

Informe final* del Proyecto HA025

Actualización de la base de datos de la Colección nacional mexicana de cultivos microbianos

Responsable: M en C. Jovita Martínez Cruz
Institución: Centro de Investigación y de Estudios Avanzados
Instituto Politécnico Nacional
Departamento de Biotecnología y Bioingeniería
Colección microbiana
Dirección: Av. Politécnico Nacional # 2508, México, D.F., 07000
Correo electrónico: jmartine@cinvestav.mx
Teléfono, fax +52 (55) 5747 3903; Fax +52 (55) 5747 3975
Fecha de inicio: Enero 29, 2010
Fecha de término: Febrero 21, 2014
Principales resultados: Base de datos, Informe final.
Forma de citar el informe final y otros resultados:** Martínez Cruz, J., 2014. Actualización de la Base de Datos de la Colección Nacional Mexicana de cultivos microbianos. Centro de Investigación y de Estudios Avanzados. Instituto Politécnico Nacional. **Informe final SNIB-CONABIO, proyecto No. HA025**, México, D.F.

Resumen:

La información relacionada a microorganismos es una de las más importantes para cualquier país, debido a que mantiene una estrecha relación con relevantes áreas esenciales como: salud, medicina, agricultura e industria, razón por la cual es necesario mantener un control estricto, minucioso y valido de toda esta información, ya que tiene diferentes aplicaciones.

Por otra parte, alrededor del mundo constantemente surgen nuevos conocimientos científicos relacionados a microorganismos; esto involucra una fuerte necesidad de mantener actualizada toda esta información.

Este proyecto pretende hacer una revisión de nuestra información, además de incorporar conocimiento actualizado para las diferentes cepas relacionado a cambios taxonómicos, nuevas características, nuevos estudios y aplicaciones, etc. Adicionalmente se incorporarán 600 nuevos registros de dos colecciones de microorganismos externas, a las cuales se les asignara un acrónimo CDBB para su reconocimiento a nivel internacional.

La incorporación de esta nueva información requiere de una modificación parcial de nuestra base de datos, porque algunas características de esta nueva información no fueron consideradas en el diseño original. La modificación nos permitirá enriquecer el conocimiento con propósitos de investigación en biotecnología, así mismo permitirá integrar nueva información a la red REMIB a la cual pertenece nuestra colección para su explotación.

-
- * El presente documento no necesariamente contiene los principales resultados del proyecto correspondiente o la descripción de los mismos. Los proyectos apoyados por la CONABIO así como información adicional sobre ellos, pueden consultarse en www.conabio.gob.mx
 - ** El usuario tiene la obligación, de conformidad con el artículo 57 de la LFDA, de citar a los autores de obras individuales, así como a los compiladores. De manera que deberán citarse todos los responsables de los proyectos, que proveyeron datos, así como a la CONABIO como depositaria, compiladora y proveedora de la información. En su caso, el usuario deberá obtener del proveedor la información complementaria sobre la autoría específica de los datos.

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional



Reporte Final del Proyecto:

**“HA025 Actualización de la Base de Datos de la
Colección Nacional Mexicana de Cultivos Microbianos”.**

Integrantes del Proyecto:

Responsable: Profa. Jovita Martínez Cruz – Cinvestav.

Responsable Externo: Dr. Joaquín Sergio Zepeda Hernández – UAM-C.

Curador: Biol. Juan Carlos Estrada Mora – Cinvestav.

Capturista: Tec. Armando Sánchez Chavarría – Cinvestav.

Informático: Ing. Israel Piñones Pozos.

Asesor: Dr. Sergio Víctor Chapa Vergara – Cinvestav.

19-julio-2013

México, D.F

Resumen

El presente reporte final ofrece una descripción detallada de todo el trabajo realizado durante el proyecto “HA025 Actualización de la Base de Datos de la Colección Nacional Mexicana de Cultivos Microbianos”, el cual fue apoyado por la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (Conabio). El proyecto fue desarrollado en dos vertientes: *Base de datos e información microbiana*, las cuales fueron trabajadas de manera paralela a fin de cubrir los tiempos comprometidos.

La primera vertiente cubre de manera esencial el diseño del nuevo modelo de datos, así como los detalles de implementación, descripción de tablas y campos. La segunda vertiente relacionada a la información microbiana, contempla una descripción del proceso de *análisis, revisión, verificación, validación e incorporación* de 1000 registros nuevos respecto al informe anterior, estableciendo la cantidad de 1800 cepas en total (*información de 200 cepas adicionales a las comprometidas*) y un reporte detallado sobre los datos contenidos en las tablas.

Parte de la gran importancia de este proyecto, es que anexa nueva información de colecciones externas de microorganismos como: La Colección de *Histoplasma capsulatum* perteneciente a la Universidad Autónoma de México (UNAM), y la Colección de Levaduras Marinas del Centro de Investigaciones Biológicas del Noroeste (CIBNOR). Lo que convierte a nuestra Colección en referente nacional sobre información microbiana.

La actualización de información y diseño de la base de datos rebasan las expectativas iniciales y se mantiene la perspectiva de que servirán para el desarrollo de nuevos proyectos y la adecuación para compatibilidad con nuevas tecnologías emergentes.

1. Introducción

Las colecciones de cultivos microbianos son entidades en donde se realizan una serie de actividades cuyo objetivo fundamental es: obtener, preservar, clasificar, estudiar y documentar de manera completa y accesible un acervo de cultivos de microorganismos auténticamente puros, estables y bien clasificados de interés específico, los cuales generan toda una serie de información especializada de importante relevancia en diferentes ámbitos como: salud, medicina, industria, agricultura, biotecnología, investigación y docencia.

En México desde 1972 se fundó la Colección Nacional de Cepas y Cultivos Microbianos perteneciente al Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav), desde su fundación la Colección brinda los servicios de: *Depósito, Conservación, Aislamiento, Selección, Identificación, Clasificación, Caracterización y Tratamiento para procesos industriales*. Adicionalmente a nivel nacional e internacional se ofrece: *Asesoramiento técnico, Cursos de entrenamiento y Cursos de actualización*.

La Colección es el único organismo reconocido oficialmente a nivel nacional dedicado al mantenimiento, distribución y depósito de cultivos microbianos para actividades de investigación de diferentes instituciones educativas nacionales. En el ámbito internacional ocupa un lugar preponderante al ser reconocida como una autoridad internacional en el mantenimiento, depósito y conservación bajo resguardo de cepas microbianas puras y establecer colaboración con colecciones microbianas de diferentes países. Además fue la primera colección en Latinoamérica en contar con un catálogo de cepas en línea e incorporarse a federaciones mundiales de bases de datos microbianas como: World Data Center for Microorganisms (WDCM), Agricultural Research Service (ARS) y el portal StrainInfo de la Universidad Gent en Bélgica.

Actualmente la colección contiene información de cuatro diferentes tipos de cepas que son: *Bacterias, Hongos filamentosos, Levaduras y Micro-algas*, donde cada tipo cuenta con características particulares y especializadas todas ellas de gran importancia para el desarrollo del país.

2. Antecedentes del proyecto

En 1977, la Colección Nacional de Cultivos Microbianos fue aceptada por el World Data Center for Microorganisms (WDCM) con el acrónimo: *Colección del Departamento de Biotecnología y Bioingeniería (CDBB)* y el número de registro 500 para su identificación a nivel internacional. Para el año 1981, quedó formalmente afiliada a la World Federation for Culture Collections (WFCC) y en 1992, la WFCC postuló a la colección a formar parte del Comité de Educación Internacional de Colecciones de Microorganismos.

A finales de 1993, la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), solicita al Cinvestav el desarrollo de un sistema de datos prototipo para el manejo de información microbiana a nivel nacional y establecer un acercamiento e intercambio de información entre diferentes instituciones y laboratorios nacionales que trabajaban con este tipo de información especializada.

De esta manera, en 1994 por parte de la Colección Nacional de Cepas Microbianas la Maestra Jovita Martínez Cruz junto con la sección de Computación del Departamento de Ingeniería Eléctrica del Cinvestav representado por el Dr. Sergio Chapa Vergara, definen un proyecto interdisciplinario para dar respuesta a la solicitud de Conabio que propone dar apoyo mediante el *proyecto P-153 "Colección de Cultivos Microbianos del CINVESTAV fase 1, Base de Datos", convenio FB122/E016/94*, donde los principales objetivos del proyecto fueron:

- a) Crear la primera base de datos relacionada y especializada a información de microorganismos.
- b) Administrar y poner a disposición de la comunidad la información de la base de datos.
- c) Desarrollar un sistema que facilite el intercambio de información microbiana, por medio de una red nacional entre diversas instituciones.
- d) Implementar proyectos asociados que permitan la investigación y el desarrollo tecnológico en computación y microbiología.
- e) Animar la colaboración entre diversos institutos y laboratorios especializados en cepas microbianas.

Como resultado de este primer proyecto y con el fin de cumplir los objetivos iniciales los resultados que se obtuvieron fueron los siguientes:

- Desarrollo del modelo y esquema conceptual de la base de datos con información relacionada a microorganismos denominada como CDBB-500.
- Desarrollo del primer diccionario de datos de la base CDBB-500.
- Implementación de la base de datos en Access a solicitud de la Conabio.
- Primera versión del Sistema de Información CDBB-500 V 1.1 (ver Fig. 2.1).

En 1997 nuevamente la CONABIO apoyo el *proyecto P-153 "Colección de Cultivos Microbianos del CINVESTAV fase 1, Base de Datos, parte de vista Usuarios Externos", convenio FB357/J015/94.*

Los alcances relevantes que se obtuvieron con ambos convenios fueron:

- El sistema de información CDBB-500 V 1.1. implementado en Microsoft Access, el cual contaba con pantallas de captura, consultas, reportes y modificaciones de los registros en la base de datos.
- La base de datos CDBB-500 con 1004 registros referentes a cepas microbianas con información especializada sobre: género y especie, depositante, acrónimo, características de cepa, taxa, fuente de aislamiento, medios de crecimiento, temperatura óptima de crecimiento, medios de cultivo, índice y formulas, etc.
- La distribución del sistema de información a diferentes instituciones nacionales para administrar su información propia.

