# Strengthening Genomic Sovereignty in Africa: The African Bioinformatics Center for Omics (ABCOMICS) Project

## Authors

*Gueye Khadim[1,*], Mane Karim[1,*], Mboup Pape Adama[1,*], Diop Mouhamadou Fadel[1,*], Shehu Mahfouz[1,*], Ahadzie Bankolé[1,*], Diame Mamado L.[1,*], Dieng Idrissa[1,*], Diop Ousmane[1,2], Faye Awa[1,*], Kotanmi Brezesky[1,*], Mboup Mame Ndew[1,*], Gerald Mboowa[1,*], Mohammed Nuredin[1,*], Ndour Pape Adama[1,*], Ndiaye Aminata[1,*], Cisse Badara[2], Diamanka Arfang [4], Diop Gora [4], Gueye Modou [4], Sall Khadidiatou [6], Sarr Idrissa [4,5], Thiare Ousmane[3], Mboup Souleyemane[2]*

*[1] ABCOMICS - African Bioinformatics Center for OMICS*
*[2] IRESSEF - Institute for Health Research, Epidemiological Surveillance, and Training*
*[3] CINERI - Cyber-Infrastructure nationale pour l'Enseignement supérieur, la Recherche et l'Innovation*
*[4] UCAD - Cheikh Anta Diop University of Dakar*
*[5] CURI - Centre universitaire de Recherche et de formation aux technologies de l'Internet*
*[6] UnCHK- Université Numérique Cheikh Hamidou Kane (ex UVS)*

## Abstract

Africa boasts immense human, animal, vegetal, and microbial diversity. While several communicable and non-communicable diseases result from the interactions between these entities and their molecular changes, they are little explored by local researchers and clinicians to understand and predict the rise and trajectory of diseases of major concern, including neglected tropical diseases. A strong genomic approach to overcoming this gap will rely on an independent and well-established Bioinformatics infrastructure designed to address the needs of local biomedical, climate, and environmental research. In this article, we present how the African Bioinformatics Center for Omics (ABCOMICS), an initiative that integrates Bioinformatics, AI, and data science to promote research autonomy and training, will contribute to addressing the bioinformatics-related challenges in West Africa. A core objective of the initiative is to build an African genomic data repository that supports local research and collaboration with existing initiatives in Africa and other parts of the world.

*Keywords: Genomic sovereignty, African genomic database, Bioinformatics infrastructure, Data governance, INSDC, Capacity building, Genomic surveillance, Artificial intelligence, FAIR data, Global health equity*

## Introduction

Genomics now presents an immense opportunity for Africa, particularly in public health, agriculture, and biodiversity conservation. The ability to locally sequence and analyze genomes of several organisms would enhance epidemic preparedness and response, optimize disease treatments for African populations, improve crop resilience to climate change, and deepen ecological studies of the continent's unique biodiversity [1]. Currently, genomic sequences from Africa account for only 4.9% of publicly available nucleotide sequences deposited in the International Nucleotide Sequence Database Collaboration (INSDC). In comparison, 22.2% originate from Europe, 31.1% from North America, 14% from South America, 23.6% from Asia, and 3.4% from Oceania. These data were retrieved from the European Nucleotide Archive (ENA) API on June 11, 2025, at 12:58 UTC (Figure 1A) [2]. This underrepresentation reflects the limits in the ability of African researchers to produce genomic-based research tailored to local realities, whether in precision medicine, crop improvement, disease surveillance, or biodiversity monitoring [3].

According to the representation map of global biodiversity projects available in the EBI's biodiversity database, Africa is strikingly underrepresented, with almost no genomic biodiversity data currently available from the continent (Figure 1B) [4]. This lack of representation significantly hinders Africa's ability to catalogue, understand, and preserve its vast and unique biodiversity through genomics.

