BioLedger: A Token Incentive-based Framework for the Secure Sharing of Genetic Data in Consortium Blockchain

Xin Wang†‡
*Shenzhen Technology University**
Shenzhen, 518118, China
wangxin@sztu.edu.cn

Ruixuan Lin†
*Shenzhen Technology University**
Shenzhen, 518118, China
202200202055@stumail.sztu.edu.cn

Jiaming Tang
*Shenzhen Technology University**
Shenzhen, 518118, China
202200201002@stumail.sztu.edu.cn

Ruisheng Huang
*Shenzhen Technology University**
Shenzhen, 518118, China
202200201028@stumail.sztu.edu.cn

Huankai Chen
*Shenzhen Technology University**
Shenzhen, 518118, China
202200202183@stumail.sztu.edu.cn

Lixin Liang
*Shenzhen Technology University**
Shenzhen, 518118, China
lianglixin@sztu.edu.cn

Bingding Huang‡
*Shenzhen Technology University**
Shenzhen, 518118, China
huangbingding@sztu.edu.cn

2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 979-8-3315-1557-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/BIBM66473.2025.11357073

***Abstract*—Current health data sharing frameworks encounter significant challenges in genetic data management, including issues related to privacy protection, constraints on access control, and a lack of effective incentive mechanisms. These deficiencies expose patients and researchers to risks of information leakage and result in inefficiencies in collaboration throughout the entire lifecycle of genetic data—from creation and access to sharing and revocation. Such problems have impeded the advancement of precision medicine and escalated the costs of data acquisition in medical research. Notably, the security and efficiency of data circulation become particularly acute when genetic information is shared across institutional boundaries to support the development of personalized treatment regimens or pharmaceutical research. To tackle these challenges, we propose a genetic data sharing framework based on consortium blockchain, BioLedger. In contrast to traditional centralized storage models, this framework leverages the FISCO BCOS consortium blockchain to construct a decentralized architecture. By integrating smart contracts with HCE (Hybrid CP-ABE with ECIES Encryption), it achieves fine-grained access control while ensuring a high standard of security and privacy for genetic data. Each user retains a high degree of control over their genetic data and can earn token rewards through data contribution, thereby incentivizing user participation in data sharing. Our solution aims to establish an ecosystem that not only safeguards data security and privacy but also enables efficient data exchange, with applications spanning genetic research, pharmaceutical development, and personalized medicine.**

Code Availabilnnity:

https://github.com/WangLabforComputationalBiology/GeneSharing

***IndexTerms*—genetic data sharing; consortium blockchain; smart contract; token incentive; hybrid cryptographic algorithm**

I. Introduction

In the digital healthcare environment, individuals have raised higher demands for the integrity, accessibility, and security of their health data. Genetic data, in particular, has attracted significant attention due to its high sensitivity and crucial role in precision medicine. Traditional centralized management models often result in genetic data being scattered across different institutions, forming "data silos." This not only limits the potential of medical research, personalized treatment, and health management but also increases the risks of privacy leakage and data loss [1]. Some genetic testing companies have sold data without authorization, violating data ownership and privacy rights and triggering concerns about data sovereignty [2]. Furthermore, the lack of interoperability standards complicates cross-institutional data sharing and exacerbates nnsecurity risks.

To address these challenges, blockchain technology, with its features of decentralization, immutability, and traceability, offers a new solution for genetic data sharing [3]. However, existing blockchain applications face challenges such as limited scalability and low user participation (due to insufficient privacy and compliance) [4]. We propose a consortium blockchain-based framework that utilizes FISCO BCOS, BioLedger, combined with HCE and smart contracts, to achieve secure and user-centric data sharing. Through a token incentive mechanism to reward data contributions, participation is enhanced while privacy and ownership are protected. This framework supports the reliable circulation of data for precision medicine and genetic research.

II. Background And Relatedwork

With the advancement of precision medicine and digital health, genomic data has become essential for disease diagnosis, personalized treatment, and drug development [5]. Compared to traditional electronic health records (EHRs), genomic data's highly personalized and sensitive nature imposes stricter requirements on privacy protection, access control, and data sovereignty [6]. However, current genomic data management and sharing systems face challenges such as privacy breaches, data silos, limited interoperability, and a lack of incentive mechanisms, which restrict their potential in medical research and health management [7]. This section reviews these challenges from three perspectives—genomic data storage, transmission issues, and blockchain applications—while outlining the innovative foundation of BioLedger.

A. Current Status of Genetic Data Storage

Genetic data is mainly stored by genetic testing institutions, medical institutions, or research units through centralized systems, covering genotypic data (such as SNP loci) and phenotypic data (such as health characteristics) [8]. Centralized storage facilitates management but is vulnerable to network attacks or equipment failures, increasing the risk of data leakage. Some genetic testing companies sell user data without authorization, violating data ownership and triggering trust issues [9]. Although cloud storage improves scalability, its reliance on third-party service providers exacerbates privacy risks [10]. In addition, the lack of unified data standards leads to genetic data being scattered in isolated systems, forming "data silos"[11]. For example, reports generated by individuals on platforms such as WeGene or 23andMe are difficult to integrate, requiring users to manage them manually, which significantly increases operational complexity and leakage risks [12].

