

# An Integrated Database and Web Service for Microbial Resources at KACC

Chang Kug Kim<sup>1</sup>, Young Ah Jeon<sup>2</sup>, Gyu Taek Cho<sup>2</sup>, Soon Wo Kwon<sup>2</sup>, Yong Hwan Kim<sup>1</sup> and Seung Beom Hong<sup>2\*</sup>

<sup>1</sup>Genomics Division and <sup>2</sup>National Agrobiodiversity Center, National Academy of Agricultural Science (NAAS), Suwon 441-707, Korea

## Abstract

The Korean Agricultural Culture Collection (KACC) has developed a web-based system to provide an integrated database with information updates about microbial resources. This integrated database consists of 5 major functions and contains general information, which includes identification numbers, culture media composition, image information, DNA sequences, patent information, and general forms for ordering and depositing microorganisms. In 2008, KACC started providing characterization information. KACC maintains 9,801 cultures of microorganisms, including 3,296 strains of bacteria, 4,734 fungi, 784 actinomycetes, 64 yeasts, and 923 others.

**Keywords:** KACC, microbial resource, integrated database, microorganism information

## Introduction

Microbial resources are closely associated with our daily life, and their collection and utilization are very important. In terms of microbial resources, the value of a microorganism database depends on the amount, quality, and accuracy of the information that it contains. Microbial databases have been developed in various fields, including the IMG 2.3 (Victor *et al.*, 2008), NMPDR (McNeil *et al.*, 2007), and MBGD databases (Uchiyama *et al.*, 2007). In line with the increase in the number of databases for microbial resources, various integrated web service systems have been constructed. BCCM<sup>TM</sup> (<http://bccm.belspo.be/>) consists of 4 complementary research-based service culture collections. This database maintains over 53,500 well-documented and

authenticated strains of bacteria, filamentous and yeast fungi, over 950 plasmids, and 18 DNA libraries. KCTC (<http://kctc.kribb.re.kr/>) has a wide and systematic collection of materials that comprise microorganisms, animal and plant cell lines, and patent strains from domestic and international sources. UKNCC (<http://www.ukncc.co.uk/>) offers identification by a range of techniques for many organisms, including actinomycetes, algae, animal cells, arthropods, bacteria, filamentous fungi, nematodes, protozoa, mycoplasma, and yeast. JCM (<http://www.jcm.riken.go.jp/>) supplies authentic microorganisms to researchers in the fields of life sciences and biotechnology. JCM preserves 6,600 strains of bacteria, including actinomycetes, 280 strains of archaea, and 4,000 strains of fungi, including yeast.

In Korea, the Korean Agricultural Culture Collection (KACC, <http://kacc.rda.go.kr/>) serves as an authorized organizer and the official depository (NIAB, 2008) of microbial resources. KACC maintains microbial resources, including bacteria, actinomycetes, yeast, and filamentous fungi, and has developed a database of microbial resources (Chang *et al.*, 2009). In order to collect microbial resources from international centers, KACC joined the World Federation for Culture Collection (WFCC), the World Data Center for Microorganism (WDCM), and the Asian Consortium for the Conservation and Sustainable Use of Microbial Resources (ACM). In 2008, KACC developed a web-based system to provide an integrated database, which contains updated information on microbial resources.

## Methodology

### Dataset

The database information was collected from the microorganism project (<http://kacc.rda.go.kr/>), the National Agrobiodiversity Center (NAC, <http://genebank.rda.go.kr/>), the National Academy of Agricultural Science (NAAS, <http://www.niast.go.kr/>), the Bio-Green 21 project (<http://biogreen21.rda.go.kr/>), relevant microorganism divisions of universities, and various institutes in Korea. In most instances, the strains were isolated from Korean agricultural environments, but some strains were shared with other culture collections. In addition, the database information was accumulated and maintained through several collaborating international institutes, such as the Centraalbureau voor Schimmelcultures (CBS, [\\*Corresponding author: E-mail \[sbhong@rda.go.kr\]\(mailto:sbhong@rda.go.kr\)  
Tel +82-31-299-1655, Fax +82-31-299-1657  
Accepted 2 February 2009](http://www.</a></p></div><div data-bbox=)

cbs.knaw.nl/) in the Netherlands, ARS culture collection (NRRL, <http://nrri.ncaur.usda.gov/>) in the USA, and the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, <http://www.dsmz.de/>) in Germany.

