JORNADAS SOBRE INFORMACIÓN DE BIODIVERSIDAD Y ADMINISTRACIONES AMBIENTALES

Del dato local al dato reutilizable

Integración de CromoCat y ChromoDB en GBIF

Joan Simon · Universitat de Barcelona

Barcelona · 25 de noviembre de 2026

CromoCat
Chromosome Database of the flora of the Catalan Countries

ChromoDB

GBIF

Universitat de Barcelona

¿Podemos gestionar sin conocer la diversidad genética?

- *Rutidosis lanata:* asterácea rara y amenazada de Australia.
- Poblaciones con distinta ploidía (4x, 5x, 6x): mezclarlas puede reducir su viabilidad.



CC BY-NC 4.0 Martin Bennett

Ecology and Evolution

ORIGINAL RESEARCH | Open Access

Genetic rescue in a plant polyploid complex: Case study on the importance of genetic and trait data for conservation management

Alexander N. Schmidt-Lebuhn, David J. Marshall, Brad Dreis, Andrew G. Young

First published: 25 April 2018 | https://doi-org.sire.ub.edu/10.1002/ece3.4039 | VIEW METRICS

Esquema conceptual

4x + 4x → ✓ Mezclar

4x + 5x → ⚠ Separar

La biodiversidad también existe dentro de las especies

Conocer la variación es el primer paso para priorizar y gestionar

- Río de Janeiro, 1992 (en vigor desde 1993): tres niveles de biodiversidad
- La diversidad genética permite adaptarse, reproducirse y resistir los cambios

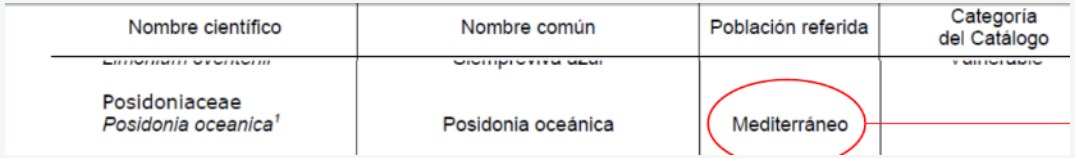


Conocer → **Priorizar** → **Gestionar**

Hoy: el primer paso

Marco legal y diversidad genética

- **Ley 42/2007**: diversidad genética, de poblaciones y de especies
- **LESPRE y CNEA**: taxones o poblaciones
- **Cataluña, 2023**: el Catálogo de Flora Amenazada incorpora información genética



Nombre científico	Nombre común	Población referida	Categoría del Catálogo
Posidoniaceae *Posidonia oceanica*[1]	Posidonia oceánica	Mediterráneo	

LESPRE: protección de una población (Mediterráneo)