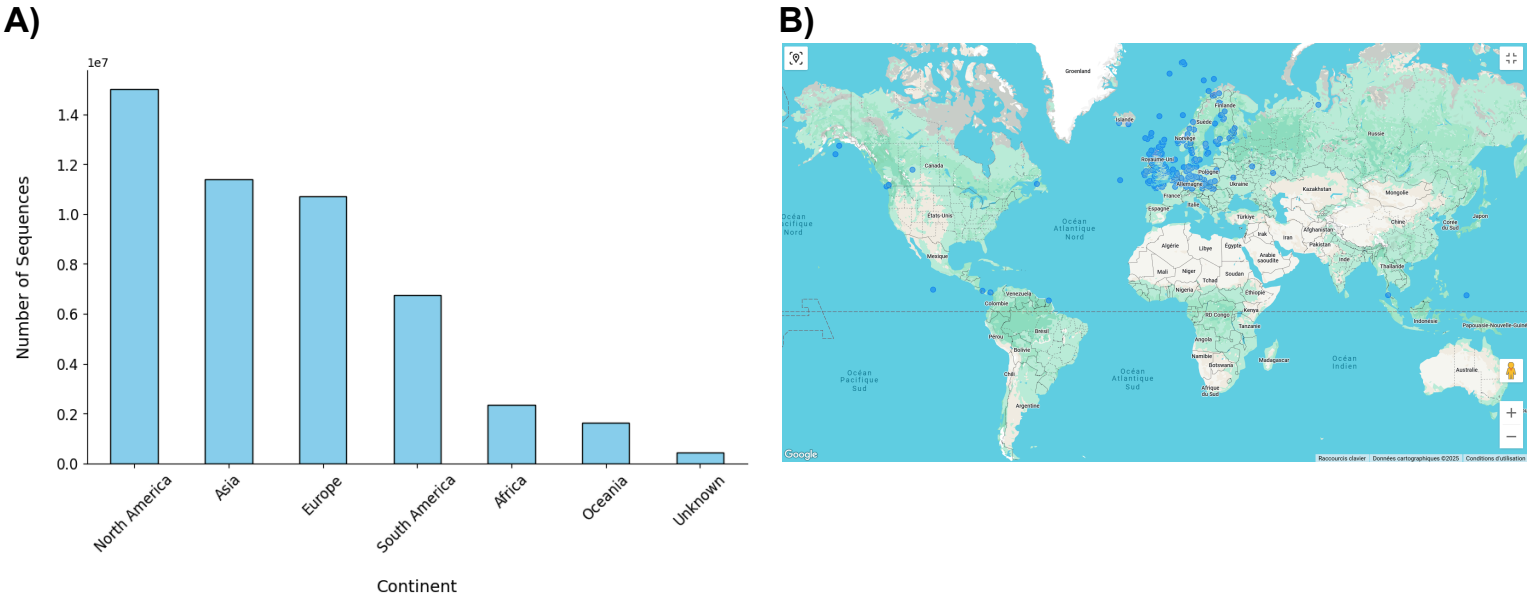


*Figure 1: Global representation of African genomics in public data repositories*

- *(A) Distribution of publicly available nucleotide sequences by continent in the INSDC database. Source: ENA API (accessed June 11, 2025)*
- *(B) Distribution of global biodiversity genomics projects by continent. Source: EMBL-EBI Biodiversity GIS Portal – https://www.ebi.ac.uk/biodiversity/gis*

In response to this gap, several African initiatives have emerged to advance genomic research, including H3Africa, H3ABioNet, and the Africa CDC Pathogen Genomics Initiative [5], [6], [7]. These programs have strengthened bioinformatics training and pathogen surveillance networks across the continent, laying critical groundwork for African genomics. However, while these initiatives have contributed to research capacity-building, they largely operate through distributed models and remain reliant on foreign infrastructures [8], restricting Africa's scientific autonomy.

Moreover, most researchers in biomedical, environmental, and agricultural sciences still do not incorporate genomic data into their work due to several structural challenges: a lack of local infrastructure, a shortage of trained bioinformatics specialists, limited access to advanced computational resources, and insufficient funding. These barriers result in dependence on external platforms and laboratories for processing African genomic data, hindering scientific progress.

In response to these challenges, the African Bioinformatics Center for Omics (ABCOMICS) has been conceptualized as a strategic initiative to establish and strengthen genomic-based research in West and Central Africa [9]. Although ABCOMICS is still in its development phase, it aims to become a leading institution, empowering African researchers to generate, analyze, and leverage their omics data independently. Key pillars of the project are the creation of an African omics data repository, providing a secure and locally governed platform for storing and managing NGS data and its related metadata, offering researchers the support and platform that aligns with their specific needs, and facilitating the integration of genomic insights into their work.