### Database design

The integrated database is designed to provide information on Korean microbial resources with a bio-database management system (Tae *et al.*, 2008). The Entity Relation Diagram (ERD) of the database is shown in Fig. 1. The schema consists of 4 major functional categories. In the first category, a general characterization table shows the identification of microorganisms by TCSP\_INFO table. The MGDTD table shows the process of registration for microorganisms, including Korean viruses, fungi, and others. The operating table and management table show the ordering and management process for microorganisms, patent strains, and gene clones. Using the collected information, a microbial resources database system was developed. The platform was developed using MYSQL and JAVA languages, and the data were stored in an Oracle Relational Database Management System (RDBMS). The logical and the physical schema of the database fol-

lowed the standard principles of relational databases using by ERWin Data Modeler software (<http://www.ca.com>), and the data were distributed into numerous tables to establish the hierarchical relationships (Chang *et al.*, 2008) between or among the datasets.

### Database contents

The integrated database now maintains 9801 cultures of microorganisms, including 3296 strains of bacteria, 4,734 fungi, 784 actinomycetes, 64 yeasts, and 923 others (mushrooms, gene clones, etc.). In 2008, a total of 1,855 strains of new microbial resources were registered, including 616 isolates of bacteria, 950 isolates of fungus (including Basidomycetes), and 289 actinomycetes. In most instances, microorganisms are industrial bacteria, such as *Lactobacillus*, *Bacillus*, and *Burkholderia*. The database contains rare fungi, such as *Ceroospora*, *Septoria*, and *Ramularia*. Table 1 shows the number of registered microbial resources by year.

### Implementation and features

The integrated database provides various information, including text record, morphological images, and DNA

