BMC Medical Ethics

SYSTEMATIC REVIEW

Open Access

A scoping review of human genetic resources management policies and databases in high- and middle-low-income countries

Hongwei Liu[1], Yin Liu[1], Yanyan Zhao[1,2], Yingqi Ma[1], Qiong Chen[1], Huifang Xu[1], Xiaoyang Wang[1], Xiaoli Guo[1], Hong Wang[1], Zelong Chen[1,2], Shaokai Zhang[1*] and Binbin Han[1*]

Abstract

Background This review examines global human genetic resources management, focusing on genetic data policies and repositories in high- and middle-low-income countries.

Methods A comprehensive search strategy was employed across multiple databases, including official government websites and Google, to gather relevant literature on human genetic resources management policies and genetic resource databases. Documents were screened for relevance, focusing on high-income countries (United States, United Kingdom, Japan) and middle-low-income countries (China, India, Kenya). Data were extracted, coded, and analyzed to identify common themes and differences in genetic resource management practices.

Results High-income countries benefit from robust legal frameworks and advanced technological infrastructures. The United States enforces the Health Insurance Portability and Accountability Act and the Genetic Information Nondiscrimination Act to protect privacy and facilitate data sharing, while Japan relies on the Act on the Protection of Personal Information and ethical guidelines. Additionally, high-income countries host a variety of genetic databases and biobanks that support scientific research. In contrast, middle-low-income countries like China, India, and Kenya are still developing their frameworks. China has regulations such as the Biosecurity Law and the Regulations on the Management of Human Genetic Resources, but still requires more unified standards. India's policies focus on genetic research and data protection through the Biological Diversity Act, while Kenya seeks to improve data management through the 2019 Data Protection Act.

Conclusion Significant disparities exist in human genetic resources management between high-income and middle-low-income countries. High-income countries have robust systems balancing privacy protection with research facilitation, supported by comprehensive and large-scale databases for scientific research. Middle-low-income countries need to enhance legal frameworks and build population-specific databases. Promoting equitable data sharing and adopting best practices from high-income countries are essential for advancing global scientific discovery and ensuring fair management of genetic resources.

*Correspondence:
Shaokai Zhang
shaokaizhang@zzu.edu.cn
Binbin Han
zlyyhanbinbin2859@zzu.edu.cn

Full list of author information is available at the end of the article

Keywords Human genetic resources, Databases, Data sharing policies

Introduction

In the era of genomic medicine, human genetic resources (HGRs) have become critically impactful in scientific research and medical advancements [1]. The rapid increase in genetic data, driven by advancements in sequencing technologies and bioinformatics, has underscored the necessity for robust management of HGRs [2], covering aspects such as individual privacy protection, data security assurance, and the promotion of the rational utilization of genetic material.

The management of HGRs encompasses critical aspects like data security, privacy protection, database management, and data sharing. Different countries have established varying policies and regulations to address these concerns. For instance, the United States has established robust frameworks for data security, privacy protection, and data sharing [3]. These include the Health Insurance Portability and Accountability Act (HIPAA) [4], which protects personal health information, the Genetic Information Nondiscrimination Act (GINA) [5], which prevents discrimination based on genetic information, and the Genomic Data Sharing Policy [6], which encourages researchers to share data obtained from NIH-funded projects. In contrast, Japan adopts a more cautious and detailed approach to HGRs management. The Act on the Protection of Personal Information (APPI) [7, 8] in Japan ensures strict controls over personal data while still supporting research initiatives. This reflects a balance between facilitating research and maintaining stringent ethical and privacy standards, showing a nuanced commitment to data quality and safety.

Understanding the global landscape of HGRs management is vital for fostering international collaboration, ensuring compliance with diverse legal frameworks, and promoting equitable and responsible use of genetic data. Additionally, establishing comprehensive databases and sharing policies is essential for advancing research and maximizing the benefits of genetic data [9, 10].

The objective of this scoping review is to summarize and compare the human genetic resources management policies and databases across selected high-income and middle-low-income countries. The review aims to identify existing policies, evaluate their implementation, and analyze the available genetic resource databases to highlight commonalities, differences, and best practices.

Methods

This review aimed to understand the nature of existing human genetic resources management policies and databases in the United States, United Kingdom, Japan, India, China, and Kenya. These countries were selected for their diversity in socio-economic status (high-income vs. middle-low-income nations) and regulatory environments, as they represent both well-established and emerging frameworks for genetic data management [11–15]. High-income countries such as the United States, United Kingdom, and Japan have long-established policies and sophisticated infrastructure for managing genetic resources, while middle-low-income countries like China, India, and Kenya offer important insights into rapidly evolving legal frameworks and data sharing practices in developing contexts. By comparing these two groups, we aim to provide a comprehensive view of both established and emerging practices in genetic resources management.

We conducted our assessment by way of document collection, review, and synthesis, using a scoping review methodology. Scoping reviews are useful for assessing a research or policy area and identifying gaps. The scoping review method we utilized was proposed by Arksey and O'Malley (2005) [16].

Data sources and search strategy

For human genetic resources management policies, data were primarily sourced from grey literature, including official government websites, reports from research institutions, and other non-peer-reviewed sources, supplemented by searches on Google. We conducted a Google search for the following search strings: (genetic OR genomic OR "genetic data" OR "genomic data" OR genome) AND (guideline OR guidance OR policy OR law OR regulation OR framework OR legislation OR act OR privacy OR ethics OR compliance OR standard OR recommendation OR code OR "soft law" OR bill OR statute) AND [country name]. For genetic resource databases, a similar search strategy was employed, utilizing Google to locate comprehensive and authoritative sources. The search strings were tailored to each country and focused on identifying key genetic databases and repositories. The specific search strings used were: (biobank OR biorepository OR bioethics OR "data sharing" OR "data protection" OR "informed consent" OR "ethical oversight" OR repository) AND (guideline OR guidance OR policy OR law OR regulation OR framework OR legislation OR act OR privacy OR ethics OR compliance OR standard OR recommendation OR code OR "soft law" OR bill OR statute) AND ([country name]).

Eligibility

Documents were eligible for inclusion if they provided guidance, policy, law, or databases focused on the management of genetic resources and data

sharing. Documents or databases not specifically related to genetic information were excluded. No date restrictions were applied to the search.

Grey literature screening

Five hundred and two documents were initially identified as potentially eligible. After removing duplicates, 89 were excluded, leaving 413 for further screening. These documents, which appeared relevant based on their titles, were reviewed to determine whether they met the eligibility criteria. The initial screening involved identifying key words to ascertain whether the documents discussed genetics or genomics research in high-income countries or middle-low-income countries. Subsequently, the documents were further screened to determine if they focused on genetic/genomic resources. This process led to the exclusion of 287 additional documents, leaving 126 for a full review. Upon careful reading, 68 documents were excluded for not meeting the criteria, resulting in 58 documents that were selected for data extraction and analysis (Fig. 1).

Data abstract, coding, and analysis

For the data abstraction, coding, and analysis phases, a systematic approach was employed to ensure a thorough examination and synthesis of the identified documents. The process began with the extraction of key information from each eligible document, focusing on aspects related

