

我国沿海常见浮游植物检索数据库的建立

高 华, 梁君荣, 高亚辉*, 骆巧琦, 李雪松, 陈长平

(厦门大学生命科学学院, 福建 厦门 361005)

摘要: 利用本实验室保存的固定和活体硅藻标本, 并采集我国沿海常见海洋浮游植物样品, 特别是赤潮生物样品, 通过数码显微摄影技术, 采集不同侧面的藻细胞形态特征图像, 同时结合其他手段收集我国沿海常见海洋浮游植物的图片和文字资料. 通过 net(#)语言、asp、net 等技术初步建立了我国沿海常见浮游植物网络检索数据库, 其网络查询界面可实现多分类参数查询. 目前该数据库已收录我国常见浮游植物 144 种, 特征图片 704 张, 以硅藻门 (93 种) 和甲藻门 (40 种) 为主. 同时对每个种类的分类特征和生态分布特点进行了归类与集成. 可通过其细胞主要特征参数或生态分布特征等参数进行详细的信息检索. 该数据库为目前国内首个沿海常见浮游植物网络分类数据库. 本文简要介绍了建库目标、系统结构、功能与特点, 并对该库应用前景作了探讨.

关键词: 浮游植物; 数据库; 分类

中图分类号: Q 94-34

文献标识码: A

文章编号: 0438-0479(2006) S-0230-04

近年来, 随着科研、生产的发展, 以及人们对环境保护意识的日益加强, 在海洋生态资源调查、赤潮监测、环境监测、水产养殖、压舱水检疫等众多方面都离不开对浮游植物的鉴定, 尤其是大规模的海洋生物生态调查、赤潮检测等方面需要对大量海区样品进行藻类鉴定工作, 此项工作难度高, 工作量大, 而有经验的鉴定专家相对匮乏^[1]. 对浮游植物鉴定的自动化是未来发展方向之一. 我国拥有丰富的海洋生物资源, 而对我国沿海常见浮游植物的特征参数、分类数据、生态数据进行建库, 是这项工作的必要基础; 同时也可以使珍贵的浮游植物分类资料得到可靠的保存.

国外生物信息数据库的发展始于上个世纪 60 年代, 比较著名的有三大数据库组成的 DDBJ/EMBL/GenBank 国际核酸序列数据库, SWISS-PROT、PRFOWL、NRL3D、TrEMBL 等蛋白质序列数据库, 以及 PDB、NDB、BioMagResBank、CCSD 等三维结构数据库等, 这些多是为分子生化研究提供信息服务的. 各国研究机构也为给相关研究人员提供生物物种的鉴定、生态资料等信息, 建立了大量的生态分类数据库. 在浮游植物方面, 较为知名的有爱尔兰国家大学建立的 algaebase(<http://www.algaebase.org>), 该库收录了 6 万多种藻类信息和 3 千多张图片, 美国自然科学院的藻类部建立的藻类图像数据库 ANSP Algae Image Data-

base(<http://diatom.acnatsci.org/AlgaeImage/>)收入了美国河流的淡水硅藻显微图片, 日本国家生态研究所建立的藻类资源数据库 Algae Resources Database(<http://shigen.lah.nig.ac.jp/algae>)收录了微生物株系收集中心收集的各种藻类的形态、生理特征、培养条件、环境特性等资料. 这些基于互联网的数据库使得全球范围内生物信息资源的共享交流成为可能, 研究人员可以方便、快捷地得到必须的资料, 极大促进了生命科学的发展.

国内生物多样性信息数据库建设起步较晚, 且多偏重于大型动植物, 如“新疆植物信息资源数据库”, “甘肃省稀有濒危植物数据库”, “山西高等植物数据库”等. 藻类方面已见报道的有中科院科学数据库系统的中国淡水藻类数字影像库(<http://brim.hb.ac.cn/08/CN/08.asp>), 目前, 尚未见针对我国沿海常见浮游植物的分类、生态特征的相关数据库的报道.

本文以厦门大学生命科学学院硅藻实验室的浮游植物固定样品、固定玻片标本和活体标本的显微拍摄采集为基础, 收集各方面资料^[1~4], 建立了我国沿海常见浮游植物数据检索系统框架, 网址 <http://algaebase.oicp.net> 该数据库可为用户提供网络检索服务, 为海洋浮游植物资料保存和科研交流提供了新途径.