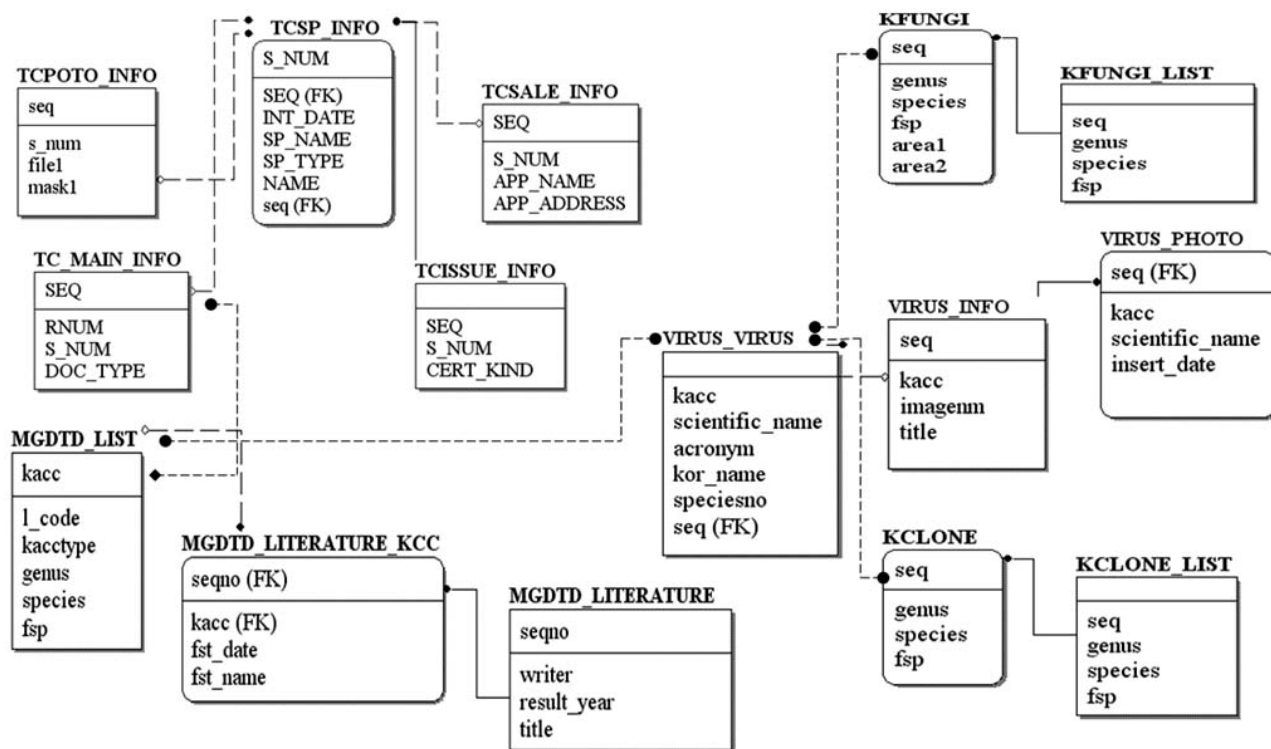


Fig. 1. Overview of the integrated database design and Entity Relation Diagram (ERD). Each box represents a separate entity, implemented as a database table. Lines between entities indicate that relationships exist.

sequence data. The database platforms consist of 5 major functions: microorganism search, ordering, deposit, patent microorganism management, and gene clone management.

The search function can perform searches by type of strain, as shown in Fig. 2. The searching of microorganisms by genus and species can be done using 2 methods: 1. A keyword search (Fig. 2A) by which users can search using partial or complete species names from 2 classifications or 6 categories; to choose a specific menu for classification and categories, users may specify one submenu from the pull-down menu. 2. Alphabetical order searches (Fig. 2B) in which the information tables of strains are alphabetically arranged according to the scientific names. Users may search for specific data by combining search parameters and can search linked information by using different search windows, such as gene clone (Fig. 2C) and media search

windows (Fig. 2D).

Screenshots of results pages from various search methods are shown in Fig. 3. In the information table of specific microorganisms, users can access an image information table by clicking the thumbnail image on the right (Fig. 3A). A detailed image table provides information that consists of title, identification number, scientific name, accession number, definition, description, image number, authors, institute, and memo record (Fig. 3B). By clicking a linked 'Media' field, users can access the information tables that contain individual media composition data. The media tables provide information on the media for growing the microorganisms (Fig. 3C). A user can acquire a specific strain by clicking on the linked 'Strain Order' object, which will then download an order form (Fig. 3D).

When using the keyword or alphabetical search, users can query different identification names or historical classification fields by using the hierarchical tree web form. The query results are linked to each detailed table through their respective strain lists. The user can then view an elements report page, and it is possible to download the detailed information, including the specific DNA sequence and image information. Users can also access a page that contains individual plant disease information that is related to a specific microorganism by clicking the 'Search for plant disease list in Korea' link. This table provides information on the 'List of Plant Diseases in Korea,' including descriptions of all related pathogens and annotations from the Korean Society of Plant Pathology (KSPP, <http://www.kspp.org/>).