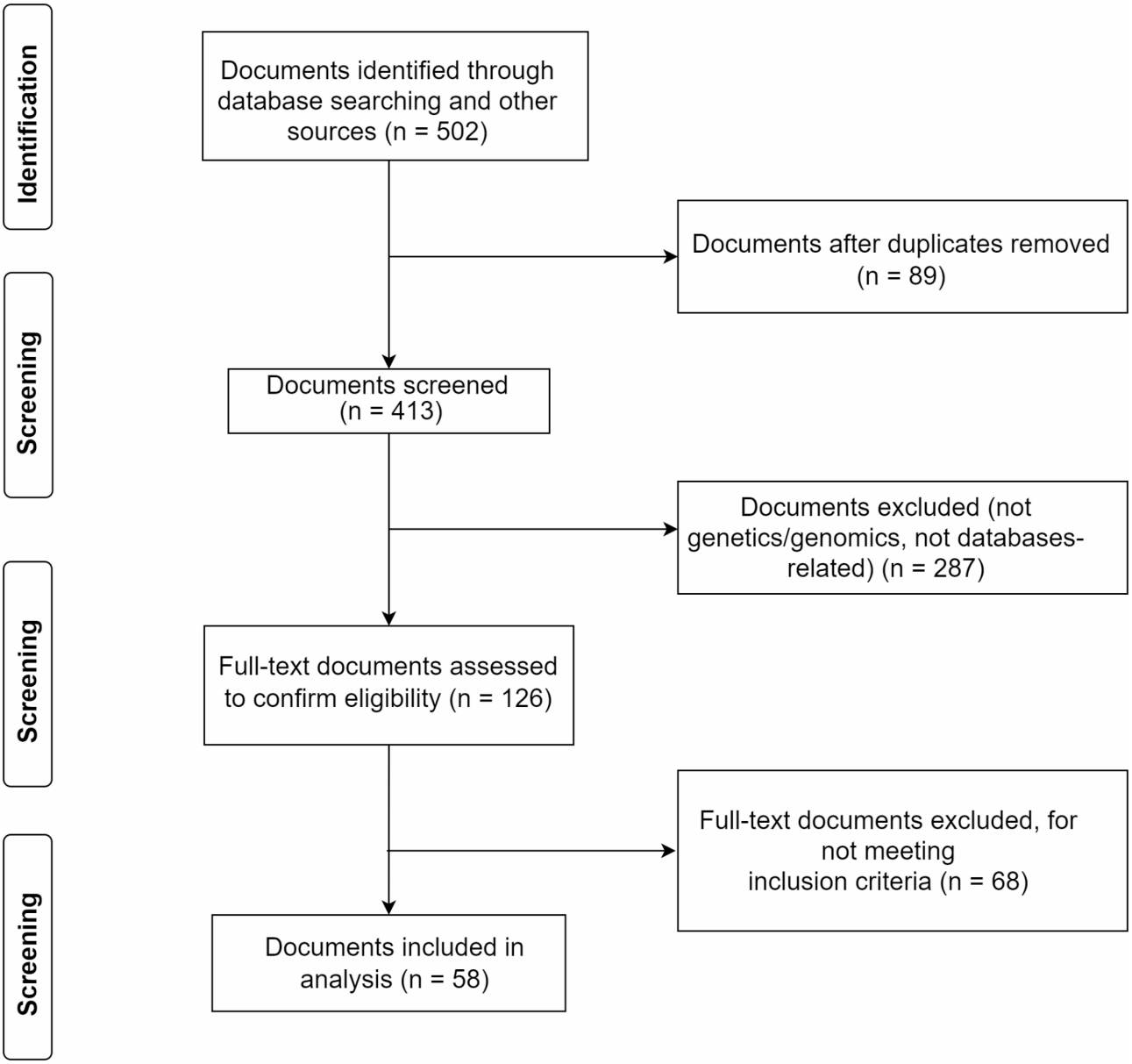

Fig. 1 PRISMA flow diagram for document selection

to genetic resources management and data sharing policies, as well as genetic databases. The extracted data included the country of origin, the name of the policy or database, and a detailed description of each.

Coding involved categorizing the documents based on their relevance to high-income countries and middle-low-income countries. Each document was assigned specific codes to indicate whether it pertained to policy or database management. Further sub-coding was conducted to highlight specific themes such as data protection, privacy, ethical considerations, and the scope of genetic data management.

Analysis was performed by comparing and contrasting the policies and databases across the two groups of countries. This comparative analysis aimed to identify commonalities and differences in the approaches to genetic resources management and data sharing between high-income and middle-low-income countries. The results were organized into two main sections: policies and databases.

In the results section, the findings are presented in two parts. The first part discusses the policies related to genetic resources management, divided into high-income and middle-low-income countries, and provides a detailed analysis of each policy. The second part focuses on the genetic databases, again divided into high-income and middle-low-income countries, with a thorough examination of each identified database. Tables were created to summarize the key information, including the country, policy or database name, and a brief description. These tables facilitate easy reference and provide a clear overview of the comparative insights from the study.

All documents that met the inclusion criteria were imported into Microsoft Excel and coded using a defined coding framework by one member of the research team (QC). The coding was then reviewed by two other members (HWL, YL) to ensure appropriate and consistent application of codes. Any coding discrepancies were discussed and resolved through consensus among the three members of the coding team. Frequencies, percentages, and data visualizations were generated using Microsoft Excel.

Results

Our findings are organized into two main sections: policies related to human genetic resources and genetic resource databases. We have included detailed analyses of countries with well-established regulatory frameworks, including three high-income countries and three middle-low-income countries. All policies and databases are comprehensively documented in Tables 1 and 2, providing easy reference and a thorough understanding of the comparative insights from our study.

Policies related to human genetic resource

High-income countries: United States, United Kingdom, and Japan

United States The management of human genetic resources in the United States is guided by several key legislative acts. The Federal Policy for the Protection of Human Subjects (The Common Rule) [17] was first established in 1991 and has since been adopted by 17 federal agencies. The Common Rule is an essential regulation that ensures ethical governance of human genetic resources and provides a legal foundation for data privacy and research integrity in the U.S. It sets clear ethical standards for informed consent, privacy protection, and research oversight, ensuring that genetic data is used responsibly and ethically for scientific progress. The Health Insurance Portability and Accountability Act (HIPAA) [18] of 1996 sets forth specific provisions for the protection of individual privacy and establishes standards for the use and public disclosure of de-identified health information [4]. This act ensures that personal health information, including genetic data, is handled with strict confidentiality, protecting individuals from unauthorized disclosure and misuse. Following HIPAA, the Genetic Information Nondiscrimination Act (GINA) [19] of 2008 was enacted to prevent discrimination based on genetic information in health insurance and employment [5]. GINA prohibits insurers and employers from requesting or using genetic information to make decisions about eligibility, coverage, underwriting, or employment, thus safeguarding individuals against genetic discrimination. More recently, the American Data Privacy and Protection Act (ADPPA) introduced in 2022 emphasizes balancing personal privacy protection with the utility of data [20]. The ADPPA aims to enhance individuals' control over their data while allowing the beneficial use of data for innovation and public interest, building on the protections established by HIPAA and GINA to ensure comprehensive privacy safeguards in the evolving digital landscape.

In addition to these management policies, specific regulations govern the sharing and protection of genetic data. The Genomic Data Sharing Policy [6], implemented by the National Institutes of Health (NIH), mandates that all NIH-funded research involving large-scale human or non-human genomic data generate comprehensive data sharing plans. This policy aims to maximize the utility of genomic data by promoting widespread data sharing while also protecting the privacy and confidentiality of research participants. Complementing this, Certificates of Confidentiality issued by the NIH protect sensitive information from forced disclosure in legal proceedings [21]. Researchers can use these certificates to refuse to disclose identifying information about research

Table 1 Policies for human genetic resources management in High-Income and Middle-Low-Income countries