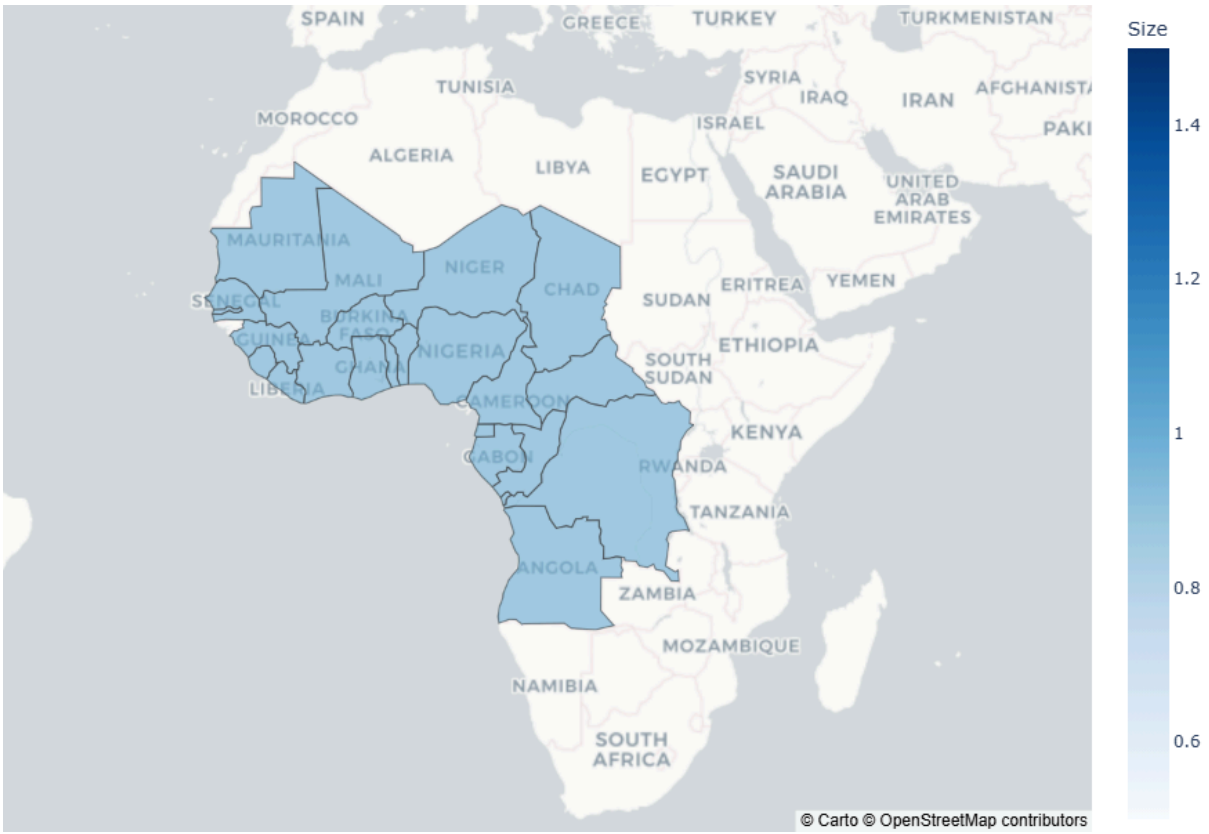

*Figure 2: ABCOMICS's Regional Scope: West and Central Africa*

ABCOMICS provides a sovereign and sustainable alternative, combining a state-of-the-art bioinformatics platform with a targeted training strategy and a collaborative institutional approach tailored to African research environments.

To ensure that we meet the needs of our partners and contribute to groundbreaking genomic-based research, ABCOMICS will offer a suite of specialized services to researchers across the region. These include genome assembly and annotation, transcriptomics and metagenomics, structural bioinformatics and proteomics, artificial intelligence-driven analysis of multi-omics data, and research software development. Additionally, ABCOMICS will contribute to capacity-building through bioinformatics training and statistical analysis, covering essential tools like R, Python, Bash, Nextflow/Snakemake, and high-performance computing. Alongside these services, the initiative fosters a multidisciplinary approach, connecting universities, research centers, and public health laboratories to strengthen regional collaboration and maximize scientific impact.

