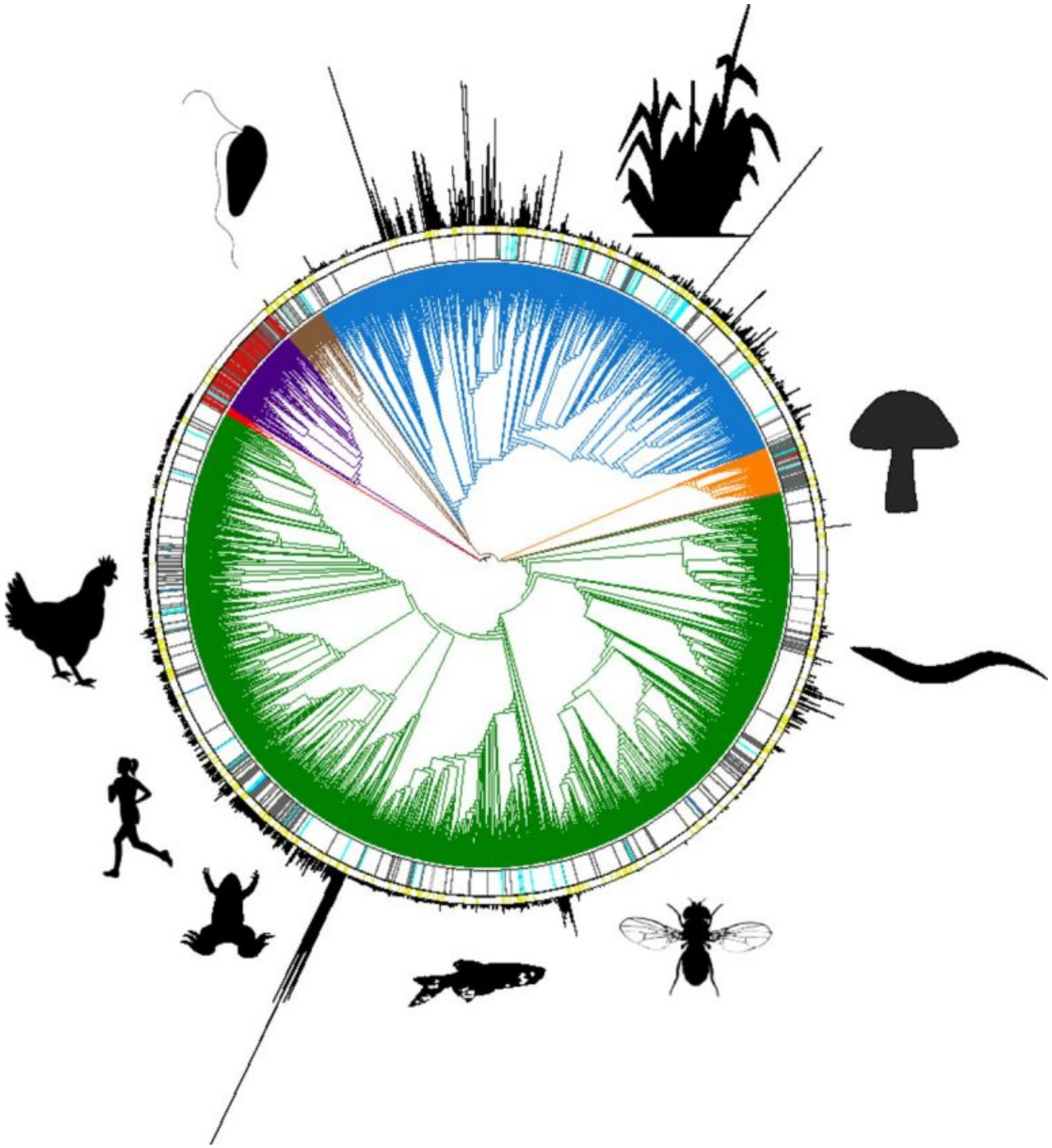
Veille citoyenne d'information sur les OGM et les semences

The genome of 1.8 million species is being sequenced

Par Eric MEUNIER

Publié le 11/12/2024, modifié le 09/03/2026

Can biological diversity escape any risk of biopiracy when part of it is digitised in computers? The answer depends on ongoing negotiations within international bodies. In the meantime, an international project to sequence the genome of all known eukaryotic species is making progress. Financed indirectly by players in the IT and artificial intelligence fields, this project even hopes to be able to bypass certain rules thanks to more powerful working tools.



