



NEXTGENDEM

Una plataforma web para la Gestión Científica de la Biodiversidad

MANUAL DE USO



MAC 2014-2020
Cooperación Territorial

Interreg

Fondo Europeo de Desarrollo Regional



EUROPEAN UNION

1. INTRODUCCIÓN	4
2. CARACTERÍSTICAS	7
2.1 FLEXIBILIDAD	7
2.2 REPLICABILIDAD	8
2.3 ARCHIVO	8
2.4 ANOTACIONES	8
2.5 CONTROL DE ACCESO	8
2.6 PROCESOS	8
2.7 RECURSOS INFORMÁTICOS	9
2.8 INTERFAZ DE USUARIO	9
2.9 VISUALIZACIONES Y VISTAS	9
2.10 ESTUDIOS DE CASO	9
2.11 COLECCIONES	10
2.12 LIMITACIONES	10
3. CONCEPTOS BÁSICOS DE LA ORGANIZACIÓN DEL SISTEMA NEXTGENDEM	12
4. INICIO DE SESIÓN Y PERFIL DE USUARIO	15
4.1 PERFIL DEL USUARIO	16
5. DISEÑO GENERAL DE LA INTERFAZ DE USUARIO	19
6. ELEMENTOS Y CONCEPTOS UTILIZADOS EN LA INTERFAZ DE USUARIO	21
6.1 OBJETOS DE DATOS	22
6.2 PERMISOS	22
6.3 LISTAS-DETALLES	25
6.4 EDITOR DE ANOTACIONES	28
6.5 BÚSQUEDAS Y FILTROS	30
6.6 NAVEGACIÓN	33
6.7 PERSISTENCIA DE PANTALLAS	33
6.8 APERTURA CON IMPORTACIÓN – PROCESO - EXPORTACIÓN -REIMPORTACIÓN	33
6.9 CONVERSIONES DE FORMATO	33
6.10 VERSIONADO	34
7. GESTIÓN DE DATOS BIOINFORMÁTICOS	36
7.1 IMPORTAR DATOS MOLECULARES	36
7.2 IMPORTAR SECUENCIAS DE NUCLEÓTIDOS	37

7.3 IMPORTAR ANÁLISIS MOLECULAR	39
7.4 CONSULTAR DATOS MOLECULARES	40
7.5 EDITAR DATOS MOLECULARES	41
7.6 ELIMINAR DATOS MOLECULARES	41
7.7 EXPORTAR DATOS MOLECULARES	41
8. EDITOR DE ALINEAMIENTO MÚLTIPLES SECUENCIAS	43
9. EDITOR DE DATOS DE ALINEAMIENTO FALTANTE	47
9.1 EDITOR DE VALORES PERDIDOS	47
9.2 SECCIÓN DE IMPORTACIÓN E HISTORIAL	47
9.3 SECCIÓN IMPORTAR/EXPORTAR	48
10. VISOR DE ÁRBOLES FILOGENÉTICOS	50
10.1 SECCIÓN DE CONTROLES	50
10.2 SECCIÓN DE COLOR	52
10.3 VISOR DE ÁRBOLES	53
10.4 GUARDAR BOTÓN DE PROGRESO	53
11. CÓDIGO DE BARRAS	55
12. GESTIÓN DE DATOS GEOGRÁFICOS	61
12.1 LISTA DE CAPAS	61
12.2 DETALLES DE CAPA Y EDICIÓN	62
12.3 IMPORTAR CAPA	64
12.4 EXPORTAR CAPAS O ASOCIARLAS A CASOS DE ESTUDIO / COLECCIONES	64
12.5 PERMISOS	64
12.6 CREACIÓN DE UNA NUEVA CAPA A PARTIR DE CAPAS EXISTENTES	64
13. VISOR GEOGRÁFICO Y MAPAS	68
13.1 CÓMO SE DIBUJA UN MAPA	68
13.2 EL PANEL DE CONFIGURACIÓN (DERECHA)	69
13.3 MAPA INTERACTIVO (PANEL IZQUIERDO)	74
14. COLECCIONES	75
15. ESTUDIOS DE CASO	78
16. VISORES	80

17. GESTIÓN DE PROCESO	82
17.1 EJECUCIÓN DE UN PROCESO	82
17.2 DETALLE DEL PROCESO	84
18. PROCESOS DISPONIBLES	88
18.1 ALINEAMIENTO MÚLTIPLE DE SECUENCIAS	88
18.2 FILOGENIAS	90
18.4 DIVERSIDAD FILOGENÉTICA	95
18.5 PICANTE*	95
18.6 IDENTIFICACIÓN DE SECUENCIA	96
18.7 BLAST+*	96
18.8 MATRICES DISCRIMINANTES	96
18.9 MOLD*	96
19. CARACTERÍSTICAS DESCARTADAS	98

INTRODUCCIÓN



NEXTGENDEM

El propósito de este documento (Manual de Uso) es proporcionar una guía completa sobre la utilización del sistema NEXTGENDEM (<https://sys.nextgendem.eu>), una herramienta poderosa y versátil desarrollada bajo el proyecto NEXTGENDEM (MAC2/4.6d /236, ver <https://www.nextgendem.eu/es>). Este sistema innovador está diseñado principalmente para profesionales de la biología y tomadores de decisiones, pero también para estudiantes, profesorado o personas interesadas en la biodiversidad, que buscan aprovechar el poder de los indicadores filogenéticos de biodiversidad y otros indicadores ecológicos a nivel geográfico para guiar el proceso decisorio. Al aprovechar los datos moleculares y geográficos, el sistema NEXTGENDEM tiene como objetivo facilitar la toma de decisiones informadas en el ámbito de la conservación de la biodiversidad y, por lo tanto, en el desarrollo sostenible (principalmente, el objetivo 15 de los ODS, que habla de "proteger, restaurar y promover el uso sostenible de los ecosistemas terrestres, gestionar de forma sostenible bosques, combatir la desertificación y detener e invertir la degradación de la tierra y detener la pérdida de biodiversidad").

El sistema NEXTGENDEM es un recurso de alto valor para obtener indicadores filogenéticos y de biodiversidad, ya que son herramientas esenciales para evaluar la salud y la resiliencia de los ecosistemas. Mediante la integración de datos moleculares de regiones codificadas por secuencias de ADN (regiones relativamente cortas del ADN conocidas por caracterizar especies de plantas terrestres en el caso del proyecto NEXTGENDEM, pero abierto a otros seres vivos), y con información geolocalizada, el sistema permite a las personas usuarias generar indicadores ecológicos que pueden guiar los procesos de toma de decisiones. Además, el sistema NEXTGENDEM permite la identificación de especies de una muestra determinada comparando su secuencia de ADN con las secuencias existentes en la base de datos. Sin olvidar que, además de esto, se pueden realizar estudios de caracteres de diagnóstico, de cara a analizar variaciones entre especies.

La interfaz del sistema NEXTGENDEM está diseñada para que las personas usuarias puedan navegar a través de los navegadores en busca de conjuntos de elementos (secuencias, alineaciones de secuencias, capas geográficas...), acceder a formularios detallados y emplear visualizadores especializados adaptados a diferentes tipos de datos. Esto permite explorar, analizar e interpretar de manera eficiente la gran cantidad de información disponible en el sistema.

Una de las claves en la creación del sistema NEXTGENDEM ha sido que se han integrado componentes y procesos ya validados para lograr su funcionalidad, y que se pueden usar computadoras externas de alto rendimiento para realizar los cálculos pesados. Al aprovechar los avances recientes en biodiversidad y bioinformática, el sistema garantiza que las personas usuarias tengan acceso a métodos actualizados.

Este manual centra su esfuerzo en explicar el uso del sistema NEXTGENDEM a través de la interfaz de usuario/a web. Por lo que ofrece instrucciones detalladas sobre cómo navegar e interactuar con las diversas características y funcionalidades de la plataforma. En cambio, otra información del sistema NEXTGENDEM no incluidos en el presente manual, está incluida en la documentación adicional que facilita el manual de administrador/a, el manual del usuario/a de API (para usar el sistema con scripts en Python y R), y el manual del desarrollador/a. Por tanto, como se puede deducir, estos recursos complementarios abordan la administración del sistema, el uso de la API desde scripts, la implementación y la mejora, respectivamente.

Por último, cabe mencionar que, tras una descripción de las características, el documento presenta las funcionalidades y luego aborda el "cómo se usa" a través de algunos casos que pueden ser útiles para completar la referencia a las funciones.

CARACTERÍSTICAS



NEXTGENDEM

2. CARACTERÍSTICAS

NEXTGENDEM es un sistema de software diseñado para gestionar y procesar datos moleculares y geográficos de diferente tipo, con aplicación a la conservación de la biodiversidad. Cuenta con un marco que proporciona capacidades generalistas, como archivado, control de versiones, anotaciones, ontologías fácilmente conectables, control de acceso, etc. que pueden volverlo reutilizable en otras áreas relacionadas con la información molecular o geográfica y su procesamiento.

Se puede utilizar a través de dos interfaces; la interfaz gráfica de usuario/a (GUI), objeto de este Manual de Uso y la interfaz para programadores, actualmente solo en lenguaje Python, que permite implementar otros flujos de trabajo, y se describe en el documento “Manual de usuario/a de la API”. Asimismo, este documento explica la funcionalidad, pero no entra en la administración (parametrización) de la herramienta, que se explica en el “Manual de Usuario Administrador”, donde se explica cómo configurar los aspectos de la herramienta que se pueden adaptar (ver Flexibilidad justo debajo).

2.1 FLEXIBILIDAD

El sistema está estructurado para permitir flexibilidad en diferentes aspectos:

- **Taxones:** aunque está pensado para plantas, está abierto al estudio de otros reinos.
- **Loci:** el sistema ha sido construido para regiones de códigos de barras. Pero está abierto a otras regiones de ADN nuclear o extranuclear.
- **Regiones geográficas:** las capas geográficas contienen información sobre una localidad, por lo que un conjunto de capas geográficas puede hacer referencia a varias regiones. Aunque un caso de estudio normalmente se refiere a la ubicación (Gran Canaria y Santiago en el caso del proyecto NEXTGENDEM).
- **Formatos:** en el aspecto molecular, hay algunos formatos recomendados para importar y exportar datos, aunque muchos otros están permitidos. Los formatos principales son Genbank y FASTA para secuencias de nucleótidos, FASTA, Clustal y Nexus para alineamiento múltiple de secuencias y Newick y Nexus para árboles filogenéticos. En el caso de datos geográficos, se admiten Shapefiles, Geopackage, CSV y GeoJSON. También se han implementado formatos especiales para la ejecución de paquetes de Diversidad Filogenética, aunque se utilizan casi exclusivamente cuando se ejecuta este tipo de procesos.
- **Procesos:** aunque el sistema se ha desarrollado albergando todos los procesos bioinformáticos necesarios para cubrir los objetivos de NEXTGENDEM, se pueden agregar nuevos procesos fácilmente (no obstante, tiene que ser a través de programación). De hecho, durante el proyecto se preparó un volumen especial de “geoprocesos”, aunque ahora mismo no está habilitado, para lanzar algoritmos que operen con capas geográficas.
- **Recursos informáticos:** una vez que se define un proceso, debe ejecutarse en una infraestructura HPC (computadora de alto rendimiento), o en un servidor estándar, según sea necesario. Desde la interfaz de usuario (para usuarios administradores) es posible configurar el acceso a cualquier número de recursos de cómputo, y asociarlos a los procesos.
- **Anotaciones:** cada uno de los objetos de datos pueden ser anotadas. Las anotaciones pueden ser anotaciones de texto enriquecido (forma libre) o campos o plantillas definidas mediante ontologías cargadas en el sistema. Las plantillas pueden ser, por ejemplo, referencias bibliográficas u otras y las anotaciones pueden incluir referencias a otros objetos de datos en el sistema.
- **Implementaciones:** el uso de *Docker* para empaquetar el *backend* y del *docker compose* para repartir los servicios complementarios, además de los archivos de configuración, facilita la personalización de la implementación en casi cualquier servidor.

2.2 REPLICABILIDAD

La replicabilidad califica cuán fácil es implementar y adaptar un sistema a diferentes contextos. Por lo que, está bastante relacionado con lo anterior, en lo que a flexibilidad se refiere, pero enfocado en cómo las diferentes réplicas del sistema pueden responder a las necesidades locales. Algunos factores adicionales que se alinean con la reproducibilidad son:

- Utiliza componentes de código abierto y es de código abierto en sí mismo. En caso de que el sistema no tenga la funcionalidad deseada, se puede modificar o mejorar a nivel de diseño y código fuente.
- Debido a que se ha colocado en contenedores con Docker, se podrían considerar tecnologías de implementación avanzadas como Kubernetes, un servicio en la nube alternativo (como Postgres en AWS o REDIS).

2.3 ARCHIVO

El archivo es uno de los dos subsistemas principales del sistema. Permite almacenar información de muy diversa índole: molecular, geográfica, metadatos, etc. y gestionarla ofreciendo a los usuarios un conjunto de operaciones para introducir información en el sistema, y sacarla al exterior para procesos no considerados en su interior.

2.4 ANOTACIONES

La herramienta de anotación administra las anotaciones asociadas con los datos almacenados. Cada objeto de datos se puede anotar por separado. El sistema comienza con algunas plantillas preestablecidas de BibTex y Darwin Core, pero los usuarios pueden crear nuevas plantillas. Las anotaciones proporcionan contexto adicional, notas descriptivas o información definida por el usuario que puede mejorar la comprensión del elemento y evitar los límites rígidos y estáticos de un esquema de base de datos.

2.5 CONTROL DE ACCESO

Se ha preparado un mecanismo para controlar el acceso tanto a la funcionalidad como a los datos. Se pueden asignar permisos directamente a los usuarios (a ambas clases de elementos, funciones y datos autorizados), pero existen otros mecanismos para facilitar la administración de permisos: roles, colecciones, tipos de permisos. Un usuario con un rol básico tendrá acceso a mucha menos funcionalidad que uno con un rol de nivel superior. De la misma manera, si un usuario no tiene un rol que permita el acceso a los datos, casi no verá información. Siendo éste el rol predeterminado, pudiendo exclusivamente abrir-objetos o colecciones de datos muy concretos.

2.6 PROCESOS

El subsistema de procesamiento (el otro subsistema central junto con el de archivo) proporciona un formulario donde los usuarios pueden ingresar la información necesaria para ejecutar los procesos bioinformáticos admitidos. Esto puede incluir la selección del tipo de análisis, los recursos informáticos, los datos de entrada y los parámetros.

Todos los análisis disponibles en el sistema son todos los procesos en el sistema para la conservación e identificación de especies. Esto implica análisis tales como alineamiento múltiple de secuencias y filogenias.

Una vez que el usuario completa el formulario y selecciona los recursos informáticos, el sistema envía el proceso al servidor remoto o al clúster de HPC para su ejecución. El usuario podrá seguir el estado del proceso y se le notifica una vez que haya finalizado. Es posible descargar, importar y visualizar los archivos de salida resultantes, verificar los parámetros seleccionados y leer los registros del proceso.

2.7 RECURSO INFORMÁTICOS

Los recursos informáticos son infraestructuras a las que se accede de forma remota, externas al sistema NEXTGENDEM, capaces de realizar cálculos pesados y en las que se han instalado uno o más procesos compatibles con el sistema NEXTGENDEM.

Estos recursos a veces se denominan computadoras de alto rendimiento (HPC), que generalmente integran hardware especializado y costoso capaz de realizar cálculos complejos, lo que permite al personal científico procesar grandes cantidades de datos y ejecutar software avanzado de manera eficiente. Este tipo de infraestructura normalmente se rige por un sistema de colas, que recibe las solicitudes para ejecutar los cálculos acompañado por la cantidad de procesadores y horas de tiempo real de espera que se le otorgar. Una vez que se reciben las solicitudes, el programador del sistema de colas encuentra la mejor opción, considerando que muchos otros usuarios pueden estar ejecutando sus cálculos y enviando sus solicitudes en paralelo.

El sistema administra una lista de recursos informáticos que los usuarios pueden seleccionar para ejecutar procesos bioinformáticos o geográficos en él, y es posible que los administradores agreguen nuevos con relativa rapidez.

2.8 INTERFAZ DE USUARIO

Al acceder a la URL del sistema (<https://sys.nextgendem.eu/>), se ingresará a la interfaz gráfica de usuario, que es el medio para operar con todas las funciones proporcionadas por el sistema. La interfaz de usuario en sí se ha construido utilizando modernas tecnologías de aplicaciones web e integra componentes avanzados existentes que se han adaptado en gran medida a las especificidades del proyecto.

2.9 VISUALIZACIONES Y VISTAS

La interfaz de usuario puede visualizar información de diferente tipo (secuencias, alineaciones, capas geográficas...) con visores especializados. Los visores son otro tipo de pantalla en la que la información se representa mayoritariamente de forma gráfica, aunque puede aparecer información textual para complementar las gráficas. Estos visores están preparados para ofrecer a los usuarios controles interactivos y que puedan cambiar la forma en que se representa la información (no la información en sí), con diferentes finalidades (exploración, revisión de resultados, análisis, mostrar resultados, publicación...). Debido a que la interfaz de usuario se usa para intercambiar pantallas, es posible guardar el estado de las visualizaciones en las vistas, que se pueden volver a abrir más tarde.

2.10 ESTUDIOS DE CASO

El subsistema de archivo puede almacenar miles de registros de cualquiera de los tipos existentes. A la hora de navegar o visualizar, esto podría ser un problema a la hora de encontrar los datos sobre los que trabajar. Sin embargo, al crear un estudio de caso, el usuario puede agregar objetos a ese dicho estudio de caso e indicar a la interfaz de usuario que solo muestre objetos en ese estudio de caso.

2.11 COLECCIONES

Al igual que los estudios de casos, administrar los permisos de muchos objetos puede ser engorroso y propenso a errores. Para mitigar esto, se pueden agregar objetos a una colección. Al hacer esto, se heredan los permisos de esa colección.

2.12 LIMITACIONES

El sistema se ha preparado para cubrir las necesidades del proyecto en el que ha sido concebido, NEXTGENDEM. Aunque el marco informático, construido desde cero para el proyecto, también proporciona otras funciones gracias a los potentes componentes y subsistemas que forman parte de él.

Sin bien existen limitaciones y éstas son:

- **Genomas completos o secuencias muy largas.** Aunque teóricamente el sistema sería capaz de ingerir genomas completos, en la práctica el sistema se cuelga (deja de funcionar) cuando se utilizan secuencias con más de 100k nucleótidos. La gestión de tal cantidad de información necesita un almacenamiento de datos diferente y algoritmos diferentes.
- **El sistema no puede administrar capas vectoriales de alta resolución también debido al tamaño.** Nuevamente, se deben considerar nuevos tipos de almacenamiento.
- **Capas ráster.** A veces es útil preparar cálculos geográficos aprovechando la capacidad de las capas disponibles como ráster (ortofotos, satélite en su mayoría). Actualmente no están en la funcionalidad, incorporar ráster implicaría un nuevo almacenamiento de datos, una nueva gestión y visualización y adaptación de algoritmos a cómo se ordena y codifica la información (principalmente la proyección).
- **Los procesos se ven afectados no solo por el uso de secuencias largas.** Sino que también, si el envío de un proceso usa una combinación de parámetros que fuerza muchas interacciones, el tiempo aumentará.
- **Control adicional de envíos a HPC.** Algunas infraestructuras de HPC permiten diferentes colas para una cuenta, por usuario, incluso para la misma institución.
- **También en el dominio geográfico, actualmente se carece de la posibilidad de almacenar y procesar imágenes ráster.** Esto requeriría agregar nuevas funciones tanto en el backend como en la interfaz de usuario.



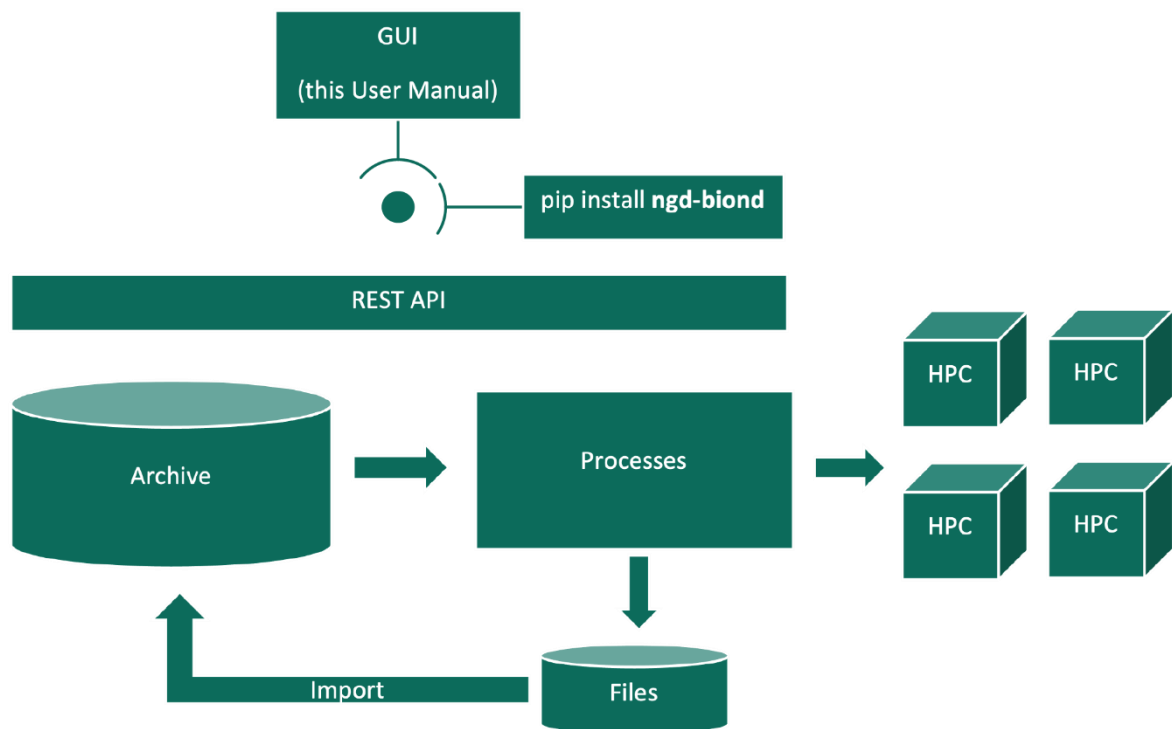
CONCEPTOS BÁSICOS DE LA
ORGANIZACIÓN DEL SISTEMA
NEXTGENDEM



NEXTGENDEM

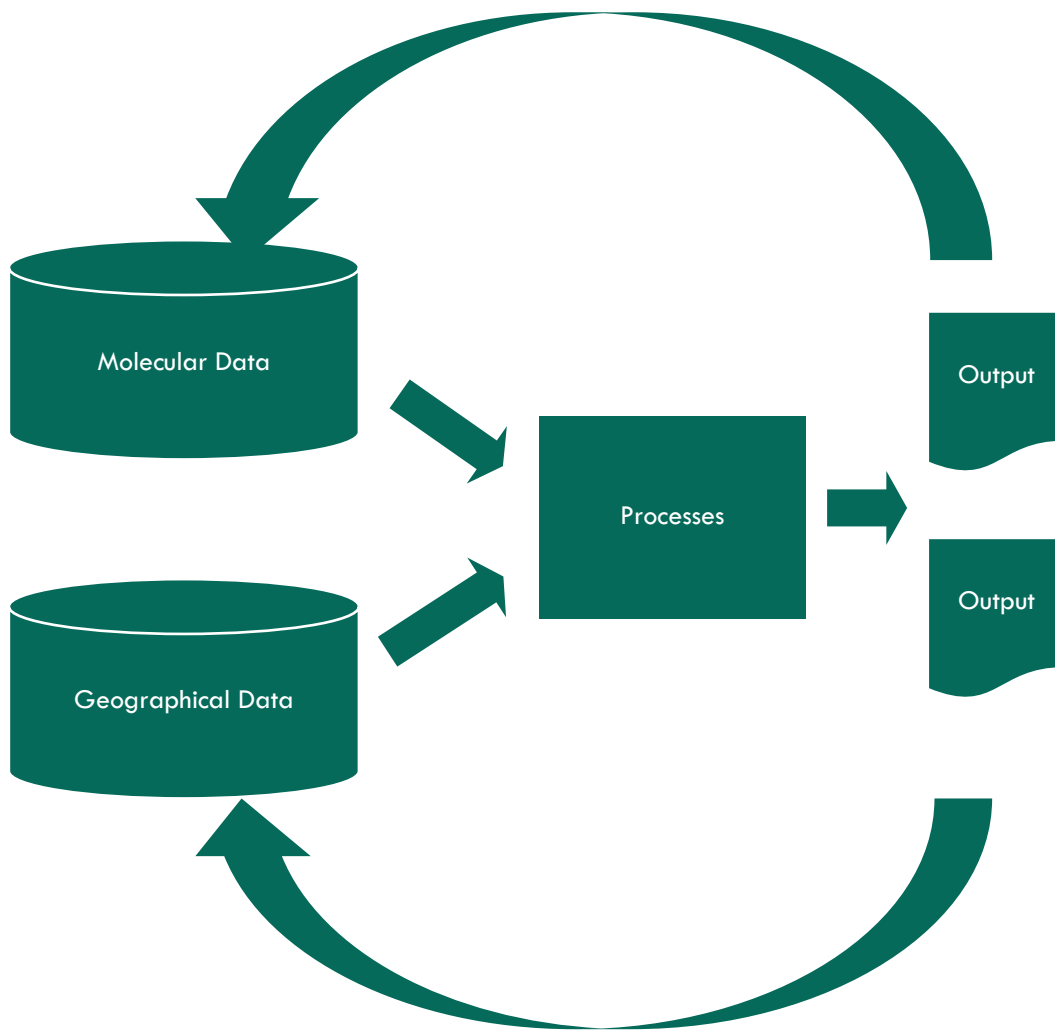
3. CONCEPTOS BÁSICOS DE LA ORGANIZACIÓN DEL SISTEMA NEXTGENDEM

NEXTGENDEM sigue una organización de arquitectura cliente-servidor. Esto quiere decir que servidor espera las solicitudes de los clientes/usuarios. El cliente puede enviar solicitudes por dos vías. Una es la interfaz de usuario explicada en este documento, y la otra es Biond, para usar el servidor desde scripts (Python o R).



Componentes principales del servidor.

Los procesos reciben entradas del Archivo. El procesamiento se produce en los HPC. Los resultados de los procesos se almacenan en un sistema de archivos temporal y se importan al Archivo.



Un esquema similar, para mostrar que los datos moleculares y geográficos se almacenan utilizando diferentes formatos.

INICIO DE SESIÓN Y PERFIL DE USUARIO



NEXTGENDEM

4. INICIO DE SESIÓN Y PERFIL DE USUARIO

Cuando una persona abre la aplicación el sistema solicita una identificación (para determinar el nivel de acceso a funciones y datos). En este caso, se muestra un cuadro de diálogo como el siguiente:

Primero seleccione un idioma. Luego, estarán disponibles diferentes opciones para iniciar sesión:

La imagen muestra una interfaz de usuario para el inicio de sesión de la aplicación NEXTGENDEM. Se presentan dos versiones del formulario: una en español ('Iniciar sesión') y otra en inglés ('Sign in').

El formulario en español ('Iniciar sesión') incluye:

- Un menú desplegable de idiomas con 'Español' seleccionado.
- Un campo de texto para 'Email'.
- Un campo de texto para 'Contraseña'.
- Un enlace: '¿Aún no tienes cuenta? Puedes crear una.'.
- Un botón 'Iniciar sesión'.
- Un enlace: '¿Olvidaste la contraseña?'.
- Un texto: 'O puedes iniciar sesión usando'.
- Botones de inicio de sesión con 'Google' e 'Invitado/a'.

El formulario en inglés ('Sign in') incluye:

- Un menú desplegable de idiomas con 'English' seleccionado.
- Un campo de texto para 'Email'.
- Un campo de texto para 'Contraseña'.
- Un botón 'Registrarse'.
- Un enlace: '¿Ya tienes cuenta? Accede a ella.'.

1. Inicie sesión como **invitado/a**.

Al hacer clic en el botón **Invitado/a**, no se necesitan credenciales ni creación previa de cuenta. Puede ser una buena opción para explorar las funciones y datos abiertos al público.

2. Inicie sesión con la cuenta **de Google**.

Seleccionando el botón de Google deberá seguir las indicaciones.

3. Inicie sesión registrando (**crear**) una cuenta.

Para iniciar sesión con un Email y Contraseña, debe seleccionar "Puedes crear una", donde después de llenar el formulario de registro, se enviará un correo de confirmación para activar la cuenta e iniciar sesión con la anterior opción.

4. Inicie sesión con **una cuenta ya existente**.

Si ya dispone de un correo electrónico y contraseña de usuario, llene los campos "Correo electrónico" y "Contraseña", para luego seleccionar "Iniciar sesión".