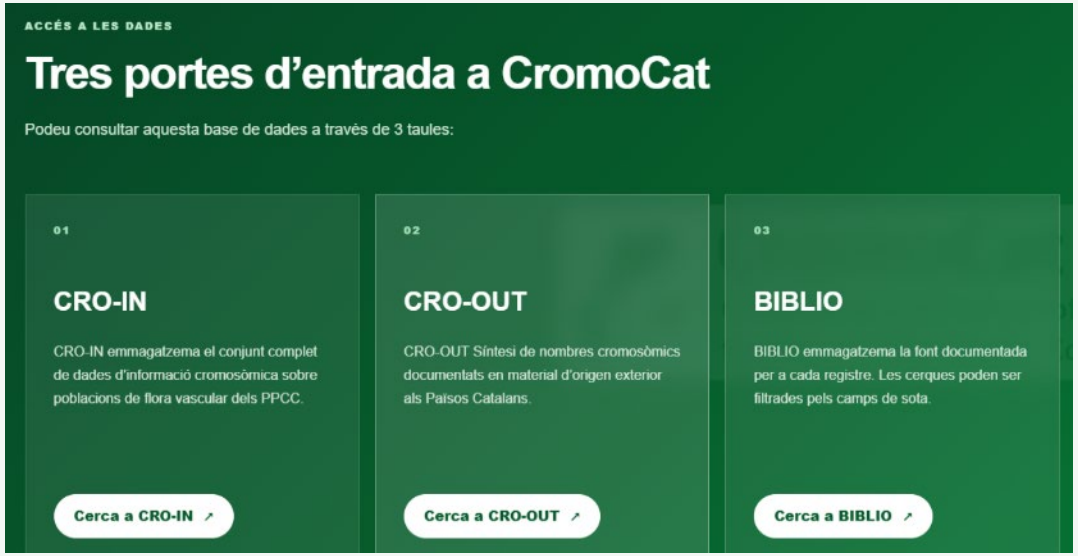
Fenotipo

Marcadores moleculares

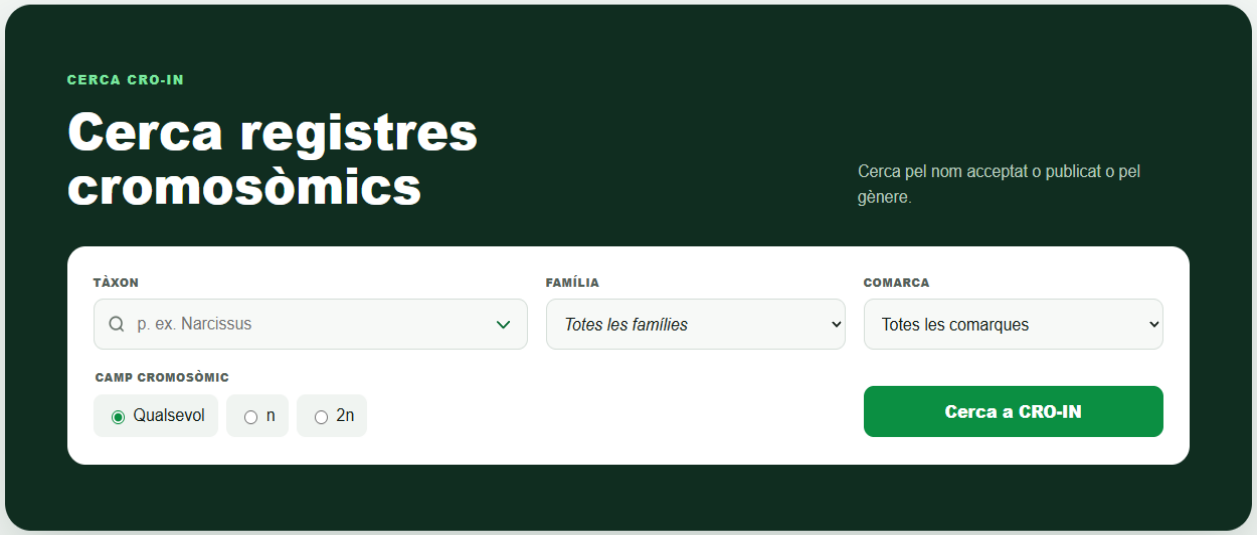
Citogenética: nivel básico
ploidía · nº cromosómico

Cromo.Cat: la base de datos regional

- *Països Catalans*, desde 1998
- Módulo de BDBC (UB – Dr. Xavier Font)
- CRO-IN: taxones del territorio
- CRO-OUT: taxones de fuera
- más de 71.500 registros y más de 5.000 taxones (estado a noviembre de 2025)







Del territorio al mundo

La misma exigencia: cada dato enlazado a su publicación

CromoCat

1998 →

Datos del territorio

→

DCDB

2006

Recuentos de *Delphinieae*

→

ChromoDB

2024 - GLOBAL

Plantas vasculares de todo el mundo

→

GBIF

2026 - ABIERTO

Publicación interoperable

La trayectoria no es una simple suma de registros: cambia la escala y la forma de compartirlos.

REPOSITORIO WEB · CHROMODB.NET

≈ 217.000

- Unidad: **1 recuento por publicación y país**
- Incluye híbridos y taxones aún sin resolver
- Cada registro enlaza con POWO y GBIF