Country	Policy Name	Issuing Authority	Year	Key Provisions
USA	Health Insurance Portability and Accountability Act (HIPAA)	U.S. Congress	1996	Establishes national standards to protect individuals' medical records and personal health information (PHI), including genetic data. Requires healthcare providers, health plans, and healthcare clearinghouses to implement administrative, physical, and technical safeguards for PHI.
USA	Genetic Information Nondiscrimination Act (GINA)	U.S. Congress	2008	Prohibits genetic discrimination in health insurance and employment. Prevents health insurers from using genetic information to determine eligibility, coverage, or premium rates. Bars employers from requesting, requiring, or using genetic information for hiring, firing, job placement, or promotions.
USA	American Data Privacy and Protection Act (ADPPA)	U.S. Congress (Proposed)	Pending	Establishes a unified national framework for personal data protection, including sensitive information such as genetic data. Introduces comprehensive requirements for data security, transparency, and individual rights over their data.
USA	NIH Genomic Data Sharing Policy	National Institutes of Health (NIH)	2014	Requires NIH-funded researchers to comply with genomic data sharing (GDS) standards to promote broad data access for biomedical research. Mandates the submission of large-scale human genomic data to controlled-access repositories such as the NIH database of Genotypes and Phenotypes (dbGaP). Ensures that researchers obtain explicit consent from participants for data sharing and follow data access policies to balance privacy and open science.
USA	Federal Policy for the Protection of Human Subjects ("Common Rule")	U.S. Department of Health and Human Services (HHS)	2017	Establishes ethical principles and regulatory requirements for human subject research, including genomic and biomedical studies. Mandates informed consent from participants, specifying how their genetic data will be used and whether it will be shared.
USA	Certificates of Confidentiality	National Institutes of Health (NIH)	2017	Grants legal protections to researchers and participants in genetic and biomedical research, preventing compelled disclosure of identifiable data. Shields research data from being subpoenaed by law enforcement, courts, or other government entities.
UK	Human Tissue Act 2004 (HTA 2004)	Human Tissue Authority (HTA)	2004	Regulates the storage and use of human tissues, including DNA and genetic material. Requires licensing for the collection, storage, and use of human biological samples. Ensures informed consent for the use of human genetic materials in research.
UK	UK Biobank Ethics and Governance Framework	UK Biobank	2006	Provides a governance model for UK Biobank genetic research, ensuring long-term data security, participant consent, and controlled data access.
UK	Research Governance Framework for Health and Social Care	UK Department of Health and Social Care	2017	Establishes governance principles for health and social care research, including genetic research. Ensures research integrity, ethical compliance, and patient/participant safety in NHS-funded research.
UK	Data Protection Act 2018 (DPA 2018)	UK Parliament	2018	Implements UK GDPR and provides additional data protection measures for specific contexts, including scientific research. Regulates personal data processing and outlines individual rights regarding genetic and health data.
UK	UK General Data Protection Regulation (UK GDPR)	UK Government	2021	Establishes rules on the collection, processing, and storage of personal data, including genetic data, to ensure privacy and security. Defines genetic data as "special category data," requiring additional safeguards for its processing. Retains key principles of the EU GDPR but with UK-specific modifications.
UK	Medical Research Council (MRC) data sharing policy	Medical Research Council (MRC)	Ongoing (latest version active)	Requires MRC-funded researchers to make data available for reuse, ensuring transparency, reproducibility, and maximizing public benefit from research investments.
Japan	Act on the Protection of Personal Information (APPI)	Japanese Government	2003 (Revised 2017, 2020)	Protects personal data, including genetic information, establishing legal frameworks for data collection, storage, and sharing. The 2020 revision strengthens cross-border data transfer regulations, requiring data exports to meet adequacy standards.
Japan	Ethical Guidelines for Human Genome and Gene Analysis Research	MEXT, MHLW, METI	2001 (Revised 2017, 2021)	Establishes ethical requirements for genomic research, ensuring privacy protection, informed consent, and data security. Requires research institutions to obtain Ethics Review Committee (ERC) approval before conducting human genetic studies.
Japan	Ethical Guidelines for Medical and Health Research Involving Human Subjects	MEXT, MHLW, METI	2014 (Revised 2017)	Governs precision medicine and personalized healthcare research, mandating ERC approval, informed consent, and data security compliance for human genetic research.

Table 1 (continued)

Country	Policy Name	Issuing Authority	Year	Key Provisions
China	Biosecurity Law of the People's Republic of China	National People's Congress (NPC)	2020	Establishes national security measures for biosafety, including regulation of biotechnology research, prevention of biological threats, and management of genetic resources.
China	Data Security Law	National People's Congress (NPC)	2021	Defines how data, including genomic and health data, should be collected, stored, processed, and transferred, ensuring national security and individual privacy.
China	Personal Information Protection Law	National People's Congress (NPC)	2021	Regulates the collection, processing, and storage of personal data, ensuring individual privacy and preventing data misuse.
China	Regulations on Management of Human Genetic Resources	Ministry of Science and Technology (MOST), State Council	2019	Governs the collection, preservation, utilization, and sharing of human genetic resources, ensuring ethical compliance and national security.
China	Implementation Rules for the Regulations on the Management of Human Genetic Resources	Ministry of Science and Technology (MOST)	2023	Provides detailed procedures for applying for approval to use human genetic resources, ensuring ethical and legal compliance in research and commercial applications.
India	Biological Diversity Act, 2002	National Biodiversity Authority (NBA), Government of India	2002	Regulates access to biological resources and associated traditional knowledge; aligns with the Nagoya Protocol to ensure fair benefit-sharing.
India	Biological Data Storage, Access and Sharing Policy of India	Department of Biotechnology (DBT), Indian Biological Data Centre (IBDC)	Latest version 2023	Provides guidelines for the ethical collection, storage, and sharing of biological and genomic data; ensures compliance with privacy laws.
India	Centre for Cellular and Molecular Biology (CCMB)	Council of Scientific and Industrial Research (CSIR), India	Ongoing	Establishes ethical and governance guidelines for cellular and molecular biology research; promotes genomic studies and personalized medicine.
India	Digital Personal Data Protection (DPDP) Act	Government of India	2023	Establishes a legal framework for the protection of digital personal data, including genetic data. Sets provisions for informed consent, data processing, data subject rights, cross-border data transfers, and data fiduciary responsibilities.
Kenya	Data Protection Act 2019	Government of Kenya	2019	Establishes regulations for the collection, processing, and storage of personal data; includes special provisions for sensitive personal data such as genetic information.
Kenya	National Biodiversity Action Plan (NBAP)	Ministry of Environment and Forestry, Kenya	2019 (latest version)	Provides strategies for protecting biodiversity, including genetic resources; aligns with the Nagoya Protocol to ensure fair access and benefit-sharing of genetic materials.
Kenya	Data Protection (General) Regulations 2021	Office of the Data Protection Commissioner, Kenya	2021	Specifies how personal and sensitive data, including health and genetic data, should be handled, ensuring data subjects' rights to access, correction, and deletion.
Kenya	KEMRI-Wellcome Trust Research Programme (KWTRP)	Kenya Medical Research Institute (KEMRI), Wellcome Trust, University of Oxford	Ongoing	Establishes ethical guidelines for conducting genomic and biomedical research; ensures compliance with local and international regulations.

participants, ensuring an additional layer of privacy protection for genetic data.

The U.S. privacy regulations create a dynamic balance between protecting individual privacy and enabling the use of genetic data for research through a multi-layered and complementary legal framework. HIPAA and GINA address data security and anti-discrimination concerns, respectively. HIPAA enforces technical safeguards, such as de-identification and access controls, ensuring compliance with the minimum necessary principle for sharing genetic data. It also allows de-identified data to be used for research, provided there is Institutional Review Board (IRB) approval [22]. This minimizes disclosure risks and provides a compliant pathway for research.

Table 2 Genetic resource databases in High-Income and Middle-Low-Income countries

Database Name	Country	Governing Policies	Access Policy	Eligible Users	Data Type
dbGaP	USA	NIH Genomic Data Sharing Policy (GDS), Common Rule, HIPAA	Controlled access, requires IRB and NIH approval	NIH-funded investigators, academic researchers	Genomic sequences, phenotype, EHR-linked data
All of Us	USA	NIH GDS, HIPAA, Common Rule	Controlled access, requires researcher registration and training	Qualified academic, non-profit, and industry researchers	Whole genome sequencing (WGS), EHR, participant-reported data
TOPMed	USA	NIH GDS, HIPAA, Common Rule	Controlled access via dbGaP	NIH-approved investigators	Multi-omics data (WGS, metabolomics, proteomics for cardiovascular, lung, blood disorders)
GTEx	USA	NIH GDS, HIPAA, Common Rule	Open-access and controlled-access tiers	Approved researchers	Tissue-specific genomic, transcriptomic, and epigenomic data
GEO	USA	NIH policies on open-access genomic data	Public access	Public	Gene expression, RNA-seq, transcriptomic datasets
GeneBank	USA	Open-access under NIH and NCBI policies	Public access	Public	Nucleotide sequences, annotated genetic information
dbVar	USA	NIH and NCBI policies on structural variation	Public access	Public	Genomic structural variation data
dbSNP	USA	NIH and NCBI policies	Public access	Public	SNP and short genetic variation data
Gene Expression Omnibus Datasets	USA	NIH and NCBI policies	Public access	Public	Genomic datasets for expression studies
RefSeq	USA	NIH and NCBI policies	Public access	Public	Curated reference sequences for genomes, genes, and proteins
SRA	USA	NIH GDS, HIPAA, Common Rule	Controlled and open access tiers	Public (open data), approved researchers (controlled data)	Raw sequencing data from NGS studies
TCGA	USA	NIH Genomic Data Sharing Policy (GDS), Common Rule, HIPAA	Open-access (summary data) and controlled-access (raw genomic and clinical data) tiers	Public (for summary data), NIH-approved researchers (for controlled data)	Multi-omics cancer genomic data, including WGS, transcriptomics, epigenomics, proteomics, clinical metadata
1000 Genomes Project	USA	Governed by IGSR policies, NIH, EBI	Public access	Public	Whole genome sequencing from global populations
EMBL-EBI	UK/Europe	Governed by EMBL policies, UK GDPR, EU GDPR	Public access	Public	Genomic, proteomic, and bioinformatics datasets
UK Biobank	UK	UK GDPR, Data Protection Act 2018, UK Biobank Ethics & Governance Framework	Controlled access, requires project approval and ethics review	Academic researchers, government agencies, industry (under strict conditions)	Genomic, phenotypic, lifestyle, health record data
100,000 Genomes Project	UK	UK GDPR, Data Protection Act 2018, NHS England Regulations	Controlled access, requires NHS approval	NHS-affiliated researchers, government-approved researchers	Whole genome sequencing (WGS) data from rare disease & cancer patients
DDBJ	Japan	Governed by MEXT policies, APPI, and NBDC regulations	Public access	Public	Genomic sequences, DNA barcoding, transcriptomics
KEGG	Japan	Governed by Kyoto University, APPI	Public access with some subscription-based tools	Public	Biological pathways, gene and protein functional annotations, metabolic networks
HGVD	Japan	Governed by Japanese Society of Human Genetics, APPI, MEXT policies	Public access	Public	Genetic variation data, SNP frequencies in the Japanese population