5. Puede recuperar/recordar su contraseña activando la opción de ("Olvidó su contraseña").

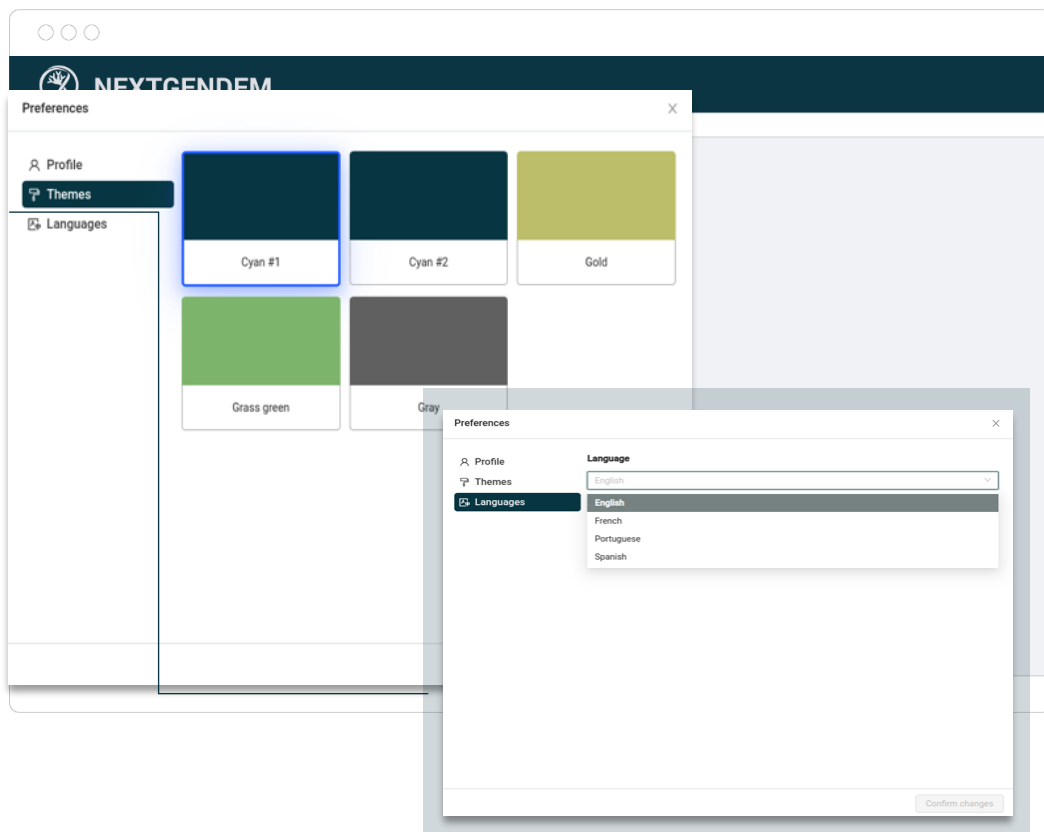
4.1 PERFIL DEL USUARIO

Una vez que un usuario inicia sesión, es posible configurar los datos de perfil, tema de color, idioma y claves API (esto último, es solo para usuarios que usen el paquete API Python).

Primero haga clic en el icono de Usuario, en la parte superior derecha de la ventana y se desplegará la ventana de preferencias con tres opciones; perfil, temas e idiomas.

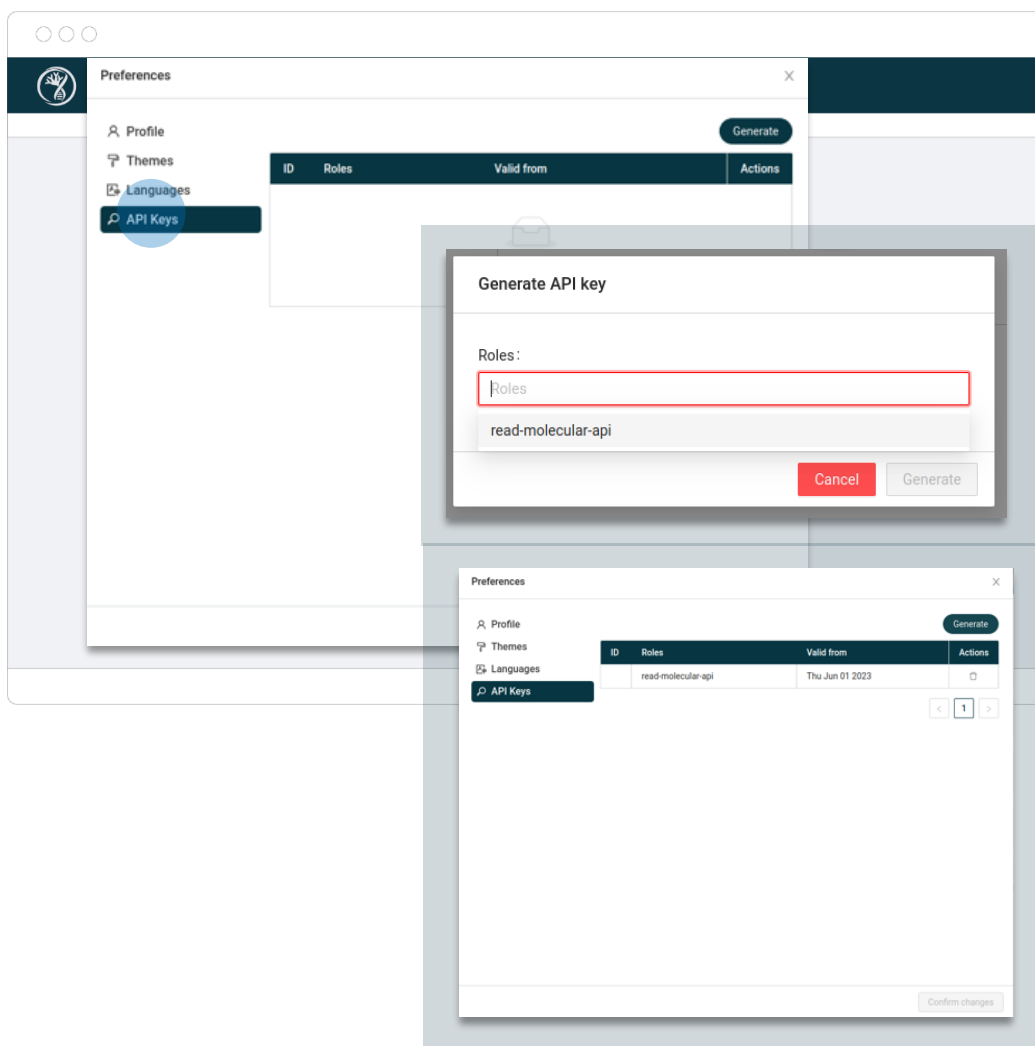
Para cambiar el color de la interfaz, haga clic en **“Temas”** y seleccione la combinación preferida.

Para cambiar el idioma (de los elementos de la interfaz de usuario, no del archivo) seleccione **“Idiomas”** y elige entre español, portugués, inglés o francés (el idioma cambia inmediatamente).



Finalmente, si un usuario ha sido asignado a uno de los roles API (solo los administradores pueden hacer esto), la opción “Claves API” estará visible. Al seleccionar esta opción, se muestra la lista de API Keys actuales. Al hacer clic en “Generar”, se abre un cuadro de diálogo para crear una nueva clave API. El cuadro de diálogo necesita especificar los roles deseados para la clave API, seleccionados entre aquellos habilitados para el usuario (hay roles de solo lectura y roles de lectura y escritura, útiles para evitar llamar a operaciones API que pueden modificar datos en el archivo cuando esa no es la intención del trabajo que se está realizando).

Cuando se crea la clave API, se genera un archivo ‘api_key.json’ y se descarga automáticamente. Este archivo debe ser guardado en un lugar seguro por el usuario. Cuando un script que usa el paquete ‘biond’ realiza un inicio de sesión, se refiere al nombre del archivo (no al contenido). Además, el sistema no tiene una forma de recuperarla, aunque es posible generar una nueva clave de API y eliminar la perdida, o tener diferentes claves de API para diferentes propósitos.



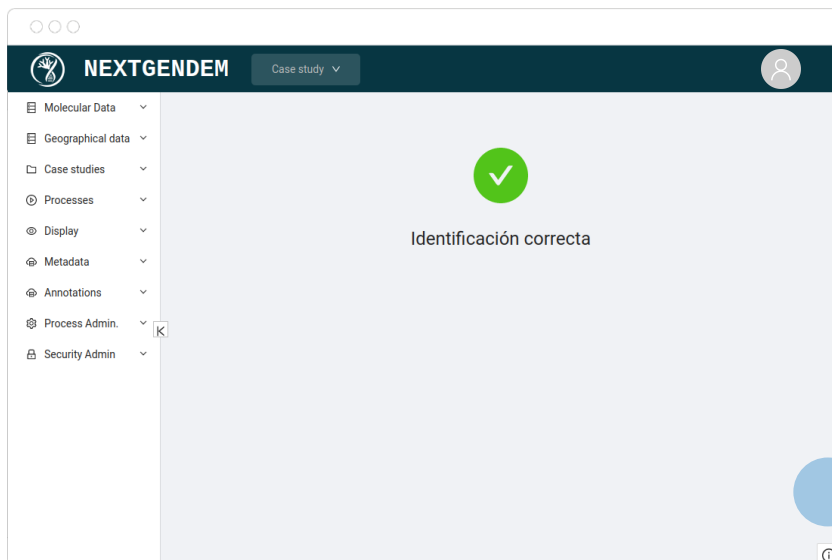
DISEÑO GENERAL DE LA INTERFAZ DE USUARIO



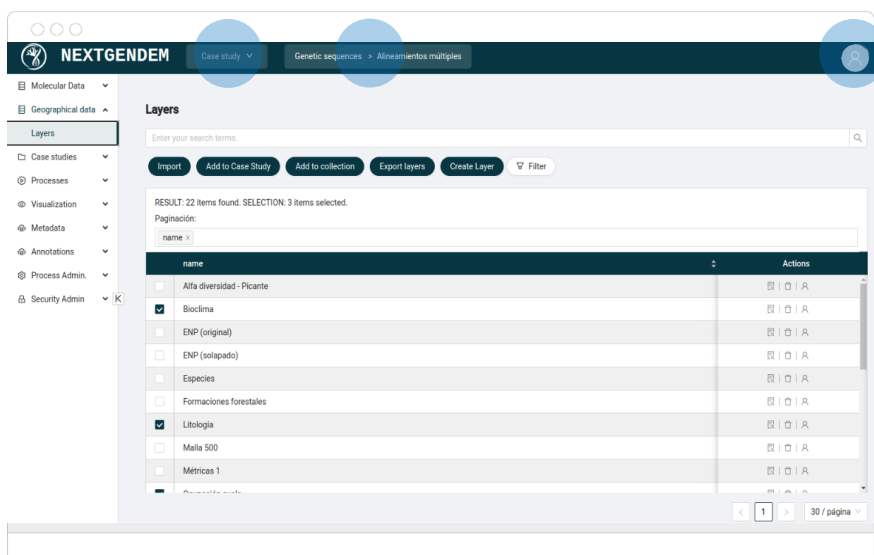
NEXTGENDEM

5. DISEÑO GENERAL DE LA INTERFAZ DE USUARIO

La interfaz gráfica de usuario está diseñada para proporcionar una experiencia de interacción fácil e intuitiva con el sistema. El diseño general consta de varios componentes clave que facilitan la navegación y el acceso a las características y funcionalidades.



En la parte superior de la interfaz se encuentra la **barra superior** que muestra, de izquierda a derecha, el **logotipo** del proyecto NEXTGENDEM sirviendo como identificador visual. Junto a él, tiene un **selector de estudios de caso** que permite a los usuarios establecer el contexto de trabajo para sus búsquedas y operaciones. Esto ayuda a los usuarios a concentrarse en datos y tareas específicos relevantes para su proyecto o estudio actual. A la derecha del selector de estudios de caso, tiene un control **de migas de pan** que permite volver a las pantallas anteriores.



En la esquina derecha de la barra superior, los usuarios pueden encontrar un ícono de avatar que brinda acceso a la sesión y las preferencias del usuario registrado. Al hacer clic en el avatar, los usuarios pueden administrar la configuración de su cuenta, personalizar su perfil y realizar otras acciones relacionadas. Consulte la sección anterior.

En el lado lateral izquierda de la pantalla, en forma de columna hay un espacio dedicado al menú de opciones donde horizontalmente se agrupan las categorías funcionales. Esta columna lateral sirve como centro de navegación, lo que permite a los usuarios navegar fácilmente a través de diferentes pantallas y secciones del sistema.

El menú está organizado en categorías siguiendo como lógica, la separación de datos moleculares y geográficos, luego otras etapas como procesamiento, visualización, administración, etc. Esta estructura ayuda a los usuarios a ubicar y acceder rápidamente a las funcionalidades específicas que necesitan.

En la esquina inferior derecha de la interfaz, hay un pequeño **ícono de exclamación** con una insignia que indica la **cantidad de notificaciones**. Los usuarios pueden hacer clic en este icono para mostrar la tabla con el registro de mensajes recientes. Esta tabla o ventana de notificaciones, alertas o mensajes relevantes, garantiza que los usuarios estén informados sobre los eventos o actualizaciones importantes del sistema.

Por último, la vista del componente principal utiliza el centro de la interfaz, ocupando la mayor cantidad de espacio posible en la pantalla. Esta área está dedicada a mostrar la pantalla actual, donde se activa la funcionalidad de la aplicación. Está diseñado para proporcionar un espacio de trabajo claro y enfocado donde los usuarios pueden interactuar con datos moleculares y geográficos, realizar visualizaciones, acceder a procesos del sistema y realizar otras tareas específicas de sus flujos de trabajo.

Al organizar la interfaz gráfica de usuario de esta manera, se pretende mejorar la productividad del usuario, simplificar la navegación y brindar una experiencia agradable a la vista que facilite la interacción eficiente con el sistema.

ELEMENTOS Y CONCEPTOS
UTILIZADOS EN LA INTERFAZ DE
USUARIO



NEXTGENDEM

6. ELEMENTOS Y CONCEPTOS UTILIZADOS EN LA INTERFAZ DE USUARIO

La interfaz de usuario del sistema NEXTGENDEM brinda a los usuarios acceso a un conjunto amplio de información con distinta tipología y características, que abarcan casi 100 pantallas diferentes. Para establecer una experiencia coherente y fácil de usar, se han aplicado patrones específicos en todo el sistema que proporciona una forma uniforme de organizar los diversos elementos de la interfaz gráfica de usuario (GUI). En el resto de la sección se enumeran estos patrones en detalle.

6.1 OBJETOS DE DATOS

El sistema está organizado internamente para operar con diferentes tipos de datos, que actualmente son: secuencias de nucleótidos, alineamiento múltiple de secuencias, árboles filogenéticos, matrices discriminantes, capas geográficas (vectoriales), especies y supermatrices. También existen objetos de datos o entidades secundarias (importantes, pero no objeto de estudio en sí mismos): colecciones, estudios de casos, anotaciones y vistas.

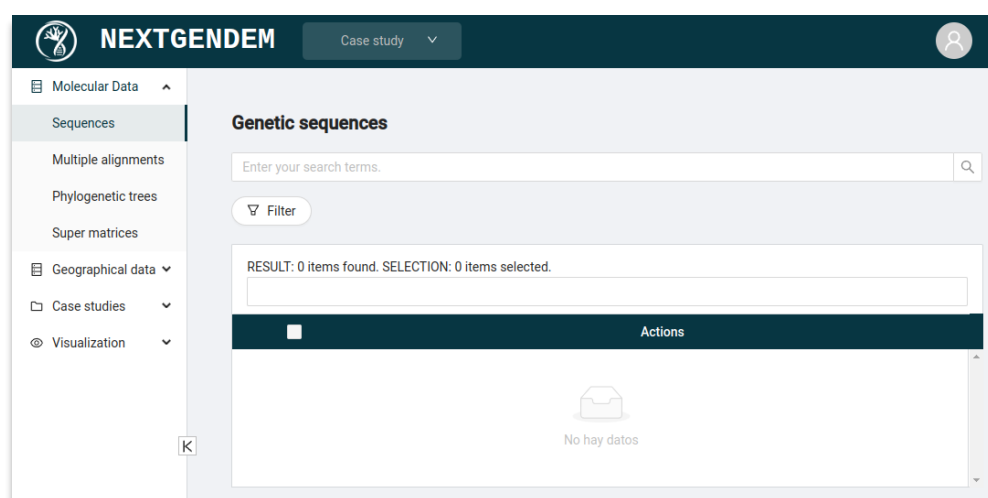
Todos ellos son objetos de datos en el sentido de que son funcionalidades que están provistas de un conjunto de operaciones y campos de información comunes, por lo que comparten parte del comportamiento: búsqueda de texto completo y búsqueda de filtro por atributo, almacenamiento de atributos, permisos y referencia de permisos y anotaciones.

6.2 PERMISOS

NEXTGENDEM tiene un sistema elaborado para controlar a qué funciones se puede acceder y cuales se pueden ejecutar.

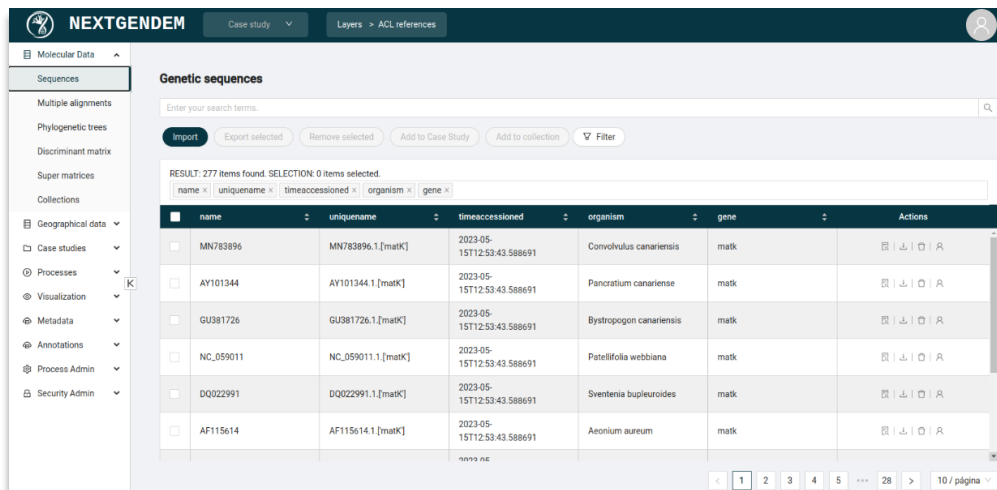
Cuando un usuario inicia sesión, la interfaz de usuario muestra las acciones que se pueden realizar. Al explorar los objetos de datos, las tablas desplegadas mostrarán datos con los que el usuario puede operar de alguna manera, desde leer hasta editar o eliminar.

El nivel de acceso a la funcionalidad de la interfaz de usuario depende del rol del usuario que inició la sesión. Por ejemplo, las siguientes figuras muestran las funciones que pueden ver un invitado y un usuario con acceso completo:



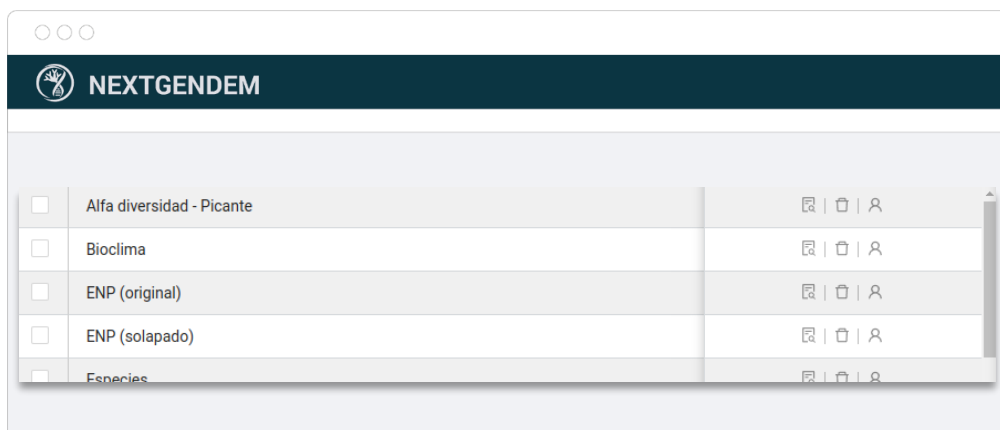
En cuanto al acceso a los datos, el sistema de permisos comienza prohibiendo el acceso a cualquier dato. Por lo tanto, teniendo en cuenta que se establecen los permisos para cada objeto de datos o entidad, y que sería poco práctico conceder acceso una por una, se han implementado tres mecanismos para adjudicar rápidamente el nivel de acceso requerido a la información:


- **Roles:** en lugar de asignar permisos a los usuarios, se asignan roles. Cuando se asigna un rol a un usuario, obtendrá todos los permisos otorgados a ese tipo de rol. Del mismo modo, se pueden otorgar permisos a organizaciones. Los roles son para funciones, mientras que las organizaciones son útiles para separar el acceso por institución.

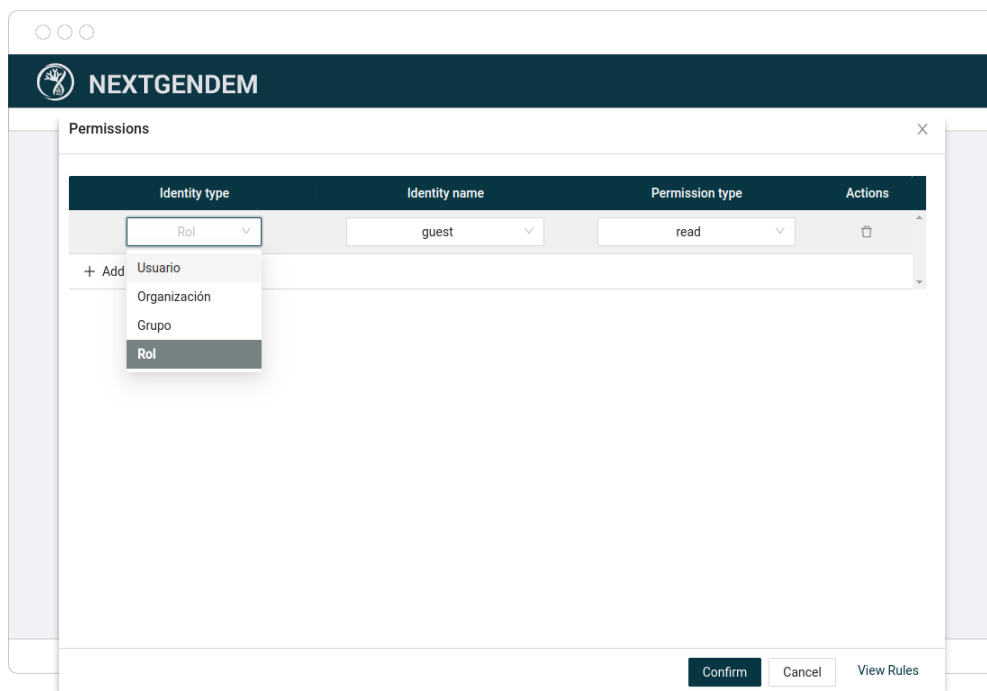


- **Objetos de referencia:** cada tipo de objeto, secuencia, alineamiento, capa geográfica, etc. tiene un objeto de referencia (esto puede ser administrado por los usuarios administradores dentro del menú "Administrador de seguridad"). Todos los objetos de ese tipo asumen los permisos establecidos para las referencias.
- **Colecciones:** el sistema permite administrar colecciones como contenedores de objetos principalmente para administrar permisos. Los permisos asignados a una colección son heredados por los objetos dentro de esa colección.

Los usuarios pueden definir permisos de objetos de datos, ya sean datos específicos o colecciones. Las siguientes capturas de pantalla muestran cómo funcionan los permisos:

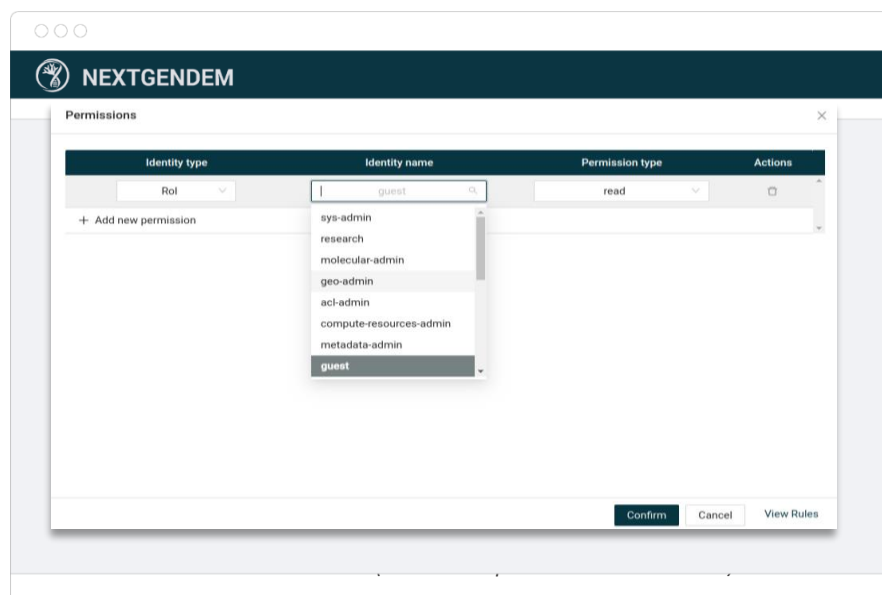


El navegador muestra iconos de acción y el icono de acción para establecer permisos es “”. Cuando se hace clic en este icono de acción, se muestra el siguiente cuadro de diálogo:

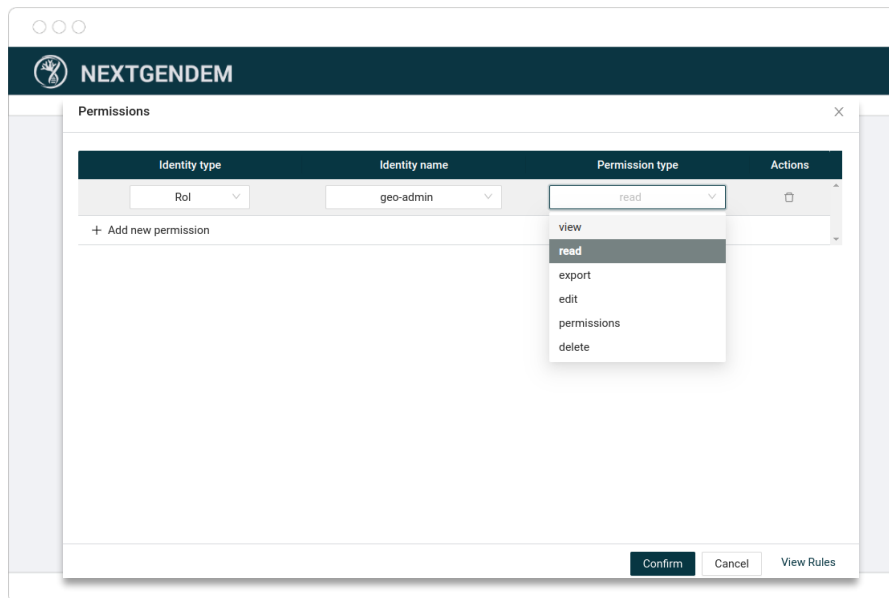


En el cuadro de diálogo se puede administrar (crear, editar, eliminar) la lista de permisos específicos para ese objeto. Como puede verse, cada permiso requiere configurar tres campos:

- **Tipo de identidad.** Puede ser Usuario, Organización, Grupo y Rol.
- **Nombre de identidad.** Dependiendo del tipo seleccionado, la lista "Nombre de identidad" mostrará una lista diferente: la lista de roles, la lista de organizaciones o la lista de usuarios.
- **Tipo de permiso.** Esta lista depende también del tipo de objeto que se defina, ya que diferentes objetos pueden tener diferentes operaciones permitidas. Los permisos habituales son: leer, anotar, crear, editar, eliminar.

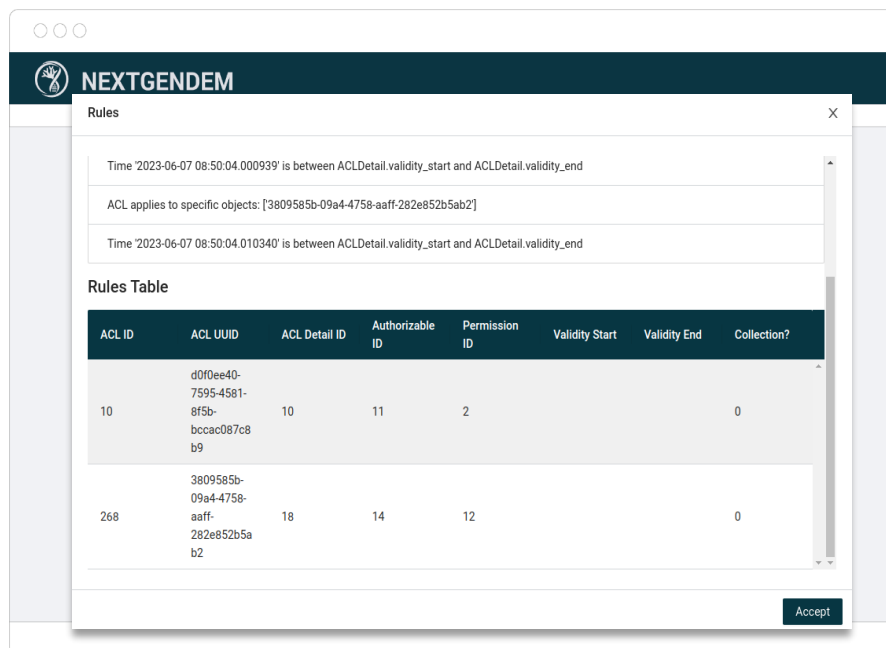


La lista de roles (cuando el tipo de identidad es rol).



La lista de tipos de permisos que se muestra al administrar capas geográficas.

