

La iniciativa catalana per a l'Earth BioGenome Project

Elisabet Tintó-Font,¹ Helga Simon-Molas,² Oriane Hidalgo,³ Roderic Guigó⁴ i Montserrat Corominas⁵

¹ Institut de Salut Global de Barcelona (ISGlobal), Hospital Clínic - Universitat de Barcelona

² Department of Experimental Immunology and Department of Hematology, Amsterdam UMC, Amsterdam, Països Baixos

³ Institut Botànic de Barcelona (IBB), CSIC - Ajuntament de Barcelona

⁴ Centre de Regulació Genòmica (CRG), Universitat Pompeu Fabra i Institut d'Estudis Catalans

⁵ Universitat de Barcelona i Institut d'Estudis Catalans

Correspondència: Montserrat Corominas. Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Institut de Biomedicina de la Universitat de Barcelona (IBUB). Av. Diagonal, 643. 08028 Barcelona. Adreça electrònica: mcorominas@ub.edu.

DOI: 10.2436/20.1501.02.210

ISSN (ed. impresa): 0212-3037

ISSN (ed. digital): 2013-9802

<http://revistes.iec.cat/index.php/TSCB>

Rebut: 08/03/2022

Acceptat: 05/04/2022

Resum

Els potents avenços en les tecnologies de seqüenciació de genomes, juntament amb la reducció de costos, permeten per primera vegada a la història utilitzar la genòmica per a ajudar a caracteritzar molecularment tant les espècies conegudes de la Terra com les que encara resten per identificar. L'Earth BioGenome Project (EBP) és una iniciativa internacional que aspira a seqüenciar, catalogar i caracteritzar els genomes de tota la biodiversitat eucariota de la Terra. Estructurada com una xarxa internacional de xarxes, vol crear una nova base per a la biologia, per a buscar les bases moleculars de molts caràcters únics a les espècies, per a trobar solucions per a preservar la biodiversitat i per a contribuir, finalment, al benestar i a la millora econòmica de les nostres societats. La iniciativa catalana per a l'Earth BioGenome Project és un dels nodes de la xarxa que té per objectiu la caracterització i la catalogació del genoma de les espècies eucariotes que viuen als territoris de parla i cultura catalanes.

Paraules clau: biodiversitat, genoma, genòmica, seqüenciació.

Introducció

Que la biodiversitat de la Terra està minvant és un fet. En els darrers quaranta anys, s'ha perdut el 60 % de la població d'espècies salvatges —un 52 % en el cas dels vertebrats— i actualment hi ha més de 35.000 espècies en perill d'extinció (https://nc.iucnredlist.org/redlist/resources/files/1630480997-IUCN_RED_LIST_QUADRENNIAL_REPORT_2017-2020.pdf). El canvi climàtic, la destrucció d'hàbitats i l'explotació d'algunes espècies amb finalitats econòmiques són només alguns exemples de com l'activitat humana és un dels principals causants de la pèrdua d'espècies. Segons l'entomòleg Edward Osborne Wilson, reconegut mundialment pels seus estudis sobre la biodiversitat, ens trobem davant de la sisena gran extinció en la història de la Terra (Wilson, 1999). En aquest escenari, és necessari que actuem per a preservar i conservar els ecosistemes, tant per al futur de la Terra com també per a la supervivència de l'espècie humana. Per

tal de donar resposta a aquesta emergència, primer cal conèixer les espècies que habiten la Terra, i una de les maneres més potents de fer-ho és estudiant-ne el genoma.

Malgrat els més de dos-cents cinquanta anys de classificació taxonòmica i més d'1,2 milions d'espècies catalogades, s'ha suggerit que un 86 % de les espècies existents en la totalitat de la Terra i el 91 % de les espècies de l'oceà encara esperen ser descrites (Mora *et al.*, 2011). Els grans avenços en la tecnologia de seqüenciació de genomes, juntament amb la informàtica, l'automatització i la intel·ligència artificial, permeten per primera vegada a la història afrontar un dels reptes científics i socials més importants actualment: augmentar la nostra comprensió de la biodiversitat i millorar-ne sensiblement l'ordenació taxonòmica.

L'Earth BioGenome Project

L'Earth BioGenome Project (EBP; www.earthbiogenome.org) és una iniciativa a escala mun-

The Catalan Initiative for the Earth BioGenome Project

Abstract

The major advances in genome sequencing technologies, together with the reduction of the costs involved, allow genomics to be used for the first time in history to help to characterise molecularly both the known species on Earth and those still remaining to be identified. The Earth BioGenome Project (EBP) is an international initiative that aims to sequence, catalogue and characterise the genomes of all the Earth's eukaryote biodiversity. Structured as an international network of networks, it seeks to create a new basis for biology in order to look for the molecular bases of many unique characters in species, to find solutions for preserving biodiversity, and to contribute to the welfare and the economic improvement of our societies. The Catalan Initiative for the Earth BioGenome Project forms one of the nodes of the network. Its goal is to characterise and catalogue the genome of the eukaryote species living in the territories of Catalan language and culture.

Keywords: biodiversity, genome, genomics, sequencing.

dial impulsada l'any 2017 per la Universitat de Califòrnia a Davis (UC Davis), la Smithsonian Institution i la Universitat d'Illinois a Urbana-Champaign, que té com a objectiu global seqüenciar, catalogar i caracteritzar el genoma de tota la biodiversitat eucariota de la Terra. L'abril de l'any 2018, la revista científica *Proceedings of the National Academy of Science* (PNAS) va publicar l'article original en què els coordinadors de l'EBP, juntament amb altres investigadors de set països, detallaven els objectius del projecte, així com les fases en què es desenvoluparia i el retorn a la societat que podria comportar (Lewin *et al.*, 2018). El títol de l'article, que ha esdevingut el lema del projecte, és prou suggestiu de les implicacions d'aquesta iniciativa: «The Earth BioGenome Project: Sequencing life for the future of life», és a dir, seqüenciar la vida per al futur de la vida. Més específicament, els objectius de l'EBP es poden concretar a: 1) revisar i aportar una nova visió de la comprensió de la biologia,

els ecosistemes, la filogènia i l'evolució; 2) permetre la conservació, la protecció i la regeneració de la biodiversitat; 3) maximitzar la rendibilitat per a la societat i el benestar humà.