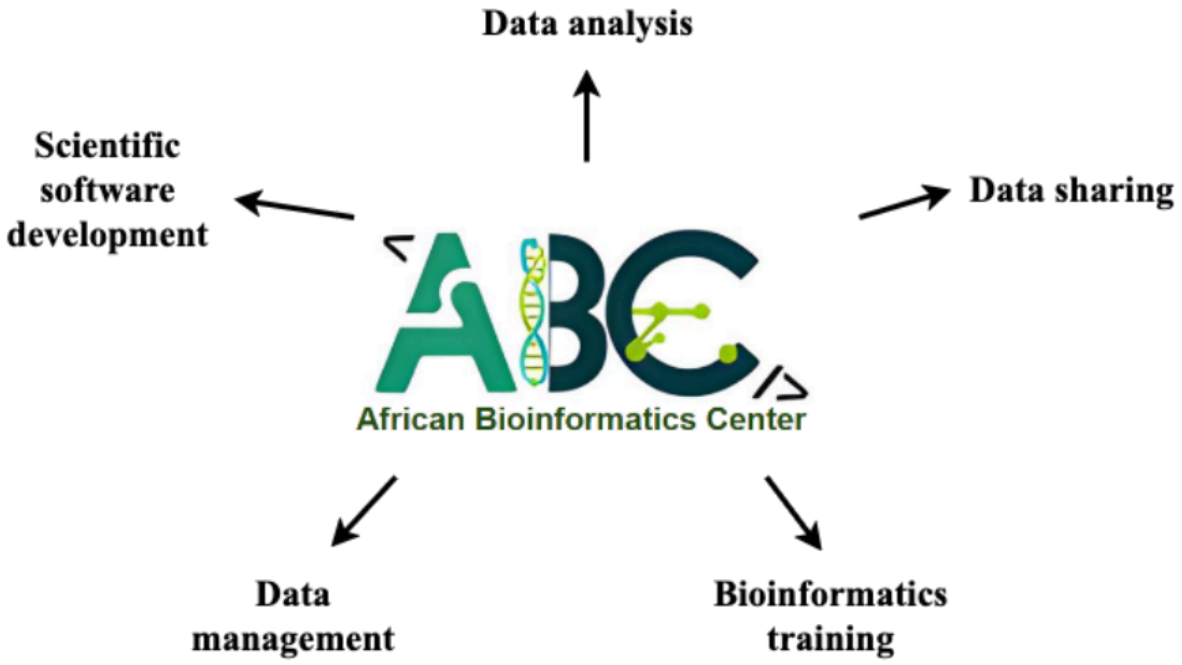


*Figure 3: A diagram showing our services.*

ABCOMICS's objectives are ambitious and designed to transform Africa's genomic landscape. By ensuring data sovereignty, reducing dependency on foreign infrastructure, facilitating African researchers' integration into global genomic databases, training a new generation of bioinformaticians, and enhancing genomic surveillance capacity, ABCOMICS seeks to reposition Africa as a key player in global bioinformatics.
Through this initiative, ABCOMICS aspires to redefine the role of West and Central Africa in global genomic research. By enabling local data management, facilitating access to bioinformatics tools, and structuring a dynamic research network, the initiative provides a concrete solution to the challenges hindering genomic advancements in Africa. With an inclusive and collaborative model, ABCOMICS represents a strategic lever to enhance

Africa's scientific sovereignty and ensure that discoveries made within the continent directly benefit African researchers and populations.

## Challenges to Overcome

The success of the African Bioinformatics Center for Omics (ABCOMICS) depends on its ability to overcome a range of complex and multidimensional challenges. These are not only technical but also political, regulatory, institutional, and ethical. By addressing these obstacles proactively, ABCOMICS will not only fill critical gaps but also contribute to redefining scientific sovereignty across the African continent.

One of the most pressing issues is the underrepresentation of African genomic data in global repositories [10]. Despite Africa's vast genetic diversity and high burden of infectious diseases, the continent's contribution to major international genomic databases remains very poor. This gap significantly limits the development of diagnostics, treatments, and interventions tailored to Africa's genetic, environmental, and epidemiological realities. More than just a scientific limitation, this imbalance represents a strategic risk—if African data is not generated and managed locally, it risks being permanently excluded or underutilized in shaping health policies, drug discovery, and future technologies.