Table 2 (continued)

Database Name	Country	Governing Policies	Access Policy	Eligible Users	Data Type
Biobank Japan	Japan	Governed by MEXT, AMED, and Japanese data privacy laws (APPI, Ethical Guidelines for Human Genome/Gene Analysis Research)	Public access for GWAS summary statistics; Controlled access for genomic, clinical, and phenotype data (requires ethics approval)	Public (GWAS summary data); Academic researchers, approved government and industry partners (for controlled data)	GWAS summary statistics (public); Full genomic, phenotypic, clinical, and lifestyle data (restricted)
GSA-Human	China	Governed by Chinese Academy of Sciences (CAS), Ministry of Science and Technology (MOST), Chinese data privacy laws	Controlled access, requires approval from CAS	Approved academic and government researchers	Whole genome sequencing (WGS), transcriptomic, and human genomic data
China Kadoorie Biobank	China	Governed by China National Health Commission, Oxford University, MOST	Controlled access, requires ethics and institutional approval	Approved academic researchers	Genotypic, epidemiological, and health data of 512,000 Chinese participants
ChinaMAP	China	Governed by National Metabolic Disease Clinical Research Center	Controlled access, requires institutional approval	Approved academic researchers and healthcare institutions	Metabolic disease biobank with 3 million samples, genomic and biochemical data for precision medicine research
PGG.Han	China	Governed by the PGG Consortium	Public access	Public	Genomic data of 137,012 Han Chinese individuals, including whole-genome sequencing (WGS), SNP variation, and population genetics data
WBBC	China	Governed by Westlake University, Chinese data protection regulations	Controlled access, requires institutional approval	Approved researchers from academic institutions	Biobank of Chinese population samples, including genomic and health data
HuaBiao project	China	Governed by Chinese Academy of Sciences (CAS), Chinese Ministry of Science and Technology	Controlled access, requires ethics approval	Approved academic researchers	Genomic data of Chinese ethnic groups for anthropological and medical research
Indian Genetic Disease Database	India	Governed by Indian Institute of Chemical Biology (IICB) guidelines	Public access	Public	Common genetic diseases afflicting the Indian populations, covering 52 diseases with information on 5760 individuals carrying the mutant alleles of causal genes.
IndiGenomes	India	Governed by CSIR-Institute of Genomics and Integrative Biology (IGIB)	Public access	Public	Mutation data for 52 genetic diseases affecting Indian populations, covering 5760 individuals; includes locus heterogeneity, mutation types, clinical and biochemical data, geographical distribution, and common mutations
Indian Cancer Genome Atlas	India	Governed by ICGA Foundation, PRIDE guidelines, Indian data privacy laws	Open access for summary data; Controlled access for detailed multi-omics data	Public (for summary data), approved researchers and clinicians (for detailed data)	Multi-omics cancer genomic data, including DNA, RNA, and protein profiles, integrated with 50 breast cancer patients.
GenomeIndia Project	India	Funded by the Department of Biotechnology (DBT), Ministry of Science and Technology, Governed by BIOTECH-PRIDE guidelines, Indian Biological Data Centre (IBDC) policies	Controlled access	Approved academic researchers, government-approved projects	Whole-genome sequencing (WGS) data for 10,000 individuals, archived at IBDC; 20,000 samples collected from 83 diverse populations across India, forming a biobank for future research

dbGaP: Database of Genotypes and Phenotypes, TOPMed: Trans-Omics for Precision Medicine, GTEx: Genotype-Tissue Expression Project, GEO: Gene Expression Omnibus, dbVar: Database of Genomic Structural Variation, dbSNP: Single Nucleotide Polymorphism Database, RefSeq: NCBI Reference Sequence Database, SRA: Sequence Read Archive, TCGA: The Cancer Genome Atlas Program, DDBJ: DNA Database of Japan, KEGG: Kyoto Encyclopedia of Genes and Genomes, HGVD: Human Genetic Variation Database, GSA-Human: The Genome Sequence Archive for Human, ChinaMAP: China Metabolic Analytics Project, PGG.Han: Han Chinese Genomes Database, WBBC: Westlake BioBank for Chinese

GINA prevents genetic discrimination in insurance and employment, reducing public concerns over misuse of genetic information and indirectly encouraging participation in genetic studies, thereby expanding the diversity and scale of data sources [23]. The NIH Genomic Data Sharing Policy and the Common Rule form a dual-track system, balancing ethical considerations and technological safeguards. The former mandates controlled data submission to databases (e.g., dbGaP [24]) with tiered access, while the latter requires informed consent and dynamic consent updates for secondary data use. Certificates of Confidentiality and the proposed ADPPA further enhance this framework by providing legal protections against compelled disclosure and, if passed, standardizing privacy laws across states, easing cross-state data flow. These policies work together to build trust, ensure data quality, and streamline data-sharing practices.

United Kingdom The UK's policy framework for human genetic resource management, privacy protection, and data sharing is comprehensive, with multiple regulations working in tandem to ensure ethical governance and data security in genomic research. The Human Tissue Act 2004 (HTA 2004) [25] forms the foundation by regulating the use, storage, and disposal of human tissues and genetic material. It mandates explicit informed consent from donors for any use of genetic resources, ensuring that participant rights and privacy are safeguarded throughout the research process. Building on the HTA, the UK Biobank Ethics and Governance Framework [26] establishes guidelines for the management of genetic data collected by the UK Biobank. This framework promotes transparency, informed consent, and data privacy, while encouraging responsible data sharing for scientific progress. It strengthens the ethical governance of genetic resources, ensuring that the use of genetic data aligns with the public interest and research integrity. Further ensuring ethical standards, the Research Governance Framework for Health and Social Care [27] provides overarching principles for research involving human participants in the UK. This framework ensures that all studies, including those involving genetic data, adhere to ethical oversight and quality standards throughout the data generation process, safeguarding participants' welfare.

The Data Protection Act 2018 (DPA 2018) [28], which incorporates the General Data Protection Regulation (GDPR) [29] into UK law, provides robust protections for personal data, including genetic information. It mandates that explicit consent must be obtained for the collection and processing of genetic data, while granting individuals the right to access, correct, and control their data. The Act also requires data controllers to implement strict security measures to protect sensitive genetic data from unauthorized access or breaches. In parallel, the UK GDPR [30] enforces stringent rules for the collection, processing, and storage of personal data. It ensures that individuals' privacy is maintained while enabling data sharing for scientific research. The GDPR guarantees that genetic data can be shared under ethical guidelines, ensuring privacy protections and encouraging global collaboration. Finally, the Medical Research Council (MRC) Data Sharing Policy [31] promotes the sharing of genetic data generated from MRC-funded research. The policy sets protocols for data access, ensuring compliance with privacy laws and ethical standards. It fosters cross-institutional data sharing, clarifying responsibilities and conditions for data access and facilitating international research collaboration.