B. Issues in Genetic Data Transmission

The need for genomic data sharing arises from healthcare, medical research, and cross-institutional collaboration [13]. However, existing transmission methods are flawed. Some users rely on insecure channels, such as email, instant messaging tools (e.g., WhatsApp), or USB drives, which are inefficient and prone to privacy breaches, violating regulations such as GDPR [14]. Even with API-based integration, incompatible protocols or non-standardized formats often lead to data loss, requiring manual verification and increasing time costs [15]. In cross-regional or cross-border scenarios, transmission complexities further underscore the need for secure, efficient, and standardized sharing mechanisms [16].

C. Blockchain Technology in Genetic Data Sharing

Blockchain technology, with its advantages of decentralization, immutability, and traceability, provides a new solution for the secure sharing of genetic data [17].

- *User Identity Management*

Effective user identity management in genomic data sharing requires decentralized control and privacy protection. MetaMask, a Web3-based wallet, utilizes the ECDSA algorithm (secp256k1 curve) to generate public-private key pairs, enabling users to autonomously manage their blockchain identities [18]. Private keys, stored locally, are used for transaction signing and authentication, while public keys derive blockchain addresses that interact with FISCO BCOS smart contracts for decentralized authentication [19]. This mechanism allows users to control data access without relying on centralized authorities, enhancing privacy and security while facilitating integration with the framework's ecosystem [20].

- Consortium Blockchain Platform Design

As an enterprise-level consortium blockchain, FISCO BCOS is suitable for genetic data sharing due to its high privacy protection and customizability [21]. It supports smart contracts, private transactions, and national cryptographic algorithms to ensure that data is only visible to authorized parties [22]. Its high-performance consensus mechanism meets the needs of concurrent access, with hash values stored on the chain and original data encrypted and stored off-chain, reducing costs while ensuring immutability [23].

- Blockchain-Based Genomic Data Storage

Blockchain's decentralized and tamper-resistant features enable secure and verifiable genomic data storage. A common approach stores encrypted data hashes on-chain while maintaining raw data in off-chain distributed file systems (e.g., IPFS), ensuring data integrity and traceability [24]. For example, genomic files (e.g., VCF or FASTQ) are encrypted and uploaded to IPFS, with corresponding hashes and access control policies recorded on the blockchain. Smart contracts automate access management, generating tamper-proof authorization logs to ensure transparency and trustworthiness in the sharing process [25].

D. Typical Application Cases

In recent years, blockchain technology has gained significant traction in genomic data sharing, with research focusing on public and consortium blockchain approaches, achieving progress in data privacy, access control, and incentive mechanisms. However, limitations persist.

Public blockchain platforms, such as Genobank, Encrypgen, and Nebula Genomics, leverage client-side encryption and smart contracts to ensure data ownership, using NFTs or token-based incentives to enhance user engagement [26]. However, inherent public blockchain limitations, including low throughput (approximately 27 TPS) and high transaction costs, restrict their ability to handle large-scale genomic data. Critically, these platforms lack dynamic permission revocation, preventing data owners from flexibly adjusting access rights and failing to meet privacy compliance needs in complex medical scenarios.

In contrast, consortium blockchain solutions offer improved performance. LifeCODE.ai, built on the Quorum platform, achieves higher processing efficiency and regulatory compliance for the Chinese market but employs coarse-

grained access control, lacking fine-grained dynamic authorization [27]. Similarly, the framework proposed by Sahi et al. [28] enhances data availability through distributed storage but fails to address dynamic permission management or incentive mechanism deficiencies. Wu et al. [29] advanced access control by integrating fine-grained permission management and dynamic revocation, yet the absence of cryptoeconomic incentives may limit user participation.

E. Typical Application Cases

Despite the potential demonstrated by these projects, existing solutions face several challenges:

1) Performance and Scalability. The large volume and high access frequency of genomic data pose challenges for blockchain platforms such as Ethereum and Hyperledger, which suffer from low throughput and high latency in high-concurrency scenarios [30]. While FISCO BCOS improves performance through the PBFT consensus mechanism, ultra-large-scale node deployment requires further optimization to support real-time processing of massive genomic data [31].

2) Privacy Protection and Compliance. The high sensitivity of genomic data demands robust privacy protection. Although existing solutions employ encryption, support for the "right to be forgotten" is limited, making it challenging to balance blockchain's immutability with dynamic authorization needs, as required by regulations such as GDPR [32][33].

3) User Participation and Incentive Mechanisms. Most frameworks focus on inter-institutional data sharing, overlooking individual users' data sovereignty needs. The lack of incentive mechanisms discourages active data sharing, resulting in low participation [34].

4) Heterogeneous systems hinder data sharing efficiency. While standards such as BioCompute Object (BCO, IEEE 2791-2020) aim to unify high-throughput sequencing (HTS) workflows and results, their integration with blockchain systems remains inadequate [35].

To address these challenges, we construct a user-centric genetic data sharing framework through the FISCO BCOS consortium blockchain, Bioledger, combined with Web3 wallets, OAuth2 protocol, HCE, local decryption clients, and ERC20 (Ethereum Request for Comments 20) token incentive mechanisms. Bioledger allows users to share data with individuals or institutions according to their wishes, ensures data ownership, and implements access control, data traceability, and privacy protection through smart contracts. It overcomes the limitations of existing solutions and provides an innovative approach for the secure and efficient circulation of genetic data.