Para ver los permisos que afectan a un objeto (objeto de referencia, colecciones y específicos), haga clic en el botón "Ver reglas", que abre el siguiente cuadro de diálogo:

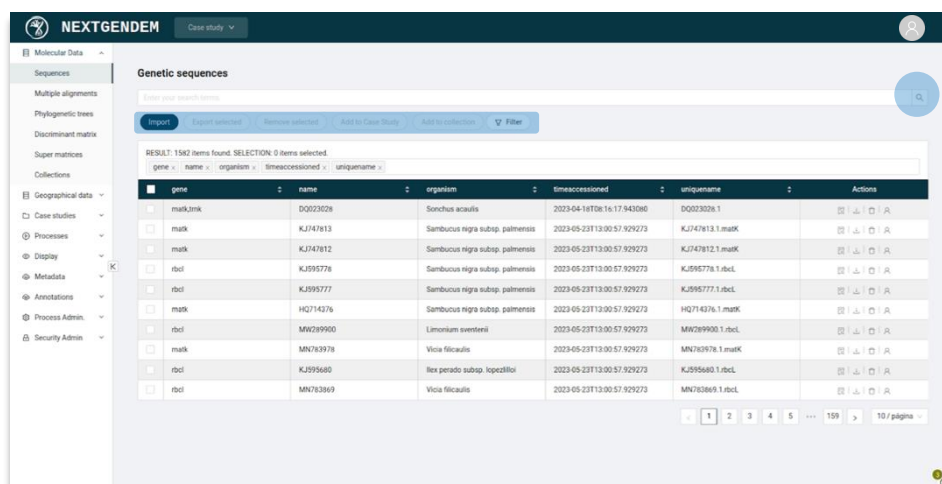


6.3 LISTAS-DETALLES

Para permitir que los usuarios vean y administren los datos almacenados en el sistema, hay dos vistas importantes, una vista de lista para enumerar todos los objetos de un tipo con los que el usuario puede operar y una vista detallada de un objeto para mostrar todos los datos sobre un objeto seleccionado de la lista, y modificarlo si es preciso.

- **VISTA DE LA LISTA**

La vista de lista es un importante componente de la interfaz gráfica de usuario que permite a los usuarios interactuar con los datos de forma estructurada. Dependiendo de la pantalla **de objetos de datos** concreta a la que se acceda (desde secuencias a datos geográficos), la vista de lista puede variar ligeramente en su configuración, pero mantiene un conjunto fijo de características para facilitar la exploración y manipulación de datos.



En la parte superior de la vista, los usuarios encontrarán una **barra de búsqueda** para realizar búsquedas **de texto completo** dentro de los datos mostrados, lo que facilita la localización de información específica de interés. Los usuarios pueden ingresar palabras clave o términos de búsqueda, y la tabla mostrará los resultados coincidentes.

Debajo de la barra de búsqueda, hay un grupo de botones que brindan algunas operaciones y opciones de navegación. Estos botones permiten a los usuarios realizar acciones como ir a la pantalla **de importar, exportar, eliminar objetos de datos seleccionados**, o incluir los seleccionados en estudios de casos o colecciones, y acceder a la **ventana de filtro de búsqueda**, que proporciona una forma alternativa de filtrar la lista. Esto es, utilizando atributos o características cuyos posibles valores se conocen, en lugar de texto libre, se facilita la obtención del resultado deseado.

La pieza central de la vista es la tabla paginada, que ocupa la mayor parte del espacio de vista. En el encabezado de la tabla, los usuarios pueden ver el recuento de resultados coincidentes y el recuento de filas seleccionadas, lo que brinda una descripción general y rápida de los datos mostrados. Además, hay disponible un selector de columnas que permite a los usuarios mostrar, ocultar y ordenar columnas según sus preferencias.

La tabla muestra los resultados de la búsqueda en un formato tabulado y en el que cada fila representa un registro distinto. La primera columna de cada fila incluye una casilla para facilitar la selección, lo que permite a los usuarios realizar acciones masivas en varias filas. Las columnas restantes muestran los campos de datos relevantes en función de las columnas seleccionadas, lo que proporciona una vista completa de los datos. La última columna de cada fila incluye pequeños botones o iconos que permiten a los usuarios realizar acciones individuales, como exportar, eliminar o editar permisos para una fila específica.

Para facilitar la navegación y la exploración de datos, la vista de tabla incluye un navegador de páginas en la esquina inferior derecha. Los usuarios pueden ajustar el tamaño de la página o navegar a través de diferentes páginas de datos, lo que permite una exploración eficiente de grandes conjuntos de datos.

Para opciones de filtrado más avanzadas, los usuarios pueden acceder a la ventana de filtro de búsqueda haciendo clic en el botón de filtro. La ventana de filtro aparece en el lado derecho de la interfaz y proporciona a los usuarios criterios de filtrado adicionales para refinar sus resultados de búsqueda. Este mecanismo de filtrado

basado en formularios permite a los usuarios establecer parámetros y condiciones específicas para restringir los datos mostrados.

Al combinar estas características, la vista de tabla permite a los usuarios interactuar y manipular datos de manera efectiva. Proporciona un conjunto integral de herramientas para buscar, filtrar, seleccionar y realizar diversas acciones en los datos mostrados, da la posibilidad a los usuarios de extraer información valiosa y realizar sus tareas de investigación con facilidad.

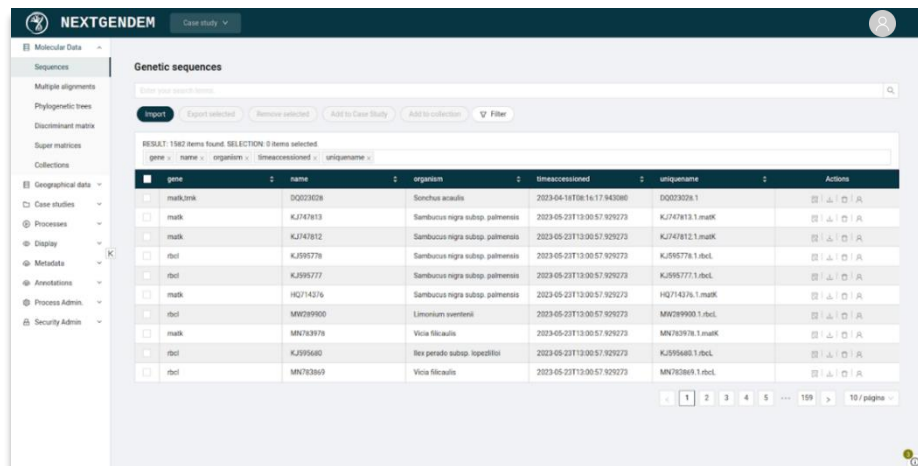
Es importante aclarar que la vista de tabla adapta sus características a los diferentes tipos de objetos de datos o funcionalidades: campos de filtro, la búsqueda de texto completo, columnas predeterminadas y columnas seleccionables, y algunos botones de acción (incluso algunos objetos tienen el botón de "ver directamente"), más la vista detallada que se abre al seleccionar un objeto de datos, que se describe a continuación.

- **VISTA ÚNICA DETALLADA**

La vista detallada única es un tipo de pantalla que proporciona una visión detallada de un elemento específico de interés dentro del sistema. Está diseñado para ofrecer información completa y para facilitar una mayor exploración e interacción con el elemento seleccionado.

En la parte superior de la vista detallada única, los usuarios encontrarán el nombre del elemento destacado como título. Esto permite a los usuarios identificar rápidamente el elemento que están viendo. Además, se proporciona un conveniente botón "Volver" en forma de flecha situado a la izquierda, concretamente a la izquierda del nombre, lo que permite a los usuarios volver fácilmente a la vista anterior.

Frente al título, ubicado a la derecha, hay un botón de exportación disponible si el elemento es exportable. Este botón permite a los usuarios exportar el elemento en un formato adecuado a sus necesidades, para brindar interoperabilidad con herramientas o sistemas externos.



El cuerpo principal de la vista detallada única está organizado en un conjunto de pestañas, cada una con un propósito específico y brindando información relevante relacionada con el elemento seleccionado.

La primera pestaña, a menudo etiquetada como "General", presenta un formulario que muestra varios campos que corresponden a las propiedades del elemento. Los

usuarios pueden ver los detalles del elemento, como sus atributos, características o metadatos. Dependiendo de los permisos del usuario, algunos campos pueden ser editables, permitiéndoles modificar ciertas propiedades directamente desde esta vista. Además, para ciertos campos, un botón puede estar disponible, proporcionando un enlace rápido a una vista más detallada específica de esa propiedad.

Además de la pestaña general, hay otras pestañas que enumeran objetos relacionados asociados con el elemento seleccionado. Por ejemplo, estas pestañas podrían incluir secciones para secuencias, individuos, análisis, organismos o cualquier otra entidad relevante. Al navegar por estas pestañas, los usuarios pueden explorar y acceder a información detallada sobre las relaciones del elemento y los datos asociados.

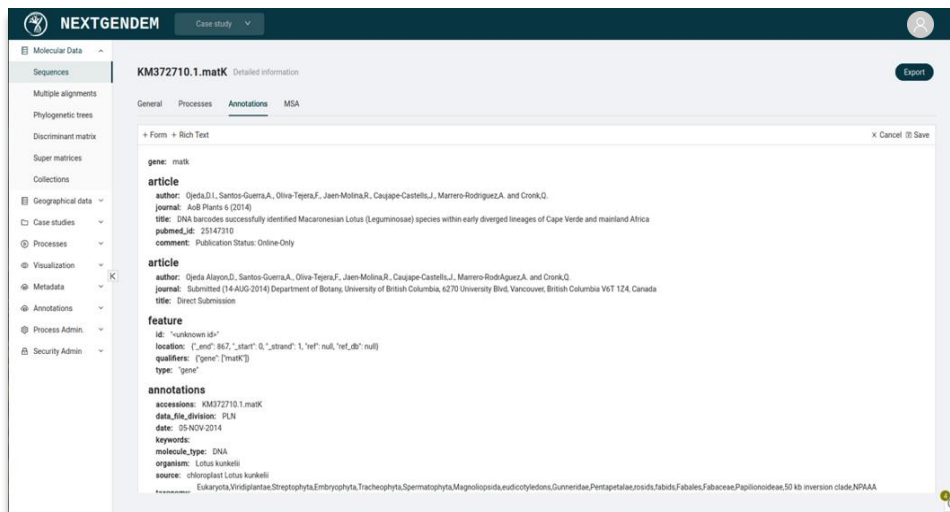
Otra pestaña importante dentro de la vista detallada única es la pestaña **Anotaciones**. Aquí, los usuarios pueden acceder y administrar las anotaciones asociadas con el elemento. Las anotaciones proporcionan contexto adicional, notas descriptivas o información definida por el usuario que puede mejorar la comprensión del elemento. Los usuarios pueden consultar y editar anotaciones dentro de esta pestaña, lo que permite la colaboración y la personalización. Al utilizar de manera efectiva la pestaña de anotaciones, los usuarios pueden acceder, crear, editar y eliminar anotaciones, lo que permite una mejor documentación e intercambio de conocimientos dentro de nuestra plataforma. El editor de anotaciones se comenta en la siguiente sección.

Además, si corresponde y está disponible para el tipo de objeto específico, la vista detallada única puede incluir pestañas adicionales que ofrecen herramientas de visualización. Estas pestañas permiten a los usuarios interactuar con representaciones visuales o visualizaciones gráficas relacionadas con el artículo. Las herramientas de visualización pueden proporcionar información valiosa y ayudar en el análisis o la interpretación de datos, según la naturaleza del elemento que se esté visualizando.

En general, la vista detallada única brinda a los usuarios una perspectiva integral y enfocada en un elemento específico dentro del sistema. Al organizar la información en pestañas y proporcionar campos editables, objetos relacionados, anotaciones y herramientas de visualización, esta vista permite a los usuarios explorar, comprender e interactuar con el elemento seleccionado de manera estructurada y eficiente.

6.4 EDITOR DE ANOTACIONES

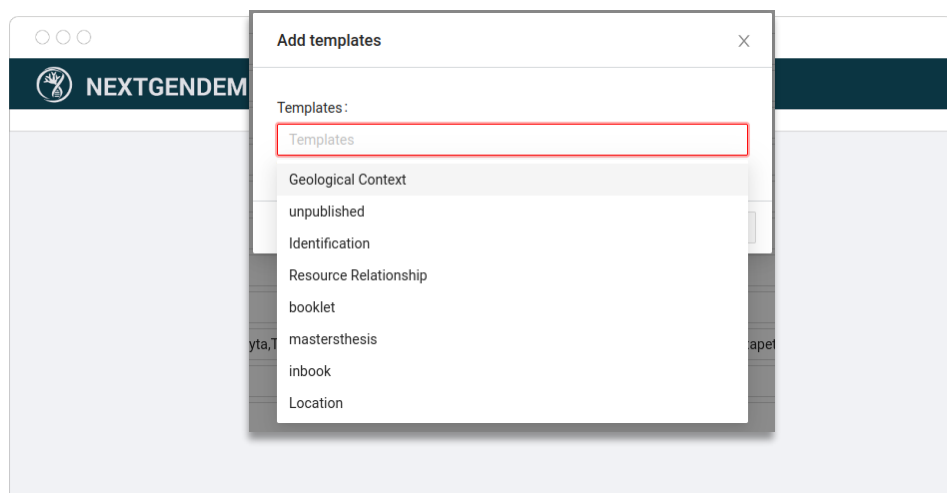
Muchas de las pantallas de detalles individuales tienen una pestaña donde se pueden administrar las anotaciones sobre el objeto actual. Estas anotaciones son piezas muy valiosas de información que proporciona contexto adicional, detalles o notas definidas por el usuario y relacionadas con el elemento. Aquí se muestran en una lista donde se pueden agregar, editar y eliminar.



Con los permisos apropiados, el usuario puede agregar nuevas anotaciones al elemento. Esto implica hacer clic en uno de los botones en la parte superior izquierda de la lista según el tipo de anotación de destino. El editor permite añadir una lista de anotaciones de tres tipos diferentes:

- **Campo estructurado único** (creado por los administradores, menú Anotaciones -> submenú Campos).
- **Plantilla creada previamente** (creada por administradores, menú Anotaciones -> submenú Plantillas). Son formularios formados por un conjunto de campos. Algunos de los valores predeterminados son plantillas BibTeX y plantillas Darwin Core.
- **Cuadro de texto enriquecido.**

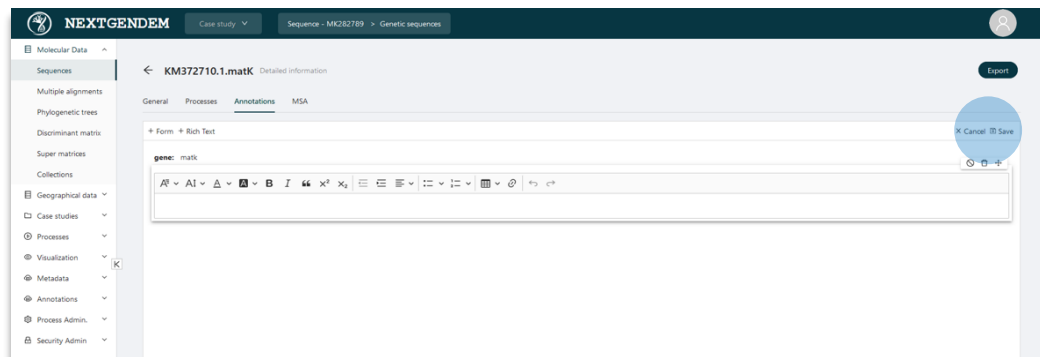
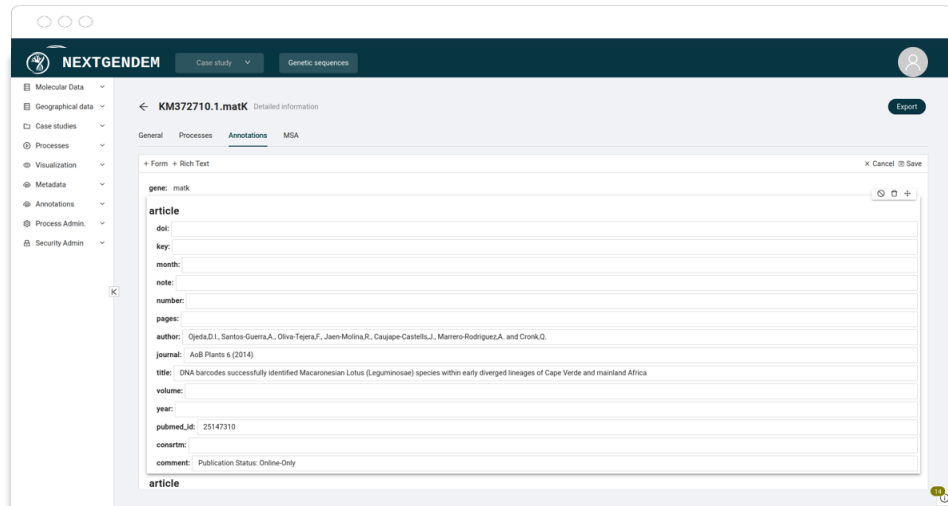
En primer lugar, aparecerá un cuadro de diálogo que permite al usuario elegir la plantilla o el campo deseado. Después de seleccionar uno, aparecerá en la parte inferior de la lista para completar.



Además de agregar nuevas entradas, las anotaciones existentes pueden ser editadas o eliminadas. La forma de acceder a cualquiera de esas acciones es hacer clic en la anotación para trabajar en ella. Esto habilitará tres botones en forma de icono, en la parte superior

derecha de la anotación visualizada: un lápiz para iniciar el modo de edición, una papelera para eliminar el enlace entre el objeto y la anotación y una cruz para cambiar su posición en la lista.

Seleccionando el botón del lápiz o doble clic sobre un elemento de anotación aparecerán todas las posibilidades de modificación. Cabe mencionar que los campos vacíos solo se muestran mientras se edita la anotación.



Finalmente, es necesario guardar los cambios haciendo clic en el botón “Guardar” en la parte superior derecha de la lista, o deshacer todos los cambios no guardados con el botón “Cancelar”.

6.5 BÚSQUEDAS Y FILTROS

La búsqueda y el filtrado de datos es una funcionalidad crucial dentro de la plataforma, ya que permite a los usuarios encontrar información relevante basada en criterios específicos (a diferencia de la búsqueda de texto completo, donde los usuarios proponen texto que está presente en los objetos de datos que se buscan).

Para iniciar una búsqueda, los usuarios pueden utilizar la barra del buscador ubicada en la parte superior del navegador. Esta barra de entrada permite a los usuarios realizar búsquedas de texto completo en los datos disponibles, en los propios datos y en las

anotaciones relacionadas. Con el uso de palabras clave o frases relevantes, los usuarios pueden recuperar resultados coincidentes, lo que ayuda a los usuarios a ir refinando sus consultas y obtener comentarios casi inmediatos.

The screenshot shows the NEXTGENDEM web interface. At the top, there is a search bar with the text 'micromeria matk'. Below the search bar, there are several action buttons: 'Import', 'Export selected', 'Remove selected', 'Add to Case Study', 'Add to collection', and 'Filter'. The main content area displays a table with the following data:

checkbox	uniqueName	timeaccessioned	organism	gene	actions
<input type="checkbox"/>	AY840178.1.matK	2023-04-05T07:30:06.017850	Micromeria pineolens	matk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	MN783937.1	2023-04-18T08:16:17.943080	Micromeria glomerata	matk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	DQ640416.1	2023-04-18T08:16:17.943080	Micromeria leucantha	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	DQ640411.1	2023-04-18T08:16:17.943080	Micromeria leucantha	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	AY840178.1	2023-03-30T11:31:49.606117	Micromeria pineolens	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	DQ640412.1	2023-03-30T11:31:49.606117	Micromeria pineolens	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	DQ640413.1	2023-03-30T11:31:49.606117	Micromeria pineolens	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	AY840178.1.trnk	2023-05-24T08:41:07.553952	Micromeria pineolens	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	DQ640413.1.trnk	2023-05-24T08:41:07.553952	Micromeria pineolens	matk.trnk	[icon] [icon] [icon] [icon]
<input type="checkbox"/>	DQ640412.1.trnk	2023-05-24T08:41:07.553952	Micromeria pineolens	matk.trnk	[icon] [icon] [icon] [icon]

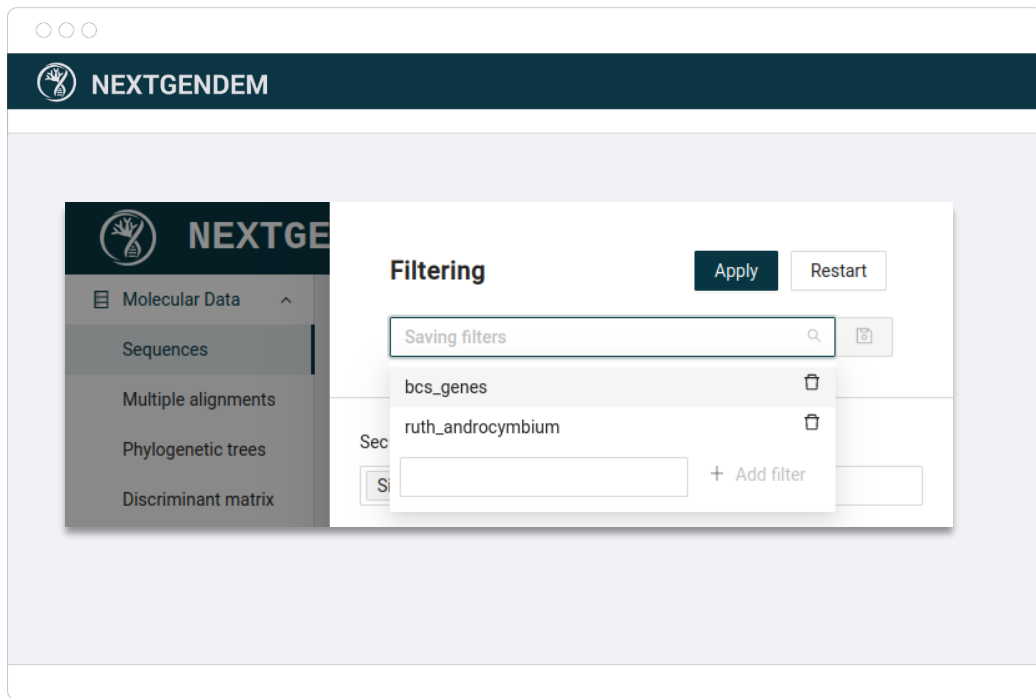
At the bottom of the table, there is a pagination control showing '10 / página' and navigation arrows.

El encabezado de la tabla proporciona información valiosa sobre los resultados de la búsqueda, como el recuento de registros coincidentes y el recuento de filas seleccionadas. Los usuarios también pueden utilizar el selector de columnas para mostrar, ocultar u ordenar columnas según sus preferencias, asegurando que los datos mostrados se adapten a sus necesidades.

Cada fila dentro de la tabla contiene una casilla de selección, lo que permite a los usuarios realizar acciones masivas en varios registros si es necesario. Además, las acciones individuales se pueden realizar en filas específicas utilizando los botones de iconos pequeños presentes en la última columna. Estas acciones pueden incluir exportar, eliminar o editar los permisos de las filas seleccionadas.

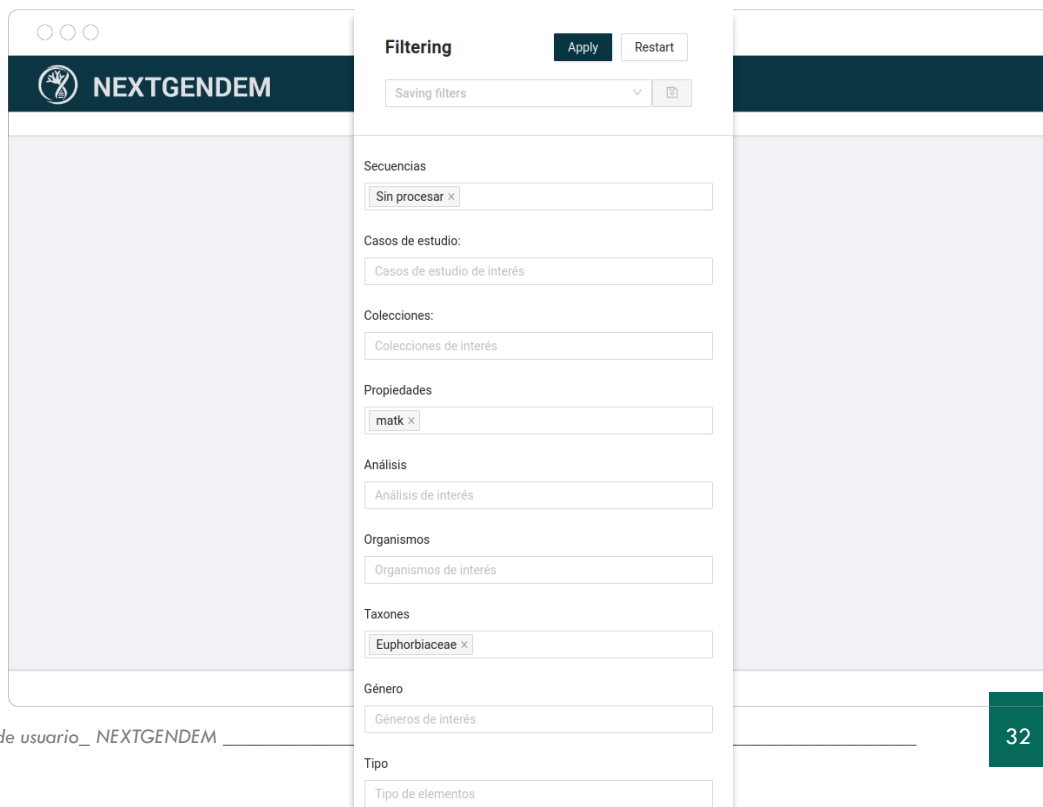
Para facilitar la navegación a través de grandes conjuntos de datos, hay disponible un navegador de páginas en la parte inferior derecha de la tabla. Los usuarios pueden ajustar el tamaño de la página o cambiar entre diferentes páginas para explorar de manera eficiente los resultados de la búsqueda.

Además de las búsquedas de texto completo, los usuarios pueden aplicar filtros para reducir aún más los resultados de la búsqueda. Al hacer clic en el botón "Filtro", aparecerá una ventana de filtro de búsqueda. Esta ventana presenta un formulario con varias opciones de filtrado adaptadas a la sección de datos o contexto específico.



El encabezado de la ventana del filtro de búsqueda se centra en la gestión del almacén de filtros. En la esquina superior derecha se pueden encontrar dos botones, uno de “Restablecer” para reiniciar los valores del formulario de filtro por defecto, y uno de “Aplicar” para realizar el filtrado por los criterios definidos. Incluye un selector de entrada para elegir una configuración de filtro existente o crear una nueva.

Dentro de la ventana de filtro de búsqueda, los usuarios pueden establecer criterios y restricciones específicas para refinar su búsqueda. Estas opciones de filtrado pueden incluir atributos, categorías, rangos de fechas o cualquier otro parámetro relevante según la naturaleza de los datos. Los usuarios pueden seleccionar varios filtros y personalizar sus combinaciones para crear consultas de búsquedas complejas.



Una vez que se establecen los filtros deseados, los usuarios pueden enviar la consulta de búsqueda, lo que hace que el sistema procese la solicitud y recupere los resultados coincidentes. La tabla paginada dentro de la vista del componente principal luego mostrará los resultados filtrados, presentando los datos relevantes en un formato estructurado y organizado.

En general, la funcionalidad de búsqueda y filtrado de la interfaz gráfica de usuario permite a los usuarios ubicar y analizar rápidamente datos específicos según sus criterios. Al combinar búsquedas de texto completo, filtros personalizables y vistas de tablas interactivas, los usuarios pueden navegar de manera eficiente a través de extensos conjuntos de datos y recuperar la información que necesitan, mejorando en última instancia la productividad y la toma de decisiones dentro de la plataforma.

6.6 NAVEGACIÓN

Una consideración importante de las interfaces de usuario es cómo los usuarios pueden moverse de una pantalla a otra, es decir, cómo se interconectan las pantallas.

En el caso del sistema NEXTGENDEM, el enfoque es conectar dos pantallas cuando en una de las pantallas hay un elemento de información que puede expandirse o verse de manera diferente en la otra pantalla, lo que permite una exploración más flexible que la organización jerárquica habitual. Debido a que se necesitan puntos de navegación iniciales, se proporcionan dos formas de navegación comúnmente utilizadas:

Menú lateral. El menú lateral (izquierdo) permite llegar a la mayor parte de la pantalla que muestra una lista de objetos de datos (datos moleculares y geográficos), además de las herramientas de Procesamiento y Visualización.

Miga de pan. Para facilitar el regreso a las pantallas visitadas anteriormente, el área superior muestra el recorrido del usuario.

6.7 PERSISTENCIA DE PANTALLAS

La interfaz de usuario tiene un mecanismo (transparente para los usuarios) para guardar el estado de las pantallas, por lo que es posible navegar a otra pantalla y volver a la anterior sin perder los cambios allí realizados. El estado también se guarda, por lo que la sesión de trabajo puede continuar en otro momento.

6.8 APERTURA CON IMPORTACIÓN – PROCESO - EXPORTACIÓN - REIMPORTACIÓN

El subsistema de archivo está preparado para importar, exportar y reimportar (sobrescribir) objetos de datos que puedan haber sido exportados, especificando “sobrescribir”. Este sencillo mecanismo permite combinar el sistema NEXTGENDEM con herramientas externas no integradas directamente.

6.9 CONVERSIONES DE FORMATO

Otro uso secundario del sistema es la capacidad de traducir datos de un formato a otro. Actualmente, esto implica importar a la base de datos, exportar usando el formato deseado y luego tal vez eliminar la información si no se pretende conservarla.