L'EBP s'estructura com una xarxa internacional de xarxes, i fomenta la implicació d'institucions i centres d'arreu del món, indispensable per tal de complir els objectius del projecte. Els nodes d'aquesta xarxa estan inclosos en el Consell Coordinador de l'EBP i poden representar una localització geogràfica (Canadà, Àfrica, etc.) o incloure comunitats i organitzacions relacionades amb tàxons (vertebrats, insectes, plantes, etc.). Entre les funcions de l'EBP hi ha: desenvolupar i promoure estàndards per a la producció escalable de genomes de qualitat de referència; fer difusió de les millors pràctiques; coordinar les activitats de seqüenciació, anotació, anàlisi de dades i formació; facilitar l'accessibilitat pública de les dades i estimular la comunicació sobre el progrés del projecte (Lewin *et al.*, 2022). El funcionament de l'EBP es vertebrava en vuit comitès: recollida i processament de mostres; seqüenciació i assemblatge; anotació de gens; anàlisi de dades; tecnologies de la informació i informàtica; qüestions ètiques, legals i socials; justícia, equitat, diversitat i inclusió, i comunicació i assumptes públics.

Cal destacar que tota la informació generada per l'EBP és pública i accessible, des del primer moment, amb l'objectiu del benefici comú. Això implica un grau elevadíssim de coordinació i col·laboració, i fa absolutament indispensable que les dades siguin compatibles i comparables des del moment inicial en què es generen. Els nodes de la xarxa implicats en l'EBP han de respectar el Protocol de Nagoya sobre Accés i Intercanvi de Beneficis (ABS, de l'anglès *access and benefit-sharing*). Aquest protocol de 2010 és un acord complementari al Conveni sobre la Diversitat Biològica (CBD, de l'anglès *Convention on Biological Diversity*) de 1992 i promou la implementació d'un dels tres objectius del CBD: el repartiment just i equitatiu dels beneficis derivats de la utilització dels recursos genètics, per a contribuir així a la conservació i l'ús sostenible de la biodiversitat.

Una vegada recollides, les mostres s'han d'acompanyar de metadades robustes i completes. Això inclou identificar-ne l'espècie i fer el dipòsit en institucions públiques d'una mostra de cada exemplar, com a referència permanent i revisable del treball. Els materials frescs s'han de processar en tubs sobre gel sec i a partir d'aquest punt, mantenir-los a -80°C o en

nitrogen líquid, malgrat que s'estan estudiant altres maneres de preservar-los en situacions en què aquest tipus de preservació no sigui possible. L'extracció de DNA d'alt pes molecular i de bona qualitat és absolutament necessària per a poder generar seqüències que segueixin els estàndards quantitius establerts pel comitè corresponent. Les directrius actuals d'assemblatge consisteixen a generar una combinació de tipus de dades que inclouen lectura llarga (metodologia PacBio HiFi i/o ONT *ultra long*), de llarg abast (Hi-C) i RNA-seq (lectures curtes d'Illumina, PacBio Iso-Seq o ONT cDNA-PCR) del mateix exemplar sempre que sigui possible. Un accés obert a les dades que respecti els principis ètics és imprescindible perquè es compleixi un dels objectius principals del projecte: contribuir al coneixement i a la ciència. Per tant, una vegada obtingudes, les seqüències dels genomes de referència s'han de dipositar en alguna de les bases de dades públiques següents: l'Arxiu Europeu de Nucleòtids, de l'Institut Europeu de Bioinformàtica del Laboratori Europeu de Biologia Molecular, ENA (EMBL-EBI) (www.ebi.ac.uk/ena/); Genbank, del Centre Nacional per a la Informació Biotecnològica (NCBI) (www.ncbi.nlm.nih.gov/genbank/), o DNA Databank of Japan (www.ddbj.nig.ac.jp). Aquestes dades de seqüència, conegudes com a *informació de seqüència digital* (DSI, de l'anglès *digital sequence information*), són clau per al progrés científic i la innovació tecnològica en camps tan diversos com la medicina, la seguretat alimentària, la producció d'energia verda i la conservació de la biodiversitat. Les regles per a accedir a la DSI, però, generalment no són clares i per aquesta raó s'ha proposat recentment un marc per a l'ús de la DSI (Scholz *et al.*, 2022) que es discutirà els propers mesos a la reunió de la CBD, on assistiran representants de l'EBP.

L'obtenció de la seqüència dels genomes de totes les espècies eucariotes tindrà una gran influència en molts aspectes de la biologia. Qüestions relacionades amb l'evolució o l'ecologia només seran abordables quan es disposi de dades de genomes sencers que permetin identificar divergències en totes les ramificacions de l'arbre de la vida o totes les espècies dels ecosistemes naturals. Tal com es conclou a l'article publicat recentment amb el títol «Why sequence all eukaryotes?» (Blaxter *et al.*, 2022), aquesta biblioteca fonamental d'informació canviarà l'economia i el creixement social del futur, fomentant l'agricultura sostenible i noves bioeconomies, accedint a una

farmacopea ampliada i promovent l'equitat i la diversitat de la societat a través de les ulleres d'una biodiversitat molt valorada.