III. TECHNICAL FOUNDATIONS

A. Blockchain, Authentication, and IPFS

Bioledger is built on FISCO BCOS, an enterprise-grade permissioned blockchain platform developed by WeBank, providing a secure and efficient data transaction environment for authorized participants. FISCO BCOS supports smart contracts—automated programs executed on the blockchain—that implement predefined logic for data operations, access control, and token distribution, ensuring transparency, immutability, and decentralized trust. Smart contracts validate and record transactions through a consensus mechanism, guaranteeing data integrity and consistency, thus forming a reliable foundation for genomic data sharing.

Users access Bioledger via MetaMask, a Web3-based digital wallet, enabling decentralized authentication and blockchain interaction. MetaMask employs the ECDSA algorithm (secp256k1 curve) to generate public-private key pairs for managing user identities and transactions on the blockchain. Private keys, stored locally, are used for transaction signing and authentication, while public keys derive blockchain addresses that interact with FISCO BCOS smart contracts. The ECDSA algorithm's efficiency and security make it widely adopted in blockchain systems, ensuring reliable authentication and tamper-proof transactions. Additionally, Bioledger integrates the OAuth2 protocol, allowing secure and compliant acquisition of phenotypic and genotypic data from trusted third-party sources based on user consent.

To address blockchain's storage limitations for large-scale data (e.g., medical images), Bioledger incorporates the InterPlanetary File System (IPFS), a peer-to-peer distributed storage system that supports permanent, immutable storage of large datasets. Users upload data to IPFS, receiving a unique hash link for efficient data retrieval. Combined with on-chain hash storage, this approach ensures data integrity and traceability.

B. Incentive Mechanism

Based on the above technologies, Bioledger adopts blockchain incentive mechanisms and refined access control to realize data sharing, privacy protection, and user participation. The incentive mechanism is centered on ERC20 tokens, which are a standardized token protocol of Ethereum that defines interfaces for token creation, transfer, and management, and are widely used in decentralized applications. As an Ethereum-compatible platform, FISCO BCOS supports ERC20 tokens, and the transparency and immutability of distribution are ensured through smart contracts. When institution personnel verified by email view the genetic data uploaded by users, smart contracts automatically distribute ERC20 tokens as rewards. Users can use tokens to redeem framework services, such as commodities around colleges and universities or data analysis, promoting participation in data sharing.

C. Hybrid Cipher Algorithm Based on CP-ABE and ECIES

Ciphertext-Policy Attribute-Based Encryption (CP-ABE) is an encryption mechanism that supports fine-grained access control [36]. Data owners use CP-ABE to define access policies, encrypting plaintext into ciphertext that only users with specific attributes can decrypt, enabling one-to-many access control ideal for secure data sharing in distributed environments. CP-ABE comprises four core algorithms: Setup generates the system public key and master key; Encrypt encrypts plaintext based on the access policy; KeyGen generates decryption keys based on user attributes; and Decrypt recovers plaintext when the key's attributes satisfy the policy. This mechanism supports dynamic and flexible access control, making it suitable for scenarios requiring granular permission management.

To further enhance the security of CP-ABE decryption keys, Bioledger introduces the Elliptic Curve Integrated Encryption Scheme (ECIES), which combines the efficiency of symmetric encryption with the security of Elliptic Curve Cryptography (ECC). ECIES ensures confidentiality, integrity, and authentication by utilizing ECC-based public-key cryptosystems, Key Derivation Functions (KDF), symmetric encryption algorithms (e.g., AES), and Message Authentication Codes (MAC) [37]. In ECIES, the encryption process takes the recipient's public key and plaintext message as input, generating a ciphertext containing the sender's ephemeral public key, encrypted message, and authentication tag. The decryption process uses the recipient's private key to recover the plaintext and verifies integrity through the MAC. ECIES is highly efficient because the key size of ECC is smaller than that of other public-key systems (e.g., RSA), providing equivalent security while reducing computational overhead. MetaMask, as a widely used Web3 wallet, automatically generates public-private key pairs for each user based on the secp256k1 curve. This key pair can be seamlessly integrated into the ECIES system, where the public key is used for encryption, and the private key securely stored in MetaMask is used for decryption. When decrypting through the local decryption client, users use the private key stored in MetaMask to decrypt the key package and obtain the CP-ABE private key for data access, thereby achieving secure and efficient key distribution.

Building on this, we propose the Hybrid CP-ABE with ECIES Encryption (HCE) algorithm, which combines the policy-driven access control of CP-ABE with the efficient and secure key encapsulation of ECIES. The process of the HCE algorithm is as follows: first, the Setup and Encrypt of CP-ABE generate encrypted data; KeyGen generates an attribute private key, and then ECIES encrypts this private key using the secp256k1 public key of MetaMask; the Decrypt process decrypts using the secp256k1 private key of MetaMask, verifies the attribute policy, and recovers the plaintext. HCE integrates the fine-grained access control of CP-ABE with the efficient encryption of ECIES, combined with the key management of MetaMask, significantly improving the security, flexibility, and user experience of genetic data sharing, and is particularly suitable for distributed scenarios in consortium blockchain environments.

