第13卷 第1期 水生生物学报 Vol. 13, No. 1
1989年3月 ACTA HYDROBIOLOGICA SINICA Mar., 1989

关于建立污水环境中原生动物数据库的研究*

卞小宇 沈韫芬 王春炜 顾曼如 施之新
龚循矩 魏印心 顾汉玥
（中国科学院水生生物研究所，武汉）

提 要

本库在 IBM PC/XT 微机上实现，以全部汉字方式进行人机对话与输出，共包含污水环境中原生动物 702 个种，涉及 32 个目，88 个科，284 个属。对每个阶元都有主要的形态特征描述。本库具有信息查询，统计计数，分类检索等功能，程序采用模块化结构，系统以"菜单"提问方式运行。本数据库的建立将有助于环保工作者快速、准确地综合分析污水环境中原生动物群落的结构变异，以便对环境污染程度作出可靠的评价，也为原生动物分类学研究者使用分类学资料提供很多方便。

关键词 数据库，原生动物群落，生物监测，微型计算机，污染水

原生动物信息是水生生物信息的组成部分，而污水环境中原生动物群落的种类、数量和分布是监测水质的重要生物学指标之一[1]，根据这些资料能够判断污染源的性质和水质的污染程度，为采取切实可行的对策提供科学依据。因此建立污水环境原生动物数据库可以帮助环保工作者快速、准确地综合分析污水环境中原生动物群落的结构变异，能从群落级水平对环境的污染程度作出正确的评价，以提高治理污染工作的效率和质量，对环境保护工作有着重要的现实意义，同时也为建立淡水原生动物数据库打下基础。

本库定名为 Data Base of Protozoa in environmental studies，简称 P. D. B。以全汉字方式进行人机对话与信息输出，共包含淡水原生动物 702 个种，涉及 32 个目，88 个科，284 个属。对每个阶元都有主要的形态特征描述。分类学资料取材于《水治理微型生物监测入门》一书[1)]，分类检索遵循形态分类法，在汉字 DBase II 上建立一批命令文件[2]，以实现分类检索、信息咨询、按类统计和数据库管理等方面的功能，系统以"菜单"提问方式运行，在键盘操作方面，P. D. B. 提供了一个简单的说明文件，并在必要处设有屏幕提示，使用者不需要对计算机本身和对 DBase II 有详细的了解，只要能掌握键盘功能就能使用本数据库，这样既可以使非原生动物学工作者在分类鉴定中减少许多查阅检索表的困难和工作量，对于分类学工作者来说，也可以为研究工作提供许多方便。

* 本文系国家自然科学基金资助项目 386—0594。
1) 中国科学院水生生物研究所沈韫芬等编。待出版。
1987年2月16日收到。

P. D. B. 的功能

1. 信息查询

要了解原生动物各门、纲、目、科、属及其中种的特性，可通过本数据库进行信息查询。本数据库具有进行向上查询、向下查询和横向查询的功能。

向上查询，可随时得到某一阶元的上一阶元及它们的特性描述。例如要查询盖虫科 Opercullariidae 属于哪一个门、纲，可通过 P. D. B. 查得它属于纤毛虫门、寡膜纲、缘毛亚纲。**向下查询**，可得到本阶元所属的所有下阶元情况，例如要查询盖虫科下有哪几个属，则通过 P. D. B. 可查得它有果盖虫属和盖虫属。**横向查询**，可得到与之具有相同分类地位的所有阶元。例如通过 P. D. B. 可查得与盖虫科具有相同分类地位的阶元有钟形科、累枝科、杯形科、睫纤科、怪游科、后游科、车轮科、鞘居科。所有输出都可随时打印记录。