At the same time, African research institutions face heavy reliance on foreign infrastructures. Many institutions now rely on external sequencing centers, cloud-based platforms, and international partners for data storage, analysis, and distribution. This reliance increases the costs related to data transfer and processing, legal and ethical concerns regarding cross-border data usage, and a loss of autonomy in setting research priorities and timelines. Additionally, researchers often submit their data to international repositories hosted outside the continent, where African-generated datasets may not be immediately accessible, properly curated, or prioritized for reuse by local scientists.

The International Nucleotide Sequence Database Collaboration (INSDC) plays a key role in the global management of genomic data [11]. It consists of three major infrastructures: the NCBI (USA), which hosts GenBank and provides tools for sequence submission and analysis; the EMBL-EBI (UK), which manages the European Nucleotide Archive (ENA) and offers advanced standards for bioinformatics; and the DDBJ (Japan), which facilitates international data exchange across Asia and other continents [8], [12], [13], [14]. These institutions ensure harmonized and open access to genetic data, yet very few African institutions actively contribute to it. This lack of engagement limits Africa's visibility and influence within the global bioinformatics landscape. ABCOMICS's long-term vision is to strategically integrate Africa into the INSDC, ensuring that the continent not only supplies data but also contributes to governance, curation, and technological innovation at an international level.

Beyond access to databases, insufficient local resources remain a fundamental challenge. Bioinformatics capacity across Africa is unevenly distributed, and while centers of excellence exist, many institutions lack high-performance computing infrastructure, stable internet connectivity, reliable electricity, and specialized expertise in programming, statistical analysis, and genomics. This creates a bottleneck: even when genomic data is generated within Africa, its analysis is often outsourced, leading to delays in results and diminished

data sovereignty. The consequence of these systemic challenges is an increasing risk of losing control over African genomic data. When genomic sequences are collected in Africa but stored and processed abroad, they become subject to foreign regulations and can be reused without appropriate consent or equitable benefit-sharing. This has resulted in numerous publications and patents where African scientists are marginalized, despite their contributions to data generation. These concerns have heightened awareness of the importance of open science ethics, equitable access, and locally governed genomic infrastructures—principles that are central to ABCOMICS's mission.

Another critical challenge is ensuring long-term financial sustainability. Initial funding is required to establish infrastructure, train personnel, develop software solutions, and build strategic partnerships.

Collaboration with local and regional authorities is another pivotal aspect of the ABCOMICS project. The absence of coherent national policies governing genomic data, biobanks, and transnational research hinders Africa's ability to lead in genomics.

Ultimately, ABCOMICS must navigate the complex task of mobilizing and aligning scientific and institutional stakeholders. African universities, public health agencies, NGOs, and research institutions must work together through inclusive governance, shared values, and complementary expertise. ABCOMICS will need to build trust, establish flexible engagement mechanisms, and ensure meaningful participation from all actors involved in genomic research.

By bridging genomics with urgent health challenges in Africa — including infectious diseases, mental health, climate-related vulnerabilities, and agricultural resilience — ABCOMICS positions itself as a discovery-driven platform advancing life, health, and wellbeing across the continent.

By tackling these challenges, ABCOMICS is not merely creating a technical infrastructure—it is shaping the future of Africa's genomic research, ensuring scientific sovereignty, equitable access, and lasting impact on both regional and global scales.

## The ABCOMICS's approach

The establishment of the African Bioinformatics Center for Omics (ABCOMICS) is based on the combination of an advanced scientific approach, a strategic plan for developing an African omics database, a strong capacity-building initiative, and the use of technologies tailored to the continent's realities. This strategy addresses key challenges in the field of genomics across the continent,

ABCOMICS is developing a robust bioinformatics platform (abcomics.org ) that centralizes and analyzes genomic data produced by researchers in Africa using advanced Bioinformatics, Statistical and AI tools. Substantial efforts will be made toward the collection and annotation of genomic data, data standardization and normalization to enhance visibility and global integration, as well as ensuring compatibility with international infrastructures. Furthermore, the development of custom bioinformatics and statistical tools and the

integration of high-performance computing technologies will support efficient genome analysis for African researchers.