El novembre de 2020 va fer dos anys del llançament de l'EBP i s'han fet avenços significatius en tots els aspectes del full de ruta (Lewin *et al.*, 2022). En aquests moments, l'EBP té quaranta-nou projectes afiliats, repartits per tots els continents i que cobreixen una diversitat de tàxons. Un dels més avançats és el Darwin Tree of Life (DTOL; www.darwintreeoflife.org/), que pretén seqüenciar els genomes de les setanta mil espècies d'organismes de la Gran Bretanya i Irlanda (The Darwin Tree of Life Project Consortium, 2022). Cal tenir en compte, però, que alguns dels països amb més representació de la biodiversitat de la Terra es troben en zones no industrialitzades. L'EBP contribuirà a destacar aquests països i a fer-los a tots més conscients de la necessitat de treballar sense oblidar ningú.

La iniciativa catalana per a l'Earth BioGenome Project

La iniciativa catalana per a l'Earth BioGenome Project (CBP, de l'anglès *Catalan Initiative for the Earth BioGenome Project*; www.biogenoma.cat) és un projecte afiliat a l'EBP que té com a objectiu seqüenciar el genoma de les més de setanta mil espècies eucariotes que es calcula que viuen als Països Catalans. La CBP es va plantejar inicialment des de l'Institut d'Estudis Catalans (IEC), concretament des de dues de les seves societats filials: la Societat Catalana de Biologia (SCB) i la Institució Catalana d'Història Natural (ICHN). La iniciativa es va presentar en el marc del congrés Genomics for Biodiversity, que va tenir lloc a la seu de l'IEC el setembre de 2019. Actualment, la CBP és un projecte col·laboratiu que compta amb el suport i la participació directa d'una trentena d'institucions i d'una xarxa d'un centenar de membres d'arreu dels territoris de parla catalana. Es tracta d'un projecte ambiciós que, per tal d'implementar-se amb èxit, té dos pilars fonamentals: la rica i extensa tradició naturalista del nostre territori i les potents infraestructures tecnològiques de què disposem.

L'àmplia diversitat climàtica, topogràfica i geològica dels Països Catalans ha fet que s'hi concentri una biodiversitat molt rica. Des dels cims pirinencs de més de 3.000 metres d'altitud fins a les costes, illes mediterrànies i la mar que les envolta, passant per boscos atlàntics i per ambients subàrids, la diversitat de formes de vida hi adquireix un dels màxims exponents. Aquest territori es troba a la intersec-

ció de les plaques europees i africanes (Casas-Sainz i Vicente, 2009), i a la cruïlla entre les regions biogeogràfiques eurosiberiana i mediterrània. Els Països Catalans representen un punt calent (*hotspot*) de biodiversitat: tot i que cobreixen menys de l'1 % del territori europeu (70.520 km², 2.500 km de vora del mar), són la llar d'aproximadament una quarta part de totes les espècies eucariotes europees conegudes. També es caracteritzen per un alt nivell d'endemisme. Per exemple, de les 7.500 plantes vasculares estimades en aquest territori, unes 5.500 són autòctones i més de 300 es consideren endèmiques (Peñuelas *et al.*, 2019). Moltes espècies endèmiques estan amenaçades (fins al 40 % en alguns dels territoris, com les Illes Balears), una tendència que s'agreuja en el futur, ja que el canvi climàtic afectarà especialment la conca mediterrània i les zones de muntanya (Cramer *et al.*, 2018; Hoegh-Guldberg *et al.*, 2019; Pepin *et al.*, 2015; Tuel i Eltahir, 2020). A Catalunya, l'informe *Estat de la natura a Catalunya 2020* (Brotons *et al.*, 2020) conclou que les poblacions de vertebrats i invertebrats autòctons de les quals es tenen dades han perdut de mitjana el 25 % dels indivi-

us en els darrers vint anys, i que els canvis en els usos del sòl són la principal causa de pèrdua de biodiversitat, malgrat que el canvi climàtic i les espècies invasores també hi tenen un paper important. Globalment, la problemàtica de la conservació de la biodiversitat a Catalunya és similar a la del conjunt d'Europa. L'informe *Natura, ús o abús? (2018-2019)* també destaca que la sobreexplotació de recursos té un gran impacte sobre el sòl, l'aire i l'aigua, i indica que els problemes de la conservació de la natura s'originen en el model econòmic actual, que és ambientalment insostenible (Peñuelas *et al.*, 2019).

Pel que fa al funcionament de la CBP, s'han definit els diferents processos implicats en l'obtenció d'un catàleg detallat dels genomes (figura 1). Aquests processos, que segueixen els estàndards de l'EBP, es poden resumir en els punts següents:

- recollida de mostres, processament i catalogació;
- conservació de les mostres en bancs de materials biològics;
- extracció i seqüenciació del DNA;
- assemblatge i anotació del genoma;