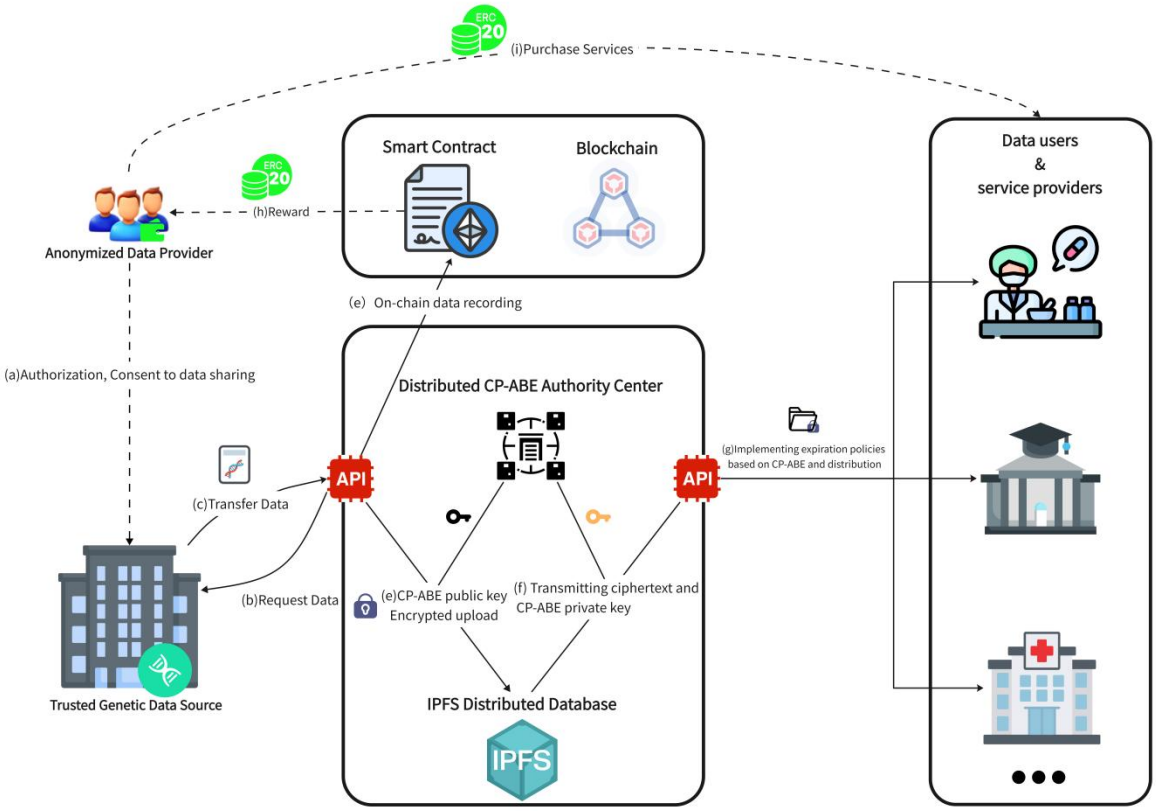

Fig 1. The overall architecture of BioLedger. The data provider authorizes their genetic data via OAuth2, and the trusted genetic data source encrypts the data using the CP-ABE master public key before uploading it to IPFS. When a data user requests data via the API, the API verifies the data provider's sharing status. If permitted, it retrieves the encrypted data from IPFS and distributes it to the data user with an embedded expiration policy. Each successful access triggers a smart contract, sending a predefined amount of ERC20 token rewards to the data provider's MetaMask wallet, incentivizing data sharing. At the same time, data providers can exchange tokens in their MetaMask wallets for value-added services (such as data analysis reports) from service providers. The solid lines indicate data flow, while the dashed lines represent authorization or incentive processes.

IV. BIOLEDGER FRAMEWORK DESIGN

This section elaborates on the entity design and architecture of Bioledger to support user-driven data circulation and privacy protection.

A. System design

The entities of Bioledger include anonymized data providers, trusted genetic data sources, data users/service providers, distributed CP-ABE authority centers, backend API services, and blockchain. The following outlines the functions and architectural design of each entity:

1) **Anonymized Data Provider.** Data owners authorize their genetic data to Bioledger via the OAuth2 protocol. Users operate on the blockchain using MetaMask's decentralized identity, achieving the decoupling of personal genetic data from specific identities.

2) **Trusted Genetic Data Source.** The trusted genetic data source (i.e., the server of the genetic testing institution) is responsible for encrypting genetic data using CP-ABE public key encryption via API and then uploading it to IPFS.

3) **Data Users or Service Providers.** Data users (e.g., hospital doctors or pharmaceutical researchers) request encrypted genetic data via an API. If the data provider authorizes sharing, the API distributes a CP-ABE private key with a timestamp and an encrypted genetic data package. Data users, acting concurrently as service providers, offer value-added services (e.g., data analysis) to data providers.

4) **Blockchain.** The blockchain works in conjunction with IPFS, where IPFS stores CP-ABE encrypted genetic data, and the blockchain records IPFS hash links, data upload records, and time-sensitive access permissions through smart contracts, enabling traceable data access.

