#### DOI: 10.20103/j.stxb.202209142628

张茜,崔继鹏,王紫瑶,李利,崔素霞.芦苇群落水体质量与芦苇转录本的相关性.生态学报,2024,44(2):819-831. Zhang X, Cui J P, Wang Z Y, Li L, Cui S X.Research on correlation between water quality and transcripts in *Phragmites australis*.Acta Ecologica Sinica, 2024,44(2):819-831.

## 芦苇群落水体质量与芦苇转录本的相关性

张 茜1,崔继鹏1,王紫瑶1,李 利1,2,崔素霞1,2,\*

1首都师范大学生命科学学院,北京 100048

2 植物基因资源与低碳环境生物技术北京市重点实验室,北京 100048

摘要:芦苇(Phragmites australis)是广生态幅多年生水生植物,在全球分布极广。芦苇能够适应并生长在极端环境,如盐碱地、 重金属污染地等,这种巨大的环境忍耐能力使得芦苇成为研究植物环境适应性的模式物种,并逐渐成为植物学和生态学等领域 的研究热点。目前,北京自然和人工湿地中大量栽植芦苇,其主要功能是净化水体、稳定水体自净能力。为深入了解水体质量 对芦苇转录本的影响及二者之间的相关性,首先对北京及河北白洋淀共计 32 个采样点的 8 项水质指标进行了检测。结果表 明,32 个采样点的水体整体处于污染状态,主要污染物为总氮(TN)和生物需氧量(BOD),并且水质均略偏碱性,且具有高氮 (N)、高磷(P)、高 BOD 和高化学需氧量(COD)等特征。主成分分析(PCA)表明 8 项水质指标间具有相关性,可被降至 3 个 维度,第一维度是 BOD/COD/高锰酸盐指数(COD<sub>Mn</sub>);第二维度是 TN/氨氮(NH<sub>3</sub>-N)/总磷(TP);第三维度是 pH/溶解氧 (DO)。对水质指标与芦苇叶片 RNA-seq 转录本进行皮尔森系数相关性分析,筛选出 247 个与水质指标相关的转录本。其中, 232 个基因(93.9%)的转录水平与水质指标呈现正相关性。经实验室单一因子或双因子验证确定了 PB.4621.1 (编码 GMP 合 酶)和 PB.12079.1 (编码 LEA14 蛋白)两个基因更适合用于监测水体中的氮和磷水平。研究为后续芦苇监测水质技术的开发 与利用奠定了理论基础。

关键词:芦苇;水质指标;水体质量;转录组;相关性

# Research on correlation between water quality and transcripts in *Phragmites* australis

ZHANG Xi<sup>1</sup>, CUI Jipeng<sup>1</sup>, WANG Ziyao<sup>1</sup>, LI Li<sup>1,2</sup>, CUI Suxia<sup>1,2,\*</sup>

1 School of Life Sciences, Capital Normal University, Beijing 100048, China

2 Key Laboratory of Plant Genetic Resources and Low-Carbon Environmental Biotechnology, Beijing 100048, China

**Abstract**: *Phragmites australis*, as a wide-ranging perennial aquatic plant, is wildly spread in the worldwide. *Phragmites australis* has adapted to adverse terrestrial habitats such as saline meadow and heavy metal pollution, which makes it become a model species for studying plant environmental adaptability, and has become a research hotspot in the fields of botany and ecology, gradually. A large number of *Phragmites australis* had been planted in natural and artificial wetlands in Beijing used for removing water pollution. However, the mechanism of monitoring water quality by *Phragmites australis* has not been clearly studied. In this study, we investigated 8 major water quality indexes of 32 sampling sites distributed all over the major water systems in Beijing and Baiyangdian of Hebei. According to "China National Surface Water Environmental Quality Standard GB3838—2002", the 8 major water quality indexes include pH, dissolved oxygen (DO), total nitrogen (TN), ammonia nitrogen (NH<sub>3</sub>-N), total phosphorus (TP), biological oxygen demand (BOD), chemical oxygen demand (COD) and permanganate index (COD<sub>Mn</sub>). The results showed that the pollution of 32 water samples were relatively severe.

收稿日期:2022-09-14; 网络出版日期:2023-10-18

基金项目:国家自然科学基金项目(31972934);首都师范大学科技创新服务能力建设-基本科研业务费项目(科研类)(19530050183)