6.10 VERSIONADO

Hay dos mecanismos que admiten el control de versiones:

- Los registros de "encabezado" del objeto de datos (la pestaña "detalle" en las pantallas de Detalle único) se mantienen automáticamente.
- Conceptualmente el mismo conjunto de datos, pero evolucionando u obtenido con diferentes procesos.

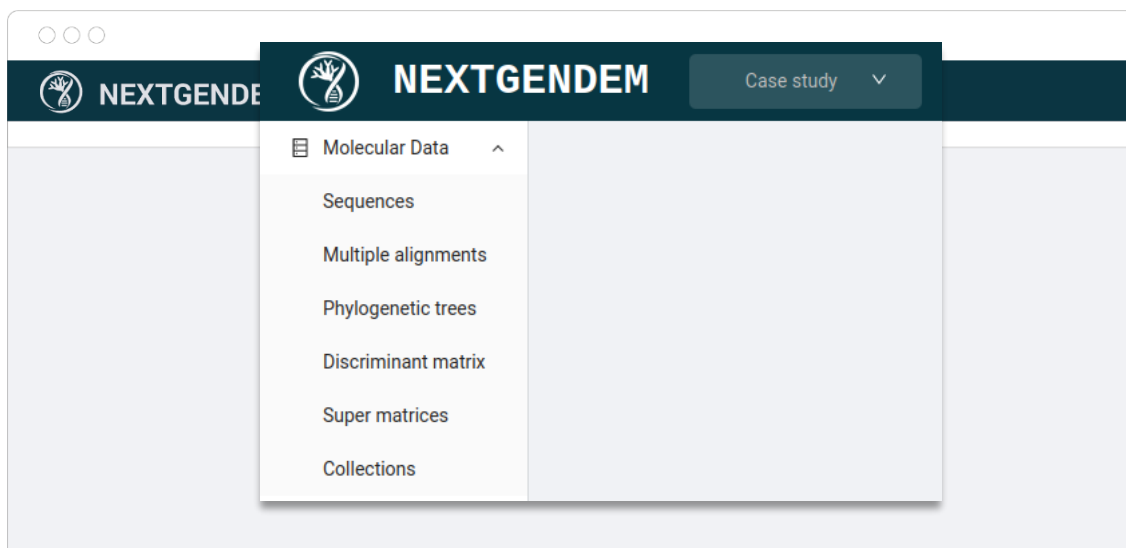
GESTIÓN DE DATOS BIOINFORMÁTICOS



NEXTGENDEM

7. GESTIÓN DE DATOS BIOINFORMÁTICOS

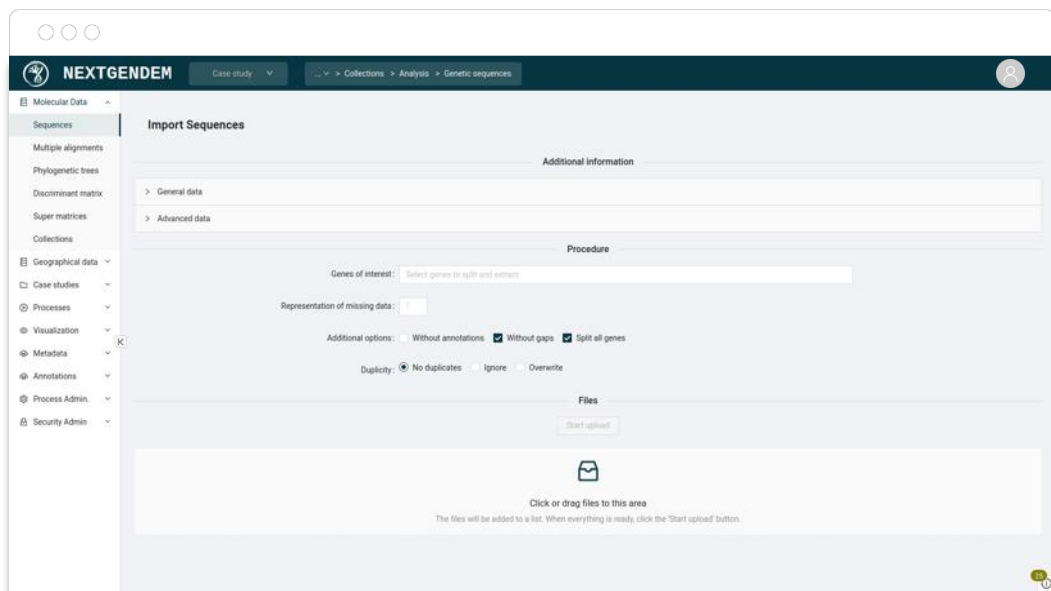
La gestión de datos bioinformáticos se erige como un pilar crucial del sistema, asegurando el manejo efectivo y la accesibilidad de los datos bioinformáticos en múltiples formatos. Estos datos sirven como un recurso vital para estudios de expertos y requieren una importación o creación adecuada para establecer relaciones y conexiones significativas dentro del sistema. Al administrar y organizar los datos, permitimos que el sistema libere todo su potencial para respaldar los procesos bioinformáticos.



7.1 IMPORTAR DATOS MOLECULARES

Cuando están autorizados a insertar y crear datos moleculares, los usuarios pueden aprovechar dos vistas distintas para el proceso de importación. La primera vista facilita la importación de secuencias de nucleótidos, mientras que la segunda vista permite la importación de análisis relacionados. Estas vistas siguen un patrón consistente y se puede acceder a ellas a través del botón de acción principal del navegador respectivo.

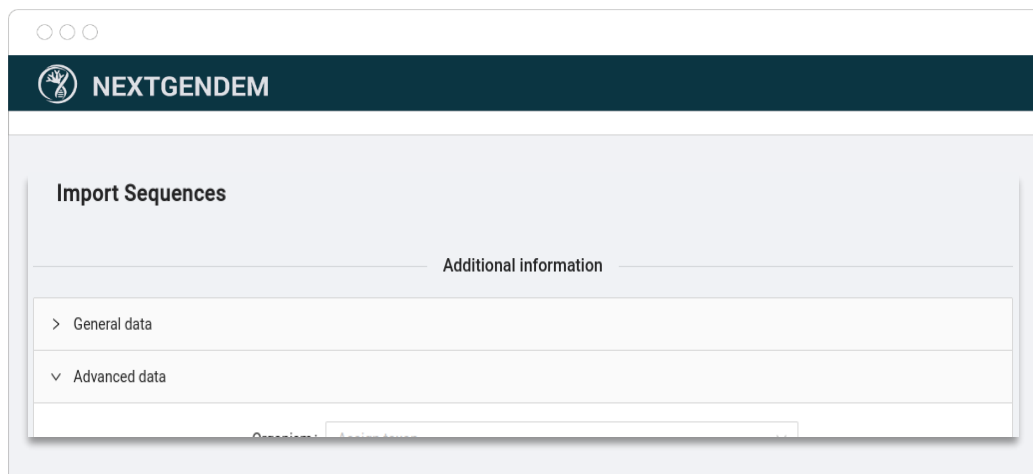
Dentro de las vistas de importación, los usuarios tienen acceso a una sección de “Información adicional”, que les permite especificar detalles que no están presentes en el archivo importado. El cuadro “Datos generales” proporciona selecciones de entrada para asociar los datos importados con estudios de casos relevantes o colecciones de interés. Mientras tanto, el cuadro “Datos avanzados” ofrece información más específica, como las regiones moleculares.



Para ofrecer un mejor control y personalización, la sección 'Procedimiento' agrupa las opciones de entrada, lo que permite a los usuarios definir parámetros y preferencias específicas para el proceso de importación. Además, la sección 'Archivos' permite a los usuarios arrastrar y soltar cómodamente uno o más archivos de secuencia y un solo archivo para cualquier análisis molecular. Esto agiliza el proceso de importación de datos y garantiza la integración dentro del sistema.

7.2 IMPORTAR SECUENCIAS DE NUCLEÓTIDOS

Entre las vistas mencionadas anteriormente hay algunos matices a tener en cuenta. En primer lugar, hay cierta información esencial que proporcionar. Los taxones y regiones moleculares son casi requeridos para cualquier consulta, estudio o ejecución que cualquier persona pueda solicitar sobre este tema, por lo que pueden especificarse en los archivos o seleccionarse en los "Datos avanzados". En caso de ser seleccionado, afectará a todos los datos importados, y como la lógica sugiere, las secuencias pueden ser uno o más genes, pero solo un organismo.



Para aquellos usuarios que deseen proporcionar información adicional dentro de los archivos, se recomienda utilizar los formatos de archivo FASTA o GENBANK como los ejemplos a continuación. Las palabras en **negrita** en los ejemplos señalan la información esencial: **identificador único de secuencia**, **organismo** y **genes**. Cualquier otra información se almacenará tanto como sea posible, al menos como anotaciones.

- GENBANK

```

LOCUS      KM372710 867 bp   DNA       linear   PLN 05-NOV-2014
DEFINITION .
ACCESSION KM372710
VERSION   KM372710.1
KEYWORDS  .
SOURCE    .
  ORGANISM Lotus kunkelii
REFERENCE 1 (bases 1 to 867)
  AUTHORS .
  TITLE   .
  JOURNAL .
FEATURES             Location/Qualifiers
     source           1..867
                     /organism="Lotus kunkelii"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:0"
     gene            <1..>867
                     /gene="matK"
ORIGIN
     1 atgcgatgcg atgctagctc gatcgatoga tcgatcgatc gatcgatoga
tcgatcgatc
     [...]
     841 acttcggctt tcatgtatta aaacttt
//

```

- FAST

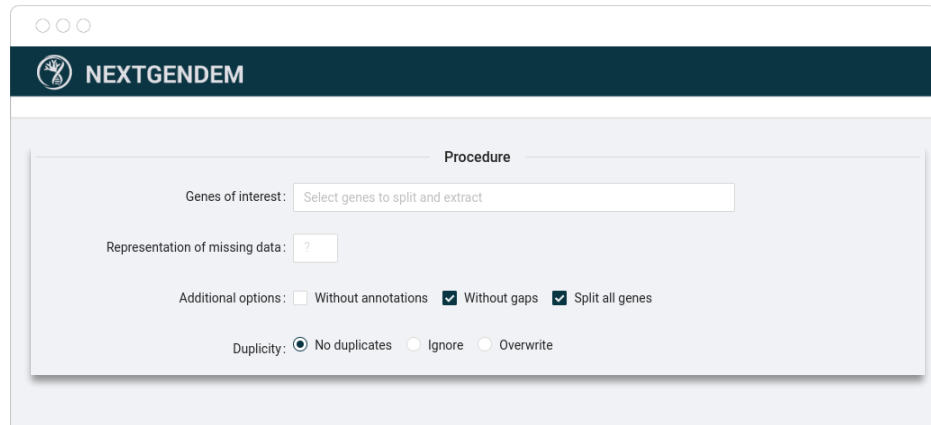
```

>KM372710.1 [organism=Lotus kunkelii] maturase K (matK) gene, partial cds;
chloroplast
CCCTTATCCTATCCATCTGGAAATCTTAGTCCAAATCCTTCGATACTGGGTGAAAGATGCCCCTTTGTTT
[...]
ACTTCGGCTTTCATGTATTAACCTTT

```

En las opciones del trámite, el usuario encontrará:

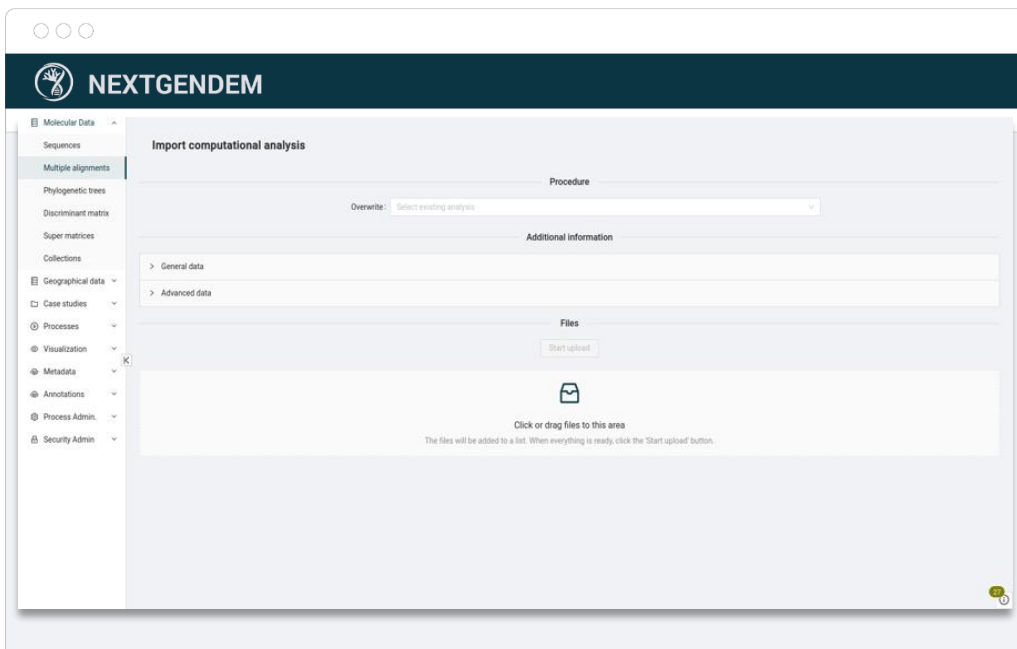
- Una selección para elegir los genes relevantes. Esto le dice al sistema que extraiga esas regiones moleculares específicas, siempre que estén debidamente marcadas. Una función relevante para tratar genomas completos y evitar la saturación del sistema.
- Una entrada para especificar el carácter que representa los valores faltantes. Si no es así, éstos serán reemplazados por razones de compatibilidad.
- Un grupo de casillas de selección:
 - **“Sin anotación”** para ignorar la información adicional en los archivos, perfecto para importar solo las secuencias de nucleótidos, tal vez para actualizarlas.
 - **“Sin gaps”** para eliminar todos los '-' de las secuencias de nucleótidos.
 - **“Dividir todos los genes”** si están debidamente marcados, sean los que sean.
- Un grupo de botones de selección para establecer la política de duplicados. En el orden mostrado son: que de fallo si se duplica, ignórelo o sobrescribalo.



7.3 IMPORTAR ANÁLISIS MOLECULAR

Esta funcionalidad facilita la importación de análisis derivados de secuencias de nucleótidos, como alineamiento múltiple de secuencias, árboles filogenéticos, matrices discriminantes y supermatrices. Se puede acceder a través del botón de acción principal del menú lateral izquierdo. No obstante, aunque se sigue la distribución descrita anteriormente, existen algunas diferencias.

En este caso, la sección de procedimiento solo contiene una selección para elegir un análisis existente en casos de sobreescritura, lo que permite a los usuarios actualizar externamente.



En “*datos generales de información adicional*” hay una selección extra “*Deriva de*” donde el usuario puede seleccionar uno o más análisis de los cuales deriva el nuevo análisis. Una situación en la que esto podría ser útil es importar el resultado de un análisis externo con un análisis interno como entrada.

Additional information

▼ General data

Derives from:

Case studies:

Collections:

> Advanced data

En los “datos avanzados de información adicional” está la misma selección de “genes”, pero un formulario de cinco campos opcionales a considerar. Estos están abiertos a criterio del usuario, pero aquí hay algunas sugerencias:

- **Nombre:** un nombre representativo para facilitar la identificación.
- **Programa:** el nombre del programa como *MUSCLE*, *MRBAYES*, *MOLD* o desconocido.
- **Versión:** la versión del programa suele ser un número como *v1.0* o *3.2.7α*
- **Algoritmo:** el tipo de proceso para obtener este resultado de análisis.
- **Descripción:** el lugar para más detalles.

Additional information

> General data

▼ Advanced data

Genes:

Name:

Program:

Version:

Algorithm:

Description:

7.4 CONSULTAR DATOS MOLECULARES

La plataforma proporciona una interfaz integral para consultar datos moleculares de manera eficiente. Los usuarios pueden navegar fácilmente a través de varias vistas y tablas de datos, utilizando funciones de búsqueda y filtros para reducir sus consultas. La vista de tabla presenta

los datos en un formato estructurado, lo que permite a los usuarios explorar información relevante, como secuencias de nucleótidos, análisis y anotaciones. Con la capacidad de personalizar la visibilidad y el orden de las columnas, los usuarios pueden concentrarse en los campos de datos específicos de interés. La función de consulta permite a los usuarios acceder y recuperar los datos moleculares necesarios de forma rápida y cómoda. (ver el apartado '6.5 – Búsquedas y filtros', para más información)

7.5 EDITAR DATOS MOLECULARES

Para garantizar la precisión y flexibilidad de los datos, la plataforma ofrece una función de edición para datos moleculares. Los usuarios con los permisos adecuados pueden modificar campos y propiedades específicos de los objetos de datos. La funcionalidad de edición proporciona una interfaz fácil de usar, lo que permite a los usuarios actualizar y refinar la información sin esfuerzo. Ya sea que se trate de modificar los detalles de la secuencia, ajustar los parámetros de análisis o actualizar las anotaciones, la función de edición agiliza el proceso y garantiza que los datos permanezcan actualizados y precisos.

7.6 ELIMINAR DATOS MOLECULARES

Cuando se hace necesaria la eliminación de datos, la plataforma ofrece un enfoque seguro y controlado para eliminar datos moleculares. Los usuarios con los permisos necesarios pueden eliminar selectivamente objetos de datos individuales u optar por la eliminación grupal, eliminando varios registros simultáneamente. La plataforma incorpora medidas de seguridad para evitar eliminaciones accidentales, por lo que requiere la confirmación del usuario antes de eliminar los datos de forma permanente. Esto garantiza que el proceso de eliminación sea deliberado y minimiza el riesgo de pérdida de datos no intencional. La funcionalidad de eliminación proporciona un método seguro para administrar y mantener la integridad de los datos dentro del sistema.

7.7 EXPORTAR DATOS MOLECULARES

Al permitir la interoperabilidad y el intercambio de datos, la plataforma incluye una función de exportación robusta para datos moleculares. Los usuarios pueden extraer cómodamente objetos de datos seleccionados o conjuntos de datos completos en varios formatos, lo que facilita la compatibilidad con herramientas y aplicaciones externas. La funcionalidad de exportación admite formatos de archivo comunes como FASTA, CSV y JSON, lo que garantiza la compatibilidad con diferentes flujos de trabajo de bioinformática. Al proporcionar opciones de exportación flexibles, la plataforma permite a los usuarios compartir datos moleculares con colaboradores, realizar análisis adicionales o integrarlos con sistemas externos, maximizando la usabilidad y accesibilidad de los datos.

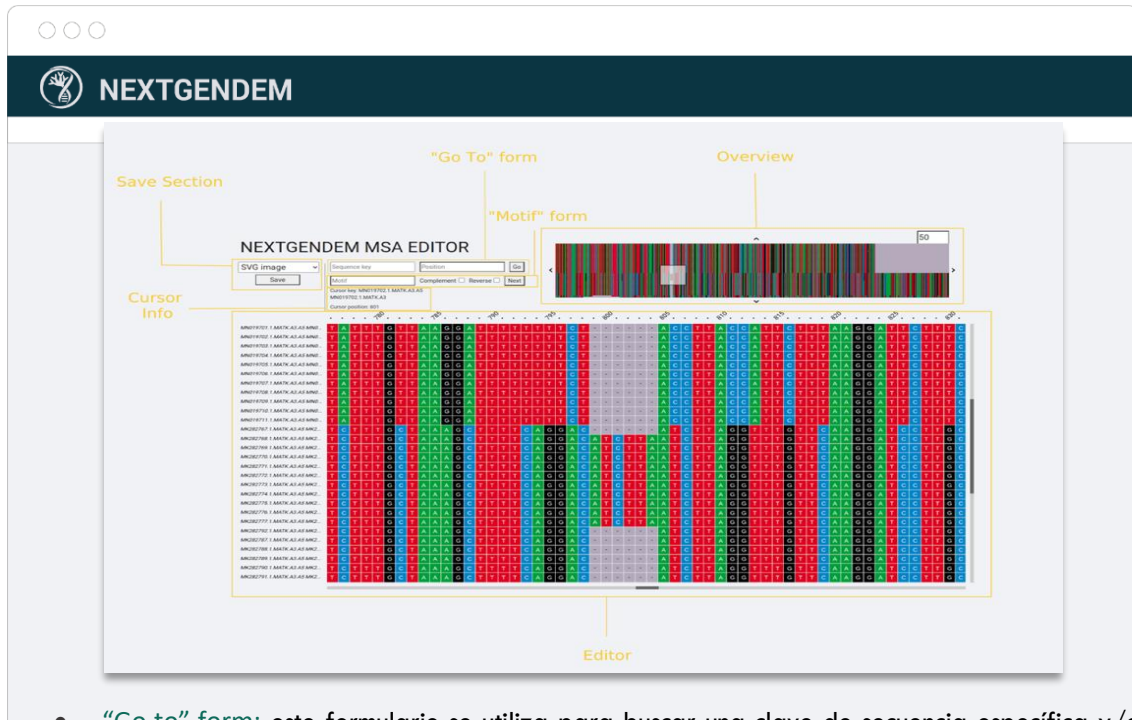
EDITOR DE ALINEAMIENTO MÚLTIPLE DE SECUENCIAS



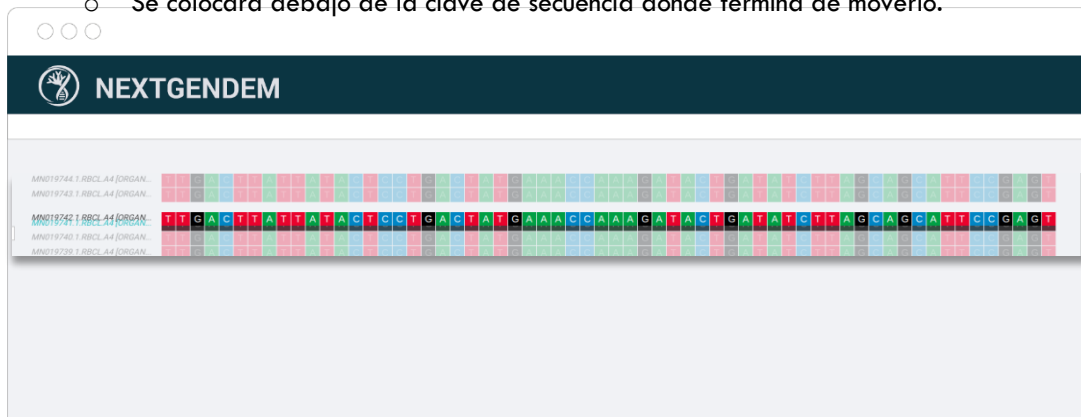
NEXTGENDEM

8. EDITOR DE ALINEAMIENTO MÚLTIPLE DE SECUENCIAS

El editor es un conjunto de herramientas para visualizar y editar correctamente un alineamiento de secuencias. Tiene la funcionalidad típica de este tipo de editores como MEGA o Bioedit pero enfocado únicamente al ADN. Aquí hay una captura de pantalla con las partes principales del editor señaladas en amarillo.



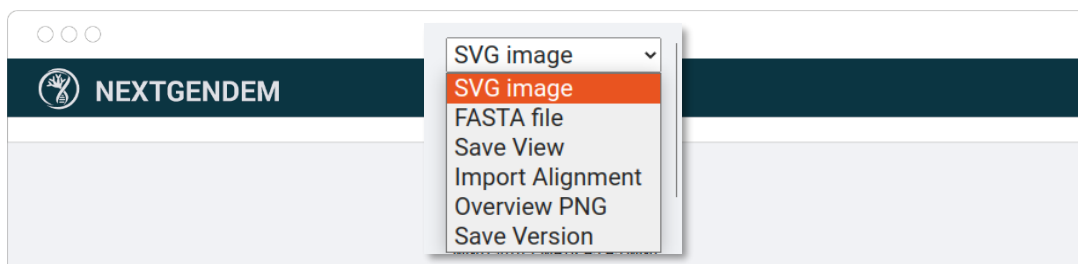
- **“Go to” form:** este formulario se utiliza para buscar una clave de secuencia específica y/o posición de la secuencia en el editor.
 - Inserte la clave de secuencia, la posición o ambas.
 - Presione el botón “Ir”.
- **“Motif” form:** este formulario se utiliza para buscar un patrón de ADN.
 - Inserte el Motivo, por ejemplo “ACCGT”.
 - Es posible buscar su complementario, inverso, complementario inverso (ambos marcados) o sin modificación (ni inverso, ni complemento marcado).
 - Una vez que se haya rellenado todo, presione el botón “Siguiente”.
- **Cursor Info:**
 - La clave del cursor y la posición se actualizan automáticamente para coincidir con el nucleótido al que apunta.
- **Reordenación de secuencias:**
 - Apunte a una clave de secuencia en el editor y arrástrela para moverla.
 - Se colocará debajo de la clave de secuencia donde termina de moverlo.



- **Overview:** la descripción general es una vista en miniatura del alineamiento completo o de una parte de ella, si tiene más de aproximadamente 150 secuencias o una longitud de secuencia de más de 800 nucleótidos. Estos máximos pueden cambiar en función del tamaño de la pantalla donde se muestra.
 - Cuando el alineamiento excede los límites que le permiten ver el *overview*, es posible mostrar el resto del alineamiento haciendo clic en las flechas que lo rodean. El número del cuadro en la esquina superior derecha es el número de nucleótidos que se moverá el *overview* al hacer clic en alguna de la flecha.
 - El cuadrado blanco es un *componente* móvil que apunta a la zona del alineamiento que muestra el editor. Si lo mueve en el *overview*, el editor cambiará automáticamente a la nueva zona señalada.
- **Barras laterales:**
 - El usuario puede moverse a diferentes partes de los alineamientos usando las barras laterales del editor en las partes inferior y derecha.
- **Save Section:** En esta sección el usuario puede importar los alineamientos editados al sistema o descargar algunos archivos. Una vez presionado el botón “Guardar” el usuario debe insertar el nombre del objeto a importar o descargar.

Estas son las opciones disponibles:

- **Imagen SVG:** descarga la vista del editor en formato SVG.
- **Fichero Fasta:** Descarga un archivo Fasta con el alineamiento editado.
- **Guardar Vista:** Importa una nueva Vista al sistema.
- **Importar alineamiento:** importa el alineamiento editado al sistema.
- **Overview PNG:** Descarga un archivo en png con la imagen.
- **Guardar versión:** guarda una nueva versión del alineamiento editado



- Editor:
 - Cuando hace clic en una celda, se selecciona. Si ya estaba seleccionada, será deseleccionada.
 - Puede seleccionar varias celdas usando **Ctrl+clic**, y lo mismo se aplica a la anulación de la selección.
 - También puede seleccionar varias celdas manteniendo presionada la tecla **Ctrl** y haciendo clic con el botón izquierdo del ratón y luego arrastrando el ratón. Lo mismo se aplica a la desección.
 - Puede seleccionar varias celdas (incluido un bloque) haciendo clic en un nucleótido y usando **Shift+clic** en el segundo.
 - Si tiene algunas celdas seleccionadas, puede deseccionarlas haciendo clic en otra celda que no esté seleccionada.
 - Al pulsar la tecla de borrar (**<--**) se borran los nucleótidos seleccionados.
 - Puede modificar el nucleótido en cada celda e incluso puede ingresar caracteres alternativos como '?'.
 ○ Al presionar la **barra espaciadora** se insertan gaps. Si tiene varios nucleótidos seleccionados usando **Shift**, los espacios no se insertarán entre cada nucleótido, sino que tendrán una longitud igual al número de nucleótidos consecutivos en la selección. Sin embargo, si los selecciona con **Ctrl**, los espacios se insertarán entre cada nucleótido.
 - Presionar **Ctrl+Z** deshace la última acción.
 - Presionar **Ctrl+Shift+Z** rehace la última acción deshecha.
 - Haciendo clic derecho sobre cualquier celda se abre un menú contextual con las siguientes opciones:
 - **Eliminar columnas:** eliminar columnas de celdas seleccionadas.
 - **Inverso:** calcula el inverso de las celdas seleccionadas.
 - **Complementario:** calcula el complementario de las celdas seleccionadas.
 - **Complementario Inverso:** calcula el complementario inverso de las celdas seleccionadas.
- Requisitos:
 - Se requiere que todas las claves de secuencia sean únicas.