1 系统结构和功能设计

1.1 软硬件结构

硬件及操作系统: 高性能 AMD Athlon 64 位微机, 硬盘 80 G, 内存 1 G. 操作系统 Windows XP Professional

收稿日期: 2005-12-30

基金项目: 国家自然科学基金 (40476055、30499340), 973 项目 (2005CB422305) 和农业部海洋与河口渔业重点开放实验室开放课题 (开-2-04-07) 资助

作者简介: 高华 (1981-), 男, 硕士研究生.

* 通讯作者: gaoyh@xmu.edu.cn

al

服务平台:通过 Windows内置的建 Web服务器软件 IIS (Internet information Server)版本 5.1建立 Web 服务平台.

数据库选择: Access2003数据库界面友好,操作简单,与操作系统结合较好,适于建立小型数据库.

软件开发: Asp. net是一个统一的 Web开发平台,能提供企业级 Web应用程序所需的各种服务,在 Net Framework支持下,具有安全性,数据库访问及管理功能,应用程序逻辑等特性^[5]. 本系统开发语言是 C#^[6]. 在此平台下开发的与 www 界面结合的接口程序可以动态地与客户端 (www 浏览器) 交换信息,访问数据库,执行用户的查询请求,将查询结果以 Web形式返回给客户^[7].

1.2 数据库结构

由于海洋浮游植物包括许多类群,而不同类群进行分类的描述有很大差异,将其刻意标准化不具现实意义. 本系统根据不同类群设计不同子库,并依其分类所依据的属性设计不同字段. 在常见海洋浮游植物中,硅藻和甲藻占绝对优势,据此,本数据库的架构也分为“硅藻子库”,“甲藻子库”,“其它藻类子库”. 而且,本数据库根据传统分类所依据的树状分类检索表和今后藻类自动鉴定所需的机器特征,设计了“检索表子库”和“机器特征子库”. 同时为人工的传统特征二叉分类的检索表以及藻类鉴定自动化提供数据库支持,其中“机器特征子库”不提供网络支持. 其结构如表 1.

表 1 中国沿海常见浮游植物数据库结构	
Tab 1 Database structure of common phytoplanktons in China coastal waters	
子库名称	包含字段
硅藻子库	种名,分类位置(纲,目,科等),大小,细胞形状,壳面形状等
甲藻子库	种名,分类位置,大小,横沟,纵沟,鞭毛数量,类型等
其它藻类子库	种名,分类位置,大小,细胞形状等
检索表子库	分类描述,分类名称,上下节点等
机器特征子库	傅立叶函数,周长,圆率等

1.3 数据库中各类浮游植物的组成

数据库共收集我国沿海常见浮游植物种类 144 种,以硅藻和甲藻为主(表 2). 硅藻门共 93 种,占 64%,其中中心纲 56 种,羽纹纲 37 种;甲藻门共 40 种,占 27%;蓝藻门共 5 种,杂色藻门 6 种.

表 2 数据库中各类浮游植物的组成		
Tab 2 Composition of different phytoplankton phyla in the database		
门类	纲	科 种
硅藻门	中心纲: 共 56 种	辐杆藻科: 2 种
		根管藻科: 4 种
		盒形藻科: 13 种
		棘冠藻科: 1 种
		角毛藻科: 4 种
		圆筛藻科: 28 种
		真弯藻科: 3 种
		直链藻科: 1 种
		等片藻科: 8 种
		菱形藻科: 7 种
		卵形藻科: 5 种
		桥弯藻科: 3 种
		曲壳藻科: 3 种
		双菱藻科: 5 种
		舟形藻科: 6 种
		羽纹纲: 共 37 种
		异沟藻科: 1 种
		鳍藻科: 6 种
甲藻门	横裂甲藻亚纲: 共 29 种	角藻科: 4 种
		膝沟藻科: 6 种
		裸甲藻科: 3 种
		角藻科: 4 种
		多甲藻科: 1 种
		多沟藻科: 1 种
		夜光藻科: 1 种
		未定: 2 种
		未定: 11 种
		纵裂甲藻亚纲: 11 种
蓝藻门	5 种	5 种
杂色藻门	针胞藻纲: 2 种	2 种
	定鞭藻纲: 1 种	1 种
	硅鞭藻纲: 2 种	2 种
	金藻纲: 1 种	1 种
	合计	144 种

1.4 客户端界面设计

Web 数据库拥有天然良好的客户端界面,用户可以用普通的 www 浏览器为工具进行查询,建议使用 IE 5.0 以上版本. 查询界面分为两块,一块为高级用户(有分类经验者)提供的二叉检索树分类界面(图 1),一块为中级用户仅通过显微镜下所观察到的藻类所具有的显著外形特点所进行的高级查询界面(图 2).