<sup>\*</sup>通讯作者 Corresponding author.E-mail: sxcui@ cnu.edu.cn

While, the contamination of 32 sampling water were able to be degraded by using microorganism. In addition, the water quality of sampling sites was slightly alkaline, and had the characteristics of high nitrogen (N), high phosphorus (P). high BOD and COD, which indicated that the contamination of sampling sites was extremely serious, and the major pollutants in water are TN and BOD. Principal component analysis (PCA) showed that the 8 water quality indexes could be divided into 3 dimensions, the first dimension was COD/BOD/COD<sub>Mn</sub>; the second dimension was TN/NH<sub>3</sub>-N/TP; and the third dimension was pH/DO. The water quality indexes within each dimension had a high positive correlation. We further screened out 247 transcripts associated with water quality indexes (Pearson correlation coefficient, PCC  $\geq \pm 0.55$ ) using Phragmites australis transcriptome data (RNA-seq data). Among them, most transcripts (93.9%) were positively correlated with the water quality indexes. The proteins encoded by 132 genes of Phragmites australis related to nitrogen and phosphorus pollution are mainly involved in photosynthesis and protein quality control systems, including protein folding, sorting and degradation; there are also proteins involved in transcription, as well as plant defense systems, signal transduction and lipid metabolism. In addition, there are a large number of unknown proteins. The proteins encoded by 89 genes of Phragmites australis related to organic matter pollution are mainly involved in protein synthesis process and the expression of detoxification genes for degradation of pollutants efficiently. Moreover, it was found that genes transcription and defense system were also response to organic matter pollution. In addition, there are a large number of unknown functional proteins involved. The combination of two genes, PB.4621.1 (encoding GMP synthase) and PB.12079.1 (encoding LEA14 protein), is expected to become a core member of genes for monitoring water quality (N and P levels). Overall, we analyzed the transcriptome changes of *Phragmites australis* leaves, in order to obtain monitoring genes that can reflect water quality, and to explore new ecological uses of *Phragmites australis*. This study laid the foundation for the subsequent development of monitoring technology for water quality by using *Phragmites australis*.

Key Words: Phragmites australis; water quality index; water quality; transcriptome; correlation

我国河流中主要污染物是氮、磷、有机化合物和重金属<sup>[1]</sup>。水体中过量的氮和磷会导致藻华<sup>[2]</sup>。因此对 于水体质量的监测非常重要。传统的检测多是使用物理、化学和微生物测试的方法。这些方法灵敏而准确, 但昂贵耗时并且波动性强<sup>[3-4]</sup>。目前,对于水质类别的评价方法主要有单因子类别和综合污染指数(P值) 两种评价方法<sup>[5]</sup>。近年来,随着科学技术的飞速发展,荧光光谱、遥感技术以及纳米生物传感器被用于评估 水体的质量,但其测试费用高昂且检测种类有限,准确性低<sup>[6-10]</sup>。水生植物生活在自然水体中,需要吸收大 量的氮(N)和磷(P)、以及各类无机物和有机物,具有污水净化和监测水体的天然优势。植物监测具有综合 性、直观性、长期性、连续性、灵敏性、经济性和保护性等优势<sup>[11-14]</sup>。

芦苇根茎发达,生长期长,适应性和抗污染能力强,已被用于处理不同类型的废水,是最常用的修复湿地 植物<sup>[15-18]</sup>。除了可以去除有机物、氮、磷以及重金属以外,还可以用来去除兽药和医药化合物、染料和杀虫剂 等<sup>[19-21]</sup>。此外,芦苇的叶片光谱信息可以在某种程度上区分水体中总氮(TN)含量的差异<sup>[22]</sup>。本研究尝试 通过低成本的 RNA-seq 技术,利用大数据分析寻找能够响应水质变化的芦苇转录本。以期建立一种基于植 物的快速而便捷的新型水质监测技术,实现对水体质量的长期综合评价。

#### 1 材料与方法

#### 1.1 植物材料与水体采样点

图 1 为采样点具体位置示意图。从图中可以看出,采样点分布在北京地区的主要水系,包括永定河水系(YongDingHe,Y)、潮白河水系(ChaoBaiHe,C)、北运河水系(BeiYunHe,B)、大清河水系(DaQingHe,D),以及河北白洋淀(BaiYangDian,HB),共计 32 个点。样点的水源多数来自城市再生水补入,有些甚至是工业废水和城镇生活污水。于 7—8 月的晴朗天气,野外采集距离岸边小于 1 米的外侧芦苇群落(Phragmites

N A 北京 4 Y02 C32 B21 1 B23 B21\_3 B22 ●B24 C31 B25 B21 2 🔬 🍋 B27 B26 Y05 06 B28 B29 D/∥1 D13 D12 HB41 BHB42 HB43 20 km 白洋淀

australis) 成熟叶片 (从顶部向下数的第5、第6叶) 的中段组织, 放入液氮储存备用, 每种芦苇群落选取3个 不同的植株。采集芦苇叶片的同时,也收集植物基部的水样。

图1 采样点示意图 Fig.1 Schematic diagram of the sampling sites

Y:永定河水系;C:潮白河水系;B:北运河水系;D:大清河水系;HB:河北白洋淀水系

#### 1.2 水质检测方法及等级的划分

按《中国国家地表水环境质量标准 GB3838—2002》,选择 8 项关键性水质指标进行检测并对采样点各类 水质进行划分,即 pH、溶解氧 (DO)、总氮 (TN)、氨氮 (NH<sub>3</sub>-N)、总磷 (TP)、生物需氧量 (BOD)、化学需氧 量(COD)和高锰酸盐指数(COD<sub>Mn</sub>)。其中,pH使用 Sigma 公司生产的 pH 试纸进行测定;DO 使用溶解氧 测量仪 (HANNA-HI9142) 进行测定; TN 参照 HJ636—2012 标准, 使用碱性过硫酸钾消解紫外分光光度法进 行测定;NH<sub>3</sub>-N 参照 HJ535—2009 标准;使用纳氏试剂分光光度法进行测定;TP 参照 GB11893—1989 标准, 使用钼酸铵分光光度法进行测定; BOD 参照 HJ505-2009 标准, 使用稀释与接种法进行测定; COD 依据 HJ828—2017标准,使用重铬酸盐法进行测定;COD<sub>Mo</sub>参照 GB11892—89标准,使用高锰酸钾进行测定。 1.3 综合污染指数及主要污染物分担率分析