A core component of ABCOMICS's initiative is the creation of an African omics data repository to centralize genetic sequences from the continent. We will ensure it is interoperable with global databases, such as INSDC, to improve the representation of African genomes in international research, while ensuring that the security and integrity are maintained according to international standards
Another key focus of the project is the development of advanced algorithms tailored for disease analysis and health concerns relevant to Africa. By prioritizing research on endemic conditions such as malaria, Lassa fever, genetic disorders prevalent in African populations, and neglected tropical diseases, ABCOMICS aims to refine disease monitoring, improve treatment strategies, and deepen insights into genetic variations unique to the region.

Moreover, ABCOMICS recognizes the continent's technological limitations and designs solutions that operate effectively in low-connectivity environments and resource-constrained settings. Unlike bioinformatics platforms built for nations with extensive computational resources, ABCOMICS optimizes its algorithms to perform efficiently on moderate-performance systems, ensuring accessibility even in regions with limited resources. ABCOMICS will promote hybrid solutions, combining local infrastructure with secure cloud platforms, adapted to the needs of regional partners. Establishing technology partnerships with global providers or regional data centers will be essential. At the same time, ABCOMICS will implement robust cybersecurity protocols to guarantee data confidentiality, integrity, and sovereignty.

Building local expertise is also a central objective, ensuring sustainable scientific independence for Africa. ABCOMICS will initiate and support bioinformatics training programs, academic collaborations, and strategic research partnerships, facilitating knowledge transfer and capacity development. Additionally, ABCOMICS will promote technological accessibility via shared analytical platforms and hands-on training in essential tools such as R, Python, Bash, Nextflow/Snakemake, and high-performance computing, empowering African scientists to leverage genomic research effectively. Additionally, ABCOMICS will leverage artificial intelligence platforms to facilitate multi-omics data exploration and automation of bioinformatics pipelines. These workflows will optimize the analysis of African genomic data, particularly for endemic diseases and regional pathogens.

In the global context, ABCOMICS aspires to integrate into leading genomic data networks, including the International Nucleotide Sequence Database Collaboration (INSDC). These databases face ongoing challenges in handling large-scale data, format diversity, and international privacy compliance [11]. African representation within these repositories remains limited, but ABCOMICS could play a strategic role by:

- Acting as a regional hub for African data submission to global repositories.
- Training local experts in INSDC-compliant data curation and submission protocols.
- Adapting tools and formats to ensure technical compatibility with Africa's genomic research ecosystem.

ABCOMICS will actively engage with Ministries of Health, Higher Education, Research, Agriculture, and Digital Economy, as well as national regulatory bodies, to harmonize legal frameworks and promote the development of ethically sound, locally relevant regulations aligned with African contexts. The establishment of ethics committees and governance bodies will be key to ensuring responsible genomic data management.

Our funding strategy must be diversified, combining international grant proposals, government support, philanthropic investments, and private-sector contributions. To remain financially viable, ABCOMICS will develop a progressive funding model, incorporating fee-based services, co-financing mechanisms, and long-term institutional commitments.

ABCOMICS's establishment follows a phased approach, structured around two key phases. In the phase (0-3 years), the project will focus on establishing bioinformatics infrastructures and launching the African omics database. This phase will also involve initial training programs and the development of strategic partnerships critical to the project's foundation. In the long term (+3 years), ABCOMICS aims to build a fully integrated African bioinformatics network, ensuring complete sovereignty over genomic data. This phase will involve expanding the project continent-wide, deepening global database integration, advancing specialized algorithms for disease analysis and regional genomic research, and strengthening institutional collaborations and scientific cooperation.

## Impact, Strategic Partnerships, and Future Vision

The African Bioinformatics Center for Omics (ABCOMICS) aims to transform bioinformatics and genomics in Africa by relying on a bold vision and strategic collaborations. Its immediate impact is reflected in strong support for research institutions and universities, particularly in genomic data analysis, scientific writing, and the publication of data. This assistance allows partners to enhance their productivity and visibility, ensuring that genomic data serves as a powerful tool for decision making in the fight against the proliferation of diseases, in preserving the environment and biodiversity, and in increasing agricultural production for food security.
ABCOMICS also stands as a key player in structuring a reference platform, making advanced bioinformatics tools more accessible while strengthening genomic data governance. By fostering sovereign and ethical data management, ABCOMICS empowers African researchers to analyse, interpret and leverage their own genetic resources independently, rather than merely supplying raw data to external initiatives.