检索树界面: 用户对照查用检索树时,根据标本特征与检索树上所记载特征进行比较. 如标本特征与记载相符合,则从该分叉逐级向下继续查询. 如其特征与检索表记载的某项号内容不符,则应查阅与该分支相对应的另一分支. 如此继续查对,便可检索出该标本的分类等级名称.

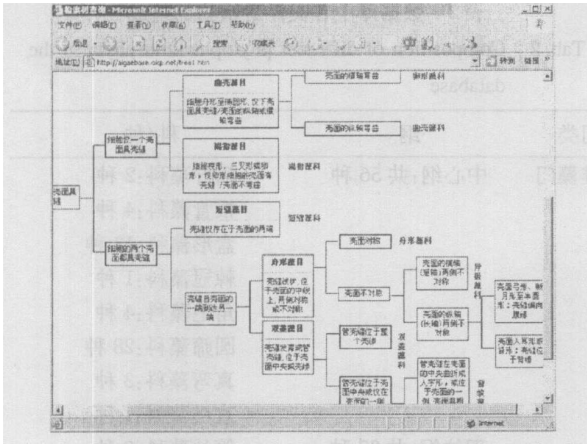


图 1 检索树查询界面

Fig 1 Search tree inquiry interface

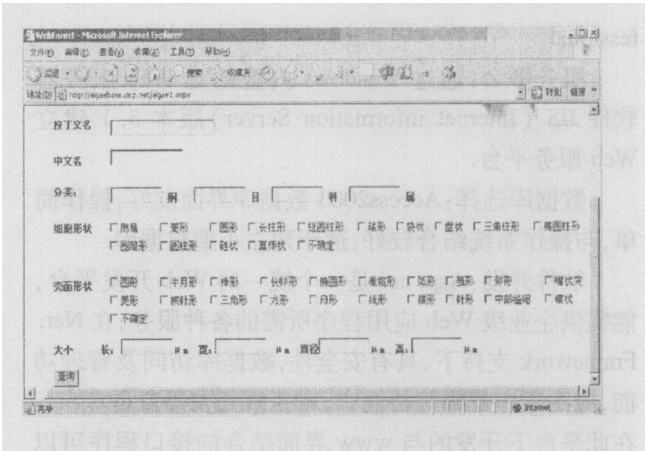


图 2 高级查询界面

Fig 2 Advanced inquiry interface

高级查询界面: 该界面的查询接口由若干输入文本框和复选按钮组成, 分别对应于拉丁文学名查询、中文学名查询、分类位置查询、细胞大小查询 (长、宽、高等)、细胞形状查询、壳面形状查询。用户可将所观察样品的已知最大信息输入相应的查询框, 点击查询, 即可获得数据库中满足查询条件的所有藻类资料的数据表 (图 3)。该查询支持单项查询, 多字段组合查询及模糊查询。出于易用性的考虑, 该表只显示出相对应藻类的简要信息, 要获取进一步的信息, 可点击“详细信息”的超链接, 即可获得该藻的图片资料, 分类描述资料, 生态分布资料, 特征性描述等信息。

关信息, 没有充分利用信息科学的优势。本文考虑到高、中级藻类分类人员的不同需求, 分别设计了检索树分类界面和高级字段查询界面。

传统的分类检索表根据排版方式的不同分为定距式 (级次式)、平行式和连续平行式等类型。其中定距式的同一特征群体的生物在空间上比较靠近, 相对直观。但由于成对特征之间的距离常常很远, 且不易于书面排版, 使用受到限制。本数据库的检索树吸收定距式的优点, 以超链接的方式克服其排版上的缺点, 用树形图指出成对特征之间的联系, 使用更为直观方便。