参照 Zhang 等<sup>[5]</sup>的方法,选取 DO、TN、NH<sub>3</sub>-N、TP、BOD、COD 和 COD<sub>Mn</sub>共计 7 项水质指标,按照以下公式 计算综合污染指数:



### 综合污染指数 = $1/n \sum_{i=1}^{n} (P_i)$

 $P_i = C_i / S_i$ 

式中, $P_i$ 为 *i* 污染物的污染指数; $C_i$ 为 *i* 污染物的实测浓度; $S_i$ 为 *i* 污染物的 III 类标准;*n* 为污染物的种类。 **1.4** 主成分分析(PCA)分析

利用 FactoMineR 包中 PCA 函数进行 PCA 数据分析,利用 R 包 corrplot 计算变量间相关性。以 R 包 factoextra 进行 PCA 可视化。

1.5 芦苇转录组 (RNA-seq) 测定

芦苇转录组的测定由诺禾致源公司(北京)进行高通量测序(Illumina HiSeq)。利用七大数据库进行基因功能注释,包括:非冗余蛋白质序列数据库(Nr),非冗余核酸序列数据库(Nt),蛋白质家族数据库(Pfam), 真核生物蛋白相邻类的聚簇/蛋白质直系同源簇数据库(KOG/COG),蛋白质序列数据库(Swiss-prot),京都基因与基因组百科全书数据库(KEGG),基因本体论数据库(GO)。使用实验室前期构建的芦苇伪基因组作为 背景库<sup>[23]</sup>,采用 RSEM 软件<sup>[24]</sup>进行 FPKM 转换,进而分析基因的表达水平。

1.6 相关性分析

通过计算 32 个叶片样本的转录本每百万映射读取的每千碱基转录片段(FPKM)值与水体中 8 项指标的 实测值的皮尔森相关系数 (PCC),以 PCC ≥±0.55 的低阈值进行筛选,筛选芦苇中与水质指标相关的转 录本。

1.7 与水质指标相关的转录本在各个不同集合中的分布分析

统计筛选出的 247 个转录本在不同集合中的分布情况, Upset 图使用 R 语言中 UpsetR 包里面的Upset() 函数进行可视化。使用 R 语言中 ggplot2 制作旋风图表征与水质指标正负相关的转录本在各个不同集合中的数量。

1.8 GO 富集分析

利用 R 语言中 clusterProfiler 包进行 GO 富集分析及可视化分析。

1.9 RT-PCR

2 结果与分析

2.1 采样点的水质指标分析

使用反转录聚合酶链反应(RT-PCR)的方法对筛选出的与水质相关的转录本进行验证。首先培养苗龄 30 天的长势均一的芦苇幼苗,并用不同浓度的含氮、含磷或氮磷复合木村培养液处理7天,其中以 NH<sub>4</sub>NO<sub>3</sub>作 为氮源,而以 KH<sub>2</sub>PO<sub>4</sub>作为磷源。RT-PCR 所使用的的引物序列如表1所示。

	Table 1Primers for RT-PCR	
基因名称	正向引物	反向引物
Gene names	Froward Primer $(5' \text{ to } 3')$	Reverse primer $(5' \text{ to } 3')$
晚期胚胎发生丰富蛋白 PB.12079.1	GGATGCAGATGCCGTTCA	TGGTTCCACCCTCTTTACTT
网格蛋白装配蛋白 PB.3207.1	CAGAAATAGCGAGGAAGC	TTGAAGTCGGAGTTGAGG
鸟苷酸合酶(GMP 合酶)PB.4621.1	TGTGGGTCCAACGGTCAT	GTCGGTGTCACTCGCATT
未知蛋白 PB.631.1	TCTTCGTCGGAGTTTATGTT	TTGGTTCGGTAGGCTTTT
TATA 结合蛋白相关因子 PB.9173.1	TATGGGATTGCCTTACTG	TTATTTATGAGCGACTGC
18S 核糖体 RNA 18S	ATGATAACTCGACGGATCGC	CTTGGATGTGGTAGCCGTTT

表1 反转录聚合酶链反应(RT-PCR)所使用的的引物序列

822

为了评估 32 个采样点的水体质量,根据《中国国家地表水环境质量标准 GB3838—2002》,选择了 8 项关键性水质指标进行检测。这些指标分别是 pH、溶解氧 (DO)、总氮 (TN)、氨氮 (NH<sub>3</sub>-N)、总磷 (TP)、生物需氧

量(BOD)、化学需氧量(COD)和高锰酸盐指数(COD<sub>Mn</sub>)。图2为采样点水体的水质指标分析,结果表明几乎 所有水样的BOD都比较高,有近50%的水样BOD值超过国家V类水标准(图2)。相比而言,多数采样点水样 中的COD和COD<sub>Mn</sub>值低于国家V类水标准(图2)。因此高BOD是芦苇生长点水质的特点之一。