Fig. 2.1.- Sistema de Información CDBB-500 V 1.1.

Después de este proyecto la Sección de Computación del Cinvestav trabajo sobre sobre la misma base de datos otro tipo de aplicaciones con el fin de facilitar la disponibilidad de los datos de una manera más interactiva dada la complejidad de la base de datos. Por su parte la Conabio crea la Red Mundial de Información sobre Biodiversidad (REMIB) e incorpora información de CDBB-500.

En el año de 2001 dados los nuevos avances tecnológicos, se empezó a desarrollar un sistema que permitiera mostrar información pública en la Web y de esta manera dar alcance a los países más avanzados que ya mostraban catálogos especializados en línea sobre información microbiana.

En el año 2002 se libera la primera versión del sistema denominado como: Micro500 el cual es el primero en Latinoamérica en ofrecer información de cepas con las que cuenta la Colección a través de la dirección Web <http://micro500.cs.cinvestav.mx> (ver figura 2.2).



Fig. 2.2 Vista principal del Sistema Micro500.

3. Objetivos

El nuevo conocimiento que ha surgido recientemente relacionado a microorganismos obliga a realizar una revisión exhaustiva de la información ya registrada, así como de incorporar nuevos datos no contemplados en el diseño original de la base de datos debido al cambio de contexto y requerimientos actuales. De esta manera los objetivos planteados y cubiertos en este proyecto son:

Objetivos generales

- a) Mantener la base de datos de microorganismos con información actualizada y especializada.
- b) Fortalecer el intercambio de información científica sobre biodiversidad microbiana, con diferentes bases de datos nacionales e internacionales.
- c) Incluir dos colecciones microbianas externas a CDBB-500: Colección de levaduras del Centro de Investigación de Baja California Norte (CIBNOR) y la colección de *Histoplasma capsulatum* de la Facultad de Ciencias de la UNAM.

Objetivos específicos

- Modificación sustancial de la base de datos CDBB-500, para incorporar los nuevos requerimientos de nuestro proyecto como: nuevas características, nuevos registros, etc.
- Actualización de información en la base de datos.
- Inclusión de nuevos campos de información no contenida en la actual base de datos.
- Ingreso de 600 nuevos registros y su respectiva información.
- Revisión, depuración, verificación y validación de información actual y nueva en la base de datos.

4. Metodología

Para alcanzar los objetivos presentados en este proyecto la metodología usada en este proyecto fue dividida en dos vertientes de trabajo paralelo: base de datos e información microbiana, dada la naturaleza interdisciplinaria que conlleva alcanzar resultados satisfactorios en este ámbito.

Respecto a la parte de la base de datos se procedió de la siguiente manera:

- Analizar y estudiar el modelo conceptual original de CDBB-500, para tener un panorama general de la base de datos.
- Obtener los requerimientos actuales de la nueva información a través de intercambio de conocimientos y enseñanza sobre caracterización de cepas en laboratorio.
- Diseñar un nuevo esquema conceptual con ayuda de los curadores de la Colección, los cuales incluyeron una serie de pláticas para entender la naturaleza semántica y compleja de la información de este tipo.
- Validar el nuevo esquema conceptual de CDBB-500 ver. 2.0, ya creado se procedió a un ciclo evolutivo de desarrollo para obtener un modelo avanzado que cumpliera con los requerimientos de la Colección.
- Crear el nuevo modelo lógico relacional de CDBB-500 ver. 2.0, según las entidades obtenidas para desarrollo en este proyecto. Cada entidad necesito de una revisión continua que permitiera la eliminación de redundancias en los datos y obtener la correspondiente relación entre datos.
- Crear el nuevo diccionario de datos una vez que se obtuvieron las diferentes tablas necesitadas, también se verifico el tipo de dato y tamaño requerido para cada dato.
- Implementar la nueva base de datos en el gestor PostgreSQL 8.4 y cotejar la correspondencia con el diccionario de datos, así como la codificación de los datos.
- Realizar la migración de los datos capturados para integrarlos a la nueva base y verificar la integridad y correspondencia.

Por la parte de información microbiana se sometió a revisión los datos contenidos en la base de datos CDBB-500 de la siguiente manera:

1ra. Fase.- Datos relacionados con Bacterias.

2ª. Fase.- Datos relacionados con Micro-Algas.

3ra. Fase.- Datos relacionados con Levaduras.

4ª. Fase.- Datos relacionados con Hongos.

Cada una de las fases mencionadas siguió el proceso siguiente para considerar la revisión completada:

- Obtener diferentes tipos de listados de las cepas con todas sus características.
- Revisar errores de sintaxis en las descripciones.
- Verificar la información con datos obtenidos de estudios y experimentos previos realizados en la Colección dados los servicios prestados en diferentes ocasiones.
- Validar información respecto a algunas cepas en laboratorio para verificar datos almacenados, usando la amplia experiencia de los curadores.
- Actualizar la información necesaria y que no corresponde o falta en la base de datos.
- Buscar la taxonomía faltante (autores y año) para los datos correspondientes a cepas microbianas.
- Buscar en la literatura y con otras colecciones los diferentes sinónimos que puede obtener una cepa específica.
- Capturar la información verificada para ser exportada a la base de datos final.
- Verificar que la nueva información proveniente de las colecciones externas (*Histoplasma*-UNAM, *Levaduras*-CIBNOR) se pudieran extraer los datos considerados para este proyecto.
- Incorporar los nuevos datos y asignar números de identificación para control interno.

5. Desarrollo del proyecto

En esta sección se presenta una descripción detallada sobre la manera en que se desarrolló todo el proceso de modelado y diseño para alcanzar el objetivo de generar la base de datos CDBB-500 ver. 2.0

Uno de los puntos más importantes en este proyecto es la culminación de la creación del modelo conceptual, el cual ha venido generándose desde hace tiempo. Este modelo es de gran relevancia debido a la dificultad de poder generar un nuevo diseño renovado dadas las condiciones y el contexto actual al que es necesario adecuarse.

La figura 5.1 nos muestra el diseño finalizado del modelo conceptual y que obtiene una abstracción de toda la información involucrada y referente a microorganismos. En este modelo podemos observar una agrupación por colores, los cuales nos indican una serie de dependencias las cuales adicionan una complejidad no trivial en la implementación final. La implementación real de todo este modelo involucra un seguimiento a largo plazo; estableciendo el objetivo principal del apoyo de Conabio, este proyecto fue enfocado en desarrollar la entidad general de descripción y desarrollar la parte referente a taxonomía y sinónimos de manera específica. En diferentes etapas adicionales fuera de este proyecto se trabajará para cubrir la información de las entidades restantes a mediano.

Una vez obtenido el modelo conceptual la siguiente etapa consistió en obtener un modelo relacional acorde con las necesidades ya ampliamente estudiadas. Para facilitar su desarrollo se conceptualizaron dos diferentes tipos de objetos: *Catálogos y tablas*.

De esta manera, los *catálogos* son tablas cuya función es obtener información especializada, detallada y sin redundancia, con el fin de eliminar la complejidad para la actualización e ingreso de información a la base de datos. Así se puede observar que los datos contenidos en ellas son únicos y con el fin de proporcionar una mayor interactividad en desarrollos de sistemas que administren y recuperen la información de manera controlada y explícita. En la figura 5.2 podemos observar el diseño final ahora convertido en la versión 2.1, ya que se estuvo mejorando este modelo relacional durante el desarrollo del proyecto. Aquí se pueden observar todas las relaciones involucradas y necesarias para describir parte de la semántica inmersa de algunas entidades del modelo conceptual.