5) **Distributed CP-ABE Authority Center.** To prevent single-point failures, the authority center adopts a distributed architecture (e.g., multi-node consensus) and is responsible for public key and private key distribution, ensuring the security of the encryption process.

B. The details of data upload and data distribution

This section describes in detail the genetic data upload and distribution process based on HCE encryption and IPFS: Data upload is encrypted using CP-ABE and stored on IPFS, with metadata recorded on the blockchain; data distribution integrates access policies and private key distribution to support dynamic sharing management.

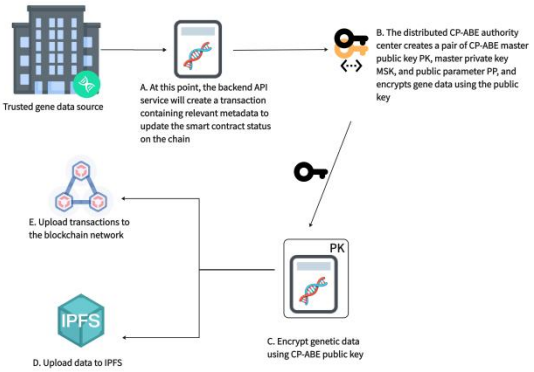

Fig. 2. The secure on-chain process of genetic data. This process includes trusted collection, CP-ABE encryption, distributed storage (IPFS), and blockchain certification (FISCO BCOS).

The following are the specific steps for uploading data:

1) **Data Acquisition:** Raw genetic data D0 is obtained from authenticated genetic data sources.

2) **Transaction Construction:** Blockchain transaction metadata is generated from the genetic data, comprising hash-linked references and access policy parameters.

3) **CP-ABE Setup:** The distributed CP-ABE authority center generates the Ciphertext-Policy Attribute-Based Encryption (CP-ABE) public key (PK), master secret key (MSK), and public parameters (PP). The MSK is securely stored by distributed key authorities, while PK and PP remain publicly accessible.

4) **Encrypted Data:** Use the generated public key PK to encrypt D0 to obtain the encrypted genetic data D1.

5) **On-Chain State Update & IPFS Upload:** Upload D1 to the IPFS distributed network to obtain a unique hash link; then record the hash link and metadata on the FISCO BCOS blockchain through a smart contract.

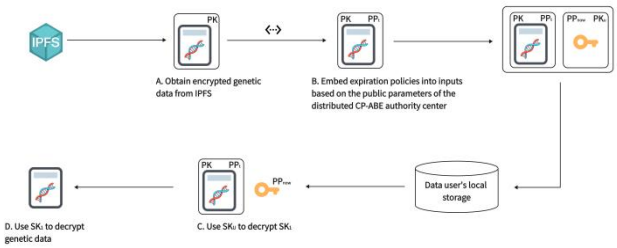

Fig. 3. The complete distribution process of genetic data. This process encompassing secure acquisition, policy embedding, encrypted key transmission, and local decryption verification.

The following are the specific steps for data distribution:

1) **Obtaining Data:** The data user obtains the PK-encrypted genetic data D1 from the IPFS network through the API.

2) **Access Policy Embedding:** The distributed CP-ABE authority center utilizes public parameters (PP) to embed the data uploader-defined access policy (e.g., expiration time) into D1 to obtain data D2.

3) **Private Key Generation & Encryption:** The authority center generates a timestamped private key (SK) using the master secret key (MSK), then encrypts SK with the data requester's MetaMask public key (PKu) via ECIES (Elliptic Curve Integrated Encryption Scheme), producing SK1.

4) **Secure Transmission:** D2 and SK1 are encrypted and transmitted to the data user's local storage through the API.

5) **Private Key Decryption:** The data user uses the MetaMask private key SKu based on ECIES to decrypt SK1 and obtain SK.

6) **Decrypt Data:** Use SK to decrypt D2. If the access policy is not satisfied (e.g., expired), the decryption fails and is recorded in the blockchain log. After successful

decryption, the smart contract is triggered and ERC20 token rewards are distributed to the data provider's MetaMask wallet.

C. System goals

BioLedger achieves the following goals:

- **Decentralization.** The system should operate without reliance on a centralized third party for data management.
- **Accountability.** The scheme should detect malicious behavior by dishonest data users.
- **User-Centered.** Data sharing decisions are determined by the data uploader.
- **Traceability.** The data sharing process is traceable and verifiable, with malicious behavior serving as evidence of accountability.
- **Revocable Access.** The system supports users in revoking access permissions at any time.
- **Dynamic Access Adjustment.** The system allows users to dynamically adjust the access scope based on needs.

V. Discussion

This section focuses on BioLedger, a token incentive-driven framework for secure gene data sharing based on consortium blockchain. The section systematically analyzes the framework's architectural design, data transmission efficiency, and security/privacy protection mechanisms, evaluates its feasibility, and discusses the impact of token incentives on user participation.

A. Architectural Core Design

Bioledger utilizes consortium blockchain as its underlying infrastructure and employs FISCO BCOS to build a permissioned control network. It verifies the identities of participants (such as users and medical institutions) through node certificate mechanisms to ensure the credibility of accessing entities. Bioledger introduces smart contracts to automate data authorization and access control, combining MetaMask with OAuth2 protocol (a third-party authorization standard). While ensuring the uniqueness and immutability of user identities, it reduces the risk of password leakage, achieves decoupling of genetic data and personal information, and enhances both security and convenience.