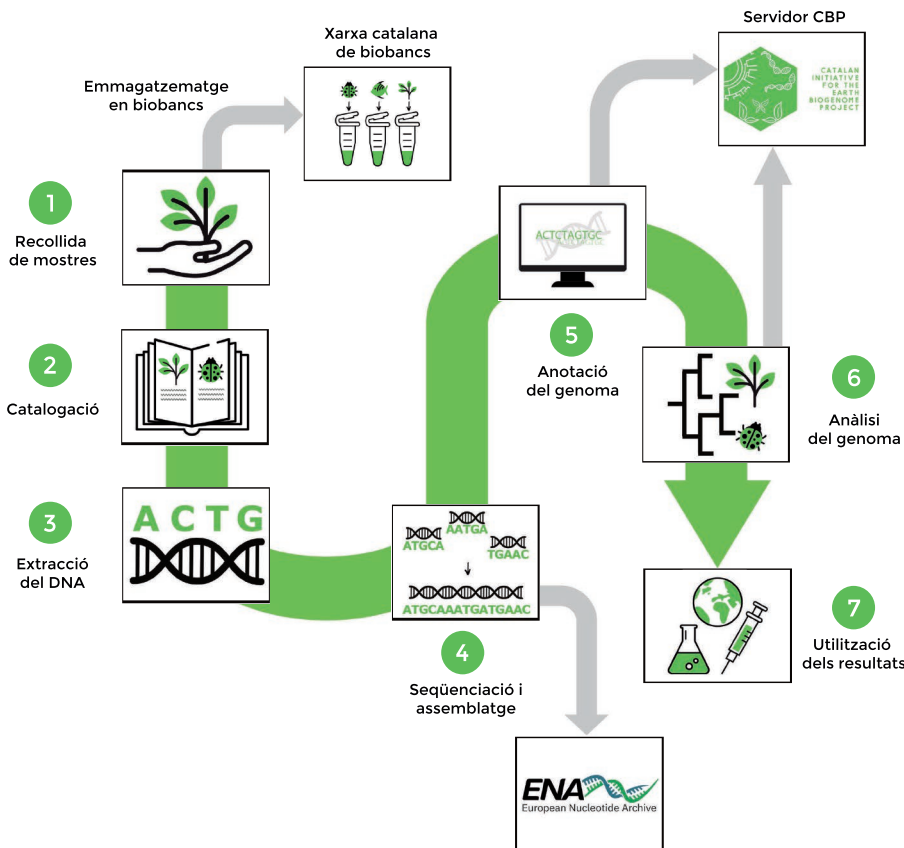
- anàlisi i adaptació de les dades per a la seva visualització;
- publicació de les dades en repositoris públics.

Implicacions i retorn social: més enllà de la biodiversitat

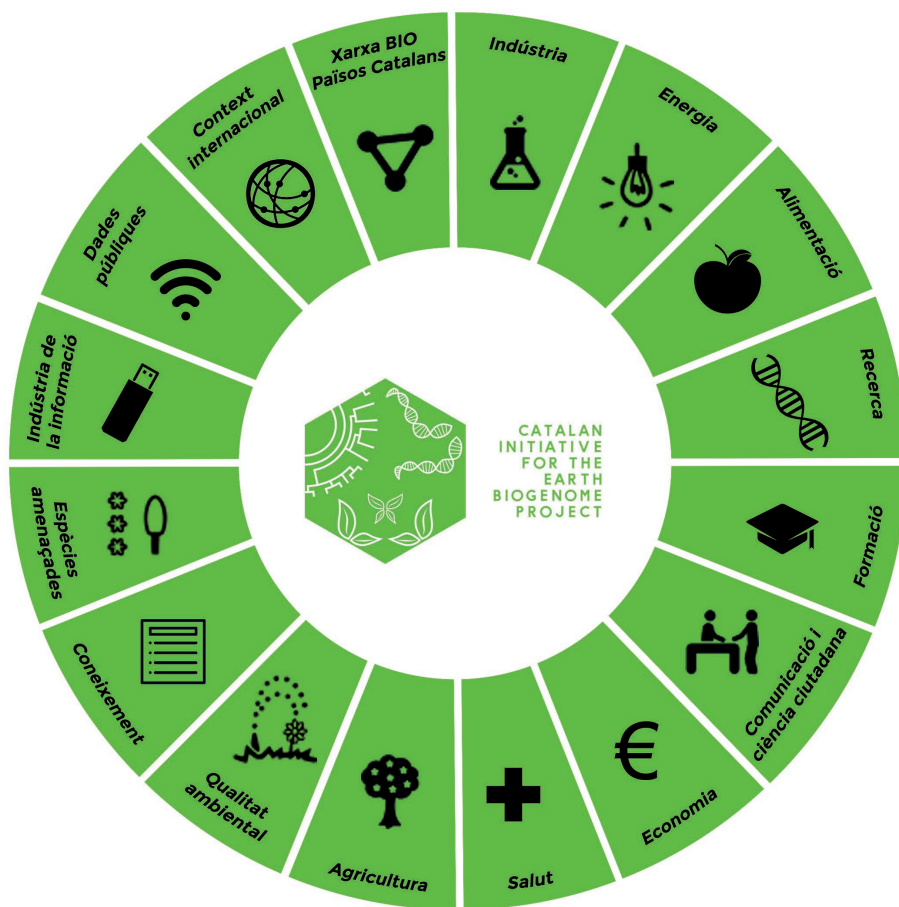
La CBP tindrà un impacte directe en el coneixement que tenim de les espècies que habiten els territoris de parla catalana. D'una banda, contribuirà a actualitzar, millorar i digitalitzar el catàleg de les espècies que hi viuen. De l'altra, entendre l'evolució genòmica d'aquestes espècies al llarg del temps i relacionar-la amb els efectes de l'activitat humana i el canvi climàtic ens farà ser més conscients de les nostres accions, dissenyar estratègies per a pal·liar-ne els efectes sobre els ecosistemes i millorar les estratègies de conservació que es duen a terme amb espècies amenaçades. Valorar la biodiversitat a través del desenvolupament d'aquesta iniciativa, i de la implicació i el compromís d'entitats —tant públiques com privades—, tindrà un efecte de presa de consciència en la població. Els ciutadans i les ciutadanes entendran com les accions de l'espècie humana han afectat els ecosistemes al llarg dels anys i com podem actuar per a preservar la biodiversitat actual.