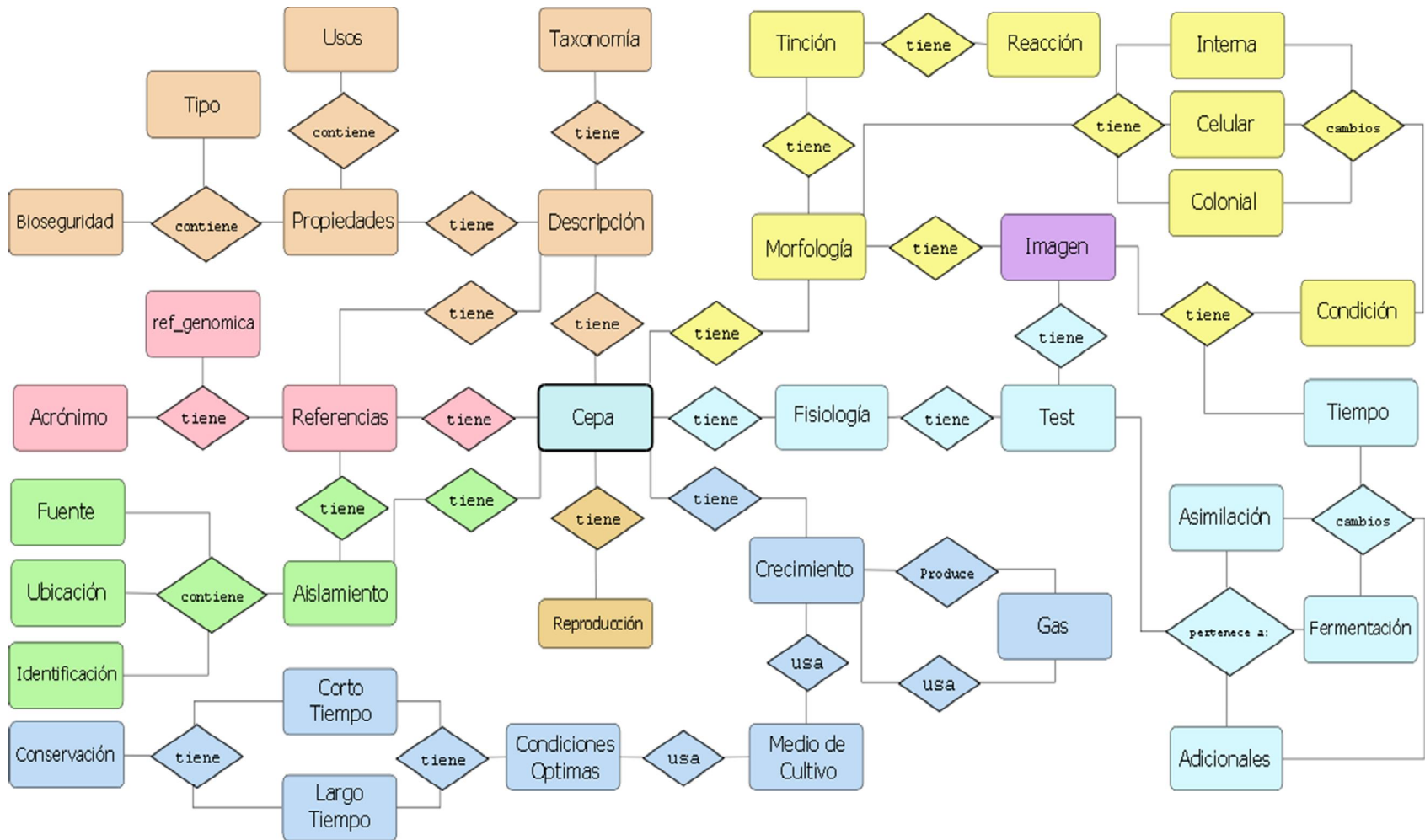


Fig. 5.1 Modelo conceptual de la base de datos CDBB-500 ver. 2.0, que contiene la semántica inmersa en información relacionada a microorganismos.

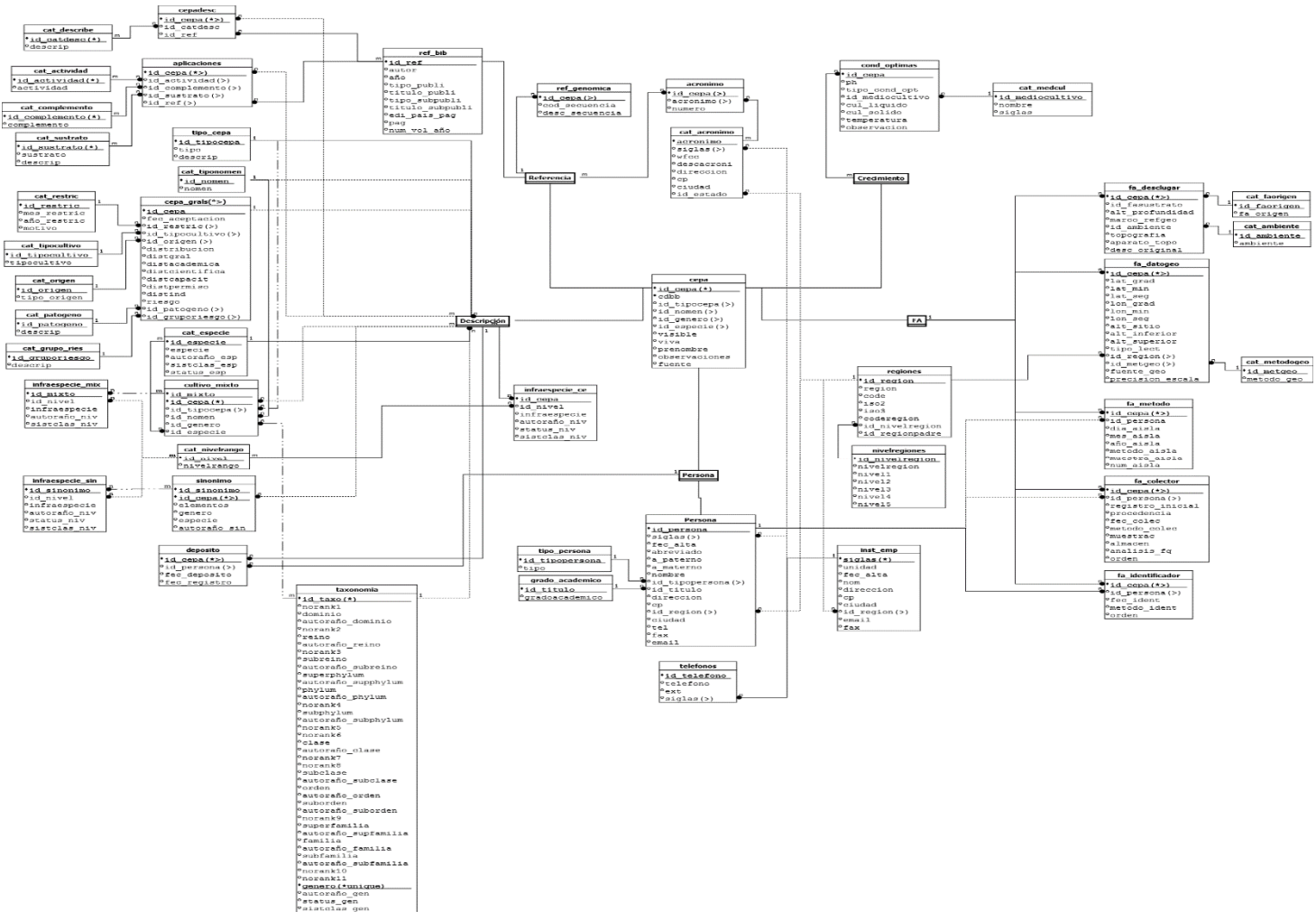


Fig. 5.2 Diseño Final del Modelo Relacional de la base de datos CDBB-500 ahora versión 2.1, dadas las mejoras obtenidas al final del proyecto.

5.1. Diccionario de la base de datos

Una vez presentado el diseño final del modelo relacional, a continuación se describe el diccionario de datos final obtenido, el cual debería contener campos compatibles y requeridos por el manual de compatibilidad de bases de datos. La definición del tamaño y tipo y algunos cambios realizados en el transcurso del proyecto fueron *verificados y validados con la gente de informática de la propia Conabio*. El diseño final del diccionario de datos, así como el script para su implementación es:

Base de datos CDBB-500 ver. 2.1

Script:

```
create database cdbb500_v2_1;
```

Catálogos

tipo_cepa

Contiene los distintos grupos de microorganismos que existen dentro de la colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_tipocepa(*)	int	-	Identificador del tipo de cepa.
tipo	char	2	Tipo de cepa (A, B, H, L, P) en español.
descrip	varchar	15	Descripción del tipo de cepa.

Script:

```
create table tipo_cepa(id_tipocepa serial not null primary key, tipo char(2) not null, descrip varchar(15) not null);
```

cat_tiponomen

Contiene información referente a los tipos nomenclaturales para una cepa

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_nomen(*>)	int	-	Identificador del tipo nomenclatural.
nomen	varchar	15	Descripción del tipo nomenclatural.

Script:

```
create table cat_tiponomen(id_nomen int not null primary key, nomen varchar(15) not null);
```

taxonomia

Almacena la jerarquía taxonómica de los microorganismos registrados en la colección desde las categorías superiores (reino) hasta las más específicas de cada género.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_taxo (*>)	int	-	Identificador de la taxonomía.
norank1	varchar	50	Nombre identificador para una categoría no oficial
dominio	varchar	50	Nombre del dominio antes llamado superreino.
autoraño_dominio	text	-	Nombre del autor o autores y año de la descripción original del nombre

			del domino antes llamado superreino
norank2	varchar	50	Nombre identificador para una categoría no oficial
reino	varchar	50	Nombre del reino.
autoraño_reino	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre del reino
norank3	varchar	50	Nombre identificador para una categoría no oficial
subreino	varchar	50	Nombre del subreino.
autoraño_subreino	text	-	Nombre del autor o autores y año de la descripción original del nombre del subreino
superphylum	varchar	50	Nombre del superphylum.
autoraño_supphylum	text	-	Nombre del autor o autores y año de la descripción original del nombre del superphylum
phylum	varchar	50	Nombre del phylum.
autoraño_phylum	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre del phylum.
norank4	varchar	50	Nombre identificador para una categoría no oficial
subphylum	varchar	50	Nombre del subphylum.
autoraño_subphylum	text	-	Nombre del autor o autores y año de la descripción original del nombre del subphylum
norank5	varchar	50	Nombre identificador para una categoría no oficial
norank6	varchar	50	Nombre identificador para una categoría no oficial
clase	varchar	50	Nombre de la clase.
autoraño_clase	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la clase
norank7	varchar	50	Nombre identificador para una categoría no oficial
norank8	varchar	50	Nombre identificador para una

			categoria no oficial
subclase	varchar	50	Nombre de la subclase.
autoraño_subclase	text	-	Nombre del autor o autores y año de la descripción original del nombre de la subclase
orden	varchar	50	Nombre del orden.
autoraño_orden	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre del orden
suborden	varchar	50	Nombre del suborden.
autoraño_suborden	text	-	Nombre del autor o autores y año de la descripción original del nombre del suborden
norank9	varchar	50	Nombre identificador para una categoría no oficial
superfamilia	varchar	50	Nombre de la súper familia.
autoraño_supfamilia	text	-	Nombre del autor o autores y año de la descripción original del nombre de la súper familia
familia	varchar	50	Nombre de la familia
autoraño_familia	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la familia
subfamilia	varchar	50	Nombre de la subfamilia.
autoraño_subfamilia	text	-	Nombre del autor o autores y año de la descripción original del nombre de la subfamilia
norank10	varchar	50	Nombre identificador para una categoría no oficial
norank11	varchar	50	Nombre identificador para una categoría no oficial
genero(*)	varchar	35	Nombre del genero
autoraño_gen	text	-	Nombre del autor o autores y año de la descripción original del nombre del genero
status_gen	int	-	Indica si el nombre del género es válido/correcto (2 = Válido/Correcto).
sistclas_gen	text	-	Autor(es) y año de publicación del sistema de calcificación, catálogo de

