[J]. EMBO Rep, 2016, 17(8): 1085-1087.

[9] Cartozzo C, Singh B, Swall J, et al. Postmortem submersion interval (PMSI) estimation from the microbiome of sus scrofa bone in a freshwater lake[J]. J Forensic Sci, 2021, 66(4): 1334-1347.

[10] 赵宝山，严程明，苏俊波，等.1960-2020年海南岛气温、降水及参考作物蒸散量变化趋势[J].节水灌溉，2023，(10):83-90.
Zhao B, Yan C, Su J, et al. Trend Analysis of the Temperature, Precipitation and Reference Crop Evapotranspiration in Hainan Island from 1960 to 2020[J].Water Saving Irrigation, 2023, (10):83-90.

[11] 吴 霞，王培娟，霍治国，等. 1961-2015年中国潜在蒸散时空变化特征与成因[J]. 资源科学，2017，39(5)：964-977.
Wu X, Wang P, Huo Z, et al. Spatio-temporal distribution characteristics of potential evapotranspiration and impact factors in China from 1961 to 2015[J]. Resources Science, 2017, 39(5):964-977.

[12] Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies[J]. Bioinformatics, 2011, 27(21): 2957-2963.

[13] Bokulich NA, Subramanian S, Faith JJ, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing [J]. Nat Methods, 2013, 10(1): 57-59.

[14] Edgar RC, Haas BJ, Clemente JC, et al. UCHIME improves sensitivity and speed of chimera detection[J]. Bioinformatics, 2011, 27(16): 2194-2200.

[15] Wang YY, Guo H, Gao XG, et al. The intratumor microbiota signatures associate with subtype, tumor stage, and survival status of esophageal carcinoma[J]. Front Oncol, 2021, 11: 754788.

[16] Hyde ER, Haarmann DP, Lynne AM, et al. The living dead: bacterial community structure of a cadaver at the onset and end of the bloat stage of decomposition[J]. PLoS One, 2013, 8(10): e77733.

[17] Pechal JL, Crippen TL, Benbow ME, et al. The potential use of bacterial community succession in forensics as described by high throughput metagenomic sequencing[J]. Int J Legal Med, 2014, 128(1):

193-205.

[18] Reject Paul MP Kapil Verma. Assessment of post mortem interval,(PMI) from forensic entomotoxicological studies of larvae and flies [J]. Entomol Ornithol Herpetol,2013,2(1):132-148.

[19] Schotsmans E, Márquez-Grant N, Forbes S. Taphonomy of human remains:forensic analysis of the dead and the depositional environment: forensic analysis of the dead and the depositional environment.2017 ,156(2):185-198

[20] Hyde ER, Haarmann DP, Petrosino JF, et al. Initial insights into bacterial succession during human decomposition[J]. Int J Legal Med, 2015,129(3):661-671.

[21] Magni PA, Venn C, Aquila I, et al. Evaluation of the floating time of a corpse found in a marine environment using the barnacle *Lepas anatifera* L. (Crustacea: Cirripedia: Pedunculata)[J]. Forensic Sci Int, 2015,247:e6-e10.

[22] Finley SJ, Benbow ME, Javan GT. Microbial communities associated with human decomposition and their potential use as postmortem clocks[J]. Int J Legal Med,2015,129(3):623-632.

[23] Cobaugh KL, Schaeffer SM, DeBruyn JM. Functional and structural succession of soil microbial communities below decomposing human cadavers[J]. PLoS One,2015,10(6):e0130201.

[24] Damann FE, Williams DE, Layton AC. Potential use of bacterial community succession in decaying human bone for estimating postmortem interval[J]. J Forensic Sci,2015,60(4):844-850.

[25] 周凯旋,鲍慧玮,朱志杰,等.复方嗜酸乳杆菌片对便秘的治疗作用及对肠道菌群的影响[J].第三军医大学学报,2020,42(20):2031-2038.

Zhou KX, Bao HW, Zhu ZJ, et al. Therapeutic effect of compound *Lactobacillus acidophilus* tablets on constipation and its impact on intestinal flora in mice[J]. J Third Mil Med Univ,2020,42(20):2031-2038.

[26] Carasso S, Fishman B, Lask LS, et al. Metagenomic analysis reveals the signature of gut microbiota associated with human chronotypes[J]. FASEB J,2021,35(11):e22011.

[27] DeBruyn JM, Hauther KA. Postmortem succession of gut microbial communities in deceased humansubjects[J]. PeerJ,2017,5:e3437.

[28] Barrios M, Wolff M. Initial study of arthropods succession and pig carrion decomposition in two freshwater ecosystems in the Colombian *Andes*[J]. Forensic Sci Int,2011,212(1/2/3):164-172.

[29] Liu RN, Wang Q, Zhang K, et al. Analysis of postmortem intestinal microbiota successional patterns with application in postmortem interval estimation[J]. Microb Ecol,2022,84(4):1087-1102.

[30] Metcalf JL, Xu ZZ, Weiss S, et al. Microbial community assembly and metabolic function during mammalian corpse decomposition[J]. Science,2016,351(6269):158-162.

[31] Derrien M, Vaughan EE, Plugge CM, et al. *Akkermansia muciniphila* gen. nov. , sp. nov. , a human intestinal mucin-degrading bacterium[J]. Int J Syst Evol Microbiol,2004,54(Pt 5):1469-1476.

[32] Svetnik V, Liaw A, Tong C, et al. Random forest: a classification and regression tool for compound classification and QSAR modeling[J]. J Chem Inf Comput Sci,2003,43(6):1947-1958.

[33] Smith BR, Ashton KM, Brodbelt A, et al. Combining random forest and 2D correlation analysis to identify serum spectral signatures for neuro-oncology[J]. Analyst,2016,141(12):3668-3678.

[34] Maguire A, Vega-Carrascal I, Bryant J, et al. Competitive evaluation of data mining algorithms for use in classification of leukocyte subtypes with Raman microspectroscopy[J]. Analyst, 2015, 140(7): 2473-2481.

[35] Lebedev AV, Westman E, van Westen GJ, et al. Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness[J]. Neuroimage Clin,2014,6:115-125.

[36] Metcalf JL, Wegener Parfrey L, Gonzalez A, et al. A microbial clock provides an accurate estimate of the postmortem interval in a mouse model system[J]. Elife,2013,2:e01104.

[37] Johnson HR, Trinidad DD, Guzman S, et al. A machine learning approach for using the postmortem skin microbiome to estimate the postmortem interval[J]. PLoS One,2016,11(12):e0167370.

[38] Wallace JR, Receveur JP, Hutchinson PH, et al. Microbial community succession on submerged vertebrate carcasses in a tidal river habitat: Implications for aquatic forensic investigations[J]. J Forensic Sci,2021,66(6):2307-2318.

[39] Zhang FY, Wang PF, Zeng K, et al. Postmortem submersion interval estimation of cadavers recovered from freshwater based on gut microbial community succession[J]. Front Microbiol,2022,13:988297.

[40] Rea K, Dinan TG, Cryan JF. The microbiome: a key regulator of stress and neuroinflammation[J]. Neurobiol Stress,2016,4:23-33.

[41] Desbonnet L, Garrett L, Clarke G, et al. Effects of the probiotic *Bifidobacterium infantis* in the maternal separation model of depression [J]. Neuroscience,2010,170(4):1179-1188.

[42] Singh H, Clarke T, Brinkac L, et al. Forensic microbiome database: a tool for forensic geolocation meta-analysis using publicly available 16S rRNA microbiome sequencing[J]. Front Microbiol, 2021, 12: 644861.

（责任编辑：周一青）