GBIF V2.5 · NOVIEMBRE 2026

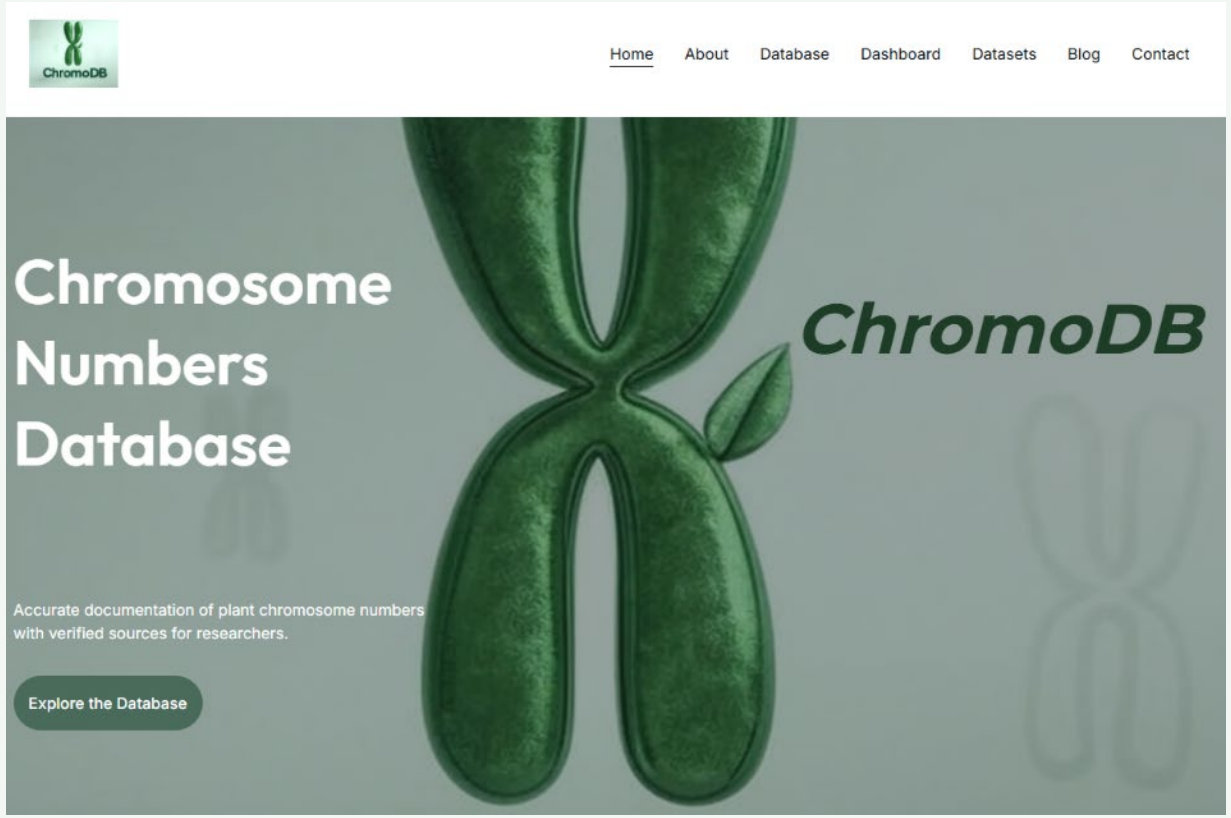
213.440

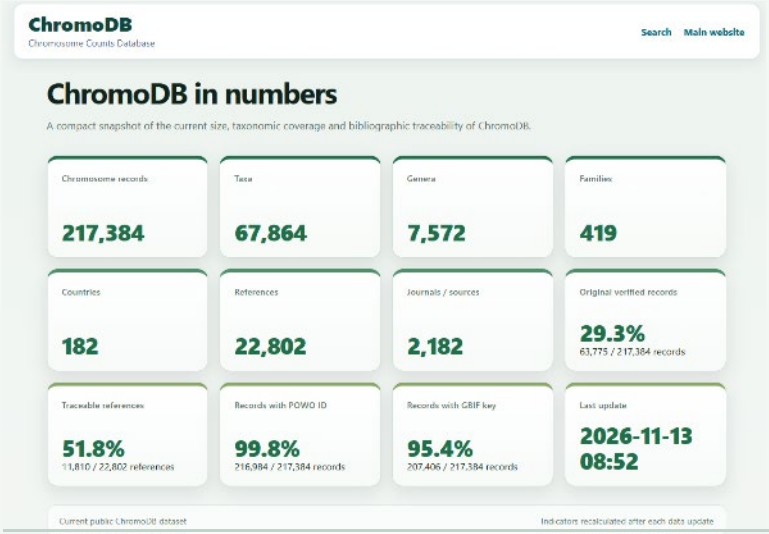
- **Código POWO** obligatorio (incluidos híbridos)
- Geografía estandarizada a nivel de país
- Nivel de confianza, DOI y enlace al registro