The sequencing and recording in computer databases of the genome of living organisms that make up biological diversity is one of the keystones of the ongoing appropriation of living organisms. This information, known as "*digital sequence information*" (DSI), is the subject of international negotiations within the Convention on Biological Diversity (CBD), the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and the World Intellectual Property Organisation (WIPO). The aim of these negotiations is to define their legal status and the conditions needed to protect biodiversity from biopiracy. Far from being a subject of fantasy, the Earth BioGenome Project[i], which has been underway since 2008, illustrates the enormous stakes involved in these ISDs.

A costly and overdue international project

In 2019, *Inf'OGM* published an article (in French) presenting the Earth Biogenome Project, an international project launched the previous year[ii]. Whether it was an announcement or not, the project seemed excessive from the outset, since its aim was to sequence the genomes of the 1.84 million known eukaryotic species (organisms with cells that have a nucleus), out of the estimated 12 to 15 million[iii]. In 2008, it was announced that the project would take 10 years to complete, at an estimated cost of 4.7 billion dollars.

The project is now four years away from the end of its theoretical timetable, and the goal of sequencing the 1.5 million genomes of known eukaryotic species still seems a long, long way off. According to the project website, the genomes of only 3,039 species have been sequenced[iv].

A network of projects

The Earth BioGenome Project is not an entity in its own right, but an umbrella project covering many other projects. There are 64 of these projects, which may differ from one another for geographical reasons or because of the organisms sequenced. The list of projects published online [v] includes the "*1,000 Chilean Genomes"*, which aims to "*decipher the genomes of animals, plants, and microorganisms endemic to Chile"*. The "*African Biogenome Project*" covers the whole of the African continent, with the aim of building "*capacity (and infrastructure) to generate, analyze and deploy genomics data for the improvement and sustainable use of biodiversity and agriculture across Africa*". The European Union is of course present with its "*European Reference Genome Atlas Umbrella*", which is "*a pan-European community initiative to coordinate the generation of reference genomes representing European eukaryotic biodiversity*".

Participating projects can also target types of organisation rather than geographical areas. In this area, we need to think big. The "*Ocean genomes*" project aims to "*accelerate and scale the production of openly accessible reference genome assemblies for marine vertebrates*" present in all oceans. The depths of the ocean have not been forgotten, as the "*Deep-ocean genomes project* " will be exploring and collecting specimens for sequencing. The "*Polar Genomes Project*" aims to " *foster research on the unique features and functions of the genome that enable unique and powerful biological processes in polar organisms*". Primates, fish, birds, insects, crabs, butterflies, ants, amphibians, bees, bats,... are also the focus of numerous dedicated projects.

A project for "*indigenous peoples and local communities*"

The "*Wise ancestors*" project is somewhat original among the affiliated projects. Its mission, as set out on its website[vi], is to co-create and develop "*species-specific research projects (ie. "Conservation Challenges") with Indigenous Peoples (IP) and Local Communities (LC) for the purpose of producing genomic information, strengthening IP and LC conservation efforts, and supporting Indigenous data sovereignty*". Put simply, this project coordinates the work of researchers and local communities to sequence the genome of species known and used by these local communities, and to gather the associated traditional knowledge. This knowledge is essential if we are to know the characteristics of a particular organism that has been harvested.

As well as being a concrete example of the risks of biodiversity appropriation through the use of ISDs and the collection of associated traditional knowledge, this project is highly original in terms of its funding.

On its website, it states that its only source of funding is the Future of Life Institute[vii]. In the European Union, this institute is registered in the Transparency Register[viii] because of its meetings with members of the European Commission and the European Parliament, between 2021 and

2024, concerning the legislation on artificial intelligence, which was adopted last July[ix]. As the institute itself explains, its financial resources come from donations received and then redistributed in the form of grants. Donors include Vitalik Buterin, co-founder of the Ethereum cryptocurrency, Jaan Tallinn, founder of Skype, and, in 2015, Elon Musk, co-founder of Space X, Tesla, PayPal and OpenAI[x].

Done in reverse, the path described here shows that players in artificial intelligence and cryptocurrencies are indirectly investing in the sequencing of the genomes of species that make up biodiversity, mobilising researchers and local communities to do so.

Risks of biopiracy?