此外,水样的 TN 含量都比较高,特别是在大清河和潮白河水体中几乎全部超过国家 V 类水标准 (图 2)。而在多数水体中,NH<sub>3</sub>-N 含量比较低,只有北运河一个样点超过国家 V 类水标准。这表明采样点水 体中的氮污染主要为硝态氮 (图 2)。至于 TP 含量,在多数水体中低于国家 III 类水标准。只是在潮白河和 大清河部分水体中,远远超过国家 V 类水标准 (图 2)。实地调查发现这是源于附近的农业和居民用水直接 排放所致。

大多数水样的 pH 均在 7 以上,偏碱性。但 DO 值在水样中的差异极大,有的进入国家 I 类水标准,有的 超过国家 V 类水标准 (图 2)。这是由于水中 DO 含量受多种因素影响的结果,包括水体中的生物数量、有机 物含量、以及水位、水温及水流速度等环境因素<sup>[25]</sup>。

因此从总体上看,选择的芦苇生长点的水体具有碱性、高氮、高磷和高 BOD 的特点。

2.2 采样点水体的污染程度分析

利用除 pH 外的 7 项水质指标,计算采样点所属水系的综合污染指数 (P 值),如图 3 所示。从图中可以 看出,这 32 个芦苇采样点水样的平均 P 值为 1.78,水体处于污染状态。其中,大清河及潮白河的 P 值分别为 2.29 和 2.12,是污染最为严重的两个水系。而其余水系污染程度相对较低 (图 3)。

进一步分析主要污染物分担率(即某一污染物对于综合污染指数的贡献度/占比)。结果表明这 32 个采 样点水样中主要污染物为 TN 和 BOD,平均分担率分别为 28.27% 和 22.92%。而 NH<sub>3</sub>-N 的分担率最低,只有 4.12%(图 3)。比较而言,不同水系中的主要污染物不尽相同。如永定河、大清河和北运河水体的主要污染 物为 TN 和 BOD;潮白河水体的主要污染物是 TN 和 TP;而白洋淀水体的主要污染物为 BOD 和 COD<sub>Mn</sub>。值 得注意的是,大清河水系总氮污染极其严重,总氮的污染分担率高达 54.19%(图 3)。

2.3 水质指标的相关性分析

图 4 为依照 8 项水质指标对 32 个芦苇采样点进行主成分分析 (PCA)的结果。从图中可以发现,采样点 基本不具有按照水系分布的特点,符合随机采样模式。前三个主成分就能够区分 80%的样点。相比之下,有 8 个样点呈现高度离散状态,即 D13、B23、C33、C34、C35、Y02、C31 和 B28。碎石图显示 COD、BOD 和 COD<sub>Mn</sub> 对第一主成分的贡献率最大,而 TN、NH<sub>3</sub>-N 和 TP 对第二主成分的贡献最大 (图 4)。

PCA 分析还显示, COD、BOD 和 COD<sub>Mn</sub>之间具有高度正相关性, TN、NH<sub>3</sub>-N 和 TP 之间的正相关性也很强, 而 pH 和 DO 之间也有一定的正相关性。因此, 所检测的 8 项水质指标可明显区分成 3 个维度, 这为后续的芦苇转录组分析方案提供了依据。不过, TN、NH<sub>3</sub>-N 和 DO 之间是负相关的 (图 4)。

2.4 水质监测基因的筛选

在获得上述 32 个采样点的芦苇转录组(RNA-seq)数据后,为了挖掘芦苇转录组中与水质指标相关的表 达基因,分别计算了转录本的 FPKM 值与水质 8 个指标之间的皮尔森相关系数(PCC)。考虑到所有样本转 录组之间的差异较小,因而仅以 PCC ≥±0.55 的低阈值进行筛选,最终筛选出与水质相关的转录本 247 个。 为了反向验证这个结果,又使用这 32 个样点的 247 个转录本的 FPKM 值进行部分转录组的主成分分析,如图 5 所示。从图中可以看出有 10 个样点相对离散,其中 8 个与 32 个样点水质指标的 PCA 分析中离散样点相一 致(图 5)。综合上述结果,可以确定这 247 个基因的转录水平能够反映出水体质量的变化,可初步视为潜在 的水质监测基因。

图 6 综合分析了这 247 个转录本与水质指标的相关性,并进行 GO 富集分析。从图中可以看出,与水质 指标呈现正相关的转录本非常多,共计 232 个,而负相关转录本只有 15 个 (图 6)。从单一指标来看,与 TP 相关的转录本最多,达 62 个;与 TN、NH<sub>3</sub>-N 和 COD<sub>Mn</sub>相关的转录本分别为 33、32 和 31 个;而与 COD 相关的 转录本只有 3 个。此外,有 44 个转录本与多种水质指标相关,并且主要体现在有相关性的指标之间。如有