2. 按类统计

对库中具有某种同样特征或相同性质的阶元可进行统计计数。例如要统计寡膜纲内共有多少个属和种，通过 P. D. B. 可立即统计出它共有 31 个属和 86 个种。

3. 分类检索

计算机提供几种可能的描述，用户选择最合适的一种，计算机通过逐项提问，缩小可能范围，最后定出来。因为用户可对照显微镜下所观察的标本的特征，根据 P. D. B. 提出的"菜单"顺序和提示进行回答，如果回答符合顺序，计算机可告知是什么种类，如果用户已知的信息不全，计算机不能回答种名，也能根据提示继续再观察标本中的特征，最后帮助鉴定出种名来。

P. D. B. 的结构和程序设计

P. D. B. 建立在汉字化的关系数据库 C DBase II 上，包括一个主数据文件(占 562 KB)其中包括 1 150 条记录，文件结构见表 1，2 两个索引文件(共占 11 个 KB)，一个暂存文件(最多占 11 个 KB)，结构同主文件，在检索过程中用以存放可能记录。一个记录文件(占3 072 个 KB) 内容见表 2。一个字组库(占 2 964 个 B)。一批命令文件(占 30 个 KB)。以上全部内容分存于两张双面双密度软盘上(360 KB × 2)。

P. D. B. 的全部数据分成 5 个数据块存贮，它们是：

原生动物亚界(占记录的 1 142—1 150)，

肉足虫门(占记录的 1—209)，

动鞭亚门(占记录的 210—316)，

纤毛虫门(占记录的 317—877)，

植鞭亚门(占记录的 878—1 141)，

表1 原生动物数据库结构

Tab. 1 Database structure of Protozoaus

文件的结构: C:MMO. DBF
记录号: 01151
前次更改的日期: 01/01/80
主区使用数据库

字 段	名 字	类 型	长 度	小 数
001	类别	C	004	
002	中文名	C	018	
003	拉丁名	C	032	
004	门	C	008	
005	亚门	C	012	
006	纲	C	008	
007	亚纲	C	010	
008	目	C	010	
009	亚目	C	012	
010	科	C	012	
011	属	C	013	
012	索引 1	C	030	
013	索引 2	C	064	
014	索引 3	C	038	
015	索引 4	C	038	
016	索引 5	C	030	
017	索引 6	C	040	
018	S7	C	002	
019	S8	C	002	
020	特性	C	070	
021	备注	C	030	
022	下	N	004	
023	Z	C	001	
总计			00489	

表2 原生动物数据变量

Tab. 2 Daea variables of Protozoans

TOTAL (N) 1
NUM (N) 0
D1 (C)TRIM(中文名+拉丁名-'——'-PRE'-'|'-'|'-亚门-'|'-纲-'|'-亚纲-'|'-目-'|'-亚目-'|'-科-'|'-属)
PRE (C) 原生动物亚界
TT3 (C) TRIM (索引1-'-'-索引2-'-'-索引3-'-'-索引4-'-'-索引5-'-'-索引6-'-') + TRIM (TTT)
总计 05UARIABLES USED 00199 BYTES USED

一旦定位于一个门或门以下的阶元，随后的检索将只在相应的块内进行，而在定位之前，它的检索仍是面向全库的，所以分块存贮不会影响信息的完整性。采用分块存贮的目的是为了缩小检索范围，提高查询速度。

程序采用模块化结构，层次清晰，便于修改与扩充，主功能以"菜单"形式给出。

程序结构框(图 1)