			autoridad, listado o diccionario del nombre del género.
--	--	--	---

Script:

create table taxonomia(id_taxo serial not null primary key, norank1 varchar(50), dominio varchar(50), autorano_dominio text, norank2 varchar(50), reino varchar(50), autorano_reino text, norank3 varchar(50), subreino varchar(50), autorano_subreino text, superphylum varchar(50), autorano_supphylum text, phylum varchar(50), autorano_phylum text, norank4 varchar(50), subphylum varchar(50), autorano_subphylum text, norank5 varchar(50), norank6 varchar(50), clase varchar(50), autorano_clase text, norank7 varchar(50), norank8 varchar(50), subclase varchar(50), autorano_subclase text, orden varchar(50), autorano_orden text, suborden varchar(50), autorano_suborden text, norank9 varchar(50), superfamilia varchar(50), autorano_supfamilia text, familia varchar(50), autorano_familia text, subfamilia varchar(50), autorano_subfamilia text, norank10 varchar(50), norank11 varchar(50), genero varchar(35) unique not null, autorano_gen text, status_gen int, sistclas_gen text);

cat_especie

Se anexaron los campos obligatorios requeridos por la Conabio donde se guarda los distintos nombres de especies.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_especie(*)	serial	-	Identificador de la especie.
especie	varchar	35	Nombre de la Especie
autorano_esp	text	-	Nombre del autor o autores y año de la descripción original del nombre de la especie.
status_esp	int	-	Indica si el nombre de la especie es válido/correcto (2 = Válido/Correcto).
sistclas_esp	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la especie.

Script:

create table cat_especie(id_especie serial not null primary key, especie varchar(35) not null, autorano_esp text, status_esp int, sistclas_esp text);

cat_nivelrango

Almacena los diferentes niveles taxonómicos que pueda tener un microorganismo por debajo de la especie.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_nivel(*)	int	-	Identificador del nivel de rango infrasub-específico.
nivelrango	varchar	50	Nivel del rango infrasub-específico.

Script:

create table cat_nivelrango(id_nivel serial not null primary key, nivelrango varchar(50) not null);

cat_restric

Integra información referente a la descripción de un grupo de riesgo que puede ser atacado por un grupo de microorganismo.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_restric(*)	int	-	Identificador único de la restricción en la estructura de datos
mes_restric	int	-	Mes de la fecha de restricción en que los datos de la cepa serán de uso público.
año_restric	int	-	Año de la fecha de restricción en que los datos de la cepa serán de uso público.
motivo	text	-	Argumentos por los cuales la información queda restringida.

Script:

```
create table cat_restric(id_restric serial not null primary key, mes_restric int, año_restric int, motivo text);
```

cat_tipocultivo

Se refiere a la información de las condiciones de pureza del microorganismo. Por ejemplo: Cultivo axénico (puro) ó Cultivo mixto (más de dos integrantes).

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_tipocultivo(*)	int	-	Identificador del tipo de cultivo.
tipocultivo	varchar	50	Tipo de cultivo.

Script:

```
create table cat_tipocultivo(id_tipocultivo serial not null primary key, tipocultivo varchar(50));
```

cat_origen

Cataloga la información sobre la procedencia del cultivo microbiano, ya sea por depósito o directamente de un proveedor.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_origen(*)	int	-	Identificador del origen.
tipo_origen	varchar	50	Tipo de origen de la cepa.

Script:

```
create table cat_origen(id_origen serial not null primary key, tipo_origen varchar(50));
```

cat_patogeno

Integra información referente a la descripción del nivel de patogenicidad que tiene un microorganismo.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_patogeno(*)	int	-	Identificador patógeno.
descrip	varchar	200	Descripción patógena.

Script:

```
create table cat_patogeno(id_patogeno serial not null primary key, descrip varchar(200));
```

cat_grupories

integra información referente a la descripción de un grupo de riesgo que puede ser atacado por un grupo de microorganismo.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_gruporiesgo(*)	int	-	Identificador del grupo de riesgo.
descrip	varchar	200	Descripción del grupo de riesgo.

Script:

create table cat_grupories(id_gruporiesgo serial not null primary key, descrip varchar(200));

cat_describe

Almacena la información referente a las descripciones que pueden tener los microorganismos registrados en la colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_catdesc(*)	int	-	Identificador del catálogo de descripción.
descrip	text	-	Descripción de la cepa.

Script:

create table cat_describe(id_catdesc serial not null primary key, descrip text not null);

cat_actividad

Contiene la información sobre las actividades de un microorganismo sobre un sustrato.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_actividad(*)	int	-	Identificador de la actividad.
actividad	varchar	100	Actividad de un microorganismo sobre un sustrato.

Script:

create table cat_actividad(id_actividad serial not null primary key, actividad varchar(100) not null);

cat_complemento

Catálogo de complementos por ejemplo (de, en, un, para, por), usados para ligar la actividad de un microorganismo sobre un sustrato en un reporte

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_complemento(*)	int	-	Identificador del complemento.
complemento	varchar	10	Complemento.

Script:

create table cat_complemento(id_complemento serial not null primary key, complemento varchar(10) not null);

cat_sustrato

Contiene la información de los sustratos que pueden ser utilizados por un microorganismo.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_sustrato(*)	int	-	Identificador del sustrato.
sustrato	varchar	150	Sustrato de una cepa.
descrip	text	-	Descripción detallado del sustrato.

Script:

```
create table cat_sustrato(id_sustrato serial not null primary key, sustrato varchar(150) not null,
descrip text);
```

regiones

Catálogo de las diferentes regiones del mundo.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_region (*)	serial	-	Identificador de la región.
region	varchar	150	Nombre de la región.
code	int	-	Código estándar numérico.
iso2	varchar	2	Código estándar para abreviación de países en dos siglas.
iso3	varchar	3	Código estándar para abreviación de países en tres siglas.
coderegion	varchar	7	Código regional.
id_nivelregion (>)	int	-	Identificador del nivel de la región.
id_regionpadre (>)	int	-	Identificador del padre de la región asignada.

Script:

```
Create table regiones(id_region int primary key not null, region varchar (150) not null, code int,
iso2 varchar(2), iso3 varchar(3), coderegion varchar (7), id_nivelregion int references
nivelregion(id_nivelregion), id_regionpadre int references regiones(id_region));
```

nivelregiones

Catálogo de los diferentes estados que contiene un país.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_nivelregion (*)	int	-	Identificador del nivel de la región.
nivelregion	varchar	50	Nombre del nivel de la región
nivel1	int	-	Indica el nivel jerárquico entre los elementos.
nivel2	int	-	Indica el nivel jerárquico entre los elementos.
nivel3	int	-	Indica el nivel jerárquico entre los elementos.

nivel4	int	-	Indica el nivel jerárquico entre los elementos.
nivel5	int	-	Indica el nivel jerárquico entre los elementos.

Script:

create table nivelregiones(id_nivelregion int primary key not null, nivelregion varchar(50), nivel1 int, nivel2 int, nivel3 int, nivel4 int, nivel5 int);

inst_emp

Se almacenan los datos generales que se requieren para una institución o empresa.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
siglas(*)	varchar	15	Siglas de la empresa o institución.
unidad	varchar	100	Unidad.
fec_alta	date	-	Fecha de alta de la empresa o institución.
nom	varchar	100	Nombre de la empresa o institución.
direccion	text	-	Dirección: calle, número y colonia
cp	varchar	40	Código postal.
ciudad	varchar	150	Ciudad.
id_region(>)	int	-	Identificador del estado.
email	varchar	50	Email.
fax	varchar	20	Fax.

Script:

create table inst_emp(siglas varchar(15) primary key, unidad varchar(100), fec_alta date, nom varchar(100), direccion varchar(400), cp varchar(40), ciudad varchar(150), id_region int references estado(id_region), email varchar(50), fax varchar(20));

tipo_persona

Cataloga los diferentes tipos de persona según su actividad profesional.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_tipopersona(*)	int	-	Identificador del tipo de persona que es.
tipo	varchar	150	Descripción del tipo de la persona.

Script:

create table tipo_persona(id_tipopersona serial not null primary key, tipo varchar(150) not null);

grado_academico

Catálogo de grados académicos.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_titulo(*)	int	-	Identificador del grado académico.
gradoacademico	varchar	70	Descripción del grado académico.

Script:

create table grado_academico(id_titulo serial not null primary key, gradoacademico varchar(70) not null);

persona

Registra todos los datos personales que se requieren de una persona para poder hacer un depósito de un microorganismo dentro de la Colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_persona(*)	int	-	Identificador de la persona.
siglas(>)	varchar	15	Siglas de la empresa o institución.
fec_alta	date	-	Fecha de alta de la persona.
abreviado	varchar	180	Nombre(s) y apellidos en formato abreviado de la persona.
a_paterno	varchar	50	Apellido paterno.
a_materno	varchar	50	Apellido materno.
nombre	varchar	50	Nombre(s).
id_tipopersona(>)	int	-	Identificador del tipo de persona.
id_titulo(>)	int	-	Identificador del grado académico.
dirección	text	-	Dirección: calle, número y colonia
cp	varchar	40	Código postal.
id_region(>)	int	-	Identificador del estado.
ciudad	varchar	150	Ciudad.
rel	varchar	30	Teléfono.
fax	varchar	40	Fax.
email	varchar	50	Email.

Script:

create table persona(id_persona serial not null primary key, siglas varchar(15) references inst_emp(siglas), fec_alta date, abreviado varchar(180), a_paterno varchar(50), a_materno varchar(50), nombre varchar(50), id_tipopersona int references tipo_persona(id_tipopersona), id_titulo int references grado_academico(id_titulo), direccion varchar(400), cp varchar(40), id_region int references estado(id_region), ciudad varchar(150), tel varchar(30), fax varchar(20), email varchar(50));

cat_faorigen

Se guarda información referente al origen de una cepa aislada.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_faorigen(*)	int	-	Identificador del origen de una cepa aislada.
fa_origen	varchar	150	Descripción del origen de una cepa aislada.