图 2 采样点水体水质指标分析

Fig.2 Analysis of water quality indexes at sampling sites

虚线指示《中国国家地表水环境质量标准(GB3838-2002)》中各项水质指标的标准限值

16个转录本与 BOD、COD 以及 COD<sub>Mn</sub>有相关性,有 5个转录本与 TN 和 TP 相关 (图 6)。

从 GO 富集分析的生物过程类群气泡图可见,与水质相关的转录本集中到了一个大的功能群,即蛋白质 合成。其中包括翻译、多肽生物合成和代谢、酰胺生物合成、蛋白质折叠以及核糖体大亚基生物合成等 (图 6)。分子功能类群显示,与水质相关的转录本是与核糖体结构和分子伴侣介导的蛋白质折叠过程相关 (图 6)。

这些结果意味着在污染的水体中,芦苇的蛋白质合成机制受到较大影响。因此,蛋白质合成相关的转录 本作为环境水体质量监测的可选靶标。并且大量水质响应转录本在当前数据库中不能被富集的现象说明植 物对水质变化的应答是多途径的,涉及的代谢过程较为复杂,需要更深入的研究。

#### 2.5 水质监测基因的验证

经筛选与分析后发现有 5 个基因的转录本与总氮 (TN) 和总磷 (TP) 均有正相关性。图 7 为在实验室 条件下,采用不同浓度的单一氮源、磷源以及复合氮磷培养液对芦苇幼苗进行处理,对上述转录本的验证和分





析。从图中可以明显看出,随着 N 浓度的增加,这些基因的转录全部有增加的趋势。尤其在 N 含量为 50 mg/L 时,这些基因的表达量达到了峰值(图 7)。这些基因对单一磷源培养液的响应也是如此。在自然水体中,总磷浓度在 0.01—1.38 mg/L 的范围。当用单一磷源处理芦苇植株时,可以看到在 P 浓度达到 1 mg/L 时,这 5 个基因的表达水平是最高的(图 7)。

在这 5 个基因中,有 2 个基因不仅对 N 和 P 浓度的变化都有较好的正响应关系,尤其是对 N+P 混合源的培养液的梯度响应强烈(图 7)。其中,一个编码晚期胚胎发生丰富蛋白(LEA, PB.12079.1);另一个编码GMP 合酶(PB.4621.1),也称为谷氨酰胺氨基转移酶。因此,这 2 个基因可能是适合监测水体中氮和磷浓度变化的靶标基因。

3 讨论

#### 3.1 采样点的水质特点

为了建立一种基于植物的快速而便捷的新型水质监测技术,实现对水体质量的长期综合评价。以北京湿地大量原生或栽植的芦苇叶片为材料,研究了约 3.1 万个基因转录本与水质指标的对应关系。现今,芦苇分布于北京各大水域,以八倍体为主,四倍体为辅,是自然河流、湿地以及人工湿地的优势物种<sup>[26]</sup>;此外,因城市发展造成的河流、湖泊水域的大量污染催生了芦苇做为水质净化植物的需求<sup>[27-28]</sup>。因此,采用芦苇材料关联水质指标的研究具有非常现实的背景和基础科学的支撑。通过实地考察,在永定河水系(YDH)、潮白河水系(CBH)、北运河水系(BYH)、大清河水系(DQH)以及河北白洋淀(BYD)共选择了 32 个样点进行水质指标和芦苇基因转录本的同步分析。

芦苇采样点的选择标准主要有两条。一是有大面积芦苇生长;二是水源主要来自城市再生水补入,或有 工业废水和城镇生活污水排入。在这种环境条件下,芦苇可以最大化发挥其对水体污染物质的清除功能;也



图 4 采样点水质指标的主成分分析

Fig.4 Principal component analysis of water quality index

虚线表示各变量的平均贡献率;Dim1:主成分1;Dim2:主成分2;COD:化学需氧量;BOD:生物需氧量;COD<sub>Mn</sub>:高锰酸盐指数;TP:总磷;DO: 溶解氧;氨氮:NH<sub>3</sub>-N;TN:总氮

可满足获得与污染因子相关的响应基因的要求。根据检测结果发现,采样点水体具有碱性、高氮、高磷、高有 机物污染的特点。

BOD/COD 的比值可以判断水体中有机物是否可以通过生物进行<sup>[29]</sup>。当 BOD/COD 比值<0.2 时,有机 物不能被微生物降解。图 8 为采样点水体的生物可降解性分析,从图中明显看出这 32 个水体样本均可被生 物降解(图 8)。这说明水体具有生物治理的可能性,水体的污染能够通过生物降解的方式得到有效治理。