**Table 1.** The accumulated number of registered microbial resources at KACC, 2005-2008

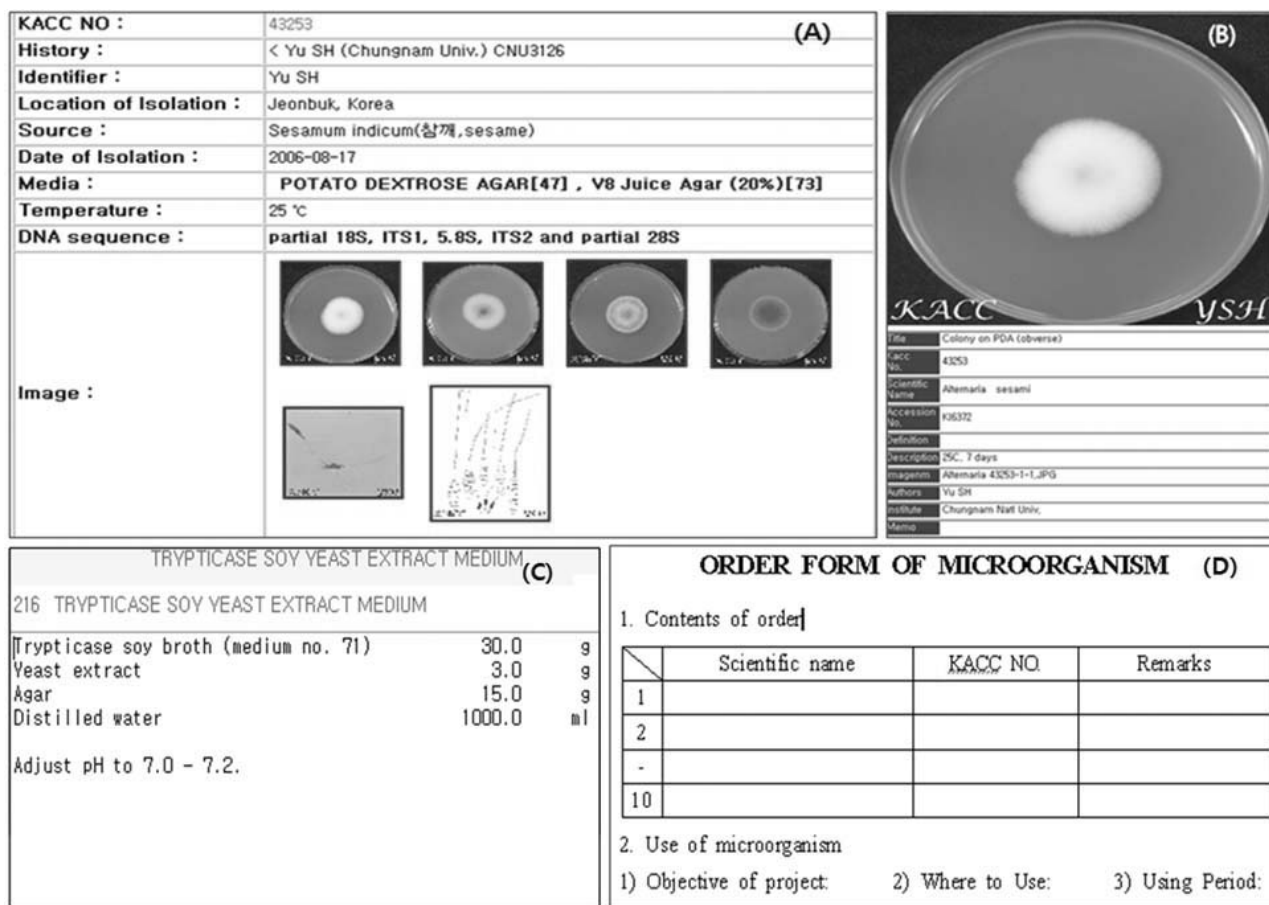
Year	Bacteria	Fungi	Actin.	Yeast	Others	Total
~2005	1,587	2,674	386	64	893	5,604
2006	519	410	55	-	26	1,010
2007	574	700	54	-	4	1,332
2008	616	950	289	-	-	1,855
Total	3,296	4,734	784	64	923	9,801

Figure 2 displays four panels of the database search interface:

- (A)** Keyword search interface. It features a 'Classification' dropdown menu set to 'Bacteria', a 'Search with' dropdown menu set to 'Scientific Name', and two input fields for 'Genus' and 'Species'. A 'SEARCH' button is located at the bottom.
- (B)** Alphabetical search interface. It shows a row of buttons labeled A through L, with 'B' highlighted. Below the buttons is a 'SCIENTIFIC NAME' section containing a list of results: *Acetobacter aceti*, *Acetobacter cerevisiae*, and *Acetobacter cibinongensis*.
- (C)** Searching gene clone interface. It includes a 'Classification' dropdown menu, a 'Search with' dropdown menu set to 'Clone Name', and an input field for 'Clone Name'.
- (D)** Media search window and report page. It shows a dropdown menu with '298. 50-fold diluted nutrient broth agar medium' selected. Below the menu is a table with the following details:
 

298 50-fold diluted nutrient broth agar medium	
Beef extract	0.2g
Peptone	0.2g
NaCl	0.1g
Distilled water	1.0L
Agar	1.5%

**Fig. 2.** Search windows of the integrated database. The individual panels show the windows for searching microorganisms. (A) Web interface for keyword search by genus and species. (B) Web interface for alphabetical search by scientific names. (C) Web interface for searching gene clone by classification. (D) A media search window and report page with detailed information on a specific element name.