Together, these policies form a cohesive system that ensures responsible management of human genetic resources. The HTA ensures the lawful and ethical use of genetic material, while the UK Biobank Framework strengthens informed consent and governance. The Research Governance Framework ensures ethical and quality standards in data generation, and the DPA 2018 and UK GDPR provide strong protections for genetic data privacy. The MRC Data Sharing Policy facilitates controlled sharing of data, enhancing global collaboration. These policies prevent the misuse of genetic resources and privacy breaches, creating a closed-loop system of "compliant collection—strict protection—controlled sharing," which enables efficient, ethical, and collaborative medical research.

Japan The Act on the Protection of Personal Information (APPI) plays a pivotal role in Japan's regulatory framework for the management of human genetic resources [8]. Enacted to enforce strict controls over the handling of personal data, including genetic information, the APPI mandates that explicit consent be obtained from individuals before their genetic data is collected, used, or shared. This act emphasizes the protection of individual privacy and requires data controllers to implement robust security measures to safeguard genetic information against unauthorized access and breaches. The APPI sets high standards for data protection, ensuring that individuals' genetic data is handled with utmost care and confidentiality. Complementing the APPI are the Ethical Guidelines for Human Genome/Gene Analysis Research [7], which provide comprehensive standards for the ethical conduct of genetic research in Japan. These guidelines mandate that all genetic research undergo ethical review to ensure compliance with ethical standards and the protection of research participants. They require informed consent from all participants, emphasizing the necessity of transparency and understanding regarding the nature, purpose, and potential risks of the research. The guidelines also stipulate secure handling of genetic materi-

als and data, ensuring that genetic information is used responsibly and ethically in research. In addition to these, the Ethical Guidelines for Medical and Health Research Involving Human Subjects provide a broader framework for the ethical conduct of biomedical research involving human participants [7]. These guidelines cover all aspects of medical and health research, including studies that involve genetic data. They mandate ethical review and informed consent, ensuring that all research involving human subjects is conducted ethically and with respect for participants' rights and welfare. The guidelines also emphasize the importance of protecting participants' privacy and confidentiality, particularly when handling sensitive genetic information.

Middle-low-income countries: China, India, and Kenya

China China has established a comprehensive regulatory framework to manage and protect human genetic resources, combining laws and regulations that ensure the ethical use of genetic information and safeguard national security. The Biosecurity Law of the People's Republic of China, enacted in 2020, includes specific provisions to ensure the safety of human genetic and biological resources [32]. It clarifies regulations related to international cooperation and clinical trials involving China's human genetic resources, emphasizing the protection of genetic data from unauthorized access and misuse. The Regulations on the Management of Human Genetic Resources issued in 2019 by the State Council specify the powers and responsibilities of regulatory authorities and detail the legal liabilities for violations, establishing a legal framework for the collection, preservation, utilization, and provision of human genetic resources to ensure their ethical use [33]. The Implementation Rules for the Regulations on the Management of Human Genetic Resources, introduced in 2023, further standardize the management of human genetic resources, providing detailed procedures and guidelines for compliance, enhancing the regulatory framework established by the 2019 regulations [34].

In terms of data protection, the Data Security Law of 2021 establishes a framework for data security management, emphasizing the protection of critical data, including genetic information [35]. It mandates the classification and protection of data based on its importance to national security, public interests, and economic development, imposing stringent requirements on data processing activities and cross-border data transfers. Complementing this, the Personal Information Protection Law (PIPL), also enacted in 2021, is China's first comprehensive national law dedicated to personal information protection, including genetic data [36]. The PIPL sets strict guidelines for the collection, use, and processing of personal data, emphasizing the necessity of obtaining explicit consent and implementing robust security measures to protect sensitive information. By regulating cross-border data transfers and imposing significant penalties for non-compliance, the PIPL ensures the protection of individual privacy rights. The combination of these laws creates a robust framework that ensures the ethical management and protection of human genetic resources in China, balancing the need for scientific advancement with national security and individual privacy.

India India's management of human genetic resources is governed by legislation that emphasizes conservation, sustainable use, and ethical research practices. The Biological Diversity Act of 2002 aims to conserve biological diversity, promote sustainable use of its components, and ensure fair and equitable sharing of benefits arising from the utilization of biological resources [37]. This act establishes a regulatory framework through the National Biodiversity Authority (NBA), State Biodiversity Boards (SBBs), and local Biodiversity Management Committees (BMCs), overseeing access to biological resources and implementing benefit-sharing arrangements. Complementing this, the Biological Data Storage, Access, and Sharing Policy of India provides a structured framework for the management of biological data, particularly from high-throughput technologies like DNA sequencing [38]. It establishes a National Biological Data Centre for standardized and secure data handling, promotes open data sharing to enhance scientific discoveries, and mandates compliance with national laws and international standards to protect privacy and ethical considerations. Together, these frameworks ensure the ethical management of genetic resources and the protection of personal data, supporting India's commitment to responsible genetic research and biodiversity conservation. The Centre for Cellular and Molecular Biology (CCMB) plays a pivotal role in genomic studies and biological resource management. It has established ethical guidelines for cellular and molecular biology research, while also promoting genomic studies and personalized medicine.

Kenya Kenya has developed a comprehensive set of regulations to protect personal data, including genetic information, and promote the sustainable use of biological resources. The Data Protection Act of 2019 establishes a framework for managing personal data to ensure privacy and security [39]. It mandates data processors and controllers to adhere to principles of lawful, fair, and transparent data processing, emphasizing the protection of sensitive data, including genetic information, and requiring robust security measures. The Data Protection (General) Regulations of 2021 provide detailed guidelines for implementing the Data Protection Act, covering data pro-

cessing principles, consent and notification requirements, data subject rights, data security measures, and breach notification protocols [40]. These regulations also govern cross-border data transfers, ensuring adequate protection safeguards are in place.

In terms of biodiversity management, the National Biodiversity Action Plan (NBAP) serves as a comprehensive framework for conserving Kenya's biodiversity, promoting sustainable use of biological resources, and ensuring equitable benefit-sharing. The NBAP integrates biodiversity conservation into national development planning and involves multiple stakeholders, including government agencies, private sector, research institutions, and local communities. It emphasizes the conservation of genetic resources and the sustainable use of biotechnology, contributing to Kenya's commitment under the Convention on Biological Diversity (CBD). These policies ensure that Kenya manages and protects its genetic resources responsibly, balancing the advancement of genetic research with the protection of individual rights and biodiversity conservation. In this context, the KEMRI-Wellcome Trust Research Programme (KWTRP) plays a critical role in the ethical governance of genetic research. KWTRP has pioneered an innovative approach to dynamic consent, which allows participants to update their consent preferences in real time. This ensures that participants retain control over their data while enabling researchers to conduct cutting-edge genomic studies. The KWTRP framework facilitates responsible data sharing by ensuring compliance with Kenya's data protection laws, as well as international ethical standards. This aligns with the National Biodiversity Action Plan (NBAP) and the Data Protection Act, ensuring that genetic data is handled with the highest level of security and privacy protection.

Genetic resource databases

Genetic sequence databases play a crucial role in the systematic storage, management, and sharing of genetic data. These databases support scientific research and innovation by providing comprehensive resources for data access and analysis.