植物学经典分类检索表的编制依据是拉马克 (Lamarck) 的二歧分类原则, 即把同一类群的植物选用明显而相关的形态特征分成相对应的两个分支, 再把每个分支中的分类群再用相对的性状分成相对应的两个分支, 依次下去, 直到将所有分类群分开为止。一般不提供直观的图片, 也不对不同类群的生物进行完整的特征描述, 用于二分的文字描述常涉及大量专业术语, 因而使用检索表需要一定专业背景并接受专门的训练, 否则很难通过检索表获得正确的结果。然而, 电子数据库具有以字段形式存储信息的特性, 高级字段查询界面使用户完全可以同时以几个特征描述 (如大小、形状、鞭毛数量等) 进行并行检索, 迅速得到结果。此外, 如果在检索过程中发现某一性状缺失或无法确定, 传统二歧分类方法就难以进行下去, 而数据库多字段查询则至少可以缩小范围, 呈现符合条件的结果。

2.2 本数据库特点

(1) 适用性广。在任何接入互联网的, 有 www 浏览器 (如 IE) 的微机上均可使用。

(2) 简明易用, 面向用户。符合不同分类经验的用户的思考习惯, 使用方便, 用户不需具备计算机专业知识都可运用自如。藻类分类不再是纯文字描述, 而是图

图 3 藻类资料数据表

Fig 3 Information table of phytoplankton data

2 讨论

2.1 检索界面设计

目前, 绝大部分藻类数据库 (如前言提及) 的检索界面均只提供学名、分类单位等少数几个检索字段, 使用人员只有针对某一已知学名的藻株才能得到该藻相

文并茂和具超链接.

(3)实用性强. 将传统的以纸张、玻片等形式保存的经典分类资料进行电子化的汇总和储存. 查询功能强大, 用户可借助该网络数据库迅速获得中国沿海有关浮游植物充分的资料. 可同时接收多个用户的查询浏览.

(4)可扩展性. 由于客户端和服务端相分离, 管理者可以随时进行必要的资料扩充和结构升级.

2 3 尚需解决的问题

目前该数据库已收录我国常见浮游植物 144 种, 特征图片 704 张, 同时对其分类特征和生态分布特征进行归类与集成. 虽已具备初步架构, 但藻类资料在未来仍需进一步增添和完善. 此外, 在对传统分类描述数据库化的过程中, 依旧存在着某些定性描述术语的不确定性给编程带来的问题, 这需要有经验的分类专家对其进行人工的标准化和代码化. 最后, 在数据库的结

构上, 有经验用户的检索树分类过程和高级查询过程还未得到有机的结合, 仍需在将来的工作中进一步改进.

参考文献:

[1] 金德祥, 陈金环, 黄凯歌. 中国海洋浮游硅藻类 [M]. 上海: 上海科学技术出版社, 1965

[2] 程兆第, 高亚辉. 硅藻彩色图集 [M]. 北京: 海洋出版社, 1996

[3] 梁象秋, 方纪祖, 杨和荃. 水生生物学 [M]. 北京: 中国农业出版社, 1996

[4] 齐雨藻. 中国沿海赤潮 [M]. 北京: 科学出版社, 2003

[5] 霍默, 萨斯曼, 李敏波. ASP. NET 1. 1 高级编程 [M]. 北京: 清华大学出版社, 2005

[6] 布拉德利, 米尔斯波, 天宏工作室. C#. NET 程序设计 [M]. 北京: 清华大学出版社, 2005

[7] 李陶深. 网络数据库 [M]. 重庆: 重庆大学出版社, 2004

Construction of a Database of Common Marine Phytoplankton in Coastal Waters of China

GAO Hua LIANG Jun-rong GAO Ya-hui*,
LUO Qiao-qiang LIXue-song CHEN Chang-ping
(School of Life Sciences, Xiamen University, Xiamen 361005, China)

Abstract A database of common marine phytoplankton in coastal waters of China was constructed with NET (C#) and ASP. NET technology. The database contained species description information and 704 characteristic images of 144 species or varieties of common phytoplankton including 93 species of diatoms and 40 species of dinoflagellates, especially red tide causative organisms. These materials were collected from different sources including living and fixed specimens available in author's lab and published papers. The specimens were photographed using microscope equipped with digital camera. The description and distribution information were compiled for the multiple parameters query with the web query interface. This is the first marine phytoplankton database in China. This article introduces briefly the purpose, system structure and function of this database. The perspective of application of this database was also discussed.

Key words phytoplankton; database; taxonomy