**Fig. 3.** Search results of the integrated database. The individual panels show samples of analytical results. (A) An information table showing records on identification number, history, identifier, location of isolation, source, date of isolation, media, temperature, DNA sequence, and image. (B) Image view with related general information. (C) A media report page with detailed information on an element. (D) An order form for acquiring specific strains of microorganisms.

In addition, KACC provides a function for the ordering and depositing of microorganisms, including patent microorganisms or gene clones. Anyone with the basic knowledge and the facilities for culturing may order and deposit microorganisms. However, orders for personal purposes may not be accommodated. A user can order a strain in 3 steps: (i) search for a strain from the database; (ii) fill-in the order form, including the ‘Materials Transfer Agreement;’ and (iii) send the request form by postal mail or email. The patent menu has 3 major functions, such as those related to ordering, depositing, and regulations. Users can view different classification tables or specific forms by using the corresponding web forms.

**Discussion**

KACC classifies and preserves diverse Korean microbial resources including bacteria, yeasts, filamentous fungi, and mushrooms and provides an integrated database.

**Table 2.** The total number of microbial resources distributed to researchers, 2006-2008

Year	Bacteria	Fungi	Actin.	Yeast	Others	Total
2006	416	847	19	45	41	1,368
2007	639	660	51	20	48	1,411
2008	1,006	982	89	25	42	2,144

This integrated database contains general information, such as identification numbers, culture media composition, image information, DNA sequence, and general forms for ordering and depositing microorganisms. In 2008, KACC, which maintains 9,801 cultures of microorganisms, started providing characterization information. With this, the number of isolates that were distributed to researchers who are involved in microorganism-related industries increased to more than 2,100 in 2008 (Table 2).

## References

- Kim, C.K., Lee, M.C., Ahn, B.O., Yun, D.W., Yoon, U.H., Suh, S.C., Eun, M.Y., and Hahn, J.H. (2008). KRDD: Korean rice Ds-tagging lines database for rice (*Oryza sativa* L. Dongjin). *Genomics & Informatics* 6(2), 64-67.
- Kim, C.K., Jeon, Y.A., Cho, G.T., Kwon, S.W., Hahn, J.H., and Hong, S.B. (2009). KACC: An identification and characterization for microbial resources in Korea. *Afr. J. Biotechnol.* 8(1), 69-72.
- McNeil, L.K., Reich, C., Aziz, R.K., Bartels, D., Cohoon, M., Disz, T., Edwards, R.A., Gerdes, S., Hwang, K., *et al.* (2007). The National Microbial Pathogen Database Resource (NMPDR): a genomics platform based on sub-system annotation. *Nucleic Acids Res.* 35, D347-D353.
- NIAB. (2008). *Annual report 200*. National Institute of Agricultural Biotechnology.
- Tae, H.S., Han, J.M., Ahn, B.Y., and Park, K.J. (2008). A bio-database management system for the monitoring and automatic FTP of public databases. *Genomics & Informatics* 6(2), 95-97.
- Uchiyama, Ikuo. (2007). MBGD: a platform for microbial comparative genomics based on the automated construction of orthologous groups. *Nucleic Acids Res.* 35, D343-D346.
- Markowitz, V.M., Szeto, E., Palaniappan, K., Grechkin, Y., Chu, K., Chen, I.M., Dubchak, I., Anderson, I., Lykidis, A., Mavromatis, K., Ivanova, N.N., and Kyrpides, N.C. (2008). The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. *Nucleic Acids Res.* 36, D528-D533.