Script:

create table cat_faorigen(id_faorigen int primary key, fa_origen varchar(150));

cat_ambiente

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_ambiente(*)	int	-	Identificador del medio donde el microorganismo fue colectado.
ambiente	varchar	150	Indica el medio donde el microorganismo fue colectado (1=Dulceacuícola, 2=Marino, 3=Terrestre, 4=Salobre, 5=Costero, 0=No Disponible).

Script:

create table cat_ambiente(id_ambiente int primary key, ambiente varchar(150));

cat_metodogeo

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_metgeo(*)	int	-	Identificador del método geográfico.
metodo_geo	varchar	150	Método de georreferencia. (1=Geoposicionador, 2=Mapa, 3=Gacetero, 4=Literatura, 5=Etiqueta, 6=No Disponible).

Script:

create table cat_metodogeo(id_metgeo int primary key, metodo_geo varchar(150) not null);

cat_acronimo

Contiene las siglas de otras colecciones de microorganismos, para su identificación internacional.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
acronimo	varchar	15	Siglas de las diferentes colecciones.
siglas(>)	varchar	15	Siglas de la empresa o institución.
wfcc	varchar	10	Registro otorgado para identificar a un miembro de la WFCC.
descacroni	varchar	200	Descripción del acrónimo.
dirección	text	-	Dirección: calle, número y colonia
cp	varchar	40	Código postal.
ciudad	varchar	150	Ciudad.
id_region(>)	int	-	Identificador del estado.

Script:

create table cat_acronimo(acronimo varchar(15) primary key not null, siglas varchar(10), wfcc varchar(10), descacroni varchar(200), direccion varchar(400), cp varchar(40), ciudad varchar(150), id_region int references regiones(id_region));

cat_medcul

Contiene información de los medios de cultivos y formulación para el desarrollo de los microorganismos.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_mediocultivo(*)	int	-	Identificador del medio de cultivo.
nombre	varchar	100	Nombre del medio de cultivo.
siglas	varchar	30	Siglas del medio de cultivo.

Script:

create table cat_medcul(id_mediocultivo serial not null primary key, nombre varchar(100), siglas varchar(30));

ref_bib

Contiene las referencias bibliográficas de la descripción de la cepa y de las aplicaciones específicas.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_ref(*)	int	-	Identificador de la referencia bibliográfica.
autor	varchar	255	Autor(es) de la publicación o subpublicación.
año	varchar	50	Indica el(los) año(s) o la fecha en que fue hecha la publicación o subpublicación.
tipo_publi	varchar	50	Tipo de la publicación. Por ejemplo: Libro, Tesis, Revista.
titulo_publi	text	-	Título de la publicación.
tipo_subpubli	text	-	Tipo de subpublicación. Por ejemplo: Capítulo, Artículo, Resumen.
titulo_subpubli	varchar	255	Título de la subpublicación.
edi_pais_pag	text	-	Entidad que llevó a cabo la edición de la publicación, Ciudad o país (sólo si el nombre de la ciudad no esta citado) donde se publico y número de paginas.
pag	varchar	20	Pagina(s) Consultadas.
num_vol_año	varchar	100	Indica el número de la publicación, el número del volumen de la publicación o páginas de la publicación o subpublicación.

Script:

create table ref_bib(id_ref serial not null primary key, autor varchar(255), año varchar(50), tipo_publi varchar(50), titulo_publi varchar(310), tipo_subpubli varchar(50), titulo_subpubli varchar(310), edi_pais_pag varchar(310), pag varchar(20), num_vol_año varchar(100));

Tablas

cepa

Muestra los datos específicos relacionados a cada estirpe microbiana, relacionando características de diferentes estudios, descripciones y fuente de aislamiento que pueda tener cada cepa depositada en la colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa*	int	-	Identificador de la cepa depositada.
cddb	int	-	Clave única de la cepa depositada.
id_tipocepa(>)	int	-	Identificador del tipo de cepa (A, B, H, L, P).
id_nomen	int	-	Identificador del tipo nomenclatural
id_genero	int	-	Identificador del catálogo "cat_taxonomia" como alias al id_taxo.
id_especie	int	-	Identificador de la cat_especie.
visible	boolean	5	Cepa visible a usuarios externos.
viva	boolean	5	Cepa viva.
prenombre	varchar	40	Nombre del aislamiento antes de ser identificado.
observaciones	text	-	Observaciones.
fuelle	varchar	5	Clave de referencia que identifica el nombre del proyecto ante la Conabio.

Script:

```
create table cepa(id_cepa serial not null primary key, cddb int, id_tipocepa int references
tipo_cepa(id_tipocepa), id_nomen int references cat_tiponomen(id_nomen), id_genero int
references taxonomia(id_taxo), id_especie int references cat_especie(id_especie), visible boolean,
viva boolean, prenombre varchar(40), observaciones text, fueulle varchar(4));
```

infraespecie_ce

Relaciona un microorganismo que tenga un nivel de rango, almacenando el nombre de la infraespecie con su sistema de clasificación y estatus de la infraespecie.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(>)	int	-	Identificador de la cepa.
id_nivel(>)	int	-	Identificador del nivel de rango infrasubespecífico.
infraespecie	varchar	50	Nombre del rango infrasubespecífico.
autorañor_niv	text	-	Nombre del autor o autores y año de la descripción original del nombre de la especie.
status_niv	int	-	Indica si el epíteto infrasubespecífico es válido/correcto. (2=Valido/Correcto)

sistclas_niv	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la infrasubspecifico.
---------------------	------	---	---

Script:

```
create table infraespecie_ce(id_cepa int references cepa(id_cepa), id_nivel int references cat_nivelrango(id_nivel), infraespecie varchar(50), autoraño_niv text, status_niv int, sistclas_niv text);
```

cepa_grals

Contiene la información general, específica y de control de un microorganismo depositado en la colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
fec_aceptacion	date	-	Fecha de aceptación del microorganismo.
id_restric(>)	int	-	Identificador único de la restricción en la estructura de datos.
id_tipocultivo(>)	int	-	Identificador del tipo de cultivo.
id_origen(>)	int	-	Identificador del origen.
distribucion	boolean	5	Distribución.
distgral	boolean	5	Distribución en general.
distacademica	boolean	5	Distribución a instituciones académicas.
distcientifica	boolean	5	Distribución a científicos.
distcapacit	boolean	5	Distribución a personal capacitado.
distpermiso	boolean	5	Distribución con permiso del propietario.
distind	boolean	5	Distribución a empresas.
riesgo	boolean	5	Es peligroso para humanos, animales, plantas o ambiente.
id_patogeno(>)	int	-	Identificador patógeno.
id_gruporiesgo(>)	int	-	Identificador del grupo de riesgo.

Script:

```
create table cepa_grals(id_cepa int references cepa(id_cepa), fec_aceptacion date, id_restric int references cat_restric(id_restric), id_tipocultivo int references cat_tipocultivo(id_tipocultivo), id_origen int references cat_origen(id_origen), distribucion boolean, distgral boolean, distacademica boolean, distcientifica boolean, distcapacit boolean, distpermiso boolean, distind boolean, riesgo boolean, id_patogeno int references cat_patogeno(id_patogeno), id_gruporiesgo int references cat_grupories(id_gruporiesgo));
```

sinonimo

Contiene la información de la sinonimia que llegara a presentar una especie.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_sinonimo(*)	int	-	Identificador del sinónimo.
id_cepa(>)	int	-	Identificador de la cepa.
elementos	int	-	Elementos.
genero	varchar	100	Nombre del Género
especie	varchar	100	Nombre de la Especie.
autoraño_sin	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la especie de un sinónimo.

Script:

```
create table sinonimo(id_cepa int references cepa(id_cepa), elementos int, id_genero int
references taxonomia(id_taxo), id_especie int references cat_especie(id_especie), sistclas_sin
text);
```

cultivo_mixto

Muestra los datos específicos relacionados a un cultivo mixto, relacionando características de diferentes estudios, descripciones y fuente de aislamiento que puede tener un cultivo mixto depositado en la colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_mixto(*)	int	-	Identificador de la cepa depositada.
id_cepa(>)	int	-	Identificador de la cepa depositada.
id_tipocepa(>)	int	-	Identificador del tipo de cepa (A, B, H, L, P).
id_nomen	int	-	Identificador del tipo nomenclatural
id_genero	int	-	Identificador de la cat_taxonomia como apodo al id_taxo.
id_especie	int	-	Identificador de la cat_especie.

Script:

```
create table cultivo_mixto(id_cepa int references cepa(id_cepa), id_tipocepa int references
tipo_cepa(id_tipocepa), id_nomen int references cat_tiponomen(id_nomen), id_genero int
references taxonomia(id_taxo), id_especie int references cat_especie(id_especie));
```

infraespecie_sin

Relaciona un sinonimo de un microorganismo que tenga un nivel de rango, almacenando el nombre de la infraespecie con su sistema de clasificacion y estatus de la infraespecie.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_sinonimo(>)	int	-	Identificador del sinónimo de una cepa.
id_nivel(>)	int	-	Identificador del nivel de rango infrasubespecífico.
infraespecie	varchar	50	Nombre del rango infrasubespecífico.
autorañor_niv	text	-	Nombre del autor o autores y año de la descripción original del nombre de la especie.
status_niv	int	-	Indica si el epíteto infrasubespecifico es válido/correcto. (2=Valido/Correcto)
sistclas_niv	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la infrasubespecifico.