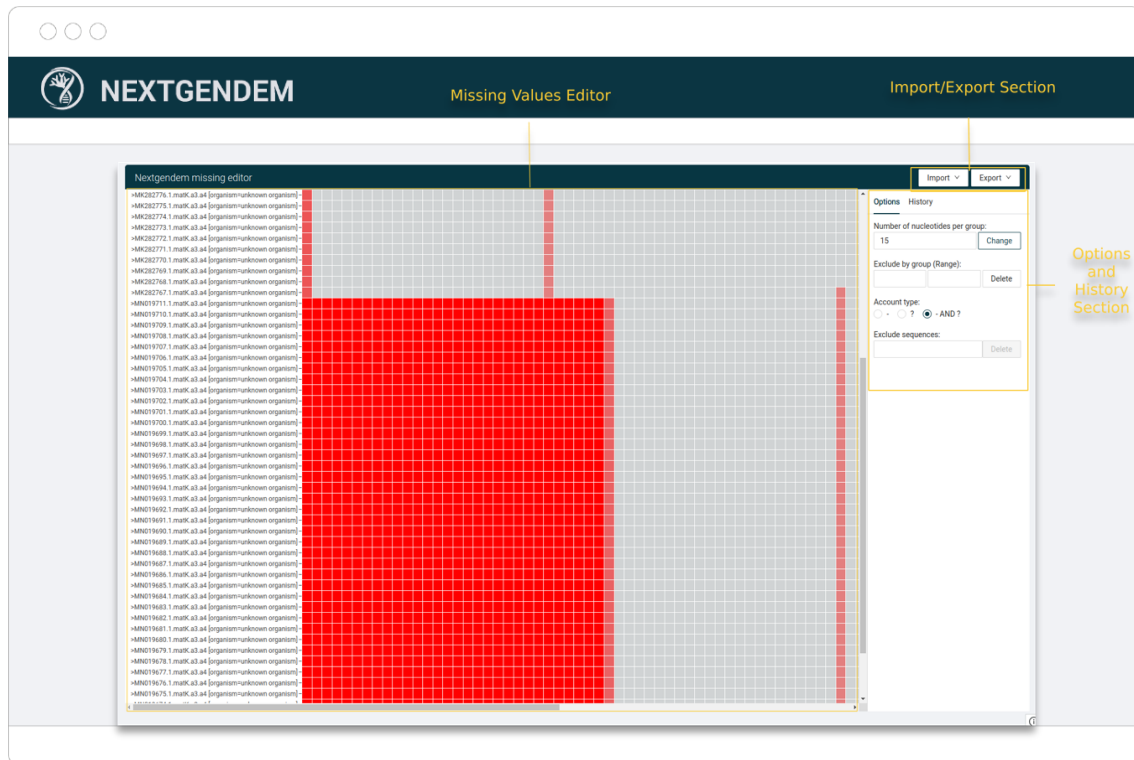
EDITOR DE DATOS DE
FALTANTES EN ALINEAMIENTO



NEXTGENDEM

9. EDITOR DE DATOS FALTANTES EN ALINEAMIENTO

El editor de datos faltantes en un alineamiento es un conjunto de herramientas para visualizar y editar correctamente los *gaps* y los valores faltantes en un alineamiento múltiple de secuencias. Está diseñado para mostrar un resumen de la distribución de *gaps* y valores faltantes, para eliminar correctamente columnas o secuencias según el criterio de los expertos. Se compone de tres secciones diferentes.



9.1 MISSING VALUES EDITOR

Esta es la vista principal del editor, un mapa de calor donde las filas son secuencias nombradas por su id y los colores representan la cantidad de valores faltantes y/o *gaps* en el alineamiento múltiple de secuencias. Es posible obtener información sobre cada celda señalándolas con el cursor.

9.2 OPTIONS AND HISTORY SECTION

En esta sección, el usuario puede personalizar la vista y editar el alineamiento múltiple de secuencias:

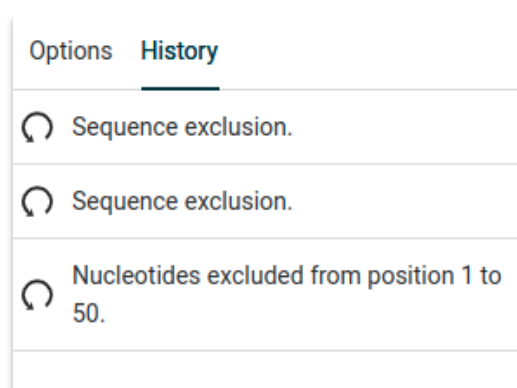
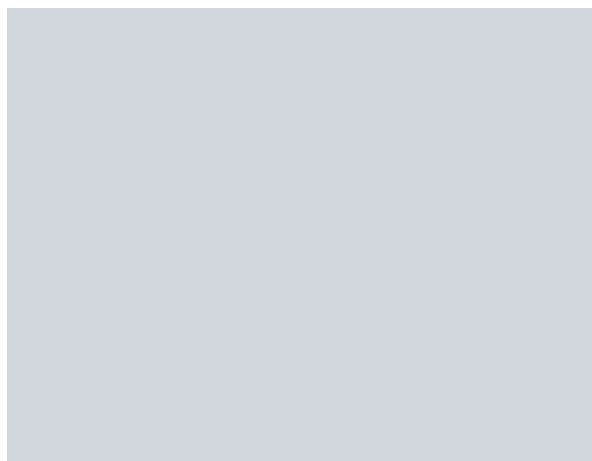
- **Número de nucleótidos por grupo:** cambia el número de sitios representados en cada celda del mapa de calor.
- **Eliminar por grupo (rango):** elimina el rango de columnas de alineamiento múltiple de secuencias.
- **Tipo de cuenta:** la intensidad del color se puede calcular contando solo '-' (*gaps*), '?' (valor faltante) o ambos.
- **Excluir secuencias:** eliminar secuencias (filas) del alineamiento múltiple de secuencias.

Si el usuario hace clic en la pestaña de historial, se mostrará el historial de cambios ya realizados. Todos ellos se pueden deshacer.

9.3 IMPORT/EXPORT SECTION

En esta sección el usuario puede:

- **Importar:** un nuevo alineamiento o versión del alineamiento existente al sistema.
- **Exportar:** el mapa de calor como png. o el alineamiento múltiple de secuencias editado como Fasta.



VISOR DE ÁRBOLES FILOGENÉTICOS

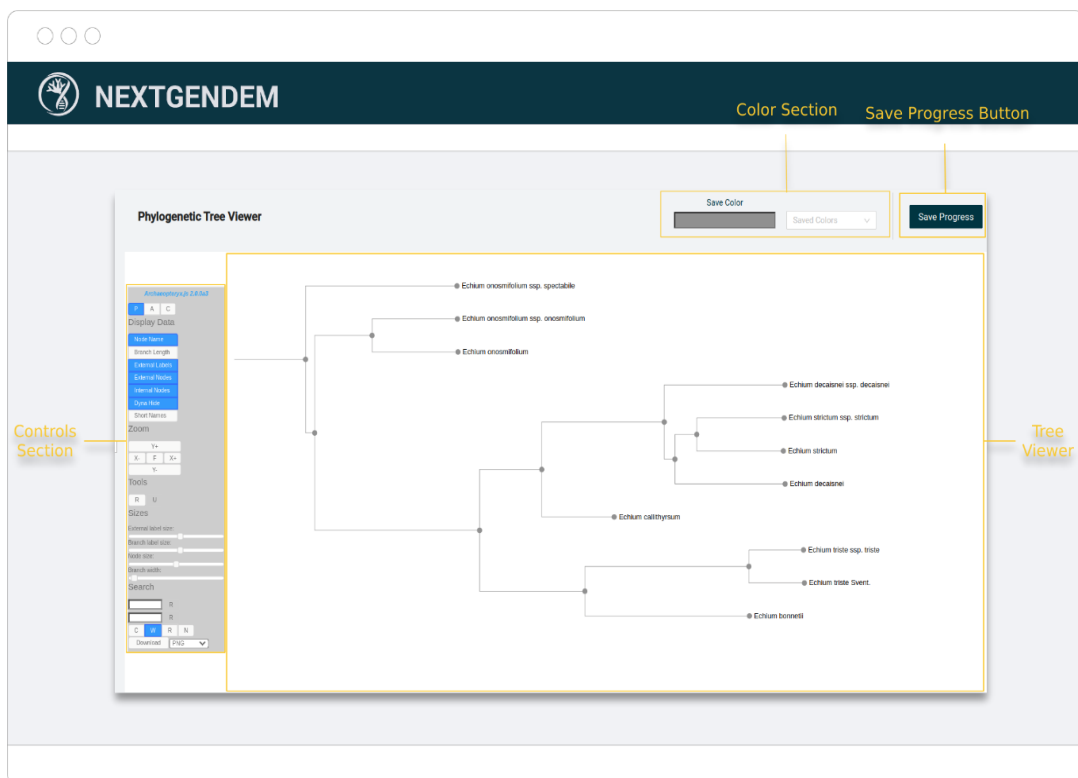


NEXTGENDEM

10. VISOR DE ÁRBOLES FILOGENÉTICOS

El visor de árboles filogenéticos es un conjunto de herramientas para visualizar correctamente los mismos. Se divide en cuatro secciones:

- **Controls Section:** en esta sección encontrará una variedad de controles que le permiten personalizar la visualización de su árbol filogenético.
- **Color Section:** personalice los colores de las ramas y los nodos para representar varios atributos, como grupos de especies, subespecies o cualquier otra clasificación de su elección.
- **Tree Viewer:** la propia visualización del árbol.
- **Save Progress Button:** le permite guardar su vista actual.



10.1 CONTROLS SECTION

A continuación, se describen los controles estándar.

- **TIPOS DE VISUALIZACIÓN DE ÁRBOLES**
 - **“P”** para visualización de filogramas (usa valores de longitud de rama).
 - **“A”** para visualización de filogramas (usa valores de longitud de rama) con etiquetas alineadas a la izquierda.
 - **“C”** para visualización de cladogramas (ignora los valores de longitud de rama).

Alternativamente, use **Alt+P** para alternar entre estos tipos de visualización.

- **MOSTRAR DATOS**

Estos ajustes controlan qué datos se muestran.

- **Node Name:** para mostrar/ocultar los nombres de los nodos (los nombres de los nodos son las claves sin tipo que se encuentran en los árboles con formato de Newick).
- **Branch Length:** para mostrar/ocultar valores de longitud de rama.
- **External Labels:** para mostrar/ocultar claves de nodos externos.
- **External Nodes:** para mostrar/ocultar los nodos externos como círculos.
- **Internal Nodes:** para mostrar/ocultar los nodos internos como círculos.
- **Dyna Hide:** para ocultar claves según la visibilidad esperada.
- **Short Names:** para acortar las claves de nodos largos.

- **ZOOM**

Estos ajustes controlan el zoom en direcciones horizontales y verticales. El zoom general (donde todo cambia de tamaño) también se puede lograr con la rueda del ratón.

- **Y+** para hacer zoom verticalmente (**Alt+Arriba** o **Mayús+rueda del ratón**).
- **Y-** para alejar verticalmente (**Alt+Abajo** o **Mayús+rueda del ratón**).
- **X+** para ampliar horizontalmente (**Alt+Derecha** o **Mayús+Alt+rueda del ratón**).
- **X-** para alejar horizontalmente (**Alt+Izquierda** o **Mayús+Alt+rueda del ratón**).
- **F** para ajustar el árbol al tamaño del visor (**Alt+C**).
- **Alt+más** y **Alt+menos** para acercar y mantener constantes todos los tamaños de fuente.
- **Shift+Alt+plus** y **Shift+Alt+menos** o **Page Up** y **Page Down** o **Shift+Ctrl+rueda del mouse** para cambiar todos los tamaños de fuente.

- **HERRAMIENTAS**

Estos ajustes controlan el árbol que se muestra.

- **O** para "ordenar" todo el árbol (**Alt+O**).
- **R1** para volver al superárbol, una rama a la vez, si está en el subárbol (**Alt+R**).
- **R** para volver al árbol completo, si está en el subárbol.
- **U** para descomprimir todo, si hay subárboles colapsados (**Alt+U**).
- **M** hasta el punto medio para volver a enraizar el árbol, si el árbol se puede volver a enraizar (**Alt+M**).

- **CONTROL DE TAMAÑO**

Estos controles deslizantes controlan el tamaño de varios elementos:

- **External label size** para controlar el tamaño de las fuentes de claves externas.
- **Internal label size** para controlar el tamaño de las fuentes de la clave interna.
- **Branch label size** para controlar el tamaño de las fuentes para confianza y longitudes de rama.
- **Node size** para controlar el tamaño de las formas de los nodos externos e internos (si está activado con "Nodos externos" y "Nodos internos").
- **Branch width** para controlar los anchos de rama (si no está configurado en el propio árbol).

- **CAMPOS DE BÚSQUEDA**

De forma predeterminada se buscan todos los campos de datos disponibles (como los nombres de secuencia).

Los términos de búsqueda se pueden introducir en los dos cuadros de entrada. Las coincidencias del término de búsqueda ingresado en el cuadro superior se resaltan

en verde, mientras que las coincidencias del término de búsqueda ingresado en el cuadro inferior están en rojo. Las coincidencias de ambos términos de búsqueda están en azul.

" R " se utiliza para restablecer (borrar) los resultados de búsqueda. Además, los botones " R " cambian de color si la búsqueda correspondiente tiene éxito (una o más coincidencias). La información sobre herramientas de los botones " R " se puede utilizar para obtener el número exacto de coincidencias.

Las opciones del modo de búsqueda son las siguientes:

- **C** para buscar distinguiendo entre mayúsculas y minúsculas.
- **W** solo para buscar palabras completas (separadas por espacios) (no se aplica a la búsqueda de expresiones regulares).
- **R** para buscar con expresiones regulares (solo para usuarios avanzados).
- **P** para buscar propiedades (ocultas) asociadas con nodos.
- **N** para negar (invertir) los resultados de la búsqueda.

- **DESCARGAR**

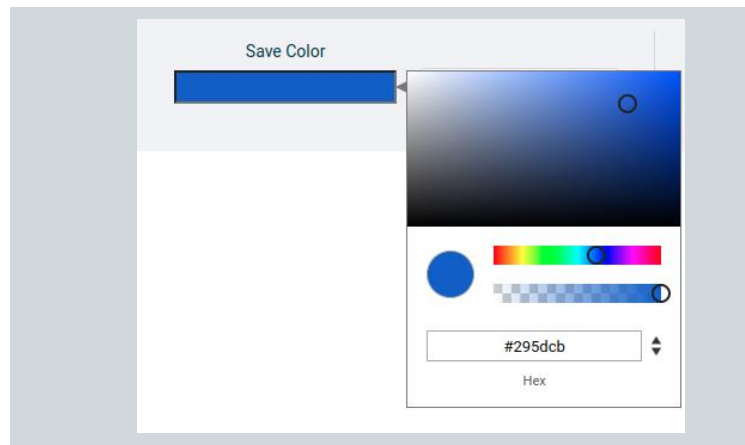
Los siguientes formatos están disponibles para su descarga:

- Imagen **PNG**.
- Imagen vectorial **SVG**.
- Archivo de árbol **phyloXML**.
- Archivo de árbol **Newick**.

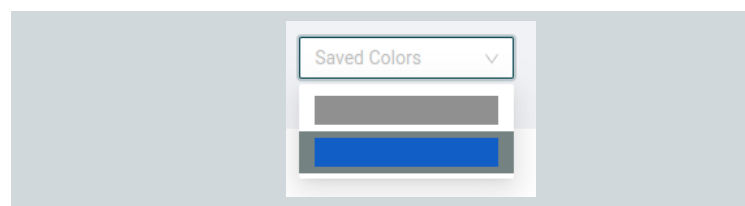
10.2 COLOR SECTION

Esta sección gestiona los colores utilizados para personalizar el árbol. Puedes realizar cuatro acciones diferentes:

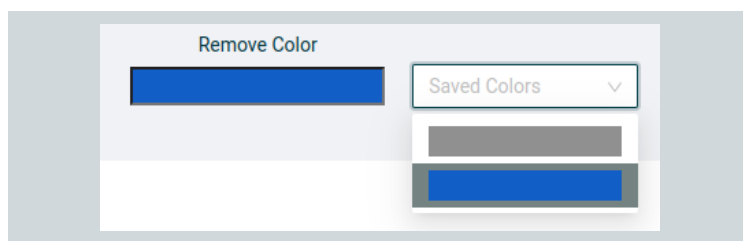
1. Seleccione un nuevo color.



2. Guardar el color seleccionado
3. Seleccione un color guardado



4. Eliminar un color guardado



10.4 TREE VIEWER

- ACCIONES DE NODO

Hacer clic derecho en los nodos permite (no todas las acciones están siempre disponibles):

- **Display Node Data:** mostrar datos del nodo.
- **Collapse:** acción de colapsar un subárbol del nodo seleccionado.
- **Uncollapse:** acción de descomprimir el subárbol colapsado en el nodo seleccionado.
- **Uncollapse all:** acción de descomprimir todos los subárboles colapsados.
- **Go to Subtree:** centra la vista en el subárbol del nodo seleccionado.
- **Swap Descendants:** intercambiar descendientes del nodo seleccionado.
- **Order Subtree:** subárbol de pedidos del nodo seleccionado.
- **Reroot:** enraizar el árbol usando el nodo seleccionado como grupo externo.
- **Delete Subtree/External Node:** eliminar subárbol del nodo seleccionado o nodo externo.
- **List External Node Data:** muestra una lista de los nombres de las hojas del subárbol del nodo seleccionado.
- **Download Ext Node Data:** para descargar datos de nodos externos que se muestran actualmente.
- **Download All Ext Node Data:** para descargar todos los datos de nodos externos.

- ACCIONES DEL RATÓN

- Colorea una rama o un nodo haciendo clic izquierdo sobre ellos.
- Mueve el árbol arrastrando el lienzo.

10.4 SAVE PROGRESS BUTTON

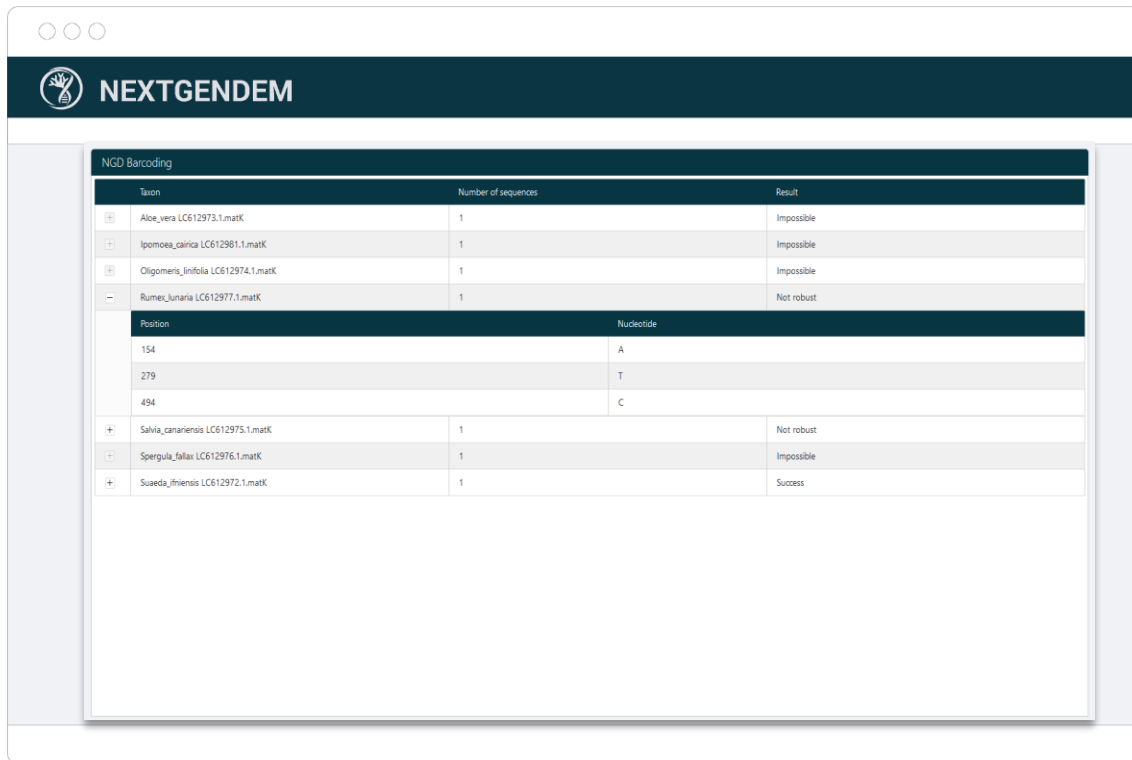
- Al hacer clic en este botón, el usuario guardará todos los cambios como una nueva vista.

NEXTGENDEM BARCODING



11. NEXTGENDEM BARCODING

Esta es una herramienta para visualizar correctamente las posiciones del diagnóstico para las regiones de los códigos de barras calculados por el software MOLD.



The screenshot shows the 'NGD Barcoding' interface. It features a table with the following columns: Taxon, Number of sequences, and Result. The table lists several taxa with their corresponding sequence counts and results. A detailed view is shown for the taxon 'Rumex_lunaria', displaying a sub-table with Position and Nucleotide columns.

Taxon	Number of sequences	Result
<input type="checkbox"/> Aloe_vera LC612973.1.matK	1	Impossible
<input type="checkbox"/> Ipomoea_cairica LC612981.1.matK	1	Impossible
<input type="checkbox"/> Oligomeris_linfofolia LC612974.1.matK	1	Impossible
<input checked="" type="checkbox"/> Rumex_lunaria LC612977.1.matK	1	Not robust

Position	Nucleotide
154	A
279	T
494	C

Taxon	Number of sequences	Result
<input type="checkbox"/> Salvia_canariensis LC612975.1.matK	1	Not robust
<input type="checkbox"/> Spargula_falax LC612976.1.matK	1	Impossible
<input type="checkbox"/> Suaeda_hibernica LC612972.1.matK	1	Success

Dentro de la tabla, hay 4 columnas:

- **Botón de expansión:** un botón para expandir u ocultar los detalles del resultado del análisis. Si el resultado es "Imposible", este botón se desactivará.
- **Taxón:** Nombre del taxón.
- **Número de Secuencias:** Número de secuencias analizadas.
- **Resultado:** resultado del análisis, que puede ser "Imposible", "No robusto" o "Éxito".

Al expandir para ver el resultado detallado, se desplegará otra tabla con dos columnas, una para posición y otra para nucleótido.

BLASTER

BLASTER

Esta es una herramienta para visualizar los resultados de un proceso BLAST. BLAST es un software que se utiliza para buscar una secuencia en una base de datos de secuencias, por ejemplo, en el caso de que el usuario no sepa a qué especie pertenece.

Se divide en tres pestañas diferentes:

- **Alignments Table:** una tabla con la información más relevante sobre la búsqueda.
- **Multiple Alignments:** un gráfico que representa un resumen de la coincidencia de la secuencia buscada (secuencia de consulta) con la secuencia más similar en la base de datos.
- **Single Alignment:** el alineamiento en detalle de la secuencia de la secuencia de consultada y la secuencia más similar en la base de datos.

The screenshot shows the BLASTER web interface. At the top, there is a dark blue header with the 'NEXTGENDEM' logo and name. Below the header, the 'Blaster' section is visible, with a search bar containing the query 'AF115614.1.BaposmatK&apos...'. There are three tabs: 'Alignments Table' (selected), 'Multiple Alignments', and 'Single Alignment'. Below the tabs, there are three buttons: 'Download as CSV', 'Download as PNG', and 'Download as JPEG'. The main content is a table with the following columns: 'Accession', 'Taxon', 'Max score', 'Total score', 'Query cover', 'E value', and 'Identities'. The table contains 15 rows of search results.

Accession	Taxon	Max score	Total score	Query cover	E value	Identities
AF115614.1	Aeonium Aureum	1772.06	1772.1	100%	0	100%
OM993524.1	Aeonium Saundersii	1679.72	1679.7	100%	0	98%
OM993521.1	Aeonium Diployclum	1674.18	1674.2	100%	0	98%
OM993526.1	Aeonium Undulatum	1663.1	1663.1	100%	0	98%
OM993522.1	Aeonium Dodrantale	1659.41	1659.4	99%	0	98%
AY082169.1	Aeonium Gorgoneum	1657.57	1657.6	100%	0	98%
OM993519.1	Aeonium Balsamiferum	1657.57	1657.6	100%	0	98%
OM993525.1	Aeonium Tabuliforme	1657.57	1657.6	100%	0	98%
MW206792.1	Aeonium Arboreum	1657.57	1657.6	100%	0	98%
NC_053947.1	Aeonium Arboreum	1657.57	1657.6	100%	0	98%
OM993518.1	Aeonium Arboreum	1657.57	1657.6	100%	0	98%
AF115585.1	Aeonium Simsii	1655.72	1655.7	100%	0	98%
AF115585.1	Aeonium Simsii	1655.72	1655.7	100%	0	98%

• ALIGNMENTS TABLE

Primero, en el campo de la esquina superior derecha, el usuario puede cambiar entre diferentes secuencias de consulta, ya que puede buscar más de una secuencia.

En la tabla tenemos las siguientes columnas:

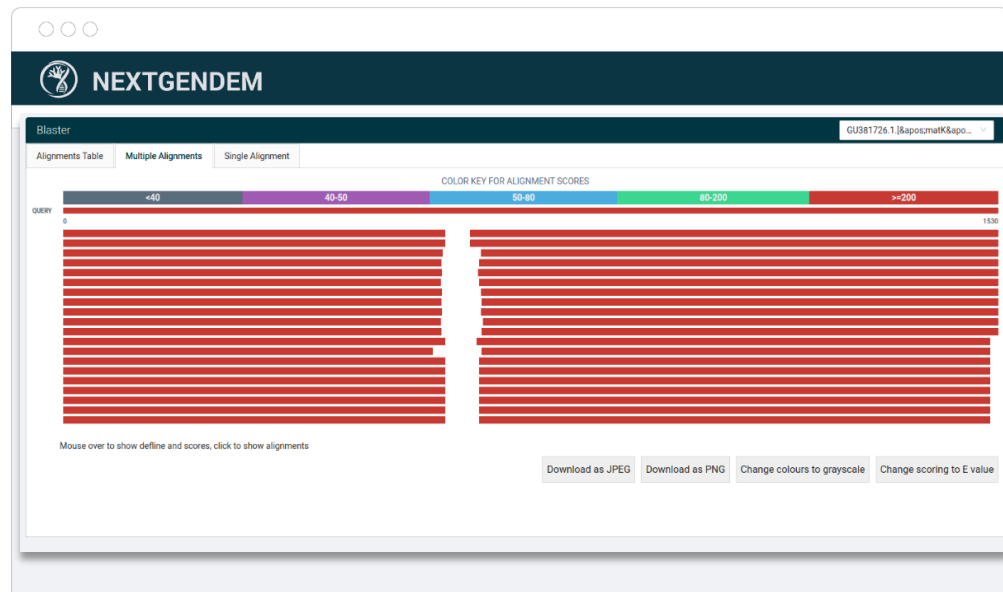
- **Accession:** ID de la secuencia en la base de datos.
- **Taxon:** la especie.
- **Max score:** la puntuación de alineamiento más alta calculada a partir de la suma de las recompensas por nucleótidos coincidentes y las penalizaciones por discrepancias y *gaps*.
- **Total score:** la suma de las puntuaciones de alineamiento de todos los segmentos de la misma secuencia de sujetos.
- **Query cover:** el porcentaje de la longitud de la consulta que se incluye en los segmentos alineados.
- **E-value:** el número de alineamientos esperados por casualidad con la puntuación calculada o mejor. El valor esperado es la métrica de clasificación

predeterminada; para alineamientos significativos, el valor E debe ser muy cercano a cero.

- **Identities:** el porcentaje más alto de identidad para un conjunto de segmentos alineados con la misma secuencia de sujeto.

Además, es posible descargar la tabla en formatos PNG, JPEG y CSV.

- **MULTIPLE ALIGNMENTS**

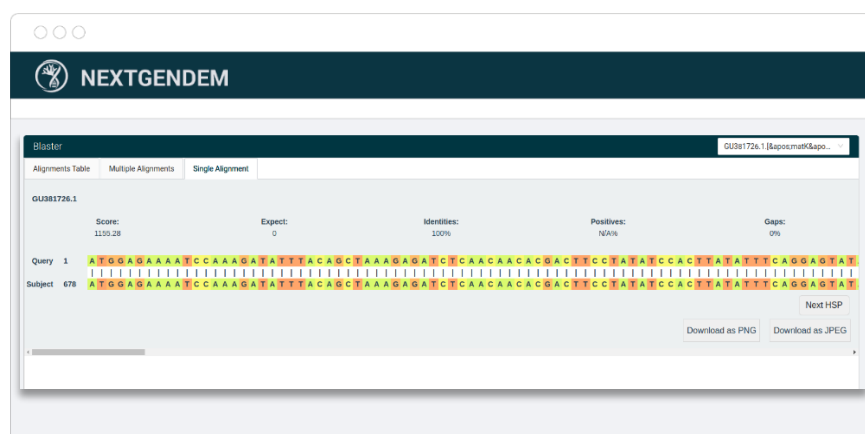


alineamiento. La clave de color asocia un color específico a un rango de puntuación concreto. La escala de colores se presenta como una barra horizontal en la parte superior de la pantalla de alineamiento.

Es posible cambiar el valor representado por los colores entre el *E-value* y la *Max Score* pulsando el botón situado en la parte inferior derecha de la pantalla. A su izquierda también hay un botón para cambiar los colores a escala de grises.

También es posible descargar el gráfico en formatos JPEG y PNG.

- **SINGLE ALIGNMENT**



Esta es una representación gráfica del alineamiento de la secuencia consultada en detalle y de la secuencia más similar en la base de datos. Hay un botón para ir al siguiente HSP. Los HSP son segmentos del alineamiento en los que la puntuación supera un cierto umbral, lo que indica un nivel significativo de similitud.

Es posible descargar el gráfico como PNG o JPEG.

GESTIÓN DE DATOS GEOGRÁFICOS



NEXTGENDEM

12. GESTIÓN DE DATOS GEOGRÁFICOS

El sistema cuenta con un subsistema de gestión de datos geográficos diseñado para ofrecer una interfaz completa y flexible capaz de gestionar datos geográficos. Esta unidad de información se denomina capa geográfica o simplemente **capa**.

El subsistema está preparado para gestionar capas geográficas de tipo vectorial. Una capa vectorial puede verse como una tabla en la que cada una de las filas tiene una columna con una forma gráfica que se puede marcar en un mapa, representando el área geográfica para la cual la otra información también es válida. A esta columna se le llama Geografía. Las columnas restantes pueden ser de diferentes tipos de datos: fecha, numérico, categoría y tiempo. En cuanto al conjuntos de datos en sentido estadístico, cada una de las columnas puede ser de una de tres opciones: una dimensión, una medida o un atributo. Ambos mundos, geográfico y estadístico, se combinan en una capa/conjunto de datos. Las columnas de dimensión se pueden usar para pivotar, las medidas son bits de información real, mientras que los atributos son metadatos que describen dimensiones o medidas.