#### 3.2 水质指标间的相关性分析

大量的研究表明不同的水质指标间可能存在相关性,但这些研究往往局限于对于两个或三个指标间的相关性分析,并没有进行综合分析。如王庆安等研究发现水中 pH 的周期性变化与 DO 的变化规律相似,经相关性分析,并没有进行综合分析。如王庆安等研究发现水中 pH 的周期性变化与 DO 的变化规律相似,经相关性回归分析发现 pH 与 DO 呈显著正相关<sup>[30]</sup>。马自伟等发现地表水中氨氮、总氮和总磷三个监测因子之间具有明显的相关关系,相关系数为0.8158—0.9953<sup>[31]</sup>。李志亮和仲跻文研究证实 BOD、COD 和 COD<sub>Mn</sub>之间具有明显的相关性,可以用其中的某一个指标去估算或预测另外两个指标。此外,还可通过这三个指标对监测数据进行预判及合理性的检验<sup>[32]</sup>。本文一共测定了 32 个样点的 8 项水质指标,并对这 8 个指标间的相关性进行了成对分析,从而综合分析水质指标间的相关性。研究结果表明,pH 和 DO 之间有一定的正相关性,皮尔森相关系数(PCC)达到0.38;TN、NH<sub>3</sub>-N 和 TP 之间的正相关性很强,其中 TN 和 TP 的 PCC 为0.57,NH<sub>3</sub>-N 与 TP 的 PCC 分别为0.27 和0.23;分析发现 COD、BOD 和 COD<sub>Mn</sub>之间具有显著正相关性,其中 COD 和 BOD、BOD 和 COD<sub>Mn</sub>以及 COD 和 COD<sub>Mn</sub>之间的 PCC 分别为0.95、0.91 和0.87。研究表明,在溶解氧高的水体中,氨氮的含量通常较低,因为在氧化环境中,氨容易转化为硝酸盐<sup>[33]</sup>。在研究中同样发现 DO 和 NH<sub>3</sub>-N 之间的负相关性最强,PCC 为-0.31。这些指标之间的相关性与大多数研究相一致<sup>[30-33]</sup>。因此,将所检测的上述 8 项水质指标降为 3 个维度,即 pH/DO、TN/NH<sub>3</sub>-N/TP 以及 COD/BOD/COD<sub>Mn</sub>,进一步综合分析研究芦苇叶片的转录组变化。

#### 3.3 芦苇中响应水质变化的基因

采用 RNA-seq 技术,获得了 32 个芦苇样本的转录组数据。根据皮尔森相关系数 (PCC),确定了一个较低的阈值 (PCC ≥±0.55)进行基因筛选。在 3.1 万个转录本中,仅有 247 个基因的转录本满足这个筛选条件。然而,对这些基因的功能富集发现,GO 只富集到 35 个基因,功能主要涉及蛋白质合成机制。认为产生这一结果的原因至少有 3 个。其一是芦苇基因组遗传背景缺乏<sup>[34]</sup>;其二可能是物种特异性基因的存在,而现有的数据库不全;其三是水质响应基因可能在功能上是高度离散的,从而导致其无法被富集到某一过程。

本研究发现芦苇中编码 LEA 和 GMP 合酶的基因对水体中的高氮和高磷有所响应。研究表明, LEA 基因家族在植物中往往具有几十个成员, 且高度保守, 功能涉及干旱、盐渍、低温等与植物脱水密切相关的多种环境胁迫的适应和调节<sup>[35-36]</sup>。尽管目前已有上千个植物 LEA 蛋白基因得以分离, 但对该类基因的生物学功



图 6 响应水质变化的芦苇转录本分布及基因本体论(GO)富集分析

能的研究仅限于少数几种植物,并且主要集中于植物脱水应答等非生物胁迫,采用的研究手段大多为植物转 基因功能分析。LEA 基因作为一类逆境胁迫的下游应答基因,其在响应非生物胁迫时会大量表达。在农作 物耐脱水遗传改良中具有普遍的应用潜力<sup>[37]</sup>。然而 LEA 基因的转录水平为何受氮和磷浓度的影响,是否为 芦苇对富营养化水体的长期适应性结果,及其作用的具体机制还需进一步探索<sup>[37]</sup>。

氮和磷都是植物体需求量很高的矿质营养素,与植物的生长发育以及产量等方面有着密切的联系。植物 不能直接利用空气中的氮气,它们一般吸收利用土壤或环境中的 NO<sub>3</sub> 或者 NH<sup>4</sup>,在相关酶的作用下合成谷氨 酰胺和谷氨酸,然后经过转氮作用形成不同的氨基酸,进而合成蛋白质<sup>[38—39]</sup>。因而,氮素的吸收和同化是植 物生长发育的必需过程。植物对磷的吸收和同化亦同样重要。研究表明,谷氨酰胺氨基转移酶可通过将谷氨 酰胺水解为谷氨酸,参与氮的初级吸收与光呼吸释放氨的再吸收,以及吸收氨的同化和固氮作用,对作物的生 长发育、氮素同化与积累、应对逆境等情况都起着调控作用。谷氨酰胺氨基转移酶在水稻叶片中表达量最高, 在水稻中过表达谷氨酰胺氨基转移酶可以促进幼苗对氮的吸收和同化,增强其对缺氮的耐受性并增加 产量<sup>[40]</sup>。

**Fig.6** Distribution of transcripts related to water-quality index in different sets and GO functional enrichment analysis 左侧旋风图显示芦苇转录本 (*n*=247) 与水质指标的相关性



图 7 芦苇中氮和磷响应基因的转录分析



本研究尝试采用廉价、快捷的 RNA-seq 技术分析 芦苇叶片的转录组变化,以期获得可以反映水体质量的 监测基因,挖掘芦苇在生态方面的新用途。

#### 4 结论

对 32 个采样点水体的 8 项关键性水质指标进行检测,结果表明芦苇生长点的水体略偏碱性且染程度较高,具有高氮、高磷、高 BOD 和 COD 的特点。主成分分析(PCA)显示, 8 项水质指标可分成 COD/BOD/COD<sub>Mn</sub>、TN/NH<sub>3</sub>-N/TP 和 pH/DO 这 3 个维度。各维度内部的水质指标均具有较强的正相关性。创新性采用植物叶片 RNA-seq 和水质指标间相关性分析的方法,筛选出了 247 个芦苇叶片的 RNA-seq 数据中可挖掘的







水质相关基因,在不同水质维度中,这些基因编码蛋白质的功能差异较大。经 RT-PCR 验证后发现 PB.