A més, el coneixement obtingut, que es posarà a l'abast de tothom, ha de comportar un retorn social en àmbits molt diversos, des de l'agricultura, l'alimentació i la salut fins al sector energètic i la indústria, sense oblidar l'impuls que suposarà per a la recerca, la indústria i l'economia. En aquest sentit, cal destacar que els objectius globals de la CBP s'emmarquen en el concepte *one health* ('una sola salut'), definit per l'Organització Mundial de la Salut (OMS) el 2017 (www.who.int/news-room/q-a-detail/one-health), i estan inclosos en l'Estratègia de Biodiversitat de la Unió Europea per al 2030 (https://ec.europa.eu/info/sites/default/files/communication-annex-eu-biodiversity-strategy-2030_en.pdf) i en l'Agenda 2030 per al Desenvolupament Sostenible de l'Organització de les Nacions Unides (ONU) (www.un.org/sustainabledevelopment/development-agenda/).

Entre les implicacions principals de la CBP en l'àmbit socioeconòmic (figura 2), es pot destacar, en primer lloc, l'impacte en l'agricultura, la ramaderia, la pesca i l'alimentació. La identificació de noves variants genètiques permetrà millorar els cultius, reduir-ne les possibilitats d'infecció per plagues, augmentar-ne la productivitat o trobar noves maneres de conrear determinades espècies. Tam-



↑ Figura 1. Diagrama de processos que se segueixen a la CBP. Elaboració pròpia.



↑ Figura 2. Esquema de les possibles implicacions de la CBP en l'àmbit socioeconòmic. Elaboració pròpia.

bé es podran recuperar espècies de cultiu tradicionals que actualment s'han perdut o trobar noves aplicacions, com, per exemple, nous biocarburants. La resiliència de les espècies davant amenaces com el canvi climàtic prové de la diversitat genètica, i una font de variació genètica que ens podria ser molt útil es troba en els parents salvatges de les nostres espècies domesticades (Castañeda-Álvarez *et al.*, 2016). La seqüència detallada del genoma facilitarà la selecció genòmica per a la producció i la conservació de races i espècies en perill d'extinció. A banda d'ampliar les possibilitats de generació de noves varietats de productes transgènics o genèticament modificats, saber més sobre les espècies conegudes actualment i descobrir-ne de noves permetrà obtenir aliments de fonts que ara mateix no s'aprofiten.

En segon lloc, cal destacar l'impacte de la CBP en la salut i la qualitat ambiental. Conèixer de manera completa el genoma de tots els eucariotes ajudarà, d'entrada, a comprendre millor moltes malalties i el sistema de transmissió entre espècies dels patògens infeccio-

sos, la qual cosa farà que el control de les pandèmies sigui més eficient. D'altra banda, es podran descobrir noves molècules terapèutiques per a millorar la salut. Cal destacar que, en el context actual, s'ha plantejat la necessitat urgent d'establir una plataforma global de biovigilància basada en la genòmica (un sistema d'intercepció de pandèmies), que afavoriria enormement la nostra comprensió del món natural i, per tant, tindria un gran valor per a la bioseguretat, la biodefensa i l'economia (Kress *et al.*, 2020). Respecte a la qualitat ambiental, l'estudi de l'evolució dels genomes ens permetrà entendre quines accions humanes, inclòs el canvi climàtic, han tingut un impacte negatiu en la biodiversitat. Fruit d'això, podrem dissenyar estratègies per a millorar la qualitat del sòl, l'aire i l'aigua, i modificar i regular les activitats humanes que es duen a terme al nostre territori i que tenen un impacte directe en la salut de tota la població.

En tercer lloc, l'impacte en l'economia i la indústria. L'anomenat *capital natural* fa referència al conjunt de recursos naturals del pla-

neta, que inclouen geologia, sòls, aire, aigua i tots els organismes vius (https://www.unepfi.org/fileadmin/documents/ncd_booklet.pdf). La inversió en capital natural és un multiplicador de l'economia i permet retornar al medi més del que se n'extreu, cosa que posa l'economia al servei de la societat, perquè invertir en natura és invertir en la millora de les condicions de vida de les persones. A més, a la nostra societat, l'energia és precursora de l'activitat econòmica. La transició energètica cap a la descarbonització és la base d'un planeta més sa i saludable i, per això, cal apostar per energies renovables o més netes. El descobriment de noves espècies aportarà informació sobre formes alternatives d'energia fins ara desconegudes. A partir del genoma dels organismes seqüenciats es podran descobrir nous biomaterials o crear-ne a través de tècniques d'enginyeria genòmica i biologia sintètica, cosa que ha de contribuir al teixit industrial. Respecte a les tecnologies de la informació, l'EBP, i de la mateixa manera la CBP, representen un repte tant pel que fa a l'anàlisi com a l'emmagatzematge d'un nombre tan important de dades. Si bé amb les possibilitats actuals el projecte és factible, caldrà donar un impuls a noves tecnologies que facilitin la visualització, comparació i classificació de les dades, i això suposarà un revulsiu per a aquest sector.

En quart lloc, l'impacte en la recerca i la internacionalització. La implicació de la bioinformàtica i les ciències òmiques és clau en aquest projecte. Aquests sectors, que tenen una importància creixent al nostre país, oferiran noves possibilitats de projectes de recerca en l'àmbit de l'EBP. La iniciativa també contribuirà a la formació, donat que estudiants predoctorals, i també aquells que cursen estudis universitaris, podran participar activament en el projecte, tant en els treballs de camp de recollida de mostres com en les etapes de seqüenciació, anàlisi i classificació de les dades. També es dissenyaran activitats específiques per a escoles i instituts amb la finalitat de donar a conèixer i aprendre a respectar i estimar el nostre gran patrimoni natural. En el context internacional, la CBP és el node EBP a casa nostra i ens emmarca en un context internacional que afavorirà la presa de decisions en col·laboració i a través d'accions globals. Els projectes científics seran més competitius a causa de la seva projecció internacional.