El resto de la sección explica cada uno de los elementos a través de los cuales se pueden gestionar las capas no visualizadas.

12.1 LISTA DE CAPAS

Al hacer clic en “Datos geográficos” y luego en “Capas”, aparece un navegador que sigue las mismas reglas de uso de otros navegadores en el sistema, lo que permite enumerar fácilmente las capas geográficas a las que se permite el acceso. Para refinar las capas que se muestran en la lista, los usuarios pueden aplicar filtros basados en valores de campo o realizar una búsqueda abierta de capas específicas.

The screenshot shows the 'NEXTGENDEM' interface for managing geographic layers. On the left is a sidebar with a navigation menu including 'Molecular Data', 'Geographical data', and 'Layers'. The main content area is titled 'Layers' and features a search bar with the placeholder 'Enter your search terms.' Below the search bar are buttons for 'Import', 'Add to Case Study', 'Add to collection', 'Export Layers', 'Create Layer', and 'Filter'. The interface displays a table of layers with the following data:

name	Actions
<input type="checkbox"/> Alfa diversidad - Picante	
<input checked="" type="checkbox"/> Bioclima	
<input type="checkbox"/> ENP (original)	
<input type="checkbox"/> ENP (solapado)	
<input type="checkbox"/> Especies	
<input type="checkbox"/> Formaciones forestales	
<input checked="" type="checkbox"/> Litología	
<input type="checkbox"/> Malla 500	
<input type="checkbox"/> Métricas 1	
<input checked="" type="checkbox"/> Ocupación suelo	

At the bottom of the table, there is a pagination control showing '1' of 30 pages.

Geographic Layers Browser

12.2 DETALLES DE CAPA Y EDICIÓN

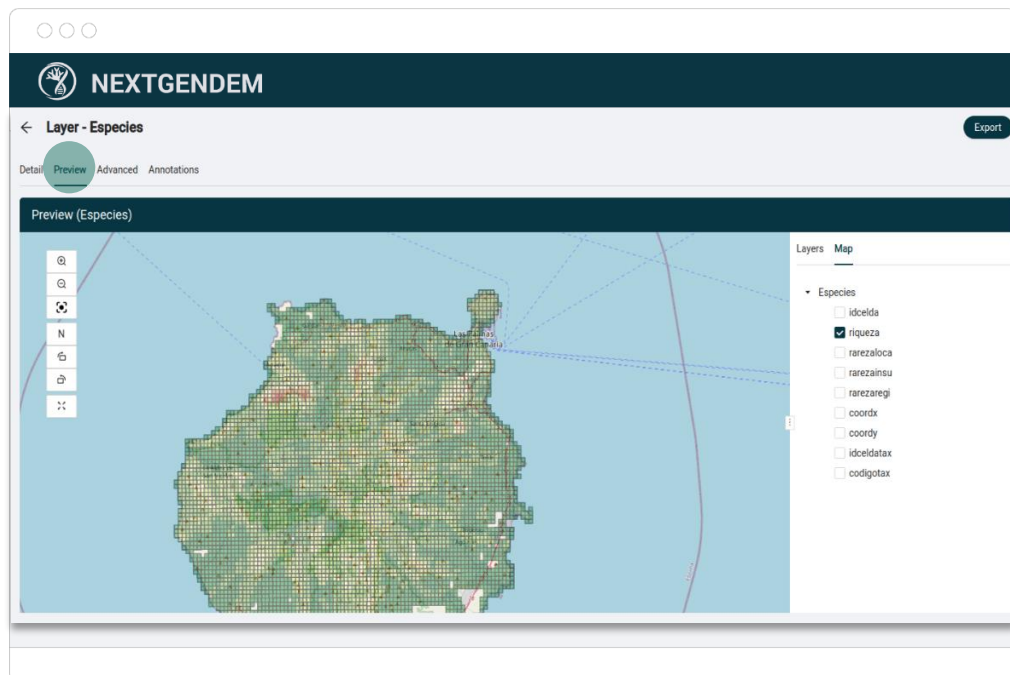
Al hacer clic en una de las capas de la lista, se abre una pantalla de detalles que muestra información detallada de dicha capa. La pantalla está organizada en las siguientes pestañas:

- La pestaña **Detalle** ofrece información general sobre la capa.

The screenshot shows the 'Layer - Especies' detail page. The sidebar on the left contains a navigation menu with categories like 'Molecular Data', 'Geographical data', 'Layers', 'Case studies', 'Processes', 'Visualization', 'Metadata', 'Annotations', 'Process Admin.', and 'Security Admin.'. The main content area has a breadcrumb trail and tabs for 'Detail', 'Preview', 'Advanced', and 'Annotations'. The 'Detail' tab is selected, showing a form with the following fields: 'Id' (2187), 'Name' (Especies), 'Tags' (Tags), 'Subjects' (Subjects), 'Authorship' (Authorship), 'Source file name' (BIOTA_Hiv202106_P1y2_T_ID.zip), 'Source' (Source), 'CRS' (CRS), and 'Creation date' (Select date). There is also a table for 'Cases studies' with one entry 'cs11'.

Pantalla de detalle de una Capa - Pestaña Detalle

- La pestaña **Vista previa** proporciona una vista de mapa simplificada de la capa:



- La pestaña **Avanzado** permite editar las propiedades de la capa, como por ejemplo cambiar la tipología de los datos:

Pantalla de detalles - Pestaña Avanzado

Name	Type	Category type	Minimum value	Maximum Value	Representable	Actions
idcelda	numeric	-	1	6373	Yes	✎
riqueza	numeric	-	1	108	Yes	✎
rarezaloca	numeric	-	0.000236967	1.94982	Yes	✎
rarezainsu	numeric	-	0.000237023	1.96514	Yes	✎
rarezaregi	numeric	-	0.0000706964	0.982267	Yes	✎
coordx	numeric	-	417500	464500	Yes	✎
coordy	numeric	-	3067500	3117000	Yes	✎
idceldatx	numeric	-	1	130764	Yes	✎

Al hacer clic en uno de los nombres de las propiedades, se abre el siguiente diálogo, en el cual es cambiar el tipo de campo, afectando desde el almacenamiento de la información hasta su visualización:

Formulario emergente para modificar el tipo de una propiedad en una Capa

Finalmente, la pestaña **Anotaciones** permite administrar las notas asociadas y las instancias de ontología como en otros objetos de datos de la interfaz de usuario.

12.3 IMPORTAR CAPA

Desde la lista de capas, al hacer clic en el botón "Importar", se abre la pantalla de importación de capas. Simplemente especificando el nombre de una capa, es posible seleccionar el archivo para importar. Los formatos permitidos son:

- GeoJSON.
- Geopackage.
- Shapefile (empaquetando los archivos en un archivo .zip previamente).
- CSV. Sin Geometría, útil para crear una nueva capa (ver más abajo) con columnas adicionales.

12.4 EXPORTAR CAPAS O ASOCIARLAS A CASOS DE ESTUDIO / COLECCIONES

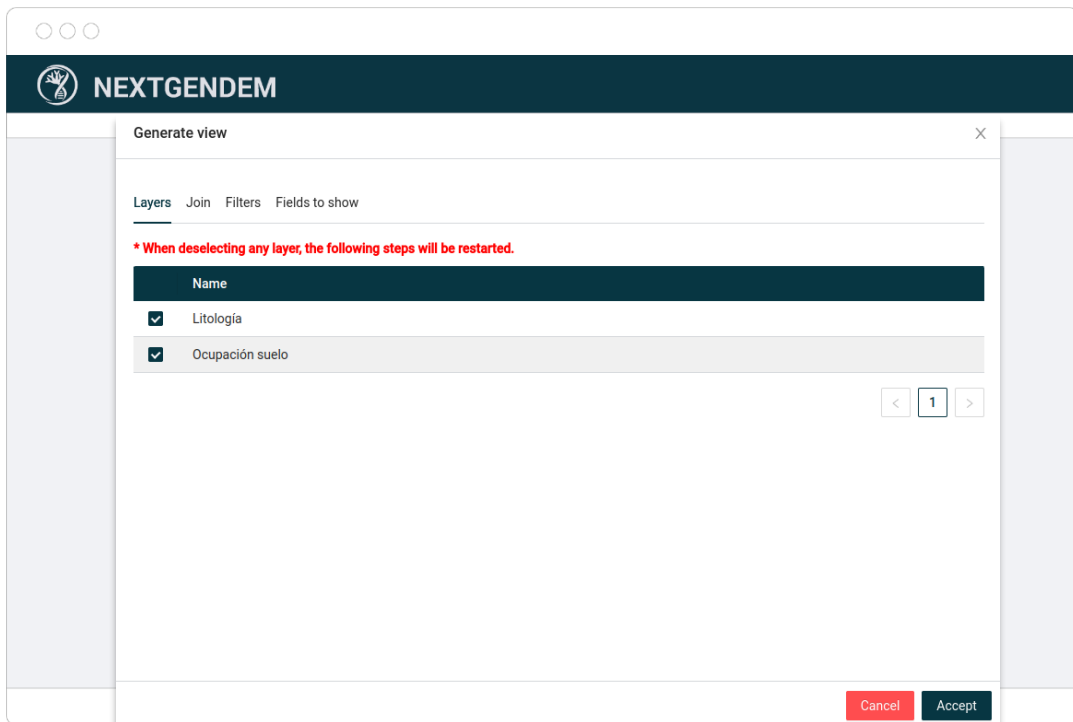
Desde la lista de capas, seleccionando una o más capas es posible exportar esas capas seleccionadas, o asociarlas con estudios de casos o colecciones.

12.5 PERMISOS

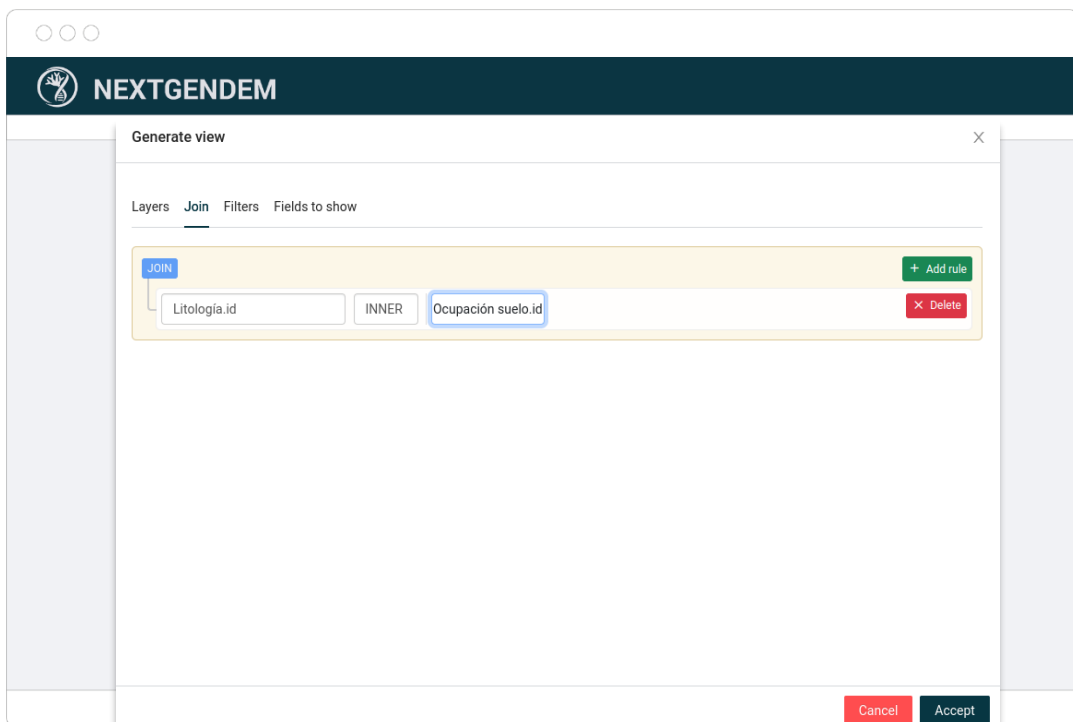
Gracias a las capacidades generales del navegador, también es posible administrar los permisos para acceder a cada capa.

12.6 CREACIÓN DE UNA NUEVA CAPA A PARTIR DE CAPAS EXISTENTES

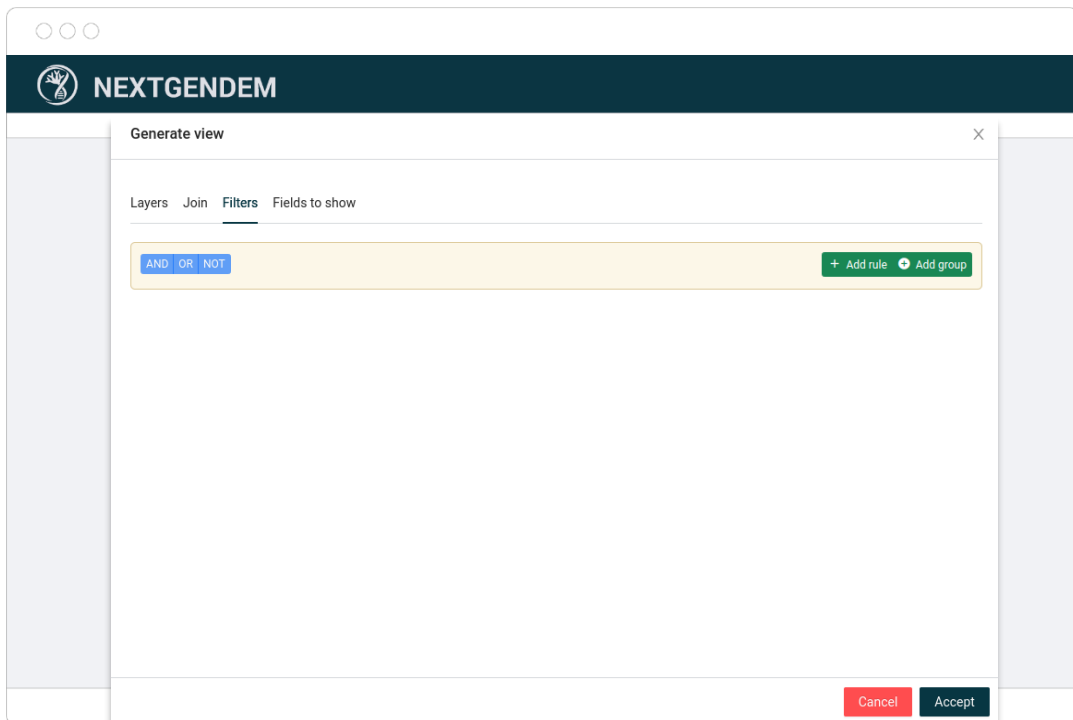
El sistema proporciona una herramienta para crear una nueva capa. Esta herramienta usa las capacidades de PostGIS en las que se puede usar un comando SQL para crear una nueva capa. El sistema NEXTGENDEM proporciona un diálogo visual para construir esta consulta SQL, mostrando los campos de las capas existentes y las operaciones válidas, especificando "FROM" (qué capas), "JOIN" (cómo conectar pares de capas), "WHERE" (cómo filtrar las filas resultantes del JOIN), y cláusulas "SELECT" (qué columnas -propiedades- tendrá la capa resultante), como muestran las siguientes capturas de pantalla:



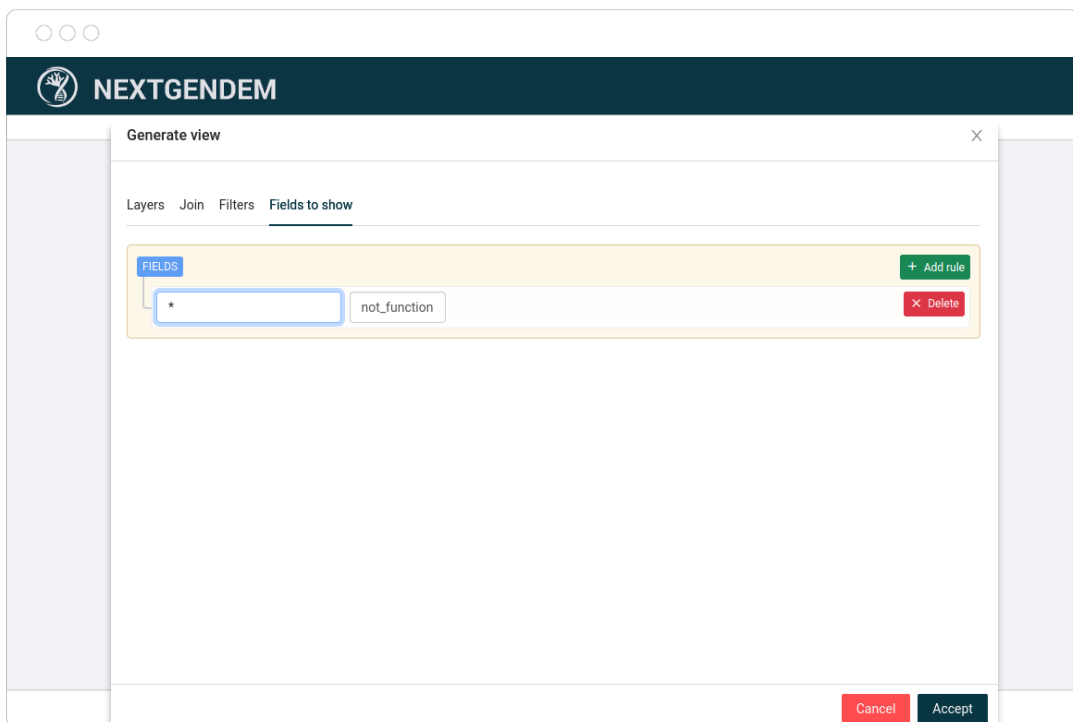
- Activa las capas para JOIN



- Especifique cómo unir las capas.



- Especifique el filtro. Se puede vaciar, como en la figura.



- Qué filtros estarán en la nueva capa.

Una vez especificado todo, al hacer clic en “Aceptar” se iniciará la creación en el servidor PostGIS (de forma transparente para el usuario) y la nueva capa aparecerá en la lista (con un nombre generado automáticamente, utilizando el ID interno).

VISOR GEOGRÁFICO Y MAPAS

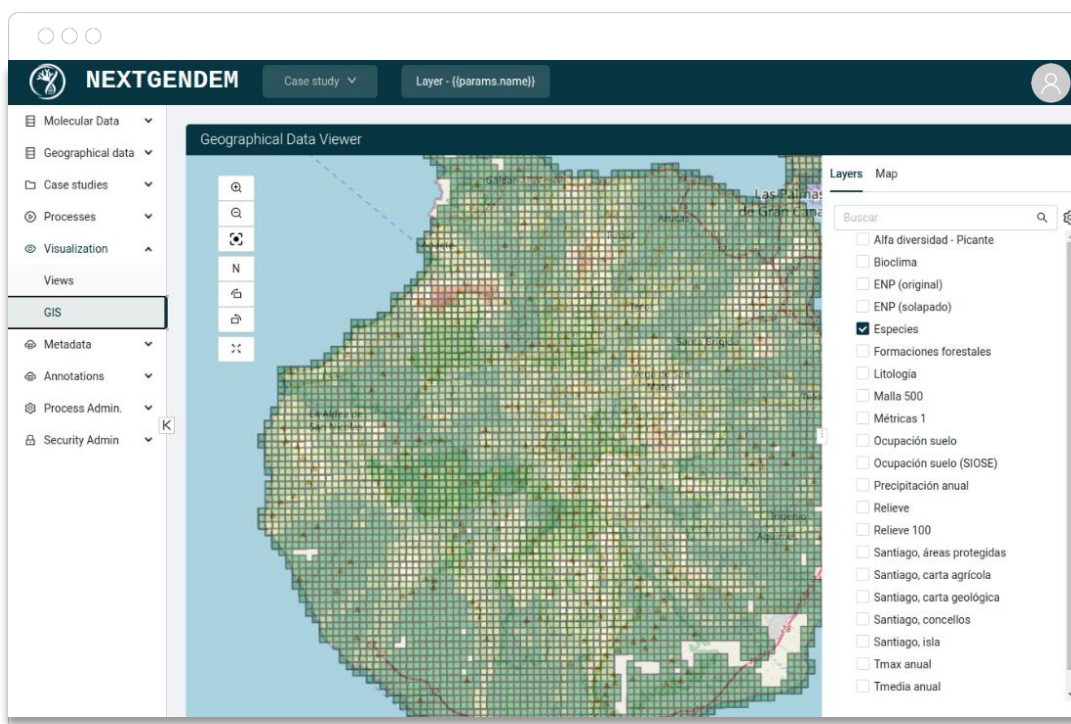


NEXTGENDEM

13. VISOR GEOGRÁFICO Y MAPAS

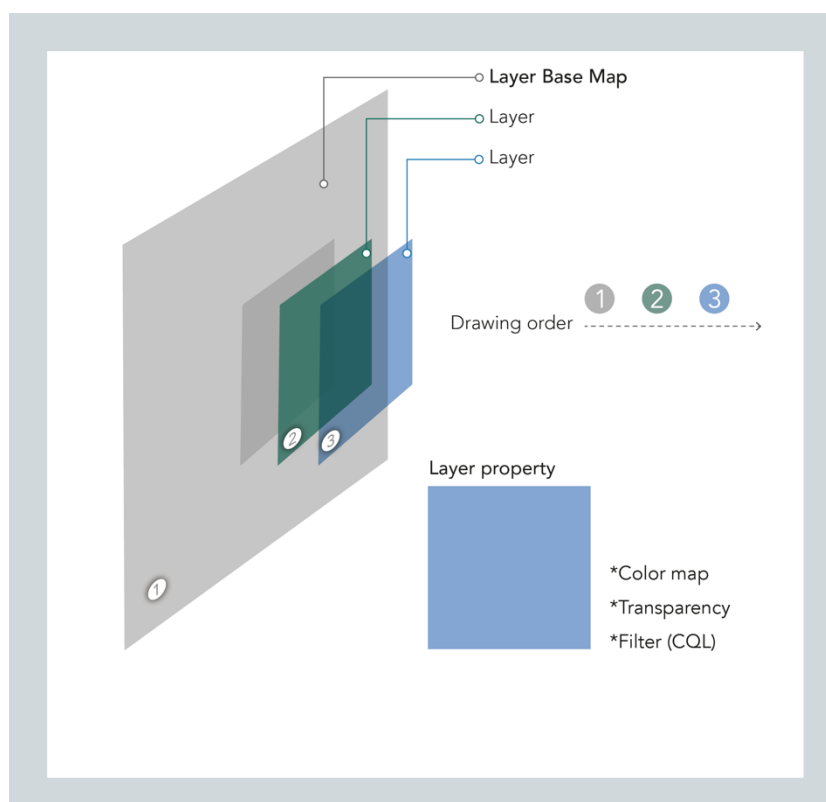
Una vez que los datos se preparan con las funciones de administración de datos geográficos, el visor de datos geográficos permite explorar datos de múltiples capas y guardar vistas para uso, explicación o referencia en el futuro.

Para abrir un nuevo visor geográfico, haga clic en **Visualización** y luego en la opción de menú **GIS**. El visor geográfico se puede dividir en dos partes:



13.1 CÓMO SE DIBUJA UN MAPA

El componente gráfico pinta un mapa capa por capa. La primera capa que se pintará es el Mapa base. Luego, cada una de las propiedades representadas, en el orden en que aparecen en la lista de la pestaña Mapas. Cada capa se dibuja usando una combinación del filtro (CQL), el mapa de colores (para mapear valores a colores) y la opacidad. El dibujo muestra esto esquemáticamente:




13.2 EL PANEL DE CONFIGURACIÓN (DERECHA)

En esta parte, tenemos dos pestañas:

○ PESTAÑA CAPAS

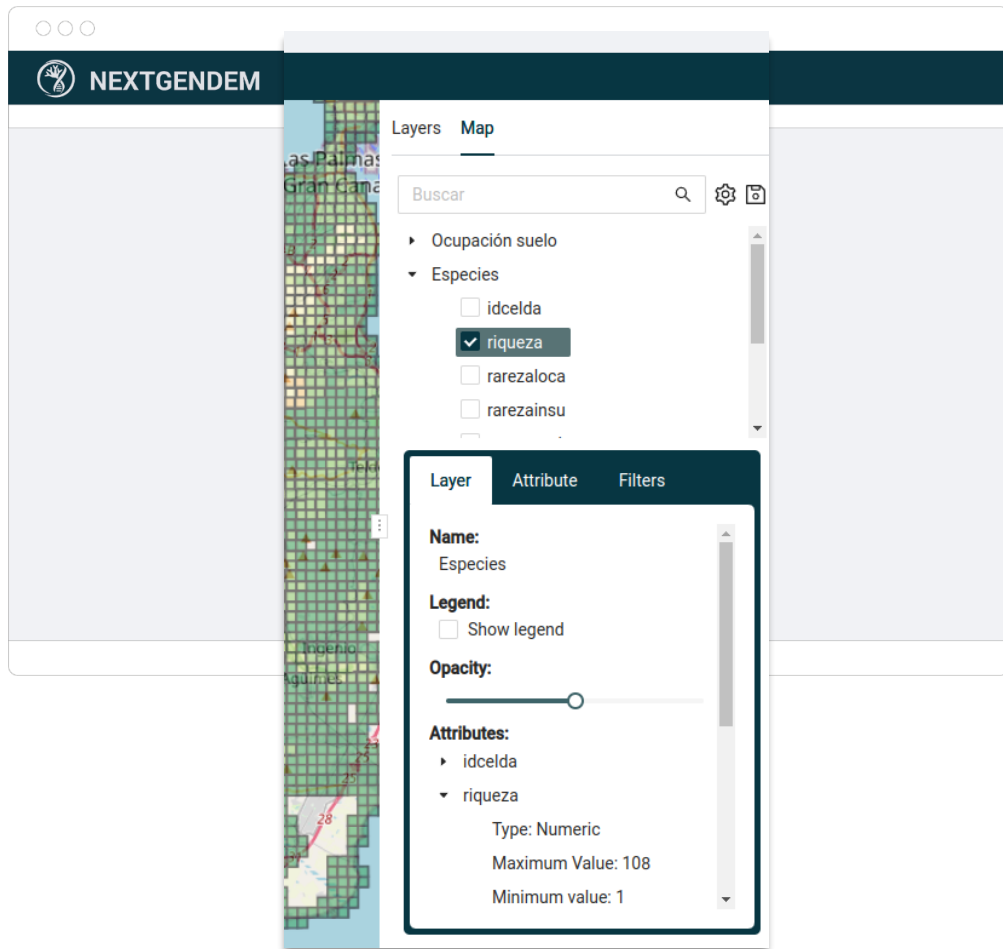
En esta pestaña, se presenta una lista de las capas accesibles. Seleccionándoles (uno o más), aparecerán en la pestaña “Mapa” para seguir trabajando con ellos. Las capas se pueden ubicar utilizando el cuadro de búsqueda, ya sea por nombre o por el contenido de los atributos que califican las capas. Además, las capas se pueden organizar

jerárquicamente haciendo clic en el icono del engranaje  .

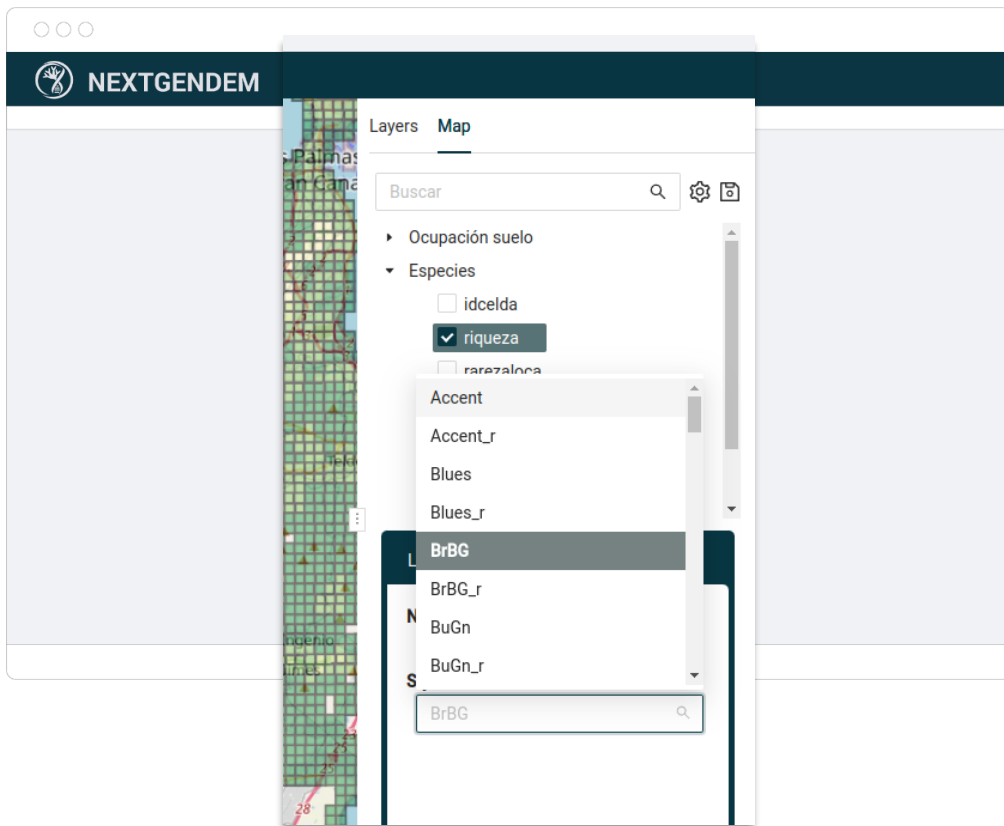
Al pulsar sobre este icono se abrirá una ventana con los atributos que acompañan a cada una de las capas, y arrastrándolos a la columna de atributos seleccionados podremos elegir cómo agruparlos jerárquicamente. El atributo ubicado en la parte superior será el primer nivel de la jerarquía y el que esté en la parte inferior será el último nivel.