In terms of data storage, Bioledger adopts an on-chain and off-chain hybrid storage architecture: genetic data is encrypted by CP-ABE and stored on IPFS, while only data hash values and access control information are recorded on-chain. This design leverages blockchain's immutability to ensure data integrity and traceability, while off-chain storage reduces on-chain pressure and improves scalability [38]. Additionally, Bioledger integrates a multi-signature mechanism, requiring critical operations (such as data access changes) to be confirmed by at least 2/3 of the nodes, thereby enhancing resistance to attacks and establishing a stronger trust foundation [39].

B. Data Transmission and Interoperability

To enhance secure sharing of genetic data between research institutions and storage platforms, Bioledger builds decentralized management based on blockchain technology, supporting distributed storage of various genetic data formats including VCF and FASTQ. By recording data hash values and access policies on the chain, data integrity and traceability are ensured; while off-chain, IPFS is utilized to store encrypted genetic data. Meanwhile, users can achieve granular access control through smart contracts, such as specifying access permissions for specific research institutions or time periods, thereby enhancing privacy protection and data security.

C. Security and Privacy Considerations

Privacy protection is the core objective of Bioledger's design. Given the extremely high sensitivity of genetic data, a balance must be achieved between security and privacy mechanisms, technological implementation, potential risk response, and compliance requirements. The following sections provide a detailed analysis from the aspects of encryption technology, blockchain characteristics, supporting mechanisms, and compliance.

First, HCE (Attribute-Based Encryption) serves as the encryption pillar of Bioledger, implementing fine-grained access control through ciphertext policy. Data providers can define access policies (such as "limited to pharmaceutical company researchers, valid for 7 days"), encrypt genetic data using the master public key, and only users who hold the corresponding private key and whose attribute set meets the conditions can decrypt it. This mechanism, combined with timestamp policy, ensures that data cannot be accessed after the permission expires, effectively preventing unauthorized reuse. The distributed CP-ABE authority center uses multi-node consensus to distribute private keys, reducing the risk of single point of failure. This not only enhances the robustness of Bioledger but also fundamentally eliminates the possibility of man-in-the-middle attacks, laying a solid foundation for the secure flow of genetic data in complex network environments.

Secondly, the immutability of blockchain provides traceability for data access. FISCO BCOS records every access log (including timestamp, accessor identity, and operation type) through smart contracts, creating immutable audit trails that enhance transparency and compliance [40]. This provides credible audit evidence for regulatory authorities, meets GDPR requirements for data usage transparency, and further strengthens privacy protection.

To further enhance data leakage prevention capabilities, Bioledger introduces digital watermarking technology. Before CP-ABE encryption, genetic data is embedded with different types of watermarks based on their sensitivity: high-sensitivity data (such as rare disease genes) use robust watermarks (resistant to compression/attacks) to trace leakage sources, while low-sensitivity data (such as health characteristics) use fragile watermarks (for tampering detection) to verify integrity. By combining watermarks with blockchain logs, illegal distribution activities can be quickly located, enabling effective tracing of malicious behaviors.

Furthermore, Bioledger incorporates a hierarchical access control mechanism that distinguishes between regular users (who can only view basic reports with access limited to their own data) and institutional users (who can access diagnostic/development features, subject to email verification). This mechanism supports dynamic access revocation, compliant with Article 17 of GDPR (the "Right to be Forgotten"). The token incentive system enhances user engagement by rewarding data sharing, while simultaneously

maintaining a balance between incentives and privacy through smart contracts that limit the maximum reward per access (e.g., 0.1 ERC20 tokens per access).

To comprehensively evaluate the advantages and disadvantages of BioLedger, we compared it with existing studies. Table I presents the comparison results based on the design goals of genetic security frameworks. The analysis shows that BioLedger exhibits significant advantages in several key aspects. First, in terms of dynamic permission revocation, BioLedger, similar to Genobank.io and Wu et al. (2023), can realize real-time adjustment of permissions, which is of practical value in meeting users' privacy needs and compliance requirements, while other studies such as Encrypgen and Nebula Genomics have not yet possessed this function. Second, FISCO BCOS provides high throughput support for BioLedger, which, similar to LifeCODE.ai, enables it to show strong performance potential in processing massive genetic data, making up for the deficiencies of frameworks such as Sahi et al. (2023). In addition, similar to Encrypgen, Nebula Genomics and Wu et al. (2023), it supports fine-grained access control of data, further improving the flexibility and accuracy of data authorization. Finally, the introduced crypto-economic incentive mechanism not only enhances user participation, but also limits the reward ceiling through smart contracts to ensure the balance between incentives and privacy, which is superior to Sahi et al. (2023) and Wu et al. (2023). In summary, BioLedger shows strong competitiveness in comprehensive performance and practicality.