Finalment, però no de menys importància, cal destacar l'impacte d'aquesta iniciativa en la comunicació científica i la ciència ciutadana. Al llarg de tot el projecte es faran dife-

rents accions de divulgació per tal que la ciutadania conegui els objectius, la metodologia i l'estat dels resultats, que es comunicaran a través del web del projecte (www.biogenoma.cat) i de les xarxes socials de les institucions implicades. Els ciutadans que ho vulguin podran col·laborar en el projecte de manera activa a través d'iniciatives que podrien incloure, per exemple, participar en la tria de les espècies que cal seqüenciar o en l'obtenció de mostres.

Estat actual (març de 2022)

L'activitat principal de la CBP s'ha centrat a estructurar la xarxa en el territori i crear tres grups de treball i un grup de coordinació, i a establir reunions periòdiques on es fan presentacions científiques i es discuteix com va avançant la iniciativa. Es va iniciar amb una fase pilot, que ha servit com a prova de concepte.

Com a part d'aquesta fase, la CBP va obrir una convocatòria de projectes la tardor de 2020, que, juntament amb el finançament addicional dels mateixos investigadors i investigadores de la Generalitat de Catalunya i del Parc Zoològic de Barcelona, va permetre començar la seqüenciació dels genomes de trenta espècies sota el paraigua de la CBP (figura 3). En aquesta fase pilot, la CBP va seguir un enfocament de baix a dalt, responent als interessos de la comunitat investigadora local, i, després d'una avaluació internacional, es van seleccionar projectes que proposaven seqüenciar espècies d'interès i icòniques.

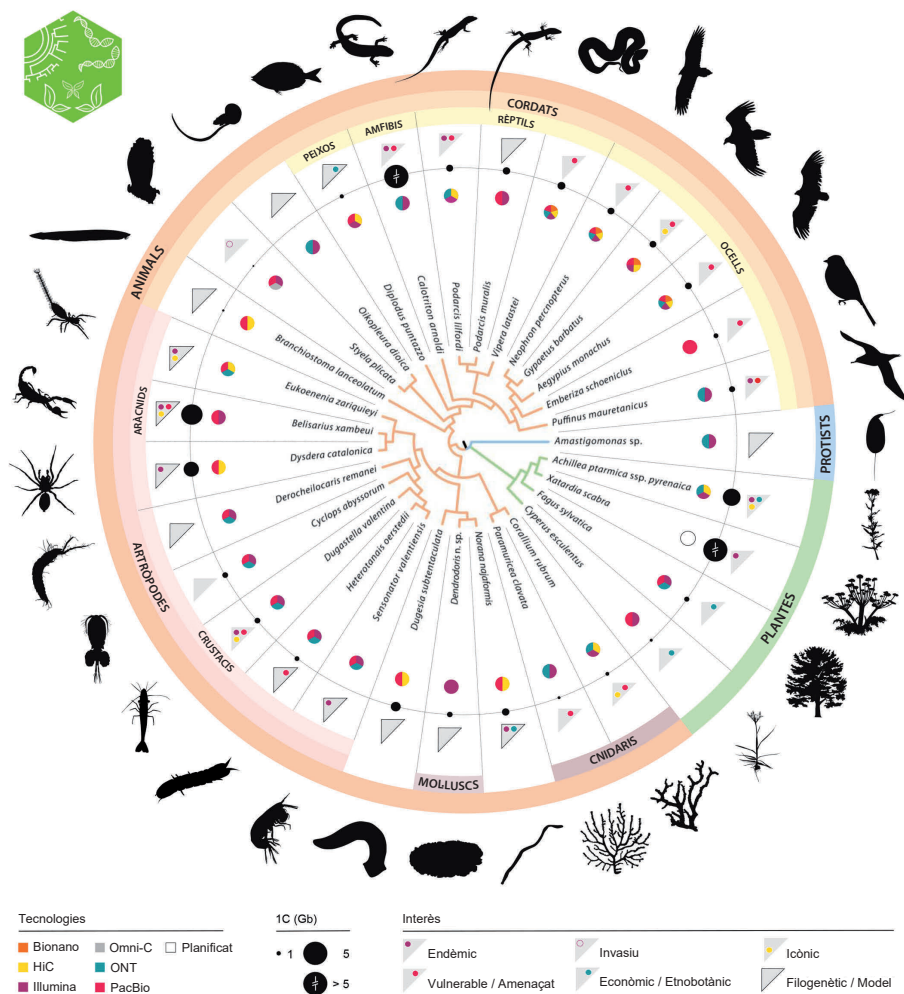
La tardor de 2021 es va obrir una segona convocatòria amb els mateixos criteris i sistema d'avaluació, amb la qual cosa el nombre d'espècies cofinançades des de la CBP són divuit avui en dia (taula 1). Entre aquestes espè-

cies hi ha tàxons poc explorats, espècies rares, endèmiques o difícils de catalogar, espècies model emergents, espècies en greu perill d'extinció o espècies medicinals. Actualment, el nombre de genomes de referència d'alta qualitat que s'estan generant és de quaranta-dos, repartits com es mostra a la figura 4. Cal destacar que el genoma de la baldriga balear (*Puffinus mauretanicus*) es troba ja disponible a la base de dades ENA (Cuevas-Caballé *et al.*, 2021). Com s'ha comentat, un aspecte crucial de l'EBP, i per tant també de la CBP, és la coordinació per tal d'evitar seqüenciar la mateixa espècie en llocs diferents. En el marc del DTOL, s'ha creat l'aplicació GoAT (<https://goat.genomehubs.org>), que utilitza la taxonomia de l'NCBI (<https://www.ncbi.nlm.nih.gov>) i serveix com a font centralitzada de metadades rellevants per al genoma per a la comunitat global. Funciona com a sistema de seguiment de seqüenciació per a la xarxa de l'EBP i conté, entre d'altres, la llista amb la informació de les espècies que se seqüencien sota el paraigua de la CBP.