Script:

```
create table infraespecie_sin(id_sinonimo int references sinonimo(id_sinonimo), id_nivel int references cat_nivelrango(id_nivel), infraespecie varchar(50), autoraño_niv text, status_niv int, sistclas_niv text);
```

infraespecie_mix

Relaciona un cultivo mixto que tenga un nivel de rango, almacenando el nombre de la infraespecie con su sistema de clasificacion y estatus de la infraespecie.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_mixto(>)	int	-	Identificador del cultivo mixto.
id_nivel(>)	int	-	Identificador del nivel de rango infrasubespecífico.
infraespecie	varchar	50	Nombre del rango infrasubespecífico.
autorañor_niv	text	-	Nombre del autor o autores y año de la descripción original del nombre de la especie.
status_niv	int	-	Indica si el epíteto infrasubespecifico es válido/correcto. (2=Valido/Correcto)
sistclas_niv	text	-	Autor(es) y año de publicación del sistema de clasificación, catálogo de autoridad, listado o diccionario del nombre de la infrasubespecifico.

Script:

create table infraespecie_sin(id_sinonimo int references sinonimo(id_sinonimo), id_nivel int references cat_nivelrango(id_nivel), infraespecie varchar(50), autoraño_niv text, status_niv int, sistclas_niv text);

deposito

Registra la cepa, la persona que realizo el depósito, la fecha del depósito y la fecha del registro en el sistema.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
id_persona(>)	int	-	Identificador de la persona depositante.
fec_deposito	date	-	Fecha del deposito.
fec_registro	date	-	Fecha del registro en el sistema.

Script:

create table deposito(id_cepa int references cepa(id_cepa), id_persona int references persona(id_persona), fec_deposito date, fec_registro date);

fa_datogeo

Almacena información referente a los datos geográficos de la fuente de aislamiento donde se obtuvo la muestra para el estudio del aislamiento.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
lat_grad	int	-	Latitud en grados (dato geográfico).
lat_min	int	-	Latitud en minutos (dato geográfico).
lat_seg	int	-	Latitud en segundos (dato geográfico).
lon_grad	int	-	Longitud en grados (dato geográfico).
lon_min	int	-	Longitud en minutos (dato geográfico).
lon_seg	int	-	Longitud en segundos (dato geográfico).
alt_sitio	int	-	Altitud del sitio o profundidad del sitio medida respecto a un marco de referencia geográfico.
alt_inferior	int	-	Límite inferior de la altitud o profundidad del sitio de colecta de un microorganismo.
alt_superior	int	-	Límite superior de la altitud o profundidad del sitio de colecta de un microorganismo.
tipo_lect	varchar	30	Tipo de lector.
id_region(>)	int	-	Identificador del municipio.
id_metgeo(>)	int	-	Identificador del método geográfico.
fuentes_geo	varchar	255	Referencia sobre la cita de la

			información geográfica, según el método de georreferencia.
precision_escala	varchar	100	Indica la precisión del geoposicionador o la escala del mapa en el cual se realizó la georreferencia.

Script:

```
create table fa_datogeo(id_cepa int references cepa(id_cepa), lat_grad int, lat_min int, lat_seg int, lon_grad int, lon_min int, lon_seg int, alt_sitio int, alt_inferior int, alt_superior int, tipo_lect varchar(30), id_region int references regiones(id_region), id_metgeo int references cat_metodogeo(id_metgeo), fuente_geo varchar(255), precision_escala varchar(100));
```

fa_desclugar

Guarda información referente a la descripción del lugar de donde se obtuvo la muestra para el estudio del aislamiento.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
id_fasustrato(>)	int	-	Identificador del origen de la cepa
alt_profundidad	int	-	Altitud o profundidad donde se colecto el microorganismo en metros.
marco_refgeo	boolean	5	Indica si el dato de la profundidad fue tomado respecto al nivel del mar o respecto a la altitud del sitio.
id_ambiente	int	-	Identificador del medio donde el microorganismo fue colectado.
topografia	text	-	Topografía del lugar.
aparato_topo	text	-	Aparato Topográfico.
desc_original	text	-	Descripción original de la ubicación del lugar de la colección.

Script:

```
create table fa_desclugar(id_cepa int references cepa(id_cepa), id_fasustrato int references cat_faorigen(id_faorigen), alt_profundidad int, marco_refgeo boolean, id_ambiente int references cat_ambiente(id_ambiente), topografia text, aparato_topo text, desc_original text);
```

fa_colector

Registra los datos del colector, métodos y fecha de colecta de la muestra para el estudio del aislamiento.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
id_persona(>)	int	-	Identificador de la persona que colecto la cepa.
registro_inicial	varchar	40	Identificador único asignado por el recolector en cada evento de

			recolección.
procedencia	int	-	Indica si el ejemplar proviene de un evento recolectado(1=Colectado)
fac_colecta	date	-	Fecha de recolecta de un microorganismo.
metodo_colect	text	-	Método de colecta.
muestrac	int	-	Número de muestras colectadas.
almacen	varchar	50	Forma de almacenamiento.
analisis_fq	varchar	120	Análisis físico-químico.
orden	int	-	Orden de importancia que tiene la persona en el evento de colecta de un microorganismo.

Script:

```
create table fa_colector(id_cepa int references cepa(id_cepa), id_persona int references
persona(id_persona), registro_inicial varchar(150), procedencia int, fec_colecta date,
metodo_colect text, muestrac int, almacen varchar(50), analisis_fq varchar(120), orden int);
```

fa_metodo

Contiene información sobre el registro del aislamiento de un cultivo microbiano.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
id_persona(>)	int	-	Identificador de la persona que aislo el cultivo microbiano.
dia_aisla	int	-	Día de aislamiento.
mes_aisla	int	-	Mes de aislamiento.
año_aisla	int	-	Año de aislamiento.
metodo_aisla	text	-	Método en como se aislo el cultivo microbiano.
muestra_aisla	int	-	Numero de muestras aisladas
num_aisla	varchar	100	Identificador otorgado al aislamiento

Script:

```
create table fa_metodo(id_cepa int references cepa(id_cepa), id_persona int references
persona(id_persona), dia_aisla int, mes_aisla int, año_aisla int, metodo_aisla text, muestra_aisla
int, num_aisla varchar(100));
```

fa_identificador

Almacena los datos del identificador, fecha y método de identificación de un cultivo microbiano.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
id_persona(>)	int	-	Identificador de la persona.
fec_ident	date	-	Fecha de la identificación de una cepa.
metodo_ident	text	-	Método de identificación del cultivo microbiano.
orden	int	-	Orden de importancia que tiene la persona en el proceso de identificación de un microorganismo.

Script:

```
create table fa_identificador(id_cepa int references cepa(id_cepa), id_persona int references
persona(id_persona), fec_ident date, metodo_ident text, orden int);
```

aplicaciones

Liga el registro del microorganismo de la tabla cepa con el registro de la tabla cat_actividad, cat_complemento, cat_sustrato y el registro de la referencia bibliográfica de la tabla ref_bib.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(>)	int	-	Identificador de la cepa.
id_actividad(>)	int	-	Identificador de la actividad.
id_complemento(>)	int	-	Identificador del complemento.
id_sustrato(>)	int	-	Identificador del sustrato.
id_ref(>)	int	-	Identificador de la referencia bibliográfica.

Script:

```
create table aplicaciones(id_cepa int references cepa(id_cepa), id_actividad int references
cat_actividad(id_actividad), id_complemento int references complemento(id_complemento),
id_sustrato int references sustrato(id_sustrato), id_ref int references ref_bib(id_ref));
```

cepadesc

Liga el registro de un microorganismo de la tabla cepa, con el registro de la descripción de la tabla cat_describe y el registro de la referencia bibliográfica de la tabla ref_bib.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
id_catdesc(>)	int	-	Identificador del catálogo de descripciones.
id_ref(>)	int	-	Identificador de la referencia bibliográfica.
pag	varchar	10	Página(s) Consultadas.
num_vol_año	varchar	100	Indica el número de la publicación, el número del volumen de la

			publicación o páginas de la publicación o subpublicación.
--	--	--	---

Script:

```
create table cepadesc(id_cepa int references cepa(id_cepa), id_catdesc int references cat_describe(id_catdesc), id_ref int references ref_bib(id_ref));
```

acronimo

Liga el registro de un microorganismo en la tabla *cepa*, las siglas incorporadas de la tabla *cat_acronimo* y su número de registro en dicha colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(>)	int	-	Identificador de la cepa.
acronimo(>)	varchar	12	Identificador del catálogo de acrónimos.
numero	varchar	50	Número de registro de las colecciones.

Script:

```
create table acronimo(id_cepa int references cepa(id_cepa), acronimo varchar(12) references cat_acronimo(acronimo), numero varchar(50));
```

cond_optimas

Se registran las condiciones óptimas de crecimiento de un cultivo microbiano de la colección.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_cepa(*>)	int	-	Identificador de la cepa.
ph	varchar	10	pH del medio.
tipo_cond_opt	varchar	60	Condición óptima con respecto a intensidad, luminosa y aireación.
id_mediocultivo(>)	int	-	Identificador del medio de cultivo.
cul_liquido	boolean	5	Cultivo líquido.
cul_solido	boolean	5	Cultivo sólido.
temperatura	int	-	Temperatura óptima en grados centígrados.
observaciones	text	-	Observaciones.

Script:

```
create table cond_optimas(id_cepa int references cepa(id_cepa), ph varchar(10), tipo_cond_opt varchar(60), id_mediocultivo int references cat_medcul(id_mediocultivo), cul_liquido boolean, cul_solido boolean, temperatura int, observaciones text);
```

ref_genomica

Referencia a base de datos genómicas de nucleótidos en National Center For Biotechnology Information (NCBI).

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
Id_cepa(>)	int	-	Identificador de la cepa.
cod_secuencia	varchar	20	Código de la secuencia genómica.
desc_secuencia	text	-	Descripción de la secuencia genómica.

Script:

```
create table ref_genomica(id_cepa int references cepa(id_cepa), cod_secuencia varchar(20), desc_secuencia text);
```

telefonos

Se registran uno o más teléfonos que pueden tener una institución o empresa.