One of the project's strongest assets lies in its strategic partnerships, which are central to its deployment and expansion. Supported by major institutions in Senegal, such as IRESSEF, UCAD, CURI, and CINERI, ABCOMICS benefits from an optimal environment for developing scientific and technological programs. The extension of this partnership to other academic, industrial, and institutional partners will further broaden the project's impact by integrating complementary expertise and necessary resources.

In line with this vision, ABCOMICS will implement robust governance frameworks, ensuring ethical and sovereign management of African genomic data. Strong security mechanisms

will protect data integrity and facilitate long-term access for African researchers, reinforcing their control over genetic resources.

By increasing its participation in international collaborations, ABCOMICS aims to position Africa not only as a data contributor but as a leading voice in genomic research governance.

ABCOMICS is more than just a bioinformatics platform—it is a structural transformation of Africa's scientific landscape, where genomic sovereignty becomes a tangible reality. Through these efforts, ABCOMICS ensures researchers gain expertise and visibility, securing scientific excellence and full strategic control over Africa's genomic data.

## References


[1] K. Theissinger *et al.*, “How genomics can help biodiversity conservation,” *Trends Genet.*, vol. 39, no. 7, pp. 545–559, Jul. 2023, doi: 10.1016/j.tig.2023.01.005.

[2] “Swagger UI.” Accessed: Jul. 09, 2025. [Online]. Available: https://www.ebi.ac.uk/ena/portal/api/swagger-ui/index.html

[3] “Scaling up genomic sequencing in Africa | WHO | Regional Office for Africa.” Accessed: May 19, 2025. [Online]. Available: https://www.afro.who.int/news/scaling-genomic-sequencing-africa

[4] “GBDP Sampling Map,” Looker Studio. Accessed: Jul. 09, 2025. [Online]. Available: http://lookerstudio.google.com/reporting/5287be84-e4c7-4ff1-90aa-87e7e86548fb/page/Pl2XE?feature=opengraph

[5] The H3Africa Consortium *et al.*, “Enabling the genomic revolution in Africa,” *Science*, vol. 344, no. 6190, pp. 1346–1348, Jun. 2014, doi: 10.1126/science.1251546.

[6] N. J. Mulder *et al.*, “H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa,” *Genome Res.*, vol. 26, no. 2, pp. 271–277, Feb. 2016, doi: 10.1101/gr.196295.115.

[7] “Africa Pathogen Genomics Initiative | Africa PGI 2.0 – Africa CDC.” Accessed: Jul. 09, 2025. [Online]. Available: https://africacdc.org/africa-pathogen-genomics-initiative-africa-pgi/

[8] NCBI Resource Coordinators, “Database resources of the National Center for Biotechnology Information,” *Nucleic Acids Res.*, vol. 44, no. D1, pp. D7-19, Jan. 2016, doi: 10.1093/nar/gkv1290.

[9] “African Bioinformatics Center (ABC) for Omics.” Accessed: Jul. 09, 2025. [Online]. Available: https://abcomics.org/

[10] “Closing the Gap: The Critical Need for African Representation in Global Genetic Databases - ODRi Media News - Breaking News, East Africa News, Sports News, Kenya News, World News.” Accessed: Jun. 12, 2025. [Online]. Available: https://www.odrimedia.co.ke/health/closing-the-gap-the-critical-need-for-african-representation-in-global-genetic-databases/

[11] “International Nucleotide Sequence Database Collaboration | Nucleic Acids Research | Oxford Academic.” Accessed: Jun. 12, 2025. [Online]. Available: https://academic.oup.com/nar/article/44/D1/D48/2503109

[12] “GenBank Overview.” Accessed: Jun. 12, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/genbank/

[13] C. O’Cathail *et al.*, “The European Nucleotide Archive in 2024,” *Nucleic Acids Res.*, vol. 53, no. D1, pp. D49–D55, Jan. 2025, doi: 10.1093/nar/gkae975.

[14] “DDBJ update in 2024: the DDBJ Group Cloud service for sharing pre-publication data - PMC.” Accessed: Jul. 09, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11701645/