La ICHN treballa, des de l'inici de la CBP, en l'actualització del catàleg de les espècies eucariotes que viuen als territoris catalans, que el desembre de 2021 inclou 26.433 tàxons. Aquest catàleg es digitalitzarà i servirà de referència per a prioritzar els genomes que cal seqüenciar durant les fases següents de la CBP. Entre els criteris de prioritització s'inclouen la posició filogenètica i la novetat, l'interès per als grups de recerca locals, el grau d'endemisme i conservació, l'interès biomèdic, agrícola i industrial.

Dins la CBP també hi ha un gran interès a desenvolupar programari de codi obert. Així, el portal de dades que s'ha implementat (<https://dades.biogenoma.cat>) conté tota la informació de les espècies, des del moment de la recollecció fins a l'obtenció de la seqüència. És un portal universal que importa automàticament totes les dades associades a projectes de seqüenciació genòmica que han estat dipositades a l'ENA. El portal és flexible i pot ser utilitzat per a mostrar les dades de qualsevol node de l'EBP. Altres desenvolupaments dins de la CBP inclouen mètodes eficients d'anotació de gens, d'alineament múltiple de gran nombre de seqüències (fins a milions) (Tommaso *et al.*, 2017) i de filogenòmica.

Més enllà dels límits dels territoris catalans, i dins de la xarxa EBP de xarxes (superposades), la CBP pretén tenir un paper central en els projectes de genòmica de la biodiversitat a Europa, que ara s'organitzen sota el European



↑ Figura 3. Projectes de seqüenciació del genoma sota el paraigua de la fase pilot de la CBP. La figura captura la posició filogenètica de l'espècie, la mida del genoma quan es coneix, les tecnologies emprades per a la seqüenciació i la raó de la selecció d'espècies. A la part superior esquerra hi ha el logotip de la CBP. Elaboració pròpia.

† Taula 1. Espècies en procés de seqüenciació del genoma que han rebut finançament de la CBP. Elaboració pròpia.

Animals	Plantes
Esgorges <i>Clathrina</i> sp.	<i>Achillea ptarmica</i> ssp. <i>pyrenaica</i>
Ctenòfors <i>Mnemiopsis</i> sp.	Protists <i>Amastigomonas</i> sp.
Cnidaris <i>Cladocora caespitosa</i> <i>Corallium rubrum</i>	<i>Caecitellus</i> <i>Paraphysomonas</i>
Equinoderms <i>Arbacia lixula</i>	
Vertebrats <i>Diplodus puntazzo</i> <i>Xyrichtys novacula</i> <i>Iberolacerta aurelioi</i> <i>Podarcis lilfordi</i>	
Artròpodes <i>Sensonator valentiensis</i> <i>Tethysbaena scabra</i> <i>Cyprideis torosa</i> <i>Belisarius xambeui</i>	
Anèl·lids <i>Norana najaformis</i>	

Reference Genome Atlas (ERGA), una iniciativa europea per a crear un atlas del genoma de la biodiversitat d'Europa (<https://www.erga-biodiversity.eu>). Així mateix, la CBP vol ser també central en els projectes dirigits a les regions geogràfiques a les quals pertanyen naturalment els territoris catalans: la península

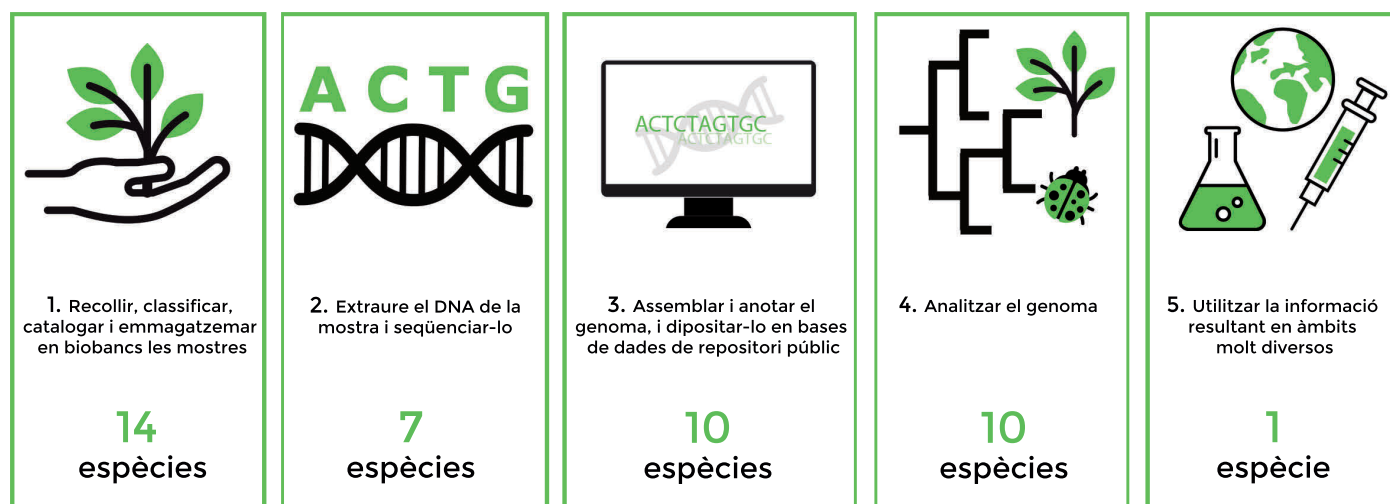
Ibèrica i la conca mediterrània. La CBP és ja un projecte associat a ERGA i, dins d'aquesta xarxa, treballem en estreta coordinació i col·laboració amb ERGA Andorra, ERGA Espanya i ERGA França.