○ PESTAÑA MAPA

En esta pestaña, se enumeran las capas previamente seleccionadas mediante la pestaña “Capas”. Cada capa se puede expandir para ver sus propiedades **representables** (que pueden no ser todas). Al hacer clic en la casilla de selección a la izquierda de cada propiedad, se muestra esa propiedad en el mapa. Actualmente, cada capa puede tener solo una propiedad (o ninguna) representada en el mapa. El orden en que se enumeran las capas en la pestaña “Ver” determina su orden de dibujo en el mapa. El orden se puede cambiar arrastrando los nombres de las capas hacia arriba o hacia abajo. Cuando se hace clic en una etiqueta de propiedad o etiqueta de capa, aparece una ventana debajo de la lista (en la parte inferior de la misma ventana) para configurar la opacidad, el mapa de colores, activar o desactivar la leyenda.

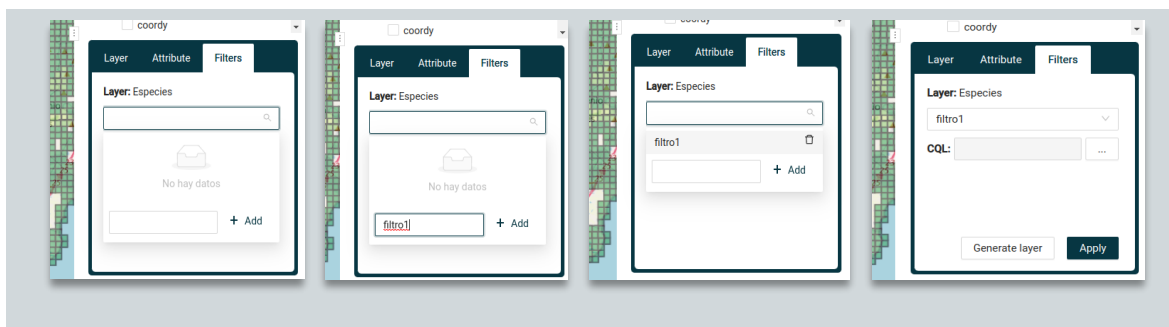


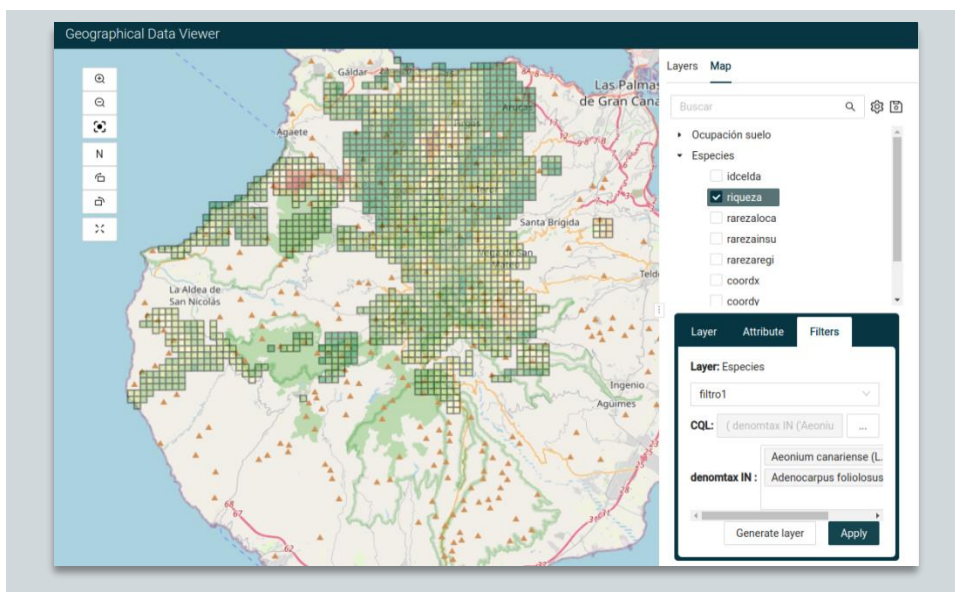
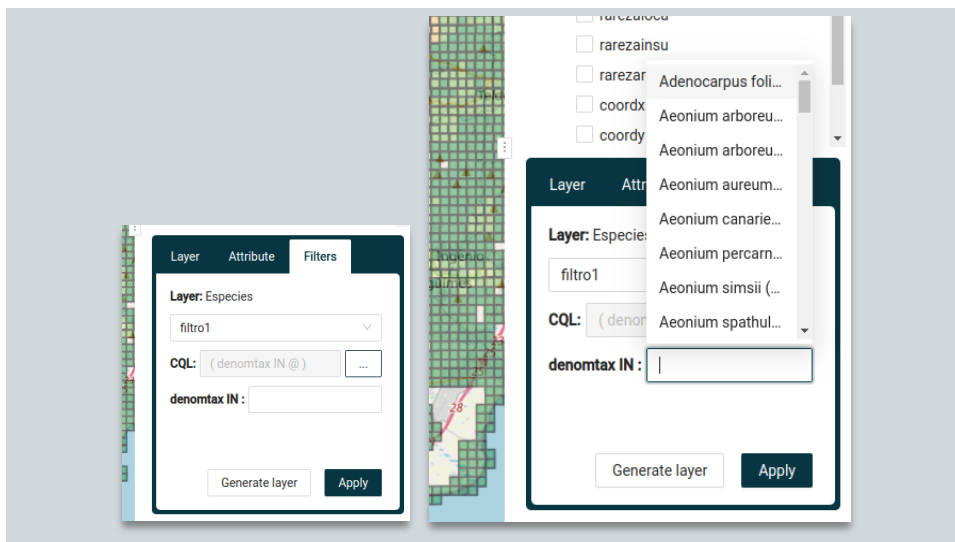
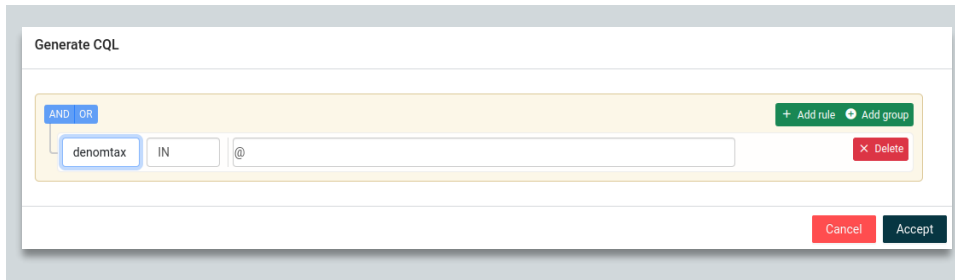
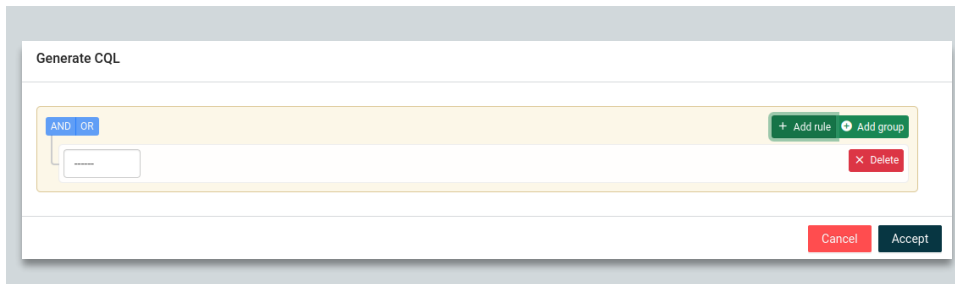
Capa: mostrar leyenda, opacidad, lista de propiedades

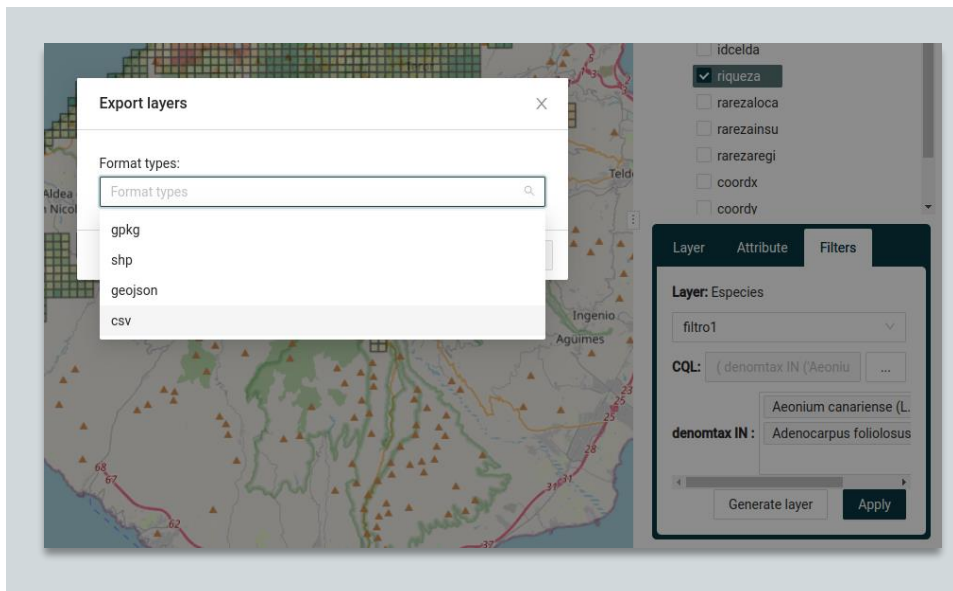


Configuración del mapa de colores asociado a una propiedad


Dentro de esta ventana "Mapa", también hay una pestaña "Filtros". En esta pestaña, se pueden crear múltiples filtros. Al seleccionar uno de ellos, se aplicará a la capa en el mapa interactivo. El filtro se establece usando CQL (Common Query Language) haciendo clic en el botón ubicado a la derecha de este campo y un asistente visual similar al "crear capa" que se muestra con anterioridad. Si en el valor de un campo se utiliza '@' como valor en alguna de las condiciones, este valor será modificable interactivamente desde la pestaña "Filtros". Además, en esta pestaña, puede exportar una nueva capa basada en estos filtros haciendo clic en "Generar capa".

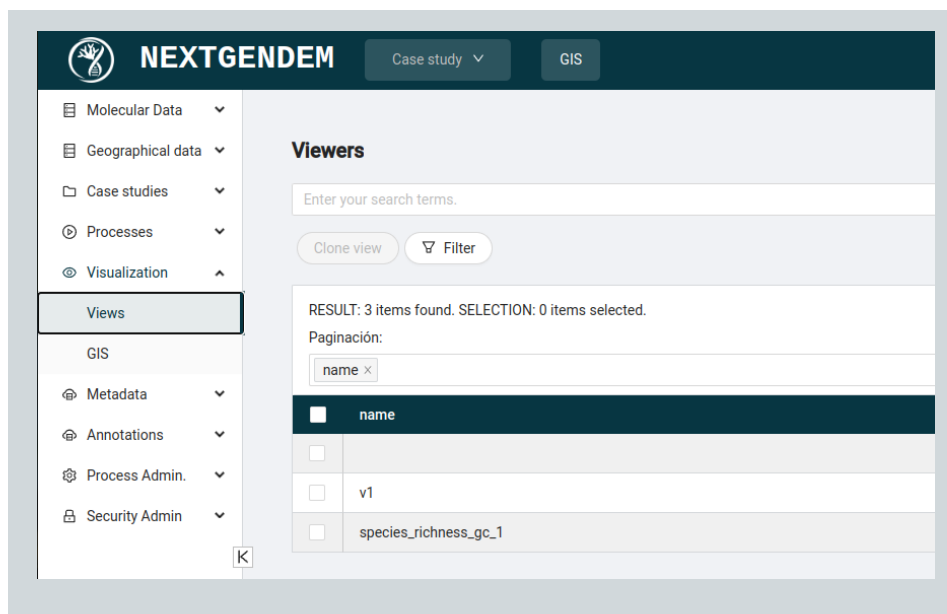
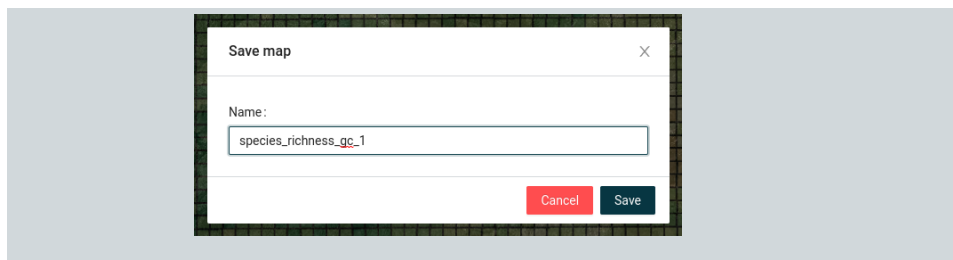







Encima de la lista de capas hay un cuadro de búsqueda para filtrar las capas por su nombre o metadatos.

El icono del disquete  se puede usar para guardar el estado actual del mapa como una Vista y reabrirlo más tarde usando la opción "Vistas" (ver la sección Vistas).





El icono del engranaje  abre una ventana de configuración donde puede cambiar la capa base. Puede elegir entre las opciones de sistema predefinidas o utilizar un servicio WMS abierto ingresando su URL. Además, la proyección del mapa también se puede configurar utilizando una de las proyecciones previamente configuradas para el sistema.



13.3 MAPA INTERACTIVO (PANEL IZQUIERDO)

Para navegar por el mapa interactivo, puede hacer clic y arrastrar el ratón a la posición deseada, o usar el botón de *scroll* para acercar o alejar. En la esquina superior izquierda, también encontrará botones para acercar/alejar, reorientar hacia el norte, girar 90° en el sentido de las agujas del reloj o en sentido contrario a las agujas del reloj y configurar la

pantalla completa. Un botón importante es el del punto de mira  que centra la vista en la medida de las capas configuradas para ser representadas. Finalmente, cuando un usuario hace clic en el mapa, se abre una ventana que muestra la información de las regiones que coinciden con el punto seleccionado, en todas las capas habilitadas que tienen información geográfica en ese punto. Este cuadro de diálogo de información se puede mover a la pestaña

"Mapa" en el panel de configuración haciendo clic en las flechas de adelante y atrás .

COLECCIONES



NEXTGENDEM

14. COLECCIONES

El sistema permite asociar un conjunto de objetos de datos: secuencias, alineamientos, árboles filogenéticos, etc. a un contenedor al que se ha denominado **colección**. Las colecciones tienen un nombre y se pueden usar para asignar permisos a los usuarios. Si a un usuario se le permite el acceso a una colección, se le otorgará acceso a todos los datos contenidos en esa colección. Cualquier objeto de datos puede ser miembro de varias colecciones. Además, al igual que con otros objetos, las colecciones pueden tener anotaciones.

The image displays the NEXTGENDEM web interface for managing collections. The main page shows a sidebar with navigation options like 'Molecular Data', 'Geographical data', and 'Case studies'. The 'Collections' section is active, displaying a search bar, a 'Create' button, and a table of existing collections. Below the main interface, three modal windows are shown:

- Create collection:** A form with a 'Name:' label and an input field containing 'Name'. A red border highlights the input field with the error message 'Enter a name for field'. 'Cancel' and 'Create' buttons are at the bottom.
- Add to collection:** A form with a 'Collections:' label and a dropdown menu showing 'New collection'. 'Cancel' and 'Accept' buttons are at the bottom.
- Collections (Filtered):** A view of the collections list with a search filter applied. The search bar contains 'name', and the results table includes 'name', 'JBCVC', 'GenBank', and 'New collection'.

ESTUDIOS DE CASO



NEXTGENDEM

15. ESTUDIOS DE CASO

De manera similar, los objetos de datos se pueden agregar a estudios de casos. **Los estudios de caso** son útiles para limitar los datos que muestra el sistema acogiéndose exclusivamente a los contenidos en el estudio de caso. Se podría decir que son mesas de trabajo donde se muestran exclusivamente los datos relativos al estudio específico diseñado por el usuario. Si bien los **casos de estudio** pueden tener permisos, los usuarios implicados en un estudio también deben tener acceso a los datos referidos para dicho estudio.

En la interfaz de usuario, el área del **estudio de caso** se puede cambiar seleccionando el control desplegable en la barra superior. Cabe mencionar que, ningún **estudio de caso** debe estar activo para volver a trabajar con todos los datos. Una consideración para tener en cuenta a la hora de crear o administrar datos que no se centran en estudios específicos.


Desde el punto de vista de la interfaz de usuario, los **estudios de caso** son como si fuesen colecciones, por lo que crear, enumerar y agregar objetos es lo mismo. Sin embargo, como ya se ha mencionado, es posible configurar un **caso de estudio** activo y esto reducirá voluntariamente los objetos que el usuario verá en los navegadores.

VISORES



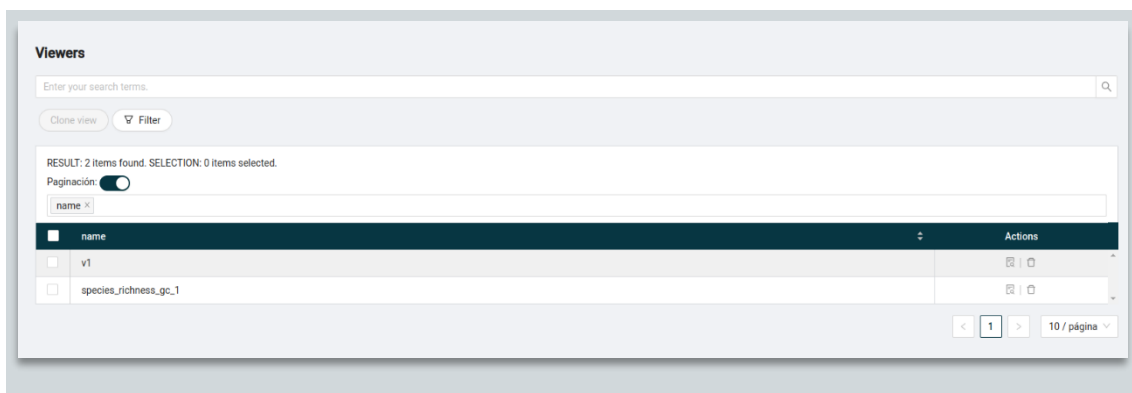
NEXTGENDEM

16. VISORES

Las visualizaciones se pueden guardar, lo que permite a los usuarios la comodidad de recuperarlas en otro momento. Por ejemplo, el visor geográfico presenta el icono de disquete  , que indica la opción de guardar.

Si se guarda una vista geográfica, esas vistas aparecerán en el navegador “Vistas”. Navegando a Visualizaciones y luego seleccionando Vistas, el usuario encontrará el repositorio de vistas, enumerando aquellas a las que el usuario ~~actual~~ está autorizado a acceder. Estas vistas se pueden reabrir con fines de reproducción o se pueden modificar para adaptarlas a las preferencias del usuario y posteriormente guardarlas con el fin de incorporar esos cambios. Las vistas revisadas se almacenan, lo que ofrece la flexibilidad de recuperarlas y aplicarlas en el futuro.

La siguiente figura muestra una captura de pantalla del navegador *Visores*. Al hacer clic en una de las filas de la vista, se abre el visor con la configuración guardada.



GESTIÓN DE PROCESOS



NEXTGENDEM

17. GESTIÓN DE PROCESOS

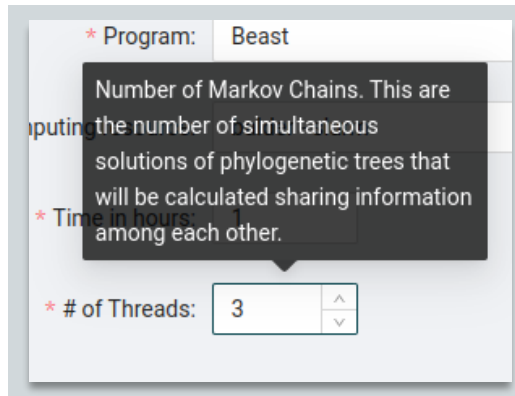
17.1 EJECUCIÓN DE UN PROCESO

La interfaz de ejecución de procesos, para cualquier análisis, sigue la estructura general que se detalla a continuación independiente del análisis seleccionado.

- Seleccione el tipo de proceso y el recurso de cómputo en el que se ejecutará el proceso: al usar una herramienta o software bioinformático, el usuario puede elegir entre diferentes tipos de análisis. Cada análisis requiere recursos informáticos específicos para ejecutarse de manera efectiva. Estos recursos informáticos incluyen campos como la cantidad de nodos (máquinas, físicas o virtuales), el tipo y la cantidad de procesadores (CPU) o procesadores gráficos (GPU) y la cantidad de tiempo solicitado en el recurso informático para completar el análisis.
- Al utilizar el sistema, primero se selecciona el tipo de análisis a realizar. Según esa selección, el sistema proporcionará una lista de recursos informáticos disponibles que tienen instalado el software necesario para ese análisis. Estos recursos informáticos pueden diferir en cuanto a su potencia informática, como la cantidad de nodos, CPU o GPU que tienen.
- A continuación, se deben especificar los recursos de cómputo que mejor se adapten al caso a ejecutar. Por ejemplo, si el análisis requiere mucha potencia computacional, ya sea varios nodos o una máquina con una gran cantidad de CPU o GPU. Para algunos procesos estos parámetros se rellenan automáticamente y para otros el usuario tendrá un mensaje de ayuda al señalar con el cursor el campo de entrada. Además, es necesario especificar la cantidad de tiempo que necesita el proceso para finalizar la ejecución.

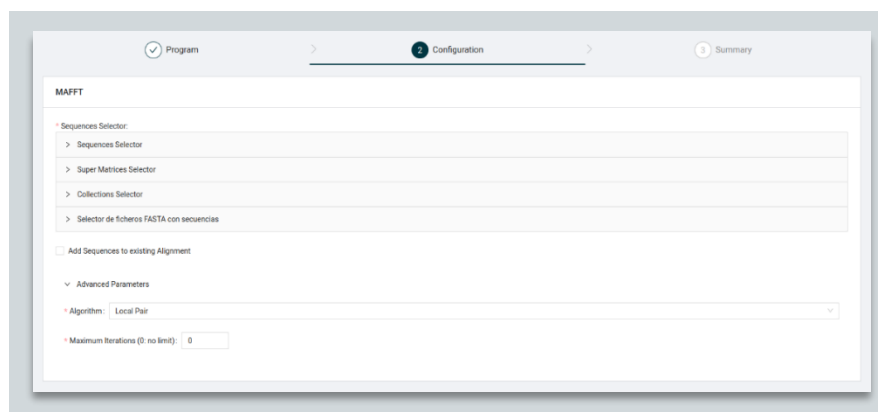
Recursos seleccionados automáticamente: se explica en el mensaje amarillo

Recursos insertados manualmente



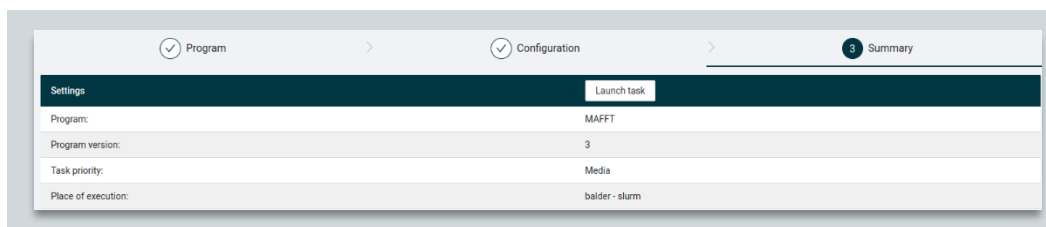
- Completar el formulario del análisis seleccionado: el usuario encontrará un formulario o un conjunto de campos de entrada para rellenar, se divide en subprocesos más pequeños. El formulario puede solicitar información específica de su análisis bioinformático. Por ejemplo, es posible que deba proporcionar los datos de entrada (p. ej., secuencias de ADN, alineamiento múltiple de secuencias) y especificar algunos parámetros específicos del análisis seleccionado (p. ej., outgroup, formato de salida).

En la mayoría de los análisis hay una sección de parámetros avanzados que permite al usuario adaptar el análisis a su pregunta de investigación específica o diseño experimental.



Página para llenar los parámetros del proceso

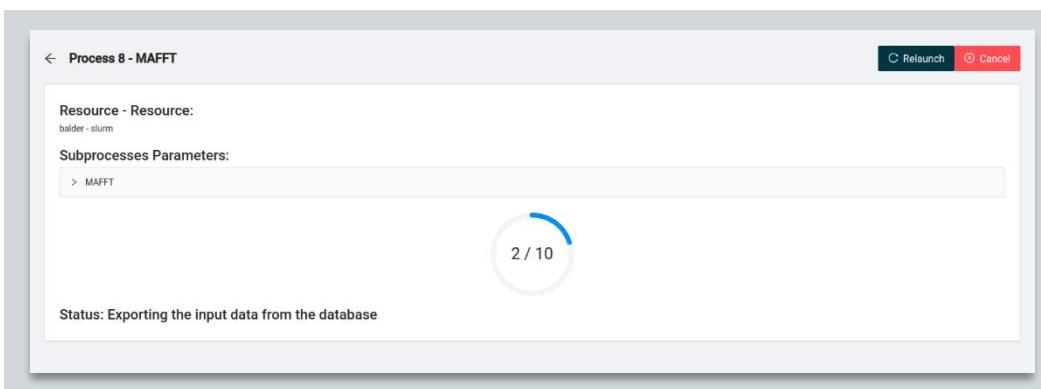
- Ver resumen y ejecutar el proceso: En el último paso el usuario tendrá una lista de los parámetros más importantes del proceso y un botón para iniciarlo.



Resumen del proceso y botón de tarea de lanzamiento

17.2 DETALLE DEL PROCESO

En la pantalla que vemos a continuación se muestra cuando se inicia un proceso. La pantalla está dividida en diferentes secciones para proporcionar una visión general y clara del proceso de ejecución del trabajo. En la parte superior de la pantalla, encontrará un encabezado o título que indica el ID del trabajo que se está ejecutando y el análisis.



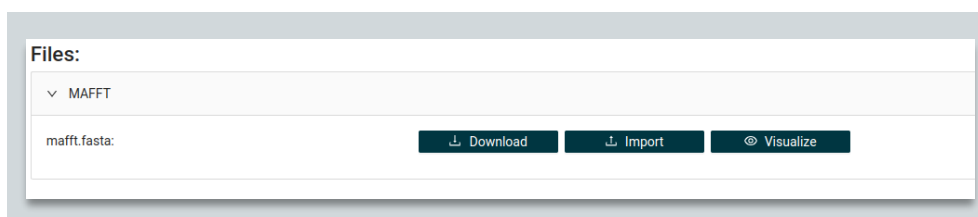
Detalle del proceso durante la ejecución

En la esquina superior derecha hay dos botones accionables:

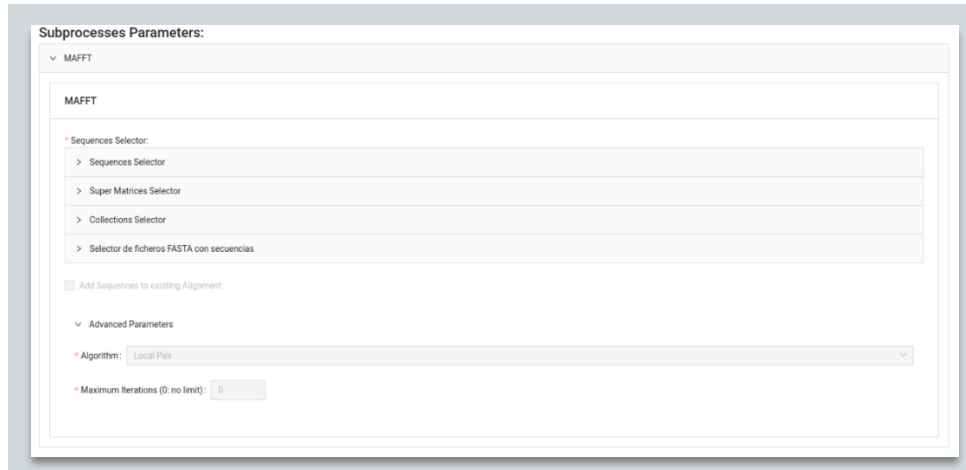
- **Cancelar:** Presione este botón para cancelar la ejecución del proceso. Solo es visible mientras el proceso se está ejecutando.
- **Relanzar:** presione este botón para volver a la página de configuración del proceso con todos los parámetros de análisis que el usuario había ya completados.

Debajo hay un área blanca con seis secciones diferentes:

- **Recurso:** muestra el nombre del recurso informático donde se ejecutó el proceso.
- **Archivos:** muestra los archivos de salida resultantes del proceso, separados por subprocessos. Hay botones para descargar y si es posible importar al sistema o visualizarlo usando uno de los visores del sistema.



- **Log:** esta sección se dedica a mostrar mensajes relacionados con la ejecución del trabajo. Estos registros pueden incluir mensajes informativos, advertencias o notificaciones de errores que ocurrieron durante la ejecución. Son útiles para solucionar cualquier problema que pueda haber ocurrido durante el análisis.
- **Parámetros de los subprocessos:** esta sección enumera los parámetros específicos elegidos para el análisis en la configuración del proceso.