Nombre	Tipo de datos	Longitud	Definición o contenido del campo
id_telefono(*)	int	-	Identificador del teléfono.
telefono	varchar	20	Teléfono de la empresa o institución.
ext	varchar	10	Extensiones.
siglas(>)	varchar	15	Identificador de la institución o empresa.

Script:

```
create table telefonos(id_telefono serial primary key, telefono varchar(20), ext varchar(5), siglas varchar(15) references inst_emp(siglas));
```

6. Mapeo de datos.

Una vez obtenido el diccionario de datos la siguiente etapa consistió en establecer el mapeo de los datos para establecer la compatibilidad entre la base de datos de la Conabio y el nuevo diseño de CDBB-500 v 2.0. *Este mapeo de datos fue validado y aceptado por las personas responsables de la propia Conabio, como se muestra en la siguiente tabla:*

Mapeo de Datos

CONABIO		CDBB		
Tabla	Campo	Tabla	Campo	Tipo de datos
Nombre				
	id_nombre	<i>cepa</i>	id_cepa	serial
	Reino	<i>taxonomia</i>	reino	varchar
	SistClasCatDiccReino	<i>taxonomia</i>	autoraño_reino	text
	divisionphylum	<i>taxonomia</i>	phylum	varchar
	SistClasCatDiccDivisionPhylum	<i>taxonomia</i>	autoraño_phylum	text
	clase	<i>taxonomia</i>	clase	varchar
	SistClasCatDiccClase	<i>taxonomia</i>	autoraño_clase	text
	orden	<i>taxonomia</i>	orden	varchar
	SistClasCatDiccOrden	<i>taxonomia</i>	autoraño_orden	text
	familia	<i>taxonomia</i>	familia	varchar
	SistClasCatDiccFamilia	<i>taxonomia</i>	autoraño_familia	text
	genero	<i>taxonomia</i>	genero	varchar
	AutorAnioGenero	<i>taxonomia</i>	autoraño_gen	text
	EstatusGenero	<i>taxonomia</i>	status_gen	int
	SistClasCatDiccGenero	<i>taxonomia</i>	sistclas_gen	text
	especie	<i>cat_especie</i>	especie	varchar
	AutorAnioEspecie	<i>cat_especie</i>	autoraño_esp	text
	EstatusEspecie	<i>cat_especie</i>	status_esp	int
	SistClasCatDiccEspecie	<i>cat_especie</i>	sistclas_esp	text
	categInf	<i>cat_nivelrango</i>	nivelrango	varchar
	infraespecie	<i>cepa</i>	nom_rango	varchar
	AutorAnioInfraespecie	<i>cat_nivelrango</i>	autoraño_niv	text
	EstatusInfraespecie	<i>cat_nivelrango</i>	status_niv	int
	SistClasCatDiccInfraespecie	<i>cat_nivelrango</i>	sistclas_niv	text
GEOGRA				
	NombreOriginal	<i>fa_desclugar</i>	des_original	text
	LatitudGrados	<i>fa_datogeo</i>	lat_grad	int
	LatitudMinutos	<i>fa_datogeo</i>	lat_min	int
	LatitudSegundos	<i>fa_datogeo</i>	lat_seg	int
	LongitudGrados	<i>fa_datogeo</i>	lon_grad	int
	LongitudMinutos	<i>fa_datogeo</i>	lon_min	int
	LongitudSegundos	<i>fa_datogeo</i>	lon_seg	int
	AltitudOProfundidadSitio	<i>fa_datogeo</i>	alt_sitio	int

	AltitudOProfundidadInferior	<i>fa_datogeo</i>	alt_inferior	int
	AltitudOProfundidadSuperior	<i>fa_datogeo</i>	alt_superior	int
	MetodoGeo	<i>fa_datogeo</i>	id_metgeo	int
	FuenteGeorr	<i>fa_datogeo</i>	fuentes_geo	varchar
	PrecisionOEscala	<i>fa_datogeo</i>	precision_escala	varchar
EJEMPLAR				
	Ejemplar	id_ejemplar	<i>cepa</i>	id_cepas
	Procedencia	<i>fa_colector</i>	procedencia	int
	DiaInicial	<i>fa_colector</i>	fec_colec	date
	MesInicial	<i>fa_colector</i>	fec_colec	date
	AnioInicial	<i>fa_colector</i>	fec_colec	date
	NumeroColectaObs	<i>fa_colector</i>	registro_inicial	varchar
	NumeroCatalogo	<i>cepa</i>	cdbb	int
	AltitudProfundidad	<i>fa_desclugar</i>	alt_profundidad	int
	MarcoRefGeog	<i>fa_desclugar</i>	marco_refgeo	boolean
	IndividuosCopias	<i>fa_colector</i>	muestrac	int
	Ambiente	<i>fa_desclugar</i>	id_ambiente	int
	Tipo	<i>cepa</i>	id_nomen	int
	DiaDeterminacion	<i>fa_identificador</i>	fec_ident	date
	MesDeterminacion	<i>fa_identificador</i>	fec_ident	date
	AñoDeterminacion	<i>fa_identificador</i>	fec_ident	date
	Fuente	<i>cepa</i>	fuentes	varchar
RelEjemDeterminador				
	Orden	<i>fa_identificador</i>	orden	int
RelEjemColectorObservador				
	Orden	<i>fa_colector</i>	orden	int
PersonaGrupo				
	idPersona	<i>persona</i>	id_persona	serial
	Abreviado	<i>persona</i>	abreviado	varchar
	ApellidoPaterno	<i>persona</i>	a_paterno	varchar
	ApellidoMaterno	<i>persona</i>	a_materno	varchar
	Nombre	<i>persona</i>	nombre	varchar
RESTRIC				
	idRestriccion	<i>cat_restric</i>	id_restric	serial
	MesRestriccion	<i>cat_restric</i>	mes_restric	int
	AnioRestriccion	<i>cat_restric</i>	año_restric	int
	Motivos	<i>cat_restric</i>	motivo	text
COLECCION				
	SiglasColeccion	<i>cat_acronimo</i>	acronimo	Varchar
	Coleccion	<i>cat_acronimo</i>	descacronimo	Varchar
	Direccion	<i>cat_acronimo</i>	direccion	Varchar
	CodigoPostal	<i>cat_acronimo</i>	cp	Varchar
	Ciudad	<i>cat_acronimo</i>	ciudad	Varchar
	Estado	<i>estado</i>	estado	Varchar

	Pais	<i>pais</i>	pais	Varchar
INSTIT				
	NombreInstitucion	<i>inst_emp</i>	nom	varchar
	SiglasInstitucion	<i>inst_emp</i>	siglas	varchar
BIBLIO				
	IdBibliografia	<i>ref_bib</i>	id_ref	serial
	autor	<i>ref_bib</i>	autor	varchar
	Anio	<i>ref_bib</i>	año	varchar
	TipoPublicacion	<i>ref_bib</i>	tipo_publici	varchar
	TituloPublicacion	<i>ref_bib</i>	titulo_publici	varchar
	TipoSubPublicacion	<i>ref_bib</i>	tipo_subpubli	varchar
	TituloSubPublicacion	<i>ref_bib</i>	titulo_subpubli	varchar

7. Informe de registros en tablas

En esta sección se ofrece una descripción detallada del número de registros que contiene cada tabla en la finalización del presente proyecto, además de presentar de manera general algunas modificaciones y cambios realizados durante la incorporación de la información, manteniendo la compatibilidad con el instructivo de bases de datos de la Conabio.

7.1 Catálogos

Nombre Catálogo	Registros	Observaciones
tipo_cepa	6	El campo tipo se modificó el tamaño del tipo de campo a 2. Se verificaron los datos que contenía este catálogo.
cat_tiponomen	10	Se anexaron los tipos de nomenclatura de acuerdo a las necesidades de la colección.
taxonomia	228	Se modificó el tamaño del tipo de campo a 50, a excepción del campo id_taxo, genero y status_gen.
cat_especie	486	El campo especie se modificó el tamaño del tipo de campo a 35.
cat_nivelrango	12	Se quitaron los campos autoraño_niv, status_niv y sistclas_niv ya que en su momento no se tenía contemplado que deberían estar con el nombre del rango y no con el nivel del rango.
cat_restric	2	Se anexo el primer registro al catálogo de restricción de acuerdo a las necesidades de la colección.
cat_tipocultivo	4	Se verificaron los datos que contenía este catálogo.
cat_origen	3	Se verificaron los datos que contenía este catálogo.
cat_patogeno	5	Se verificaron los datos que contenía este catálogo.
cat_grupories	6	Se verificaron y se anexaron datos a este catálogo.
cat_describe	210	Se verificaron y anexaron nuevos datos a este catálogo, eliminando redundancias y ambigüedad.

cat_actividad	0	
cat_complemento	0	
cat_sustrato	0	
pais	242	Se anexaron nuevos datos, cumpliendo los estándares establecidos por la ISO 3166-2.
estado	4533	Se incorporaron nuevos datos, cumpliendo los estándares establecidos por el INEGI.
municipio	6988	Se incorporaron nuevos datos, cumpliendo los estándares establecidos por el INEGI.
inst_emp	121	Se verificaron, modificaron y anexaron nuevos datos a este catálogo, eliminando redundancias y ambigüedad.
tipo_persona	4	Se verificaron los datos que contenía este catálogo.
grado_academico	13	Se anexo nueva información a este catálogo.
persona	117	Se verificaron, modificaron y anexaron nuevos datos a este catálogo, eliminando redundancias y ambigüedad.
cat_faorigen	149	Se anexo nueva información a este catálogo.
cat_ambiente	6	Se anexo nueva información a este catálogo.
cat_metodogeo	6	Se anexo nueva información a este catálogo.
cat_acronimo	619	Se verificaron, modificaron y anexaron nuevos datos a este catálogo, eliminando redundancias y ambigüedad.
cat_medcul	141	Se verificaron los datos que contenía este catálogo.
ref_bib	1134	Se verificaron, modificaron y anexaron nuevos datos a este catálogo, eliminando redundancias y ambigüedad.