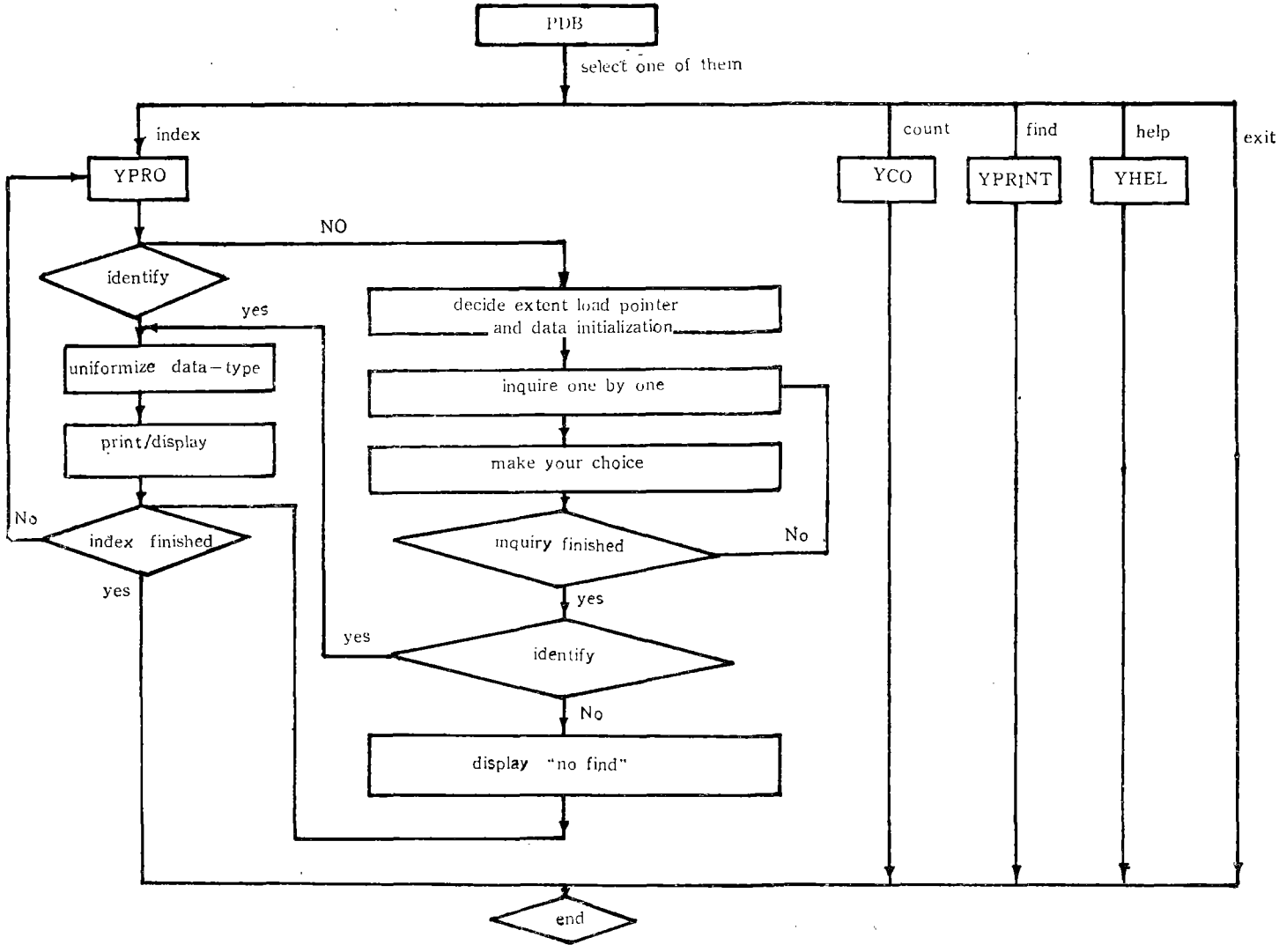


图 1 程序结构框[1)]

Fig. 1 Flowchart

1) 索引 index; 选择其一 select one of them; 定名否 identify; 对数据类型进行统一 uniformize data-type; 打印/显示 print/display; 索引结束 index finished; 决定范围设定指针及其它初值变量 decide extent, load pointer and data initialization; 逐项提问 inquire one by one; 根据回答对可能域进行删取 make your choice; 提问完毕 inquiry finished; 定名否 identify; 显示"未查到" display "no find"; 计数 count; 查询 find;帮助help;退出 exit; 结束 end.