TABLE I: QUALITATIVE COMPARISON AND ANALYSIS OF GENETIC SECURITY FRAMEWORKS

Criteria	Genobank.io	Encrypgen	Nebula Genomics	LifeCODE.ai	Sahi et al. (2023) [28]	Wu et al. (2023) [29]	BioLedger
Privacy Protection	✓	✓	✓	✓	✓	✓	✓
Dynamic Permission Revocation	✓	✗	✗	✗	✗	✓	✓
Fine-Grained Access Control	✗	✓	✓	✗	✓	✓	✓
Data Audibility	✓	✓	✓	✓	✓	✓	✓
Data Traceability	✓	✓	✓	✓	✓	✓	✓
High-Throughput Support	✗	✗	✗	✓	✗	✓	✓
Token Incentives	✗	✓	✓	✓	✗	✗	✓

D. Shortcomings and Prospects

Despite BioLedger's progress in security and privacy, several limitations need to be addressed urgently. First, key management in distributed CP-ABE authority centers faces challenges, as private key leakage may expose data. The rotation frequency needs to be optimized (e.g., from 30 days to 15 days) and security must be verified. Second, BioLedger has high computational overhead. CP-ABE encryption and IPFS storage may cause delays when processing massive genetic data, requiring algorithm efficiency optimization. Additionally, insufficient cross-chain interoperability is an issue, as the current design primarily relies on FISCO BCOS, making seamless integration with other blockchain networks difficult and limiting global collaboration. Finally, the user experience is complex. MetaMask operations and permission settings may pose a steep learning curve for non-technical users, affecting participation.

Future work directions include: (1) Conducting large-scale performance testing to evaluate the impact of key rotation and algorithm optimization on system latency; (2) Exploring cross-chain bridging technologies (such as Polkadot protocol) to enhance interoperability with other blockchains;(3) Designing adaptive watermarking algorithms to dynamically adjust embedding strength to reduce distortion; (4) Developing user-friendly interfaces to simplify MetaMask operations and verify the compliance of hierarchical access in cross-border collaboration, expanding to global gene bank sharing.

VI. CONCLUSION

BioLedger integrates the immutability of FISCO BCOS blockchain, smart contracts, and HCE (Hierarchical Cryptographic Encryption) technology to deliver a secure and efficient solution for genetic data sharing. Below, we summarize its core advantages in terms of security, efficiency, and privacy preservation.

In terms of security, the BioLedger's decentralized architecture is significantly superior to traditional centralized genetic data management solutions. FISCO BCOS ensures the trustworthiness of participants through node certificate verification. CP-ABE encryption combined with timestamp policies (such as 7-day expiration) restricts unauthorized access. ECIES asymmetric encryption can effectively prevent man-in-the-middle attacks. At the same time, MetaMask, as a decentralized identity carrier, further reduces the risk of leakage, effectively reducing data silos and security risks compared to traditional solutions.

In terms of efficiency, smart contract automated authorization eliminates the complexity of traditional cross-institutional coordination, returns data ownership to users, and accelerates the circulation of genetic data in precision medicine and research. The token incentive mechanism fosters a collaborative ecosystem between users and institutions by rewarding data sharing (e.g., 0.1 ERC20 tokens per access), significantly improving data acquisition efficiency — though the precise improvement magnitude requires experimental validation.

In terms of privacy protection, users have full control over their genetic data archives through MetaMask, and institutions only access authorized data. On-chain traceability records each access log, combined with a data expiration mechanism to ensure privacy and data sovereignty. Future work will focus on key management optimization and rigorous performance benchmarking to enhance practical deployment viability.

All code can be obtained at

https://github.com/WangLabforComputationalBiology/GeneSharing

REFERENCES

[1] Hasselgren, A., Kralevska, K., Gligoroski, D., et al. (2020). Blockchain in healthcare and health sciences—A: A scoping review. International Journal of Medical Informatics, 134, 104040.

[2] Shabani, M., & Marelli, L. (2019). Re-identifiability of genomic data and the GDPR. EMBO Reports, 20(6), e48316.

[3] Erlich, Y., & Narayanan, A. (2014). Routes for breaching and protecting genetic privacy. Nature Reviews Genetics, 15(6), 409–421.

[4] Yin, R., Yan, Z., Liang, X., Xie, H., & Wan, Z. (2023). A survey on privacy preservation techniques for blockchain interoperability. Journal of Systems Architecture, 140, 102892.

[5] R. S. Evans, "Electronic health records: then, now, and in the future," Yearbook of medical informatics, vol. 25, no. S 01, pp. S48 – S61, 2016.

[6] Shabani, M., & Marelli, L. (2019). Re-identifiability of genomic data and the GDPR. EMBO Reports, 20(6), e48316.

[7] Gordon, W. J., & Catalini, C. (2018). Blockchain technology for healthcare: Facilitating the transition to patient-driven interoperability. Computational and Structural Biotechnology Journal, 16, 224–230.

[8] Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. New England Journal of Medicine, 372(9), 793–795.

[9] Naveed, M., Ayday, E., Clayton, E. W., et al. (2015). Privacy in the genomic era. ACM Computing Surveys, 48(1), 6:1–6:44.

[10] Grishin, D., Obbad, K., & Church, G. M. (2018). Data privacy in the age of personal genomics. Nature Biotechnology, 36(11), 1012–1014.

[11] Houtan, B., Hafid, A., & Makrakis, D. (2020). A survey on blockchain-based self-sovereign patient identity in healthcare. IEEE Access, 8, 90478–90494.