United States

The United States is home to a wide array of databases that support cutting-edge research in genomics and public health. dbGaP (Database of Genotypes and Phenotypes) [24] is a well-known resource that provides controlled access to genomic and phenotype data from various NIH-funded studies. All of Us [41] is another initiative that offers whole genome sequencing, EHR, and participant-reported data, promoting research on personalized medicine. TOPMed [42] (Trans-Omics for Precision Medicine) offers multi-omics data (including WGS, metabolomics, and proteomics), primarily focused on cardiovascular, lung, and blood disorders. In the cancer field, The Cancer Genome Atlas Program (TCGA) [43] offers multi-omics data, including WGS, transcriptomics, and clinical metadata, available under open access for summary data and controlled access for raw genomic and clinical data. For tissue-specific data, the Genotype-Tissue Expression (GTEx) project [44], provides detailed genomic, transcriptomic, and epigenomic data across different tissues. Additionally, GeneBank [45] is a public resource that provides nucleotide sequences and annotated genetic information, offering open access for users. dbVar [46] focuses on structural variations in the genome, and Gene Expression Omnibus (GEO) [47] is a repository of gene expression datasets, both of which are publicly accessible. RefSeq [48] is another open-access resource that provides curated reference sequences for genomes, genes, and proteins. The Sequence Read Archive [49] (SRA) hosts raw sequencing data from next-generation sequencing (NGS) studies, available through both controlled and open access tiers. Finally, the 1000 Genomes Project [50] offers whole genome sequencing data from global populations, available for open access, providing critical insights into genetic diversity.

United Kingdom

The United Kingdom boasts several leading databases that facilitate genetic research. UK Biobank [51], one of the largest genomic and health databases, includes data from approximately 500,000 participants. It offers genomic, phenotypic, lifestyle, and health record information and provides controlled access to researchers after ethics review. Genomics England, through its 100,000 Genomes Project [52], offers whole genome sequencing data from individuals with rare diseases and cancer, with access governed by NHS approvals. The European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI) [53] is another crucial UK-based resource, offering genomic, proteomic, and bioinformatics datasets to the public.

Japan

Japan's genomic research landscape is represented by several notable databases. The DNA Data Bank of Japan (DDBJ) [54] provides a wealth of genomic sequences, available to the public for scientific use. The Human Genetic Variation Database (HGVD) [55] offers data on genetic variations and SNP frequencies in the Japanese population, which is accessible for public research. Biobank Japan (BBJ) [56] contains genomic, clinical, and phenotypic data from a large cohort, contributing to the understanding of genetic factors in disease. KEGG [57] (Kyoto Encyclopedia of Genes and Genomes) provides annotated biological pathways, gene and protein

functional information, and metabolic networks, which are essential for systems biology research.

China

China has a rapidly growing collection of genomic databases that are central to advancing public health and biomedical research. The China Kadoorie Biobank (CKB) [58], with approximately 500,000 participants, provides comprehensive data on genetics, health, and lifestyle, facilitating research on a variety of diseases. ChinaMAP [59] (China Metabolic Analytics Project) offers genomic data from three million samples, with a focus on metabolic diseases. The Genome Sequence Archive for Humans (GSA-Human) [60] provides large-scale genomic and transcriptomic data for human health research, supporting a wide range of genomic studies. PGG.Han [61], with data from over 137,000 Han Chinese individuals, includes whole-genome sequencing, SNP variations, and population genetics data, offering valuable insights into the genetic diversity of the Han Chinese population. The Westlake BioBank for Chinese (WBBC) [62], a biobank with genomic and health data, focuses on the Chinese population and requires institutional approval for access. Similarly, HuaBiao [63], with genomic data from various Chinese ethnic groups, is intended for both anthropological and medical research.

India

India is making significant strides in genomic research with several national initiatives aimed at improving health outcomes. The GenomeIndia Project [64] aims to create a genomic database from 10,000 individuals, representing a diverse range of Indian populations, to advance precision medicine and research. The Indian Cancer Genome Atlas (ICGA) [65] provides multi-omics data, including genomic, transcriptomic, and proteomic data, for cancer research, with a focus on the Indian population. IndiGenomes [66] offers mutation data for 52 genetic diseases affecting the Indian population, and the Indian Genetic Disease Database [67] (IGDD) provides curated information on mutations in common genetic diseases, helping to improve understanding of inherited disorders. While some data is available publicly, many of these resources require controlled access for research purposes, with institutional approvals and ethical compliance.

Discussion

The management of human genetic resources presents challenges that span scientific research, privacy rights, and international law. This study has examined the national approaches to human genetic resource management in the United States, United Kingdom, Japan, China, India, and Kenya, highlighting the diverse regulatory frameworks and challenges inherent to each geopolitical landscape. Our findings reveal several key points: First, there is a need to establish systematic and comprehensive regulatory systems. These systems should be similar to those in the United States and the United Kingdom, where various regulations collaborate and provide checks and balances, covering all aspects of human genetic resource management. Second, each country should establish national data management centers similar to NCBI, EMBL-EBI, and DDBJ, with unified management by the state, and adopt international data standards and norms. Third, there is a need to strengthen data protection during data sharing to prevent the loss of genetic resources. To effectively manage and safeguard human genetic resources, a coordinated approach that encompasses comprehensive legal frameworks, standardized data platforms, and robust data protection measures is essential for fostering global collaboration and advancing scientific discovery.

The United States, United Kingdom, Japan, China, India, and Kenya each have their own robust, distinct frameworks for managing human genetic resources, reflecting the unique societal values, technological capacities, and legal traditions of each region. High-income countries like the U.S., UK, and Japan have developed comprehensive policies and advanced databases that ensure both the protection and utility of genetic data. For example, the U.S. has implemented laws such as HIPAA and GINA that protect privacy and prevent discrimination while facilitating genetic data sharing under well-defined conditions. However, the U.S. faces specific challenges in balancing privacy with research facilitation, especially in light of public concerns about genetic discrimination [68]. Despite GINA's protections, there remain societal fears surrounding genetic data use, particularly with employers, insurers, and within vulnerable populations [69]. The challenge, therefore, lies in navigating the tension between individual privacy rights and the societal benefits of genomic research. Additionally, the fragmented legal system in the U.S., with varying state laws, complicates nationwide data-sharing efforts, necessitating a harmonized approach to facilitate cross-state research and international collaborations.

In middle-low-income countries like China, India, and Kenya, the frameworks for managing human genetic resources are evolving. China has established the Regulations on HGRs and their detailed implementation rules, but it could benefit from enacting a dedicated Human Genetic Resources Management Law similar to the United States. Such a law would significantly enhance the protection of human genetic resources and provide clearer guidelines for international collaborations. The societal challenge here lies in balancing national security concerns with the need for global scientific cooperation.

The increasing reliance on genetic data from Chinese populations for international research must be handled carefully, considering the country's growing regulatory restrictions on data sharing. Additionally, while China's technological advancements are impressive, the rapid pace of progress in biotechnology presents challenges in developing coherent and consistent legal standards for genetic data management. India, similarly, faces societal challenges regarding informed consent, public trust, and the equitable distribution of benefits derived from genetic research [70]. There is also a pressing need to develop more cohesive and clear legal frameworks to guide genetic data usage, particularly as the nation is poised to expand its genomics research efforts. Kenya, a low-income country, faces significant technological barriers in genomic data management, such as limited infrastructure for sequencing technologies and bioinformatics. Moreover, public understanding of the benefits and risks of participating in genetic research is still developing. Societal concerns, including fears of exploitation and inadequate compensation for participants, also create challenges to building trust in genetic data sharing. While institutions like KEMRI are making strides [71], Kenya must address these challenges to fully benefit from the genetic research opportunities presented by international collaboration.

The potential for international collaboration and data sharing in the management of human genetic resources presents both significant opportunities and challenges, particularly when bridging the regulatory and cultural divides between high-income and middle-low-income countries. Opportunities arise from the complementary strengths of these nations: high-income countries often possess advanced technological infrastructure and robust legal frameworks, while middle-low-income countries contribute diverse genetic datasets that are critical for understanding global genetic diversity and addressing region-specific health challenges. For instance, collaborative initiatives like the Human Heredity and Health in Africa (H3Africa) [72] project demonstrate how equitable partnerships can enhance genomic research capacity in low-resource settings while generating globally relevant data. However, challenges are multifaceted. Regulatory disparities, such as differing standards for data privacy (e.g., GDPR in the UK vs. less stringent frameworks in some middle-low-income countries), can hinder data sharing due to concerns over compliance and ethical risks [73]. Cultural contexts also play a critical role; for example, some communities may have strong preferences against the commercialization of genetic data or may require culturally sensitive approaches to informed consent [74]. Additionally, the lack of harmonized international data standards exacerbates issues of interoperability and trust [75]. To address these challenges, adopting international data-sharing norms—such as the FAIR [76] (Findable, Accessible, Interoperable, Reusable) principles—could provide a common framework for data management. Furthermore, establishing bilateral or multilateral agreements that respect local regulations while ensuring data security and ethical use could facilitate collaboration. For example, the Global Alliance for Genomics and Health (GA4GH) [77] has developed frameworks like the Data Sharing Lexicon and Consent Clauses to bridge regulatory gaps. Ultimately, fostering international collaboration requires not only technical and legal harmonization but also a commitment to equitable partnerships that prioritize the needs and values of all stakeholders, particularly those in middle-low-income countries [78]. This approach would not only enhance global genomic research but also ensure that the benefits of genetic data sharing are distributed fairly across diverse populations.