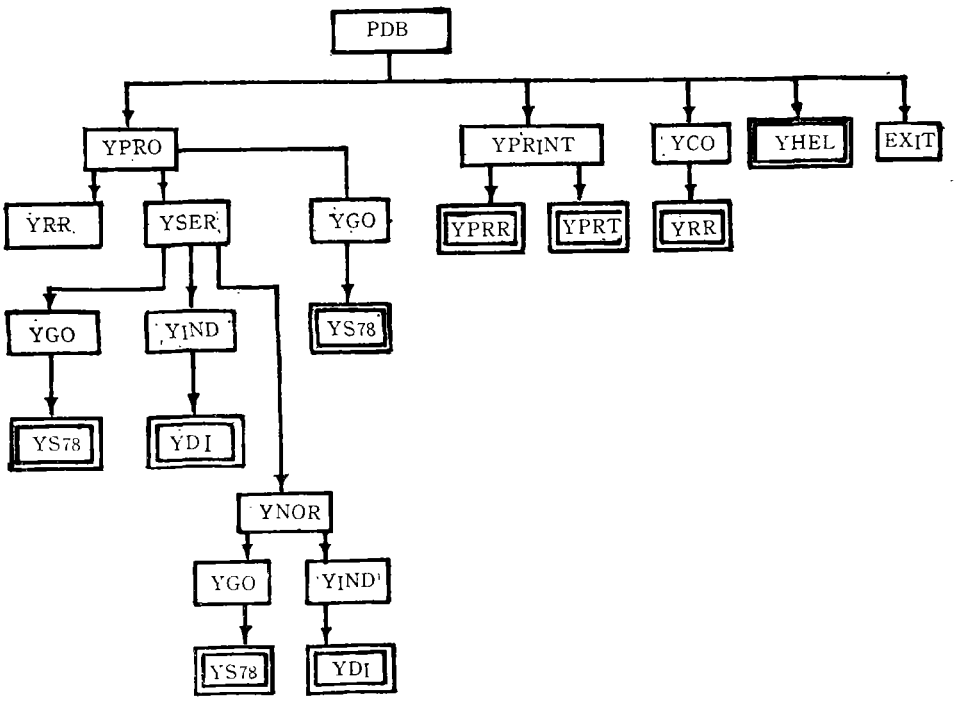


图 2 模块调用过程

Fig. 2 Module Call process

模块功能简要说明

总控: P. D. B. (Protozoa-Data-Base)

打印类: YPRINT 打印总控; Y PRR 打印范围限定; YPRT 打印装配及打印

分类类: YPRO 和 YSER 分类总控; YIND 和 YDI 判断内容的定位; YNOR 判断后的处理; YGO 查到后的处理及打印; YS 78 变量字段的处理

其它: YHEL 简单的使用说明

YCO 统计计数

模块调用情况(图 2)

在程序设计过程中,为解决数据稀疏问题,引用了变量型数据,以节省大量空间,尽量使模块功能专一化,利于修改和扩充; 采用索引文件,对索引机率较高的中文名和拉丁名这二个段使用索引文件后,定位时间小于 2 秒;数据结构上设立了一个终止标记项和一个指针项,以提高检索速度;采用填空式对话方式,可进行屏幕编辑,使操作方便,便于修改。库中还留有一定空记录以便增补。

使 用

1. 环境要求

(1) IBM-PC/XT 机或兼容机; (2) C C DOS 下运行; (3)至少一个硬盘(10 兆),一个软盘驱动器。

2. 方法

(1) 开机　启动 C C DOS (汉字操作系统),屏幕上出现汉字系统提示符 A＞;(2)将主数据文件 mmo. DBF 从二张盘上拷入硬盘 (C:)。

步骤: ①键入 RESTORE A:C:mmo · DBF↵; ②当出现提示时,根据提示相继插入 01 盘和 02 盘; (3)用一张 DBase II 盘调用数据库,即键入 DBase↵,屏幕上将出现 DBase II 提示符; (4) 在 C:上生成两个索引文件 (NDX)。(为了节省软盘空间,我们没有把 NDX 文件存在软盘上,按照以下步骤,用户很容易自己生成该文件)。