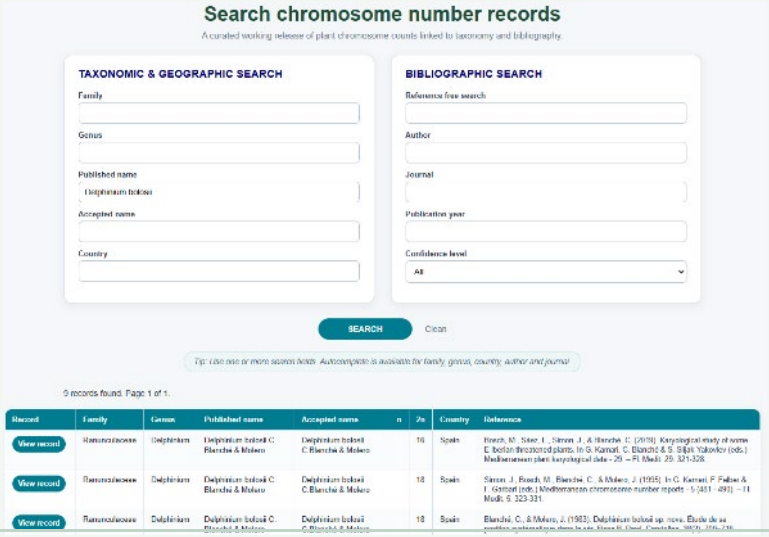
Simon et al., 2023 · DCDB / ChromoDB

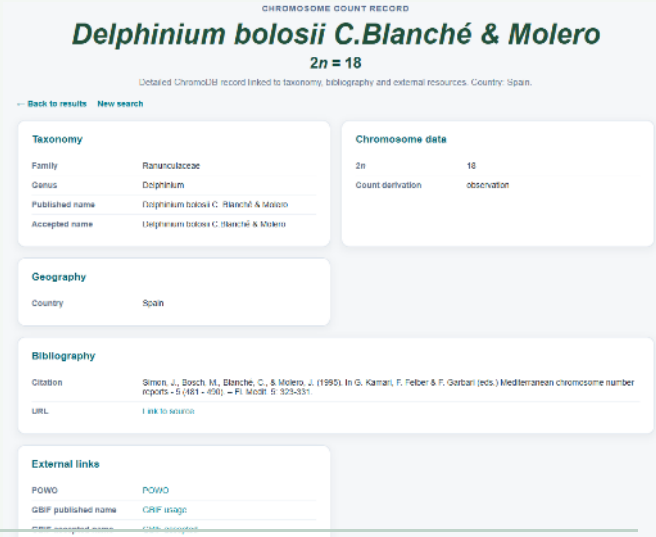
ChromoDB.net hoy

- Nombre publicado y aceptado
- Valor tal como aparece en la publicación
- Cada registro verificado enlaza con su fuente bibliográfica original
- Incorpora tamaño de genoma y genomas ensamblados









ChromoDB: chromosome counts database for vascular plants

Published by Botany Section, Department of Biology, Health and Environment (University of Barcelona)

Simon J • Bosch M • Goncalves E • Blanché C chromodb.net Licence CC BY 4.0

2 citations 213,440 occurrences

About | Project | Metrics | Download

Description

Publication date: 13 November 2026

ChromoDB is a continuously curated database of published chromosome numbe
Universitat de Barcelona from the experience of two earlier chromosome databas
Catalan Countries, and DCDB, for tribe Delphinieae (Ranunculaceae). Each record
result for one taxon and one bibliographic source and, when reported, one count
preserve the identification under which the count was published and link it to the
Online (POWO), normally its currently accepted taxon. They also document the b
and a confidence level that distinguishes records verified against the original pub
yet verified. This dataset is a periodic, versioned release of the ChromoDB public
web database and includes only records meeting the taxonomic, bibliographic ar
release. Every GBIF record links to its record page in ChromoDB, where correctior
The release contains 213,440 records for 64,421 taxa, 7,427 genera and 386 famil
published between 1882 and 2026; 28.2% of the records have been verified agair

Other

Record licence	CC BY 4.0
Institution code	UB-BSMA-BOT Institution match none
Institution ID	https://www.gbif.org/publisher/26c1df8b-7ba5-4121-8566-c25bf90bb4c7 Institution match none
Dataset name	ChromoDB: chromosome counts database for vascular plants
Identifier	chromodb-343417
Modified	2026-11-13
References	https://database.chromodb.net/record.php?id=chromodb-343417
GBIF ID	6430064321

Reference

Source	http://www.herbmedit.org/flora/10-381.pdf
Bibliographic citation	Tahiri, H., & Cubas, P. (2000). In G. Kamari, F. Felber & F. Garbari (eds.) Mediterranean chromosome number reports - 10 (1201 - 1207). – Fl. Medit. 10: 415-419.