Why does the Earth BioGenome Project raise questions? One of the issues is intellectual property. Genome sequencing feeds computer databases that will store ISDs. These ISDs are in themselves a source of information that can be used by some to develop an 'invention' and claim an associated patent. In the field of life, patents have the particularity of covering the information constituting an invention and any organism containing this information. The recent case of a KWS patent on cold-resistant maize plants, unsuccessfully challenged before the European Patent Office (EPO) and threatening the commercial exploitation of the work of breeders who had selected cold-resistant maize prior to the patent application, shows this[xi]. This is also demonstrated by the recent adoption within the WIPO of an international treaty in which every effort has been made to ensure that the origin of the '*material*' (in our case, organisms or ISDs) does not have to be provided, thereby making it easier to extend the scope of a patent beyond the sole '*invention*' claimed[xii].

In the case of the Earth BioGenome Project, recommandations have been set out to reassure that all sampling of terrestrial biodiversity is carried out in compliance with international rules. The general approach of this project is that "*sample collectors should ensure that all local and national permissions for collection are in place, and that there is a record of these permissions that can be referred to if any questions arise as to whether a specimen was legally obtained*"[xiii]. The conditional clause used in this sentence, as in the rest of the recommandations for collecting samples, nevertheless raises questions, as compliance with national or international rules is rarely an option. This conditional clause is only slightly in keeping with the reminder to collectors that the Nagoya Protocol imposes an established procedure for the collection and transport of specimens: obtaining prior consent and a benefit-sharing agreement, while overlooking the essential material transfer agreement. This omission may be explained by the fact that this guide to good practice includes a hope that the work can be simplified by artificial intelligence tools. In fact, it is detailed that the Earth Biogenome Project "*encourage activities that simplify and streamline standard operating procedures and protocols to make it easier to "containerize" extraction, sequencing, and assembling activities. Containerization means a world in which a simple portable lab could house everything needed to go from sample to sequence [...]. This relieves pressure on Nagoya Permit requirements and the budget spent on expensive shipping costs to maintain the cold-chain*".

This desire is based on the absence of a legal definition and status for CIO information. Numerous negotiations are underway at international level, and some players, particularly in Europe and North America, are pushing for ISDs not to be subject to the rules governing the exploitation of biodiversity. While today it is necessary to obtain an agreement for the transfer of material in order to take an organism out of the country where it is located, several countries consider that this should not be the case for ISDs. From this angle, the Earth BioGenome Project raises the question of the risks of biopiracy in an even more exacerbated way. This is all the more true given that access in violation of any material transfer agreement obligation is already irreversible in the case of many freely accessible electronic ISD databases, as well as for public collections of genetic

resources that do not document the origin of the physical organisms sequenced or made available in physical form, just like many private collections.

i Website of the Earth Genome Project: http://www.earthbiogenome.org/

ii Frederic PRAT, « Séquencer le génome de l'ensemble des êtres vivants sur Terre », *Inf'OGM*, 27 august 2019 (in French).

iii The figures given here for the total number of genomes to be sequenced were given in 2020 in the following article:
H.A. Lewin *et al.*, « The Earth BioGenome Project 2020: Starting the clock », *Proc. Natl. Acad. Sci. U.S.A.*, 119 (4) e2115635118, 2022.
In 2018, the figures given were 1,5 million of eucaryotes' genomes.

iv GoaT, « Earth Biogenome Project (EBP) ».

v GoaT, « pages for the Earth Biogenome Project (EBP) and EBP-affiliated projects ».

vi Wise Ancestors, « Our Mission ».

vii Website of the Future of Life Institute: http://futureoflife.org/

viii According to the European Union, "*the Transparency register is a database listing* 'interest representatives' *(organisations, associations, groups and self-employed individuals) who carry out activities to influence the EU policy and decision-making process. It is designed to show the public which interests are being represented at EU level, by whom and on whose behalf – as well as the resources devoted to these activities (including financial support, donations, sponsorship, etc.)*". EU, Transparency Register, "organisation detail for Future of Life Institute".

ix Parlement européen, Observatoire législatif, « 2021/0106(COD) – Législation sur l'intelligence artificielle ».

x FLI, « Finances ».

xi Denis Meshaka, "EPO throws cold water on opposition to KWS patent", *Inf'OGM*, 20 november 2024.

xii K.M. Gopakumar, "WIPO opens more widely the door to biopiracy", *Inf'OGM*, 6 november 2024.

xiii Earth BioGenome Project, Mara K. N. Lawniczak *et al.*, « Best Practice Guidance for Earth BioGenome Project Sample Collection and Processing: Progress and Challenges in Biodiverse Reference Genome Creation », october 2024.

Adresse de cet article : https://infogm.org/en/the-genome-of-1-8-million-species-is-being-sequenced/