键入 USE C:mmo↵

INDEX___ON 拉丁名___To___C:LD↵

INDEX___ON___中文名___To___C:ZW↵

(5) 使用本库,键入 DO___PDB↵屏幕显示菜单

检　索	显示或打印　已知:	计　数	帮　助	退　出
21 中文名	10 记录范围 15 其它	C	H	X
22 拉丁名	某纪录的所有上阶元: 11 中文名 12 拉丁名			
23 记录号	某纪录的所有下阶元: 13 中文名 14 拉丁名			
24 不知道	同分类地位阶元: 16 中文名 17 拉丁名			

(6) 根据用户需要选择某一标号。

- 选择 21—24 时可进行检索,这时用户只需根据屏幕上的提问,决定最合适的一条,键入其标号即可,此期间内对每一个定了名的阶元都可打印输出或退回主菜单。
- 选择 11—15 时进行显示或打印。

其中,选择"只要名字"将只给出中文名和拉丁名,选择"要打印"将在打印机上输出。对"15"(即其它)需键入一个条件表达式,可参阅"帮助"部分或 DBase II 用户手册[2]。

- C 时进行计数,如: 计算动鞭毛纲内共有多少种,多少属等。
- 选择 H 将给出简单说明。
- 选择 X 将退出 P. D. B. 进入 DBase II 系统,对熟悉 DBase II 的人员, 可在此状态中对数据库内容进行操作。

讨　　论

本数据库较充分地利用了 DBase II 的索引功能,并用命令文件扩展了原始功能,根据"菜单"的提示,不用经过专门培训也可利用 P. D. B. 的大部份功能,对使用人员能较快地进行操作,而且人-机对话时间每次不超过 1 分钟,速度比我们最早用的 TRS-80 I 型机大大加快,提高了检索效率,加之本库全部用汉字输出,因此通俗,方便,易于掌握。

在国外建立各种类型数据库是比较普遍的, 但在水生生物研究方面建立生物信息数

据库的报道不多，在国内，至今建立起原生动物数据库的还未见正式报道。我们对建立污水环境中常见的原生动物数据库作了初步探讨，用现代化的计算机技术与传统的分类学结合，提高工作效率，贮备科技信息，将有利于分类学的发展，也为今后在水生生物学其它领域推广应用微机技术打下了基础。

在建立这个信息检索库的过程中，不管是在数据结构的设计上，还是在程序模块的设定与联接上，我们都注意到为其今后的扩展留下余地，使这个库现已积累起来的数据信息能得到进一步的开发与利用。但由于目前原生动物主要还是按形态进行分类，形态分类性状繁杂，必须在一大批数据中不断筛选，对计算机速度带来一定的影响，因此如何提高检索的速度，还有待于从采用的机型和软件两个方面进一步研究，以满足更高的要求。

参 考 文 献

[1] 王家楫、沈韫芬、龚循矩，1983。废水生物处理微型动物图志。1—246 页。中国建筑工业出版社。
[2] 张福炎、蒋新儿、李滨宇，1984。微型计算机 IBM-PC 的原理与应用。1—528 页。南京大学出版社。

PRELIMINARY STUDY OF ESTABLISHING DATABASE OF PROTOZOA IN POLLUTED WATER

Bian Xiaoyu, Shen Yunfen, Wang Chunwei, Gu Manru Shi Zhixin, Gong Xunju, Wei Yinxin and Gu Hanyue

(Institute of Hydrobiology, Academia Sinica, Wuhan)

Abstract

Establishing an information database of Protozoa in polluted water is essential for biological assessment of the effects of waste discharges on aquatic ecosystems. This database contains 702 species belonging to 32 orders, 88 families, 284 genera, with a short morphological description for each taxon and also posseses the functions of information inquiry, counting, classification, etc.

Developed on IBM-PC/XT microcomputer, a model formation was used in the programming of the system, man-machine interaction can be made by a "menu" displayed in Chinese character. It may provide some conveniences to the reseachers dealing with Protozoa.

Key words Database, Protozoan communities, Biological monitoring, Microcomputer, Polluted water