A escala local, la CBP representa una oportunitat excel·lent per a reunir comunitats de

recerca, des de la història natural fins a la genòmica, que tradicionalment han estat aïllades les unes de les altres. Aquests sinergismes dotaran les institucions d'història natural de la infraestructura d'última generació i dels recursos humans necessaris per a garantir la documentació, la catalogació i la preservació d'exemplars, teixits i DNA per a les generacions futures. Caldrà, doncs, estructurar una xarxa de biobancs a tot el territori. La CBP reactivarà la investigació taxonòmica en un moment en què es troba en el seu mínim històric (p. ex. Crisci *et al.*, 2020). A més, atesa l'heterogeneïtat dels paisatges catalans, la CBP necessitarà la contribució de l'àmplia xarxa d'associacions naturalistes que coneixen els ecosistemes locals per a facilitar l'accés a les mostres biològiques. Més enllà d'un esforç científic pur, veiem aquesta iniciativa com a part d'un moviment transformador mundial que augmenta la consciència social sobre l'amenaça que suposa la pèrdua de biodiversitat per al benestar humà i que implica la societat a escala mundial en una relació diferent i més equilibrada amb la natura.

Agraïments

La CBP ha rebut finançament de l'IEC (a través del llegat Asenjo i dels projectes propis de recerca PRO2019-SO2, PRO2020-SO2 i PRO2021-SO2), del Departament d'Agricultura, Ramaderia, Pesca i Alimentació de la Generalitat de Catalunya, del Parc Zoològic de Barcelona i d'Andorra Recerca + Innovació. Agraïm a Mercè Rocaembosch l'ajuda en l'elaboració de la figura 3 i a Teresa Garnatje, Josep Germain, Manel Niell i Tomàs Marquès-Bonet, la lectura crítica del manuscrit.



† Figura 4. Resum de l'estat actual dels projectes de seqüenciació del genoma sota el paraigua de la CBP. Elaboració pròpia.

Bibliografia

- BLAXTER, M. [et al.] (2022). «Why sequence all eukaryotes?». *Proc. Natl. Acad. Sci. USA*, 119 (4): 1-9.
- BROTONS, L. [et al.] (2020). *Estat de la natura a Catalunya 2020*. Barcelona: Generalitat de Catalunya: Departament de Territori i Sostenibilitat.
- CAMARASA, J. M.; CASASSAS, O. (2020). *Cent anys de la Societat Catalana de Biologia, la primera societat filial de l'Institut d'Estudis Catalans*. Barcelona: Institut d'Estudis Catalans. ISBN: 978-84-9965-556-7.
- CASAS-SAINZ, A. M.; VICENTE, G. de (2009). «On the tectonic origin of Iberian topography». *Technophysics*, 474: 214-235.
- CASTAÑEDA-ÁLVAREZ, N. P. [et al.] (2016). «Global conservation priorities for crop wild relatives». *Nat. Plants*, 2: 1-6.
- CRAMER, W. [et al.] (2018). «Climate change and interconnected risks to sustainable development in the Mediterranean». *Nat. Clim. Chang.*, 8: 972-980.
- CRISCI, J. V. [et al.] (2020). «The end of botany». *Trends Plant Sci.*, 25: 1173-1176.
- CUEVAS-CABALLÉ, C. [et al.] (2021). «The genome of the Balearic shearwater (*Puffinus mauretanicus*), a critically endangered seabird: a valuable resource for evolutionary and conservation genomics». *BioRxiv*, 2021.12.17.473171.
- FOLCH I GUILLÉN, R. (dir.) (1984-2012). *Història natural dels Països Catalans*. Barcelona: Enciclopèdia Catalana.
- HINCHLIFF, C. E. [et al.] (2015). «Synthesis of phylogeny and taxonomy into a comprehensive tree of life». *Proc. Natl. Acad. Sci. USA*, 112: 12764-12769.
- HOEGH-GULDBERG, O. [et al.] (2019). «The human imperative of stabilizing global climate change at 1.5°C». *Science*, 365: aaw6974.
- KRESS, W. J. [et al.] (2020). «Opinion: Intercepting pandemics through genomics». *Proc. Natl. Acad. Sci. USA*, 117 (25): 13852-13855.
- LEWIN, H. A. [et al.] (2018). «The Earth BioGenome Project: Sequencing life for the future of life». *Proc. Natl. Acad. Sci. USA*, 115: 4325-4333.
- (2022). «The Earth BioGenome Project 2020: Starting the clock». *Proc. Natl. Acad. Sci. USA*, 119 (4): e2115635118.
- MORA, C. [et al.] (2011). «How many species are there on Earth and in the ocean?». *PLoS Biol.*, 9 (8): e1001127.
- PEÑUELAS, J. [et al.] (2019). *Natura, ús o abús? (2018-2019): Ús i abús de la natura, impactes i propostes de gestió. El cas de Catalunya com a paradigma*. Barcelona: Institut d'Estudis Catalans. ISBN: 978-84-9965-457-7.
- PEPIN, N. [et al.] (2015). «Elevation-dependent warming in mountain regions of the world». *Nat. Clim. Chang.*, 5: 424-430.
- SCHOLZ, H. A. [et al.] (2022). «Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation». *Nat. Commun.*, 13 (1): 1086.
- THE DARWIN TREE OF LIFE PROJECT CONSORTIUM (2021). «Sequence locally, think globally: The Darwin Tree of Life Project». *Proc. Natl. Acad. Sci. USA*, 119 (4): e2115642118.
- TOMMASO, P. di [et al.] (2017). «Nextflow enables reproducible computational workflows». *Nat. Biotechnol.*, 35: 316-319.
- TUEL, A.; ELTAHIR, E. A. B. (2020). «Why is the Mediterranean a climate change hot spot?». *J. Clim.*, 33: 5829-5843.
- WILSON, E. O. (1999). *The diversity of life*. Cambridge, Massachusetts: Harvard University Press. ISBN: 9780674058170.