ChromoDB en GBIF

Statistics

Occurrence records	213,440
Species	58,080

Data richness

Number of occurrences

Identified to species	213,069
With coordinates	0

Principios FAIR

F **Encontrable:** DOI y GBIF

A **Accesible:** chromodb.net y GBIF para descarga

I **Interoperable:** Darwin Core y taxonomía vinculada

R **Reutilizable:** licencia, cita, versión y fuente

Citation: *Please use this citation in publications*,
Simon J, Bosch M, Goncalves E, Blanché C (2026). ChromoDB: chromosome counts database for vascular plants. Version 2.5. Botany Section, Department of Biology, Health and Environment (University of Barcelona). Occurrence dataset https://doi.org/10.15470/ninpvc accessed via GBIF.org on 2026-11-16. https://gbif.org/occurrence/6430064319

Más de un siglo de literatura botánica no se revisa a mano



Miles de artículos históricos, fuentes dispersas y datos difíciles de comparar exigen una nueva escala de revisión.

EL RETO

- **Volumen de literatura histórica**
- Fuentes y formatos heterogéneos
- Errores de origen que revisar

EL CAMBIO DE ESCALA

Revisión retrospectiva sistemática

Posible a escala de la bibliografía

CAPACIDAD La IA amplía el alcance de la revisión documental.

La IA trabaja; las personas deciden

La automatización no delega la responsabilidad científica.

Reparto claro de responsabilidades

LA IA

CAPACIDAD DE LECTURA

Localiza · lee · extrae · compara

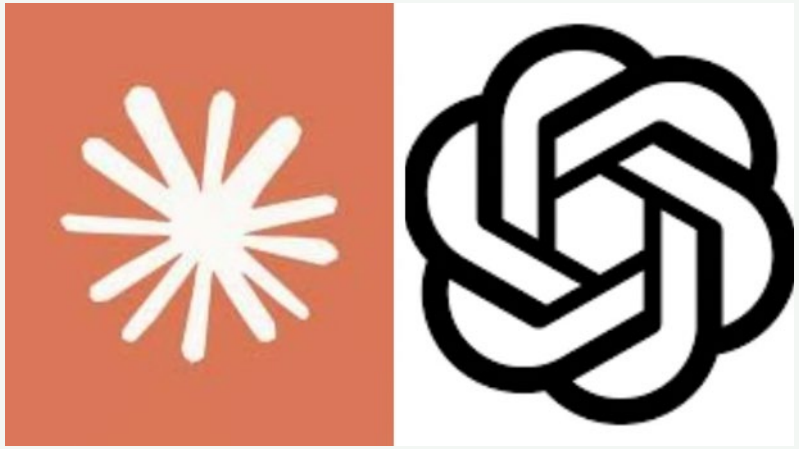
LAS PERSONAS

CRITERIO CIENTÍFICO

Fijan reglas · validan y deciden

≈ 30 % de ChromoDB revisado en el original

La fuente original sigue siendo la autoridad.



Extracción con IA (cuatro capas de verificación)

PDF · HTML
texto, OCR
si hace falta

→

IA
tabla
estructurada

→

Tabla
valor · país ·
referencia

→

POWO · GBIF
nombres

- Lotes pequeños (límite: 200-300 páginas)
- Valor literal; discrepancias anotadas siempre en *remarks*
- Sinonimias: POWO y GBIF, mediante scripts y API

1. **Dos sistemas de IA (los últimos modelos de ChatGPT y Claude), por separado**
2. **Cruce con otras fuentes: lo que no concuerda, se revisa**
3. **Muestra revisada a mano [10 % de los registros]**
4. **Campo de incidencias de la fuente**

Por qué necesitamos IA: errores y escala

La fuente también se equivoca

- Revista *Taxon*: resumen en papel y suplemento electrónico
- En una quincena de números, casi todos con incongruencias

Papel resumen ≠ **Suplemento** electrónico

⚠ **Números cromosómicos**

⚠ **Coordenadas inexistentes**

⚠ **Pliegos de herbario**

Lagascalia (series)

35 años (≈ 850 registros)

Vaciado completo; con revisión sistemática (*dataset* en Zenodo)

Fedorov (recopilación)

70 años (≈ 79.000 reg.)