Parámetros seleccionados en la configuración del proceso

- **Círculo de progreso:** La barra circular indicará el progreso de la ejecución del proceso, hay diez pasos. El nombre de cada paso se muestra en el mensaje de estado en la parte inferior de la pantalla.

Los pasos son:

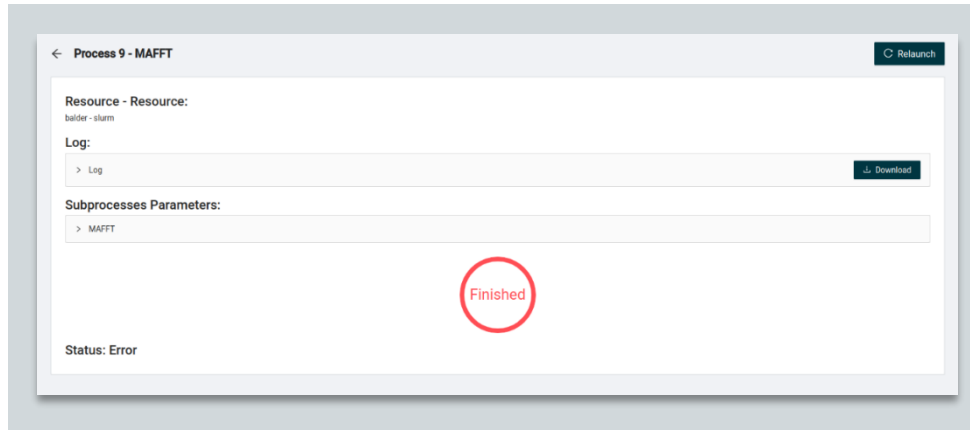
1. Proceso enviado.
2. Preparación del espacio de trabajo en el servidor remoto.
3. Exportación de los datos de entrada de la base de datos.
4. Transferencia de datos al servidor remoto.
5. Envío del proceso al servidor remoto.
6. Esperando a que el servidor ejecute el proceso enviado.
7. Esperando a que el servidor termine de ejecutar el proceso enviado.
8. Obtener resultados del proceso desde un servidor remoto.
9. Guardado de resultados en la base de datos.
10. Limpieza del espacio de trabajo del servidor remoto.

También existen tres estados de terminación diferentes:

1. Terminado.
2. Error.
3. Proceso cancelado.

Y pasos adicionales, activados cuando sucede algo inesperado:

1. Gestión de un error.
2. Proceso de cancelación.



Proceso completado con error

PROCESOS DISPONIBLES



NEXTGENDEM

18.PROCESOS DISPONIBLES

18.1 ALINEAMIENTO MÚLTIPLE DE SECUENCIAS

El alineamiento múltiple de secuencias (MSA) es una técnica bioinformática utilizada para alinear tres o más secuencias biológicas, como secuencias de ADN, ARN o proteínas, para identificar regiones conservadas y detectar relaciones evolutivas. Permite la comparación de secuencias y la identificación de patrones o motivos comunes, proporcionando información sobre la estructura, la función y la historia evolutiva de las secuencias.

CLUSTALW*

ClustalW (1) consta de tres etapas: alineamiento de secuencias por pares para calcular una matriz de distancia, cálculo de árbol guía a partir de la matriz de distancia y alineamiento de secuencias progresiva basada en el orden de ramificación en el árbol guía. Las puntuaciones de alineamiento por pares se calculan en función del número de k-tuplas coincidentes y una penalización fija para cada gap. El árbol guía se crea mediante el método Neighbor-Joining (2) y los pesos (3) se asignan a las secuencias en función de su distancia desde la raíz del árbol. El alineamiento progresivo implica el uso de programación dinámica para alinear grupos de secuencias cada vez más grandes siguiendo el orden de ramificación en el árbol guía.

PARÁMETROS

- **Sequences Selector:** Selector de secuencias
- **Output alignment format:** cambia el formato del archivo de salida. **Los valores posibles son:**
 - FASTA format: formato FASTA (predeterminado)
 - PHYLIP format: formato PHYLIP
 - Native Clustal output format: formato de salida nativo de Clustal

PARÁMETROS AVANZADOS

- **Output order:** se refiere al orden en que se escriben las secuencias en la alineación de salida. Valores posibles:
 - Aligned: orden en el que se alinearon las secuencias (a partir del árbol guía/dendrograma), agrupando así automáticamente las secuencias estrechamente relacionadas
 - Same order as input file: mismo orden que el archivo de entrada
- **Output complete alignment (or specify part to output):** seleccione para alinear las secuencias completas o parte de ellas. Los valores posibles son:
 - Complete alignment: Alineamiento completo.
 - Only part of the alignment: Solo parte del alineamiento. Si se selecciona esta opción, aparecerán dos nuevos campos para indicar el punto de inicio y el punto final de las secuencias de entrada.

* REFERENCIAS

1. Thompson, J D et al. "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." *Nucleic acids research* vol. 22,22 (1994): 4673-80. doi:10.1093/nar/22.22.4673 Saitou, N. and Nei, M. (1987) *Mol. Biol. Evol.* 4, 406-425.
2. Julie D. Thompson, Desmond G. Higgins, Toby J. Gibson, Improved sensitivity of profile searches through the use of sequence weights and gap excision, *Bioinformatics*, Volume 10, Issue 1, February 1994, Pages 19–29, <https://doi.org/10.1093/bioinformatics/10.1.19>
3. Eugene W. Myers, Webb Miller, Optimal alignments in linear space, *Bioinformatics*, Volume 4, Issue 1, March 1988, Pages 11–17, <https://doi.org/10.1093/bioinformatics/4.1.11>

MAFFT*

El paquete de software MAFFT (1) incorpora nuevas técnicas para el alineamiento progresivo e iterativo de secuencias de entrada utilizando el algoritmo Fast Fourier Transform (2). El paquete de software utiliza un árbol guía basado en la comparación de todos los pares de secuencias, que se construye a partir de una matriz de distancia utilizando el método UPGMA (3). MAFFT incluye tres métodos: FFT-NS-1, que alinea progresivamente las secuencias de entrada a lo largo del árbol guía; FFT-NS-2, que realinea las secuencias de entrada según el árbol guía inferido de FFT-NS-1; y FFT-NS-i, que mejora aún más el alineamiento obtenido por FFT-NS-2 utilizando la técnica de "particionamiento restringido dependiente de árbol" (4).

PARÁMETROS

- **Sequences Selector:** Selector de secuencias
- **Add Sequences to existing Alignment:** si se selecciona esta opción, aparecerá un selector de alineamiento y las secuencias seleccionadas en el selector de secuencias se agregarán a ese alineamiento.

PARÁMETROS AVANZADOS

Algorithm: la forma en que se alinean las secuencias. Los valores posibles son:

- Local Pair: todos los alineamientos por pares se calculan con el algoritmo de Smith-Waterman. Adecuado para un conjunto de secuencias localmente alineables.
- Global Pair: todos los alineamientos por pares se calculan con el algoritmo de Needleman-Wunsch. Adecuado para un conjunto de secuencias alineables globalmente.
- Generalized Affine Pair: todos los alineamientos por pares se calculan con un algoritmo local con el costo de gap afin generalizado (Altschul 1998). Adecuado cuando se esperan grandes espacios internos.
- FASTA Pair: Todos los alineamientos por pares se calculan con FASTA (Pearson y Lipman 1988).
- 6mer Pair: la distancia se calcula en función del número de 6mers compartidos. Aparecerán dos nuevos campos si se selecciona esta opción:
 - Repeat Guide Tree: Número de veces que se construye el árbol guía en la etapa progresiva.
 - PartTree Algorithm: utilice un método rápido de construcción de árboles (PartTree, Katoh y Toh 2007) con la distancia de 6mer. Recomendado para una gran cantidad (> ~10,000) de secuencias de entrada. Posibles opciones:
 - PartTree.
 - DPPartTree: El algoritmo PartTree se utiliza con distancias basadas en DP. Ligeramente más preciso y lento que PartTree.

*REFERENCIAS

1. Kazutaka Katoh, Kazuharu Misawa, Kei-ichi Kuma, Takashi Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Research*, Volume 30, Issue 14, 15 July 2002, Pages 3059–3066, <https://doi.org/10.1093/nar/gkf436>
2. Press,W.H., Teukolsky,S.A., Vetterling,W.T. and Flannery,B.P. (1995) *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edn. Cambridge University Press, Cambridge, UK.
3. Sokal,R.R. and Michener,C.D. (1958) A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin*, 28, 1409–1438.
4. Hirose,M., Totoki,Y., Hoshida,M. and Ishikawa,M. (1995) Comprehensive study on iterative algorithms of multiple sequence alignment. *Comput. Appl. Biosci.*, 11, 13–18.

18.2 FILOGENIAS

Las filogenias son diagramas ramificados que representan las relaciones evolutivas entre organismos o grupos de organismos. Muestran la ascendencia común y la evolución divergente de las especies, ilustrando cómo las diferentes especies se relacionan entre sí a través del proceso de evolución. Las filogenias se construyen sobre la base de características compartidas o información genética, y proporcionan una representación visual de la historia evolutiva y la relación de los organismos.

PAUP*

PAUP utiliza el método de parsimonia para reconstruir árboles filogenéticos. La parsimonia se basa en el principio de que la explicación más simple que da cuenta de los datos observados es la que tiene más probabilidades de ser correcta. En el contexto del análisis filogenético, esto significa que la historia evolutiva que requiere la menor cantidad de cambios evolutivos (como mutaciones, eliminaciones o inserciones) es la que tiene más probabilidades de ser correcta.

Para construir un árbol filogenético, PAUP toma como entrada un conjunto de secuencias de ADN alineadas para un grupo de organismos. Luego, PAUP busca el árbol que requiere la menor cantidad de cambios evolutivos para explicar los datos observados. Lo hace comparando diferentes árboles posibles y calculando una "puntuación de parsimonia" para cada árbol en función del número de cambios evolutivos necesarios.

PARÁMETROS

- **Alignment Selector:** Selector de alineamiento.
- **Taxset:** es un componente que realiza acciones sobre grupos de taxones. Se divide en dos bloques:
 - En el primer bloque es posible seleccionar uno o más taxones para ser el grupo externo y uno o más taxones para eliminar del alineamiento.
 - En el segundo bloque es posible seleccionar uno o más grupos monofiléticos. Estos grupos se tienen que añadir en orden, empezando por los grupos más pequeños y terminando con los más grandes.

The screenshot shows the PAUP software interface. At the top, there is a 'Taxset' section with a '+ Add' button. Below it is a table with columns: Key, Taxons, Type, Validation, and Actions. The table contains two rows: one for 'myOutgroup' with taxons 'Rumex lunaria, Salvia canariensis, Cymodocea nodosa' and type 'outgroup', and another for 'key_2' with taxons 'Ipomoea cairica, Sonchus oleraceus' and type 'delete'. Below the Taxset section is the 'Monophyletic Groups' section, which has a table with columns: Key, Type, Converse, Validation, and Actions. It contains one row for 'monophyletic_groups' with type 'monophyly'. Below this table is a 'Groups' section with a table containing one row for 'Spargula fallax, Oligomeris linifolia'.

PARÁMETROS AVANZADOS

- **Outgroup Rooting:** enraza el árbol en su grupo externo. Estas son las posibles opciones:
 - **Polytomy:** grupo externo en una politomía basal.
 - **Paraphyl:** exogrupo como grupo parafilético al endogrupo.

- Monophyl: el grupo externo es monofilético en relación con el grupo interno.
- Gap mode:
 - Missing: los espacios de alineamiento se tratan como datos faltantes.
 - New State: los espacios de alineamiento se tratan como un estado de carácter adicional (quinta base).
- Analysis method:
 - Simple Heuristic: método heurístico que intenta maximizar la parsimonia. Los métodos heurísticos no garantizan encontrar el óptimo, pero generalmente requieren mucho menos tiempo de computación. Si se selecciona este campo, aparecerá el siguiente campo:
 - Consensus Tree Type: Tipo de árbol de consenso. Los árboles de consenso son resúmenes jerárquicos de la información común a un conjunto de árboles "rivales". Tiene las siguientes opciones:
 - Strict: los árboles de consenso estricto contienen solo aquellos grupos que aparecen en todos los árboles rivales
 - Semistrict: este método corresponde al consenso de "componentes combinables" (Bremer, 1990)
 - Majority-rule: a diferencia del consenso estricto, puede ser de interés encontrar grupos que aparecen en un cierto porcentaje preestablecido de los árboles rivales.
 - Adams: El método de Adams (1972, 1986) (ver también Swofford, 1991) fue el primer método de consenso propuesto. Los árboles de Adams suelen conservar más estructura que los métodos estrictos.
- Bootstrap: el método consiste en muestrear el conjunto de datos original con reemplazo para construir una serie de réplicas de arranque del mismo tamaño que el conjunto de datos original. Cada uno de estos se analiza, y la variación entre estas estimaciones replicadas se toma como una indicación del error involucrado en hacer estimaciones a partir de los datos originales. Si se selecciona este campo, aparecerán los siguientes campos:
 - Number of replicas: Número de réplicas
 - Search Strategy: Estrategia de búsqueda. El método de búsqueda de árbol realizado para cada remuestreo de arranque de los caracteres. Las opciones son:
 - Heuristic: Lo mismo que la heurística simple anterior, pero utilizando todas las réplicas.
 - Branch and Bound: el método de ramificación y poda, que se utiliza con frecuencia para resolver problemas de optimización combinatoria, aparentemente fue aplicado por primera vez a los árboles evolutivos por Hendy y Penny (1982).
 - Fast Step: las búsquedas de árboles en cada replica se realizan añadiendo una secuencia aleatoria y sin intercambio de ramas
- Jackknife: este método consiste en formar nuevas muestras omitiendo, a su vez, una o más de las observaciones (caracteres observados) de los datos originales. Para cada una de las submuestras generadas, se puede reestimar el estimador

en estudio (topología de árbol), y la distribución de probabilidad así obtenida permitirá sacar conclusiones sobre la sensibilidad del estimador a las observaciones individuales. Idealmente, el resto de la filogenia debería permanecer sin cambios de submuestra a submuestra. Si se selecciona este campo, aparecerán ahora los mismos campos que aparecen si se selecciona *bootstrap*.

Los siguientes campos solo aparecerán si se selecciona una estrategia heurística ya sea en el campo *Analysis method* o *Search Strategy*:

- **Addition Sequence:** especifica la forma en que se seleccionan los taxones para añadirlos al árbol en el nodo actual del árbol de búsqueda. Las opciones posibles son:
 - **Simple:** primero se calcula la distancia entre cada taxón y un taxón de referencia; Farris llamó a esta distancia un índice de avance. Luego se agregan los taxones en orden de "avance" creciente.
 - **Closest:** inicialmente, se evalúan las longitudes de todos los árboles de tres taxones posibles, formados al unir un triplete de taxones terminales en un solo nodo interno. Los tres taxones que producen componen el árbol inicial. En cada paso sucesivo, se consideran todos los taxones no colocados restantes para conectarlos a cada rama del árbol, y se elige la combinación taxón-rama que requiere el menor aumento en la longitud del árbol.
 - **As is:** los taxones simplemente se agregan en el mismo orden en que se presentan en la matriz de datos, comenzando con los primeros tres y agregando secuencialmente el resto. Este método no suele ser muy eficaz.
 - **Random:** se utiliza un generador de números pseudoaleatorios para obtener una permutación de los taxones que se utilizará como secuencia de adición.
 - **Furthest:** lo opuesto a *closest*.

- **Branch Swapping:** una técnica para mejorar la estimación inicial de los árboles mediante la realización de conjuntos de reordenamientos predefinidos. Las opciones posibles son:
 - **No branch swapping:** Sin intercambio de ramas.
 - **Tree bisection-reconnection (TBR):** el árbol se divide en dos a lo largo de una rama, lo que produce dos subárboles disjuntos. Luego, los subárboles se vuelven a conectar uniendo un par de ramas, una de cada subárbol. Se evalúan todas las posibles bisecciones y reconexiones por pares.
 - **Subtree pruning-regrafting (SPR):** se poda un subárbol del árbol (por ejemplo, el subárbol que contiene los nodos terminales A y B como se indica). Luego, el subárbol se vuelve a injertar en una ubicación diferente en el árbol. Se evalúan todas las posibles eliminaciones de subárboles y puntos de reinserción.
 - **Nearest-neighbor interchange (NNI):** cada rama interna del árbol define una región local de cuatro subárboles conectados por la rama interna. Intercambiar un subárbol de un lado de la rama con uno del otro constituye un NNI

- **Número de árboles que se mantendrán:** el número de árboles cuya longitud es óptima que se guardarán para la siguiente ronda.

Los siguientes campos solo aparecerán si se selecciona Regla de la mayoría en el campo tipo de árbol de consenso:

- **Retain less than 50% of the trees:** si se especifica LE50 = Sí, los grupos que se encuentran en menos del 50 % de los árboles se retienen en el consenso si son compatibles con los grupos que ya están en el árbol.
- **Percentage of the trees on which a group must appear in order to be retained:** El valor entero especifica el porcentaje de los árboles en los que debe aparecer un grupo para ser retenido en el consenso de *Majority-rule*. Un grupo debe aparecer en más de este porcentaje de los árboles antes de que se retenga.

*REFERENCIAS

1. Swofford, DL 2003. PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Versión 4. Sinauer Associates, Sunderland, Massachusetts.

MRBAYES*

MrBayes es un paquete de software para la inferencia filogenética bayesiana. La inferencia filogenética bayesiana es un método estadístico que utiliza la teoría de probabilidad bayesiana para estimar la probabilidad de diferentes árboles evolutivos dado un conjunto de datos moleculares. MrBayes funciona utilizando un algoritmo Markov Chain Monte Carlo (MCMC) para tomar muestras de la distribución de probabilidad posterior de árboles posibles. La distribución posterior representa la probabilidad de cada árbol posible dados los datos moleculares y la información previa sobre la evolución de los organismos.

PARÁMETROS

- **Alignments Selector:** Selector de alineamientos.
- **Outgroup:** el outgroup o grupo externo es un taxón que está estrechamente relacionado con el ingroup o grupo interno pero que se encuentra fuera del grupo que se está estudiando. Se utiliza para enraizar el árbol y determinar la condición ancestral.

PARÁMETROS AVANZADOS

- **Number of substitution types:** Número de tipos de sustitución.
 - 1: Esto representa el JC69.
 - 2: Esto representa el K80.
 - 3: Esto representa el F81.
 - 4: Esto representa el HKY85.
 - 6: Esto representa el GTR (General Time Reversible).
- **Among-site variation model:** Modelo de variación entre sitios. Permite la incorporación de heterogeneidad en las tasas de sustitución en diferentes posiciones en el alineamiento, aceptando la suposición de que las tasas evolutivas pueden variar entre los sitios.
 - equal: asume una tasa única para todos los sitios en el alineamiento.
 - gamma: asume una distribución gamma de tasas entre sitios.
 - Inorm: asume una distribución logarítmica normal de tasas entre sitios
 - adgamma: asume una distribución gamma ajustada de tasas entre sitios.

- propinv: asume una proporción de sitios invariantes (sitios que no cambian) y una distribución gamma inversa de tasas para los sitios restantes.
 - invgamma: asume una distribución gamma inversa de tasas entre sitios.
 - kmixture: asume un modelo K-mixture de tasas entre sitios.
- **Number of generations:** número de generaciones o iteraciones en un análisis filogenético bayesiano. Es un parámetro que especifica el número total de generaciones de cadena de Markov Monte Carlo (MCMC) que se realizarán durante el análisis.
 - **Sample frequency:** especifica con qué frecuencia se registran y guardan los valores de los parámetros como salida durante la ejecución de MCMC.
 - **Burnin fraction:** la fracción de generaciones de MCMC que se descartan como *burnin* o quemados durante un análisis filogenético bayesiano.

* REFERENCIAS

1. Ronquist, F., and J.P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-157

BEAST2*

BEAST2 (Bayesian Evolutionary Analysis Sampling Trees 2) es un paquete de software utilizado para el análisis filogenético bayesiano. Es una poderosa herramienta que permite a los investigadores inferir relaciones evolutivas, estimar tiempos de divergencia y reconstruir estados ancestrales. BEAST2 implementa métodos computacionales avanzados, como el muestreo de cadena de Markov Monte Carlo (MCMC), para estimar parámetros e incertidumbre en modelos filogenéticos. BEAST2 se usa ampliamente en el campo de la biología evolutiva para abordar diversas preguntas de investigación y generar árboles filogenéticos sólidos e hipótesis evolutivas.

PARÁMETROS

- **Multiple Alignments Selector:** Es una modificación del *Alignment Selector* donde es posible seleccionar diferentes alineamientos para diferentes regiones genéticas ya sea un gen de cloroplasto o un gen ITS.
- **Monophyly:** Es un componente donde puedes crear grupos monofiléticos.

Este proceso es un pipeline de diferentes procesos creados en un gran proceso de investigación para encontrar la mejor manera de construir filogenias de la flora canaria. Este pipeline consta de los siguientes pasos:

1. Utiliza ModelTest-NG para seleccionar el modelo de evolución que mejor se ajuste a las alineaciones de ADN y proteínas.
2. Genera el código Beast para Path Sampling y Coupled MCMC.
 - a) El Path Sampling es un método computacional utilizado en las estadísticas bayesianas para estimar la probabilidad marginal o la evidencia de un modelo dados los datos. Implica integrar sobre un rango de diferentes valores de parámetros a lo largo de un camino continuo desde la distribución anterior a la distribución posterior. El Path Sampling proporciona una manera de comparar diferentes modelos y evaluar su ajuste relativo a los datos al cuantificar el soporte para cada modelo.

- b) El Coupled MCMC es una técnica utilizada para mejorar la eficiencia de los algoritmos de inferencia bayesianos. Implica ejecutar múltiples cadenas MCMC simultáneamente, donde las cadenas están "acopladas" o conectadas a diferentes temperaturas. Al permitir que las cadenas intercambien información entre estados de alta y baja temperatura, el MCMC acoplado mejora la exploración del espacio de parámetros y puede conducir a una convergencia más rápida y un mejor muestreo de la distribución posterior.

Ambos se crean dos veces, uno para cada uno de los siguientes modelos evolutivos:

- a) Birth Death Model: asume un proceso continuo de nacimiento y muerte en el que nuevas especies (nacimientos) surgen de especies existentes y las especies se extinguen (muertes) con el tiempo.
 - b) Yule Model: asume una tasa de especiación constante, donde cada especie tiene la misma probabilidad de dar lugar a dos nuevas especies en cada unidad de tiempo. No tiene en cuenta la extinción o los cambios en las tasas de especiación a lo largo del tiempo.
- i. Calcula el Coupled MCMC para ambos modelos evolutivos. Se detiene cuando se logra la convergencia o se ha ejecutado un gran número de iteraciones sin convergencia.
 - ii. Genera el árbol de consenso utilizando TreeAnnotator y selecciona las anotaciones de interés que son las que el usuario puede ver en el árbol filogenético resultante.
 - iii. Ejecuta Path Sampling para ambos modelos evolutivos y calcule el factor bayesiano para comparar ambos modelos. El usuario recibirá los dos árboles finales y un mensaje indicando cuál de los modelos evolutivos encaja mejor.

* REFERENCIAS

1. Drummond A. J. & Bouckaert R. R. (2015) "Bayesian evolutionary analysis with BEAST", Cambridge University Press.
2. Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular Biology and Evolution*, 37(1), 291-294. doi.org/10.1093/molbev/msz189

18.4 DIVERSIDAD FILOGENÉTICA

La diversidad filogenética se refiere a una medida de la biodiversidad que tiene en cuenta las relaciones evolutivas entre las especies. Cuantifica la cantidad de historia o linaje evolutivos único representado por un conjunto de especies o comunidades. Considera no solo el número de especies sino también su relación evolutiva.

18.5 PICANTE*

Picante es un paquete popular en R para analizar datos ecológicos y evolutivos. Proporciona una amplia gama de funciones para calcular diversas métricas filogenéticas y de diversidad, incluida la diversidad filogenética (diversidad alfa) y las medidas de distancia de la comunidad (diversidad beta).

PARÁMETROS

- **Phylogenetic Tree Selector:** Selector de árbol filogenético
- **Geolayer Selector:** Selector de capa geográfica

* REFERENCIAS

1. Steven W. Kembel, Peter D. Cowan, Matthew R. Helmus, William K. Cornwell, Helene Morlon, David D. Ackerly, Simon P. Blomberg, Campbell O. Webb, Picante: R tools for integrating phylogenies and ecology, *Bioinformatics*, Volume 26, Issue 11, June 2010, Pages 1463–1464, <https://doi.org/10.1093/bioinformatics/btq166>

18.6 IDENTIFICACIÓN DE SECUENCIA

Es un método utilizado para encontrar genes similares, identificar dominios funcionales en proteínas o determinar el origen taxonómico de una secuencia.

18.7 BLAST+*

BLAST (Herramienta básica de búsqueda de alineamiento local) es una herramienta de software ampliamente utilizada para comparar secuencias biológicas, como el ADN, con una gran base de datos de secuencias para identificar las más similares.

Además de la búsqueda de secuencias, BLAST+ también incluye herramientas para la creación de bases de datos, el formateo y la manipulación de secuencias y otras funciones relacionadas con el análisis de secuencias.

PARÁMETROS

- **Sequences Selector:** Selector de secuencias
- **Database:** base de datos de secuencias donde consultar, solo está disponible la del Jardín Botánico Viera y Clavijo.

PARÁMETROS AVANZADOS

- **Program:** algoritmo utilizado para el alineamiento de secuencias y la búsqueda de similitudes. Las opciones disponibles son:
 - Megablast: diseñado para la comparación rápida de secuencias de nucleótidos muy similares.
 - DC-Megablast: es una variante del algoritmo Megablast que está diseñado para manejar secuencias de nucleótidos con coincidencias discontinuas.

*REFERENCIAS

1. Camacho, Christiam et al. "BLAST+: *arquitectura y aplicaciones*". *BMC bioinformática* vol. 10 421. 15 de diciembre de 2009, doi:10.1186/1471-2105-10-421

18.8 MATRICES DISCRIMINANTES

La matriz discriminante consta de coeficientes de función discriminante que representan la contribución de cada posición del código de barras de ADN para distinguir entre diferentes especies. Estos coeficientes se calculan en base a secuencias de códigos de barras de ADN conocidas de diferentes especies.

18.9 MOLD*

El programa MOLD (MOlecular Diagnoses) construye diagnósticos basados en ADN a partir de un alineamiento de secuencias de ADN atribuidas previamente a taxones. MOLD se puede utilizar para diagnosticar taxones de género a especie o incluso subespecies. MOLD se subdivide funcionalmente en dos módulos. El primer módulo identifica múltiples "combinaciones

mínimas de nucleótidos de diagnóstico" (mDNC) para el taxón de consulta. Solo se utilizan dos tipos de caracteres de ADN para compilar mDNC: caracteres de Tipo 1 (sitio de diagnóstico puro, (2)), cada uno correspondiente a un mDNC, o Tipo 5 ("caracteres" en el análisis de agregación de población (3).) caracteres que generan mDNC compuestos. Ambos tipos de caracteres no varían entre los especímenes de consulta, y esto es esencial para la operatividad de un diagnóstico. El segundo módulo transforma el catálogo de mDNC en un conjunto de "combinaciones de nucleótidos de diagnóstico redundantes" (rDNC) y calcula una puntuación para cada uno de ellos.

PARÁMETROS

- **Alignment Selector:** Selector de alineamiento

* REFERENCIAS

1. A.E. Fedosov, Guillaume Achaz, Andrey Gontchar, Nicolas Puillandre. MOLD, a novel software to compile accurate and reliable DNA diagnoses for taxonomic descriptions. *Molecular Ecology Resources*, 2022, 5, pp.2038-2053. [ff10.1111/1755-0998.13590](https://doi.org/10.1111/1755-0998.13590)[ff. fahal-03663253f](https://doi.org/10.1111/1755-0998.13590)
2. Sarkar I.N., Planet P.J., DeSalle R. 2008. CAOS software for use in character-based DNA barcoding. *Molecular Ecology Resources*. 8, 1256-1259.
3. Davis J.I., Nixon K.C. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Systematic Biology*, 41: 421–35.

CARACTERÍSTICAS DESCARTADAS



NEXTGENDEM

19. CARACTERÍSTICAS DESCARTADAS

Durante la vida del proyecto NEXTGENDEM, se desarrollaron y descartaron varios módulos, lo cual es un proceso natural en la creación de una herramienta con una visión evolutiva. En este apartado se mencionan dichas experiencias, por si pudiera ser de interés para retomar el desarrollo en otras iniciativas (proyectos, estudios...).

- Integración con **Proyecto Galaxia**. El primer tipo de recurso desarrollado fue este, debido a la flexibilidad y comunidad detectada en el momento del desarrollo inicial.
- Integración de **JBrowse** como visor de alineamiento múltiple de secuencias. El sistema actual presenta un potente editor mucho más adecuado para el propósito del proyecto NEXTGENDEM que las funciones proporcionadas por JBrowse, que pueden encajar más en el análisis de genomas completos y/o en genética de poblaciones.
- Una **imagen de Docker** creada para ejecutar geoprosesos. Esta imagen se creó en otro contexto, el cambio climático, y no ha sido necesaria hasta ahora, donde el foco está mayormente centrado en la ecología



NEXTGENDEM



MAC 2014-2020
Cooperación Territorial

Interreg

Fondo Europeo de Desarrollo Regional



www.plataforma-nextgendem.com