小质相关基因,在不向小质维度中,这些基因编码蛋白质的功能差异较入。经 R1-PCR 验证后发现 PB. 12079.1 (编码 LEA14 蛋白)和 PB.4621.1 (编码 GMP 合酶)有望成为水质 (N和 P的水平)监测基因的核心成员。

本研究创新性利用芦苇叶片中的基因进行水质监测则发掘了芦苇的生态价值,并且有助于理解芦苇基因 表达变化与水体中污染物浓度之间的关系。这为后续芦苇监测水质技术的开发与应用奠定了理论基础。

**致谢:**感谢首都师范大学生命科学学院周华健、王安安、裘天航、袁敏、王蕊和齐振煜同学在样品采集中给予的 帮助。

#### 参考文献(References):

- [1] 李勇,单雅洁,程浩,吴旭雨,钱佳宁,冯家成,徐鸿,李娜,楼睿焘.城市重污染河流沉积物营养盐和重金属分布及潜在生态风险.河海大学学报:自然科学版,2022,50(3):31-38.
- [2] 聂煜东. 三峡库区嘉陵江回水区中碳氮磷利用关键酶对藻华的响应及其作用机制研究[D]. 重庆: 重庆大学, 2016.
- [3] Huang X J, Feng M, Ni C S, Xie D T, Li Z L. Enhancement of nitrogen and phosphorus removal in landscape water using polymeric ferric sulfate as well as the synergistic effect of four kinds of natural rocks as promoter. Environmental Science and Pollution Research, 2018, 25 (13): 12859-12867.
- [4] Rehman U, Vesvikar M, Maere T, Guo L, Vanrolleghem P A, Nopens I. Effect of sensor location on controller performance in a wastewater treatment plant. Water Science and Technology, 2015, 71(5): 700-708.
- [5] Zhang L. Different methods for the evaluation of surface water quality: the Case of the Liao River, Liaoning Province, China. International Review for Spatial Planning and Sustainable Development, 2017, 5(4): 4-18.
- [6] Radiul S M, Chowdhury J, Goswami A, Hazarika S. Fluorescence spectroscopy based characterisation method for aggregation behaviour of rhodamine B (RhB) in water, ethanol, and propanol. Laser Physics, 2022, 32(7): 075602.
- [7] Gani M A, Sajib A M, Siddik M A, Moniruzzaman M. Assessing the impact of land use and land cover on river water quality using water quality index and remote sensing techniques. Environmental Monitoring and Assessment, 2023, 195(4): 449.
- [8] Carstea E M, Bridgeman J, Baker A, Reynolds D M. Fluorescence spectroscopy for wastewater monitoring: a review. Water Research, 2016, 95: 205-219.
- [9] 岳隽, 王仰麟, 李贵才, 吴健生, 谢苗苗. 不同尺度景观空间分异特征对水体质量的影响——以深圳市西丽水库流域为例. 生态学报, 2007, 27(12): 5271-5281.
- [10] Singh S, Henderson R K, Baker A, Stuetz R M, Khan S J. Online fluorescence monitoring of RO fouling and integrity: analysis of two contrasting recycled water schemes. Environmental Science: Water Research & Technology, 2015, 1(5): 689-698.
- [11] 次仁央金,尼玛卓玛,普布央金.浅谈生物监测及其在环境影响评价中的应用. 西藏科技, 2015(4): 36-37.
- [12] 李凌,周智来,陈辉.水生植物在水污染控制中的生态效应研究.绿色环保建材,2018(9):18,20.
- [13] 贾佳敏. 植物在环境污染治理中的应用. 皮革制作与环保科技, 2022, 3(21): 122-124.
- [14] 周秀云.环境保护的生物监测与治理.农业与技术, 2020, 40(13): 146-147.
- [15] Fraser L H, Carty S M, Steer D. A test of four plant species to reduce total nitrogen and total phosphorus from soil leachate in subsurface wetland microcosms. Bioresource Technology, 2004, 94(2): 185-192.
- [16] Liu Y B, Li X R, Liu M L, Cao B, Tan H J, Wang J, Li X J. Responses of three different ecotypes of reed (*Phragmites communis* Trin.) to their natural habitats: leaf surface micro-morphology, anatomy, chloroplast ultrastructure and physio-chemical characteristics. Plant Physiology and Biochemistry, 2012, 51: 159-167.
- [17] Saltonstall K. A rapid method for identifying the origin of North American Phragmites populations using RFLP analysis. Wetlands, 2003, 23(4): 1043-1047.
- [18] Zhang C B, Wang J, Liu W L, Zhu S X, Ge H L, Chang S X, Chang J, Ge Y. Effects of plant diversity on microbial biomass and community metabolic profiles in a full-scale constructed wetland. Ecological Engineering, 2010, 36(1): 62-68.
- [19] Almeida C M R, Santos F, Ferreira A C F, Gomes C R, Basto M C P, Mucha A P. Constructed wetlands for the removal of metals from livestock wastewater-Can the presence of veterinary antibiotics affect removals? Ecotoxicology and Environmental Safety, 2017, 137: 143-148.
- [20] Carvalho P N, Basto M C P, Almeida C M R. Potential of *Phragmites australis* for the removal of veterinary pharmaceuticals from aquatic media. Bioresource Technology, 2012, 116: 497-501.