The comparison between high-income and middle-low-income countries reveals several insights. High-income countries tend to have more comprehensive and integrated systems for managing genetic data, supported by robust legal frameworks and advanced technological infrastructures. These systems are often characterized by clear privacy protection laws, such as GINA in the U.S., which are balanced with the promotion of genomic research. However, middle-low-income countries, while making significant progress, often face challenges in establishing unified standards and ensuring comprehensive data protection. These countries also experience difficulties in translating technological advances into accessible infrastructure for data management. The potential for these countries lies in adopting and adapting best practices from high-income countries, including developing dedicated legal frameworks and enhancing technological capabilities for data management.

While this paper provides a comprehensive comparative analysis of human genetic resource management across major regions, it does have some limitations. The countries were limited to the United States, the United Kingdom, Japan, China, India, and Kenya, potentially overlooking the diverse regulatory approaches of developing countries or regions with unique cultural and ethical views on genetic data. However, the chosen countries provide a comprehensive view of both high-income and middle-low-income nations, offering insights that can be generalized to other regions with similar socio-economic and regulatory contexts, thus maintaining a certain degree of external validity. Additionally, the paper could benefit from deeper case studies or practical examples that more concretely illustrate the impact of these regulations on daily research and international collaborations. Furthermore, the rapid advancement in genomic technologies may soon outdate some discussions, as the study

might not fully capture the latest or emerging technologies that could significantly influence HGRs management practices. Addressing these areas in future research could enhance the robustness and relevance of the findings, ensuring that the policy recommendations remain applicable and effective in a rapidly evolving field.

Conclusion

This scoping review highlights the significant disparities in the management of human genetic resources between high-income countries and middle-low-income countries. High-income countries have developed robust systems with strong legal frameworks and advanced technologies, balancing privacy protection with research facilitation. Middle-low-income countries need to adopt unified standards, strengthen data protection, and promote equitable data sharing. A coordinated global approach is essential for advancing scientific discovery and ensuring fair management of genetic resources.

Acknowledgements
We acknowledge all our colleagues.

Author contributions
SKZ and BBH conceived the idea for the paper. HWL and YL wrote the first draft. YYZ, YQM, and QC were responsible for data collection. HFX and XYW performed data verification. XLG and HW created the figures and visualizations. ZLC contributed to the intellectual content and provided critical revisions. All authors reviewed and approved the final version of the paper.

Funding
This work was supported by the Henan Province key research and development project (grant number 231111312200).

Data availability
All the data generated or analysed during this scoping review have been included in this manuscript.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Clinical trial number
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou 450008, China
[2]Henan Province Engineering Research Center of Artificial Intelligence and Internet of Things Wise Medical, Zhengzhou, China

Received: 30 August 2024 / Accepted: 5 March 2025
Published online: 15 March 2025

References

1. Davies K. The era of genomic medicine. Clin Med (Lond) Dec. 2013;13(6):594–601. https://doi.org/10.7861/clinmedicine.13-6-594.
2. Pereira R, Oliveira J, Sousa M. Bioinformatics and computational tools for Next-Generation sequencing analysis in clinical genetics. J Clin Med Jan. 2020;3(1). https://doi.org/10.3390/jcm9010132.
3. Wolf LE, Hammack CM, Brown EF, Brelsford KM, Beskow LM. Protecting participants in genomic research: Understanding the web of protections afforded by federal and state law. J Law Med Ethics Mar. 2020;48(1):126–41. https://doi.org/10.1177/1073110520917000.
4. Act A. Health insurance portability and accountability act. Public Law. 2023.
5. Feldman EA. The genetic information nondiscrimination act (GINA): public policy and medical practice in the age of personalized medicine. J Gen Intern Med. 2012;27:743–6.
6. Sim I, Stebbins M, Bierer BE, et al. Time for NIH to lead on data sharing. Science. 2020;367(6484):1308–9.
7. Eba J, Nakamura K. Overview of the ethical guidelines for medical and biological research involving human subjects in Japan. Jpn J Clin Oncol May. 2022;31(6):539–44. https://doi.org/10.1093/jjco/hyac034.
8. Act on the Protection of Personal Information. the Japanese government. Accessed. April, 2024. https://www.cas.go.jp/jp/seisaku/hourei/data/APPI.pdf
9. Vlahou A, Hallinan D, Apweiler R, et al. Data Sharing Under the General Data Protection Regulation: Time to Harmonize Law and Research Ethics? Hypertens Apr. 2021;77(4):1029–35. https://doi.org/10.1161/hypertensionaha.120.16340.
10. Coppola L, Cianflone A, Grimaldi AM, et al. Biobanking in health care: evolution and future directions. J Transl Med May. 2019;22(1):172. https://doi.org/10.1186/s12967-019-1922-3.
11. Chen Y, Song L. China: concurring regulation of cross-border genomic data sharing for statist control and individual protection. Hum Genet Aug. 2018;137(8):605–15. https://doi.org/10.1007/s00439-018-1903-2.
12. Majumder MA. United States: law and policy concerning transfer of genomic data to third countries. Hum Genet Aug. 2018;137(8):647–55. https://doi.org/10.1007/s00439-018-1917-9.
13. Taylor MJ, Wallace SE, Prictor M. United Kingdom: transfers of genomic data to third countries. Hum Genet Aug. 2018;137(8):637–45. https://doi.org/10.1007/s00439-018-1921-0.
14. Tanjo T, Kawai Y, Tokunaga K, Ogasawara O, Nagasaki M. Practical guide for managing large-scale human genome data in research. J Hum Genet. 2021;66(1):39–52.
15. Olono A, Mitesser V, Happi A, Happi C. Building genomic capacity for precision health in Africa. *Nature Medicine*. 2024/07/01 2024;30(7):1856–1864. https://doi.org/10.1038/s41591-024-03081-9
16. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci Sep. 2010;20:5:69. https://doi.org/10.1186/1748-5908-5-69.
17. Health UDo, Services H. Federal policy for the protection of human subjects ('common Rule'). DHHS Version. 2014;45.
18. Health Insurance Portability and Accountability Act. US Department of Health and Human Services. Accessed. April, 2024. https://www.cdc.gov/phlp/publications/topic/hipaa.html
19. Genetic Information Nondiscrimination Act. US Department of Health and Human Services. Accessed. April, 2024. https://www.hhs.gov/
20. American Data Privacy and Protection Act. United States Congress. Accessed. April, 2024. https://www.congress.gov/bill/117th-congress/house-bill/8152/text
21. Wolf LE, Beskow LM. Certificates of confidentiality: Mind the gap. Utah Law Rev. 2021;2021(4):937–50. https://doi.org/10.26054/0d-80m8-7v82.
22. Moore W, Frye S. Review of HIPAA, part 1: history, protected health information, and privacy and security rules. J Nucl Med Technol. 2019;47(4):269–72.
23. Green RC, Lautenbach D, McGuire AL. GINA, genetic discrimination, and genomic medicine. N Engl J Med. 2015;372(5):397–9.
24. Tryka KA, Hao L, Sturcke A, et al. NCBI's database of genotypes and phenotypes: DbGaP. Nucleic Acids Res. 2014;42(D1):D975–9.
25. Human Tissue Act. 2004. UK government. Accessed April, 2024. https://www.legislation.gov.uk/ukpga/2004/30/introduction
26. Biobank U. UK biobank ethics and governance framework. UK BIOBANK. 2007;3.
27. Taylor M. Research governance framework for health and social care. Health Soc Care Commun. 2002;10(1):6–9.
28. UK G. Data Protection Act 2018. 2018.