[12] Sharma, Y., Balachandran, B. M., & Bhuyar, D. (2019). A Hyperledger Fabric-based framework for secure and interoperable electronic health records management. International Journal of Network Management, 29(5), e2069.

[13] Mamoshina, P., Ojomoko, L., Yanovich, Y., et al. (2018). Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. Oncotarget, 9(5), 5665–5690.

[14] Villarreal, E., García-Alonso, J., & Berrocal, J. (2023). GDPR compliance in blockchain-based healthcare systems: A systematic review. Journal of Cybersecurity and Privacy, 3(2), 234–256.

[15] Yin, R., Yan, Z., Liang, X., Xie, H., & Wan, Z. (2023). A survey on privacy preservation techniques for blockchain interoperability. Journal of Systems Architecture, 140, 102892.

[16] Lehner, R., Ennöckl, C., & Schmied, C. (2022). Health data interoperability through blockchain technology: Challenges and opportunities. Frontiers in Blockchain, 5, 879336.

[17] Mohanta, B. K., Jena, D., Panda, S. S., et al. (2019). Blockchain technology: A survey on applications and security privacy challenges. Internet of Things, 8, 100107.

[18] Feulner, S., Sedlmeir, J., Schlatt, V., et al. (2022). Exploring self-sovereign identity in market applications. Journal of Cybersecurity and Privacy, 2(3), 441–459.

[19] Dunphy, P., & Petitcolas, F. A. P. (2018). A first look at identity management schemes on the blockchain. IEEE Security & Privacy, 16(4), 20–29.

[20] Kakarlapudi, P. V., & Mahmoud, Q. H. (2021). Consent management in blockchain-based healthcare systems. Journal of Medical Systems, 45(4), 46.

[21] Li, H., Chen, Y., Shi, X., et al. (2023). FISCO-BCOS: An enterprise-grade permissioned blockchain system with high-performance. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis.

[22] Zhang, A., & Lin, X. (2018). Towards secure and privacy-preserving data sharing in e-health systems via consortium blockchain. Journal of Medical Systems, 42(8), 140.

[23] Fan, K., Ren, Y., Wang, Y., et al. (2018). Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5G. IET Communications, 12(5), 527–532.

[24] Bender, D., & Sartipi, K. (2013). HL7 FHIR: An agile and RESTful approach to healthcare information exchange. Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, 326–331.

[25] Saripalle, R., Runyan, C., & Russell, M. (2019). Using HL7 FHIR to achieve interoperability in patient health record. Journal of Biomedical Informatics, 94, 103188.

[26] Dagher, G. G., Mohler, J., Milojkovic, M., et al. (2018). Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. Sustainable Cities and Society, 39, 283–297.

[27] Alghazwi, Mohammed, et al. "Blockchain for genomics: a systematic literature review." Distributed Ledger Technologies: Research and Practice 1.2 (2022): 1-28.

[28] Sahi, Niranjan, et al. "A blockchain-based architecture for interoperable healthcare data exchange." 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2023.

[29] Wu, Xingyu, Liang Zhang, and Honglan Huang. "Patient-Centered Data Sharing and Revision Framework Based on Redactable Blockchain." 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2023.

[30] Zheng, Z., Xie, S., Dai, H. N., et al. (2018). Blockchain challenges and opportunities: A survey. International Journal of Web and Grid Services, 14(4), 352–375.

[31] Li H, Chen Y, Shi X, et al. Fisco-bcos: An enterprise-grade permissioned blockchain system with high-performance[C]//Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. 2023: 1-17.

[32] O'Donoghue, O., Vazirani, A. A., Brindley, D., & Meinert, E. (2019). Design choices and trade-offs in health care blockchain implementations: Systematic review. Journal of Medical Internet Research, 21(5), e12426.

[33] Agbo, C. C., Mahmoud, Q. H., & Eklund, J. M. (2019). Blockchain technology in healthcare: A systematic review. Healthcare, 7(2), 56.

[34] Hölbl, M., Kompara, M., Kamišalić, A., & Zlatolas, L. N. (2018). A systematic review of the use of blockchain in healthcare. Symmetry, 10(10), 470.

[35] Simonyan V, Goecks J, Mazumder R. Biocompute Objects—A Step towards Evaluation and Validation of Biomedical Scientific Computations. PDA Journal of Pharmaceutical Science and Technology. 2017;71(2):136–146.

[36] J. Bethencourt, A. Sahai and B. Waters, "Ciphertext-Policy Attribute-Based Encryption," 2007 IEEE Symposium on Security and Privacy (SP '07), Berkeley, CA, USA, 2007, pp. 321-334.

[37] SECG, "SEC 1: Elliptic Curve Cryptography, Version 2.0," Standards for Efficient Cryptography Group, 2009. [Online]. Available: https://www.secg.org/sec1-v2.pdf, [Accessed: Jul. 31, 2025].

[38] Bethencourt, J., Sahai, A., & Waters, B. (2007). Ciphertext-Policy Attribute-Based Encryption. In IEEE Symposium on Security and Privacy (SP '07).

[39] Buterin, V. (2014). Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform.

[40] Zyskind, G., Nathan, O., & Pentland, A. (2015). Decentralizing Privacy: Using Blockchain to Protect Personal Data. In IEEE Security & Privacy