Total catálogos	Registros
27	15,044

7.2 Tablas

Nombre de Tabla	Número de Registros	Descripción
cepa	1803	Se quitaron los campos id_nivel y nom_rango y se reubicaron en la tabla infraespecie_ce incorporando autoraño_niv, status_niv y sistclas_niv y relacionandola con el id_cepa.
cepa_grals	1803	Se verificaron y modificaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
sinonimo	10789	Se eliminaron los campos id_nivel y nom_rango y se reubicaron en la tabla infraespecie_sin incorporando autoraño_niv, status_niv y sistclas_niv y relacionandola con el id_cepa. Tambien se verificaron, modificaron y anexaron nuevos datos a este catálogo.
cultivo_mixto	0	Se eliminaron los campos id_nivel y nom_rango y se reubicaron en la tabla infraespecie_mix incorporando autoraño_niv, status_niv y sistclas_niv y relacionandola con el id_cepa.
deposito	1803	
fa_datogeo	1803	Se anexaron nuevos datos a esta tabla.
fa_desclugar	1803	Se anexaron nuevos datos a esta tabla.
fa_colector	2130	Se anexaron nuevos datos a esta tabla.
fa_metodo	1803	Se anexaron nuevos datos a esta tabla.

fa_identificador	1803	Se anexaron nuevos datos a esta tabla.
aplicaciones	0	
cepa_desc	1118	Se verificaron y modificaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
acronimo	4022	Se verificaron y modificaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
cond_optima	2568	Se verificaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
ref_genomica	66	
telefonos	81	Se anexaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
infraespecie_ce	367	Se verificaron y modificaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
infraespecie_sin	2411	Se verificaron y modificaron nuevos datos a esta tabla, de acuerdo a las recientes necesidades de la colección.
infraespecie_mix	0	

Total tablas	Registros
19	36,474

8. Resultados obtenidos

A continuación se describen todos los resultados obtenidos durante el desarrollo de este proyecto:

- *Nuevo modelo conceptual.*- Este nuevo modelo incluye diversas características hasta antes no incluidas en modelos similares, lo cual permite ser una de las más avanzadas a nivel internacional (ver figura 1).
- *Nuevo modelo relacional.*- El presente proyecto fue enfocado en resolver sólo una parte del modelo conceptual, debido al tiempo aplicado para implementar la totalidad del modelo conceptual. Esta parte a desarrollar fue considerada debido a las expectativas requeridas por la Conabio, las entidades restantes serán consideradas en una etapa futura que están fuera por el momento del alcance de este proyecto.
- *Nuevas tablas.*- Mucha de la información actual no tenía un lugar donde pudiera ser integrada, debido a esta necesidad se generaron nuevas tablas que soporten estos nuevos datos (ver figura 2).
- *Nuevas relaciones entre tablas.*- Debido al cambio del modelo han surgido nuevas relaciones las cuales incrementan la complejidad en la implementación, para sortear este problema se ha verificado y probado a detalle cada una de las propuestas de diseño generando un arduo trabajo para alcanzar la implementación presentada en este reporte.
- *Rediseño de tablas.*- Todas las tablas iniciales tuvieron que ser sometidas a un proceso de evaluación y rediseño para producir mejoras sustanciales y una mayor integración con el nuevo modelo.
- *Nuevos catálogos.*- Al obtener la creación de nuevas tablas y el rediseño de otras para eliminar ambigüedad y redundancia, como solución se generaron catálogos que permiten una mayor verificación de la integridad de datos y permiten un avance sustancial para la interacción con la información.
- *Nuevo gestor de base de datos.*- Para una mejor integración con las tecnologías la nueva base de datos se ha implementado en el gestor de base de datos PostgreSQL 8.4, bajo el sistema operativo Linux. Esto nos permite una mayor capacidad de intercambio de información con bases de datos internacionales del mismo tipo.
- *Capacidad de escalabilidad de la base datos.*- El enfoque del nuevo modelo de datos fue permite escalar el sistema y extenderlo para atender las nuevas necesidades sin causar conflictos con los componentes y módulos de información ya implementados.

9. Discusión

Debido a que el modelo original fue realizado en el año de 1995 es evidente ya es obsoleto y no cumple con las expectativas para las nuevas necesidades, es por ello que el principal enfoque dado en este proyecto, es obtener un nuevo diseño e implementación que permita estar acorde con los nuevos paradigmas de disponibilidad e intercambio de información actual.

Una limitante encontrada con los datos relacionados a microorganismos, es la vasta complejidad semántica involucrada en ellos, esto dificulta la capacidad de obtener abstracciones por medios tradicionales, ya que el tipo de información es muy dinámica y contiene una serie de características que la hacen completamente diferente e involucran cierta complejidad relacional, que las técnicas y herramientas comunes no consideran. La única manera de alcanzar este objetivo es el diseño, desarrollo, implementación y prueba de pequeñas partes del modelo y verificar su integración con el modelo general propuesto.

En base a este nuevo contexto y complejidades técnicas se ha desarrollado una nueva versión de la base de datos la cual denominamos como **CDBB-500 v2.1**, la cual contiene una serie de innovaciones incluidas que representan un valioso avance en la abstracción de este tipo de información a nivel nacional e internacional.

Gran parte del éxito obtenido se debió a la estrecha colaboración de la Colección Nacional de Microorganismos y la Universidad Autónoma Metropolitana Unidad Cuajimalpa que mantuvieron una coordinación de actividades para el diseño e implementación de la nueva base de datos y la curación de toda la información. A su vez, la Universidad Nacional Autónoma de México y el Centro de Investigaciones Biológicas del Noroeste colaboraron con información de sus colecciones para integrarlas a nuestra nueva base de datos.

10. Conclusiones

El presente proyecto representa un avance sustancial para el país, ya que es el referente actualizado de la información microbiana en México y su importancia es reflejada a nivel internacional, ya que comparte información con federaciones de bases de datos mundiales especializadas en microorganismos.

Cabe destacar que los resultados obtenidos rebasan las expectativas iniciales comprometidas con la Conabio, ya que inicialmente se pensaba obtener sólo la información de 1600 cepas, pero al final se tienen *1803 registros 203 más de los comprometidos*. Otro importante rasgo a resaltar es la actualización de la taxonomía de las cepas, la cual es de vital importancia en la investigación, industria y medicina, además de contar con un amplio catálogo de la Biodiversidad existente a nivel nacional.

5. Referencias

Taxonomía y sinónimos.

<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>

<http://www.indexfungorum.org/Names/Names.asp?pg=2>

<http://www.mycobank.org/DefaultPage.aspx>

<http://tolweb.org/tree/>

<http://wdcn.nig.ac.jp/hpcc.html>

<http://www.bacterio.cict.fr>

<http://zipcodezoo.com/Bacteria>

<http://www.indexfungorum.org/>

<http://www.straininfo.net>

Referencias bibliográficas.

- [1] Dawynt P, Vancanneyt M, Janssens D. "Knowledge Management Systems for Microbial Information," *BCCM Newsletter*, Edition 19, Spring 2006.
- [2] Jovita Martínez Cruz y Sergio V. Chapa Vergara Informe Final: "Proyecto Colección de Cultivos Microbianos del CINVESTAV-IPN" *Base de Datos Convenio FB122/EO16/94*, México D.F. Abril de 1996.
- [3] Miguel Castaño, Mario G. Piattini, "Fundamentos y Modelos de Bases de Datos," *Alfaomega RA-MA editores*, 2a. Edición, 1999.
- [4] J. Sergio Zepeda, Sergio V. Chapa, "Micro500, Un Sistema para búsqueda de Conocimiento en Microorganismos" *32nd Latin America Conference on Informatics (CLEI06)*, Santiago de Chile, 2006.
- [5] J. Sergio Zepeda, Sergio V. Chapa, "Diseño e Implementación del Sistema Micro500 basado en una Arquitectura con Extension de Capacidades," *9na. Conferencia de Ingeniera Eléctrica*, Cinvestav, Mexico, 2003.
- [6] Cornelia B. Osmond, "The Universal Virus Database ICTVDB," *Computing in Science and Engineering*, pp. 16-25, May, 2003.
- [7] D. Wheeler, T. Barret, D. Benson, S. Bryan, "Database Resources of the National Center for Biotechnology Information," *Nucleic Acids Research*, Vol. 34, Database issue, D173-D180, 2006.
- [8] Michael L. Brodie, Frank Manola, "Database Management: A Survey," *Fundamentals of Knowledge Base Management Systems*, Springer Verlag, New York, 1988.
- [9] Evelina Lamma, Fabrizio Riguzzi, Sergio Storari, "Discovering Validation on Microbiological Data," *New Generation Computing*, 21, pp 123-133, 2001.
- [10] Peter Dawyndt, Marc Vancanneyt, Hans De Meyer, Jean Swings, "Knowledge Accumulation and Resolution of Data Inconsistencies during the Integration of Microbial Information Sources," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, No. 8, August, 2005.

- [11] Joaquín Sergio Zepeda Hernández, Jovita Martínez Cruz, Juan Carlos Estrada Mora, Armando Sánchez Chavarría, Sergio V. Chapa Vergara, "MicroXplore: A Prototype for Exploration on Databases with Microbial Information", *12th International Conference on Culture Collections (ICCC12)*, Florianópolis, Santa Catarina, Brazil, 2010.
- [12] Jovita Martínez Cruz, , Juan Carlos Estrada Mora, Armando Sánchez Chavarría, Joaquín Sergio Zepeda Hernández, Sergio V., "WDCM CDBB-500, Mexican Culture Collection, Importance for biotechnology development in Mexico", *12th International Conference on Culture Collections (ICCC12)*, Florianópolis, Santa Catarina, Brazil, 2010.