29. Chico V. The impact of the general data protection regulation on health research. Br Med Bull Dec. 2018;1(1):109–18. https://doi.org/10.1093/bmb/ldy038.
30. Halawi L, Makwana A. The GDPR and UK GDPR and its impact on US academic institutions. Issues Inform Syst. 2023;24(2).
31. the Medical Research Council (MRC) Data Sharing Policy. Accessed. April, 2024. https://www.ukri.org/publications/mrc-data-sharing-policy/
32. Wang L, Wang F, Zhang W. Bioethics in China's biosecurity law: forms, effects, and unsettled issues. J Law Biosci Jan-Jun. 2021;8(1):lsab019. https://doi.org/10.1093/jlb/lsab019.
33. Regulations on management of human genetic resources. State Council. https://english.www.gov.cn/policies/latest_releases/2019/06/10/content_281476708945462.htm
34. Implementation Rules for the Regulations on the Management of Human Genetic Resources. The State Council. https://english.www.gov.cn/news/202306/02/content_WS64792331c6d0868f4e8dc8f2.html
35. Daoli H, Wenhua H, Situation. the People's Republic of China (Draft). J Beijing Univ Aeronaut Astronaut Social Sci Ed. 2020;33(6):9–17. Dilemma and Countermeasures of China's Data Security Legislation: Comment on the Data Security Law of.
36. Calzada I. Citizens' data privacy in China: the state of the Art of the personal information protection law (PIPL). Smart Cities. 2022;5(3):1129–50.
37. Gadgil M. India's biological diversity act 2002: an act for the new millennium. J Biosci Mar. 2003;28(2):145–7. https://doi.org/10.1007/bf02706212.
38. Ministry of Science & Biological Data Storage, Access, and Sharing Policy, Technology. https://dbtindia.gov.in/latest-announcement/biological-data-storage-access-and-sharing-policy-india
39. THE DATA PROTECTION ACT. National Council for Law Reporting. https://www.kentrade.go.ke/wp-content/uploads/2022/09/Data-Protection-Act-1.pdf
40. The Data Protection. (General) Regulations of 2021. National Council for Law Reporting. https://www.odpc.go.ke/wp-content/uploads/2024/03/THE-DATA-PROTECTION-GENERAL-REGULATIONS-2021-1.pdf
41. Investigators AoURP. The all of Us research program. N Engl J Med. 2019;381(7):668–76.
42. Burgess DJ. The topmed genomic resource for human health. Nat Rev Genet. 2021;22(4):200–200.
43. Tomczak K, Czerwińska P, Wiznerowicz M. Review the Cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncology/Współczesna Onkologia. 2015;2015(1):68–77.
44. Consortium G. The GTEx consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369(6509):1318–30.
45. Benson DA, Cavanaugh M, Clark K, et al. GenBank. Nucleic Acids Res. 2012;41(D1):D36–42.
46. Lappalainen I, Lopez J, Skipper L, et al. DbVar and DGVa: public archives for genomic structural variation. Nucleic Acids Res. 2012;41(D1):D936–41.
47. Clough E, Barrett T. The gene expression omnibus database. Stat Genomics: Methods Protocols. 2016:93–110.
48. O'Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44(D1):D733–45.
49. Leinonen R, Sugawara H, Shumway M, Collaboration INSD. The sequence read archive. Nucleic Acids Res. 2010;39(suppl1):D19–21.
50. Siva N. 1000 Genomes project. Nat Biotechnol. 2008;26(3):256–7.
51. Allen N, Sudlow C, Downey P, et al. UK biobank: current status and what it means for epidemiology. Health Policy Technol. 2012;1(3):123–6.
52. Turnbull C, Scott RH, Thomas E et al. The 100 000 genomes project: bringing whole genome sequencing to the NHS. BMJ. 2018;361.
53. Emmert DB, Stoehr PJ, Stoesser G, Cameron GN. The European bioinformatics Institute (EBI) databases. Nucleic Acids Res. 1994;22(17):3445–9.
54. Tanizawa Y, Fujisawa T, Kodama Y, et al. DNA data bank of Japan (DDBJ) update report 2022. Nucleic Acids Res Jan. 2023;6(D1):D101–5. https://doi.org/10.1093/nar/gkac1083.
55. Higasa K, Miyake N, Yoshimura J, et al. Human genetic variation database, a reference database of genetic variations in the Japanese population. J Hum Genet. 2016;61(6):547–53.
56. Nagai A, Hirata M, Kamatani Y, et al. Overview of the biobank Japan project: study design and profile. J Epidemiol. 2017;27(SupplementIII):S2–8.
57. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res Jan. 2000;1(1):27–30. https://doi.org/10.1093/nar/28.1.27.
58. Chen Z, Chen J, Collins R, et al. China kadoorie biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol. 2011;40(6):1652–66.
59. Cao Y, Li L, Xu M, et al. The ChinaMAP analytics of deep whole genome sequences in 10,588 individuals. Cell Res. 2020;30(9):717–31.
60. Zhang SS, Chen X, Chen TT, et al. GSA-Human: genome sequence archive for human. Yi Chuan Oct. 2021;20(10):988–93. https://doi.org/10.16288/j.yczz.21-248.
61. Gao Y, Zhang C, Yuan L, et al. PGG. Han: the Han Chinese genome database and analysis platform. Nucleic Acids Res. 2020;48(D1):D971–6.
62. Zhu X-W, Liu K-Q, Wang P-Y, et al. Cohort profile: the Westlake biobank for Chinese (WBBC) pilot project. BMJ Open. 2021;11(6):e045564.
63. Hao M, Pu W, Li Y, et al. The HuaBiao project: whole-exome sequencing of 5000 Han Chinese individuals. J Genet genomics = Yi Chuan Xue Bao. 2021;48(11):1032–5.
64. The GenomeIndia Project. Accessed. January, 2025. https://cbr-iisc.ac.in/genomeindia/
65. Muddabhaktuni BMC, Koyyala VPB. The cancer genome atlas. Indian J Med Pediatr Oncol. 2021;42(04):353–5.
66. Jain A, Bhoyar RC, Pandhare K, et al. IndiGenomes: a comprehensive resource of genetic variants from over 1000 Indian genomes. Nucleic Acids Res. 2021;49(D1):D1225–32.
67. Pradhan S, Sengupta M, Dutta A, et al. Indian genetic disease database. Nucleic Acids Res. 2010;39(suppl1):D933–8.
68. Lewis A, Green R, Prince A. Long-awaited progress in addressing genetic discrimination in the united States. Genet Sci. 2021;23(3):429–31.
69. Joly Y, Dalpe G. Genetic discrimination still casts a large shadow in 2022. Eur J Hum Genet. 2022;30(12):1320–2.
70. Patra PK, Sleeboom-Faulkner M. Informed consent in genetic research and biobanking in India: some common impediments. Genomics Soc Policy. 2009;5:1–14.
71. Simiyu K, Masum H, Chakma J, Singer PA. Turning science into health solutions: KEMRI's challenges as Kenya's health product pathfinder. BMC Int Health Hum Rights. 2010;10:1–9.
72. Peprah E, Wiley K, Sampson U, Narula J. A new age for african-driven genomics research: human heredity and health in Africa (H3Africa). Global Heart. 2017;12(2).
73. Kosseim P, Dove ES, Baggaley C et al. Building a data sharing model for global genomic research. *Genome Biol*. Aug 11. 2014;15(8):430. https://doi.org/10.1186/s13059-014-0430-2
74. Colledge F, Elger B, Howard HC. A review of the barriers to sharing in biobanking. Biopreserv Biobank. 2013;11(6):339–46.
75. Townend D. Conclusion: harmonisation in genomic and health data sharing for research: an impossible dream? Hum Genet Aug. 2018;137(8):657–64. https://doi.org/10.1007/s00439-018-1924-x.
76. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data Mar 15. 2016;3:160018. https://doi.org/10.1038/sdata.2016.18.
77. Rehm HL, Page AJ, Smith L et al. GA4GH: international policies and standards for data sharing across genomic research and healthcare. Cell Genomics. 2021;1(2).
78. Chaturvedi S, Crager S, Ladikas M, Muthuswamy V, Su Y, Yang H. Harmonizing policy on human genetic resources and benefit sharing. Nat Biotechnol Dec. 2012;30(12):1169–70. https://doi.org/10.1038/nbt.2441.

Publisher's note