- [21] Haddaji D, Bousselmi L, Saadani O, Nouairi I, Ghrabi-Gammar Z. Enzymatic degradation of azo dyes using three macrophyte species: Arundo donax, Typha angustifolia and Phragmites australis. Desalination and Water Treatment, 2015, 53(4): 1129-1138.
- [22] 赵睿,卜红梅,宋献方,高融瑾.再生水补给河道内芦苇的光谱特征及其对水体氮和磷含量的响应.植物学报,2020,55(6):666-676.
- [23] 裘天航. 两种芦苇生态型适应性的分子机制和进化模型研究[D].北京:首都师范大学, 2020.
- [24] Li B, Dewey C N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics, 2011, 12: 323.
- [25] 殷浩. 扬州市典型水体溶解氧的时间分布特征与相关分析. 江苏水利, 2017(1): 45-50.
- [26] 张茜, 裘天航, 王安安, 周华健, 袁敏, 李利, 白素兰, 崔素霞. 北京地区芦苇资源状态及其多样性. 植物学报, 2020, 55(6): 693-704.
- [27] Bhatia M, Goyal D. Analyzing remediation potential of wastewater through wetland plants: a review. Environmental Progress & Sustainable Energy, 2014, 33(1): 9-27.
- [28] Rezania S, Park J, Rupani P F, Darajeh N, Xu X, Shahrokhishahraki R. Phytoremediation potential and control of *Phragmites australis* as a green phytomass: an overview. Environmental Science and Pollution Research, 2019, 26(8): 7428-7441.
- [29] Fountoulakis M S, Terzakis S, Chatzinotas A, Brix H, Kalogerakis N, Manios T. Pilot-scale comparison of constructed wetlands operated under high hydraulic loading rates and attached biofilm reactors for domestic wastewater treatment. Science of the Total Environment, 2009, 407(8): 2996-3003.
- [30] 王庆安,黄时达,孙铁珩. 多藻浅水体中 pH 值和溶解氧协同周期性变化初探. 四川环境, 2001, 20(2): 4-7, 29.
- [31] 马自伟,吴佳宁,陈明,谢丽章. 地表水中氨氮、总氮和总磷的污染程度及相关性分析. 广州化学, 2017, 42(5): 7-11, 31.
- [32] 李志亮,仲跻文.生化需氧量、化学需氧量、高锰酸盐指数三者关系简析.水利技术监督,2015,23(1):5-6.
- [33] Nakamura C H, Salla M R, Oliveira de Jesus J A, Ribeiro da Silva G H. Calibration of mathematical water quality modeling in a river basin under critical conditions. Water Environment Research, 2019, 91(12): 1678-1691.
- [34] Oh D H, Kowalski K P, Quach Q N, Wijesinghege C, Tanford P, Dassanayake M, Clay K. Novel genome characteristics contribute to the invasiveness of *Phragmites australis* (common reed). Molecular Ecology, 2022, 31(4): 1142-1159.
- [35] Abdul Aziz M, Sabeem M, Mullath S K, Brini F, Masmoudi K. Plant group II LEA proteins: intrinsically disordered structure for multiple functions in response to environmental stresses. Biomolecules, 2021, 11(11): 1662.
- [36] Hibshman J D, Goldstein B. LEA motifs promote desiccation tolerance in vivo. BMC Biology, 2021, 19(1): 263.
- [37] 张美, 张会. 胚胎发育晚期丰富蛋白(LEA 蛋白)与植物抗逆性研究进展. 生物资源, 2017, 39(3): 155-161.
- [38] Masclaux-Daubresse C, Reisdorf-Cren M, Pageau K, Lelandais M, Grandjean O, Kronenberger J, Valadier M H, Feraud M, Jouglet T, Suzuki A. Glutamine synthetase-glutamate synthase pathway and glutamate dehydrogenase play distinct roles in the sink-source nitrogen cycle in tobacco. Plant Physiology, 2006, 140(2): 444-456.
- [39] 徐洪超, 商靖, 刘铭荟, 张飞, 李玥莹. 氮代谢相关酶的研究进展. 安徽农业科学, 2022, 50(4): 17-20.
- [40] 南京农业大学. 一种水稻谷氨酰胺氨基转移酶基因 OsGAT2 的基因工程应用:中国,202210011580.9. 2022-05-13.