Vaciado completo; aún pendiente de revisión sistemática

Iranian Journal of Botany

50 años (≈ 800 reg.)

Revisión histórica del corpus (*dataset* en Zenodo)

Botanicheskii Zhurnal

110 años (≈16.000)

En proceso de extracción / comprobación

¿Y CRO-IN? Una integración avanzada... todavía en revisión

≈ 90 % en GBIF

La segunda auditoría con IA busca enriquecer la ficha, no inventar nuevos recuentos

Mayor trazabilidad

- Valor tal como lo da la publicación
- Tipo de evidencia
- Nivel de confianza
- Enlace a la publicación original

Segunda revisión de >500 artículos

- **fotografías de cariotipos**
- **cariogramas e idiogramas**
- **localidades y contexto original**

Caso 1: *Delphinium bolosii*

- Endemismo catalán raro
- 2n = 18: disploidía excepcional en el género (tesis doctoral Blanché, 1985)
- 2n = 16 en tres poblaciones (Bosch et al., 2019)
- ¿Mezclar o mantener separadas?

Consulta en ChromoDB	
2n = 18	localidad · fuente ↩
2n = 16	localidad · fuente ↩



Esperó de Bolòs (*Delphinium bolosii*) – Maria Bosch

Caso 2: citotipos

- *Allium oleraceum*: 10 registros, 3 niveles de ploidía x = 8 (**4x, 5x**, 6x) por citometría de flujo
- Alta Ribagorça, Pallars Sobirà, Cerdanya, La Selva, Ripollès, Vallès Or. y Oc. (CRO-IN, estado en 2025)
- CromoCat ubica comarcas; ChromoDB detecta vacíos globales de muestreo





- Duchoslav M., Šafářová L., Krahulec F. (2010). Ann. Bot. 105: 719-735.
- Duchoslav M. et al. (2020). Front. Plant Sci. 11: 591137.
- Duchoslav M., Šafářová L., Jandová M. (2013). Genome size variation in six cytotypes of A. oleraceum.

(cc) BY-NC jansson

Caso 3: el límite del recuento

- *Saxifraga geranioides* y *S. genesiana*: mismo número (2n ≈ 44)
- El tamaño del genoma no las distingue, pero sí detecta híbridos
- ChromoDB guarda recuentos y otras evidencias, marcadas



⚠ **Nº cromosómico: igual**

⚠ **Tamaño del genoma: igual**



Butlletí de la Institució Catalana d'Història Natural, 78: 51-61. 2014 ISSN 2013-3987 (online edition): ISSN: 1133-6889 (print edition)

GEA, FLORA ET FAUNA

Relacions entre diverses espècies de *Saxifraga* dels Pirineus i de les muntanyes Catalanídiques

Gemma Mas de Xaxars Giner*, Joan Vallès Xirau*, Teresa Garnatje Roca**, Arnau Mercadé López***, Joan Martín Villodre*, Pere Barnola Echenique**** & Josep Vigo Bonada*****





Publicado no equivale a verificado

La apertura y el control de calidad son compromisos distintos

PUBLICAR

Hacer el dato FAIR (encontrable, accesible, interoperable y reutilizable).

VERIFICAR

Comprobar el valor y el contexto frente a la fuente original.

Y conservar las dos escalas: evidencias globales + detalle de cada población.

Integrar no es copiar datos.

Es conectar lo local con lo global sin perder la fuente.

CromoCat	ChromoDB	GBIF
detalle poblacional	síntesis mundial	acceso y reutilización

1. **Un recuento puede revelar diferencias poblacionales decisivas para conservar.**
2. **La calidad y la continuidad requieren recursos y financiación estables.**
3. **Son necesarios datos fiables, actualizados y útiles para la gestión de nuestro patrimonio.**

Gracias

- https://cromo.cat
- https://chromodb.net
- DOI en GBIF: 10.15470/ninpvc
- Autores
  - Joan Simon
  - Cèsar Blanché
  - Maria Bosch
  - Estefany Goncalves

