Title

bDNA Medium: Secure Conversion of Raw Genomic Sequencing Data to Verifiable Cryptographic Objects

Authors
Shoel Lowy
Blockchain DNA Foundation

Date
February 2026

Abstract
The bDNA Medium is a secure conversion protocol and device architecture that transforms raw genomic sequencing output (FASTQ, BAM, and related formats) into cryptographically verifiable digital objects. The system captures sequencing data at the point of intake from laboratory instruments, processes it within a tamper-resistant isolated environment, and outputs encrypted artifacts and integrity proofs suitable for verification, audit, and downstream research use. Raw genomic data is never exposed outside the controlled environment, and the resulting objects support trusted provenance and authenticity checks without disclosure of sensitive biological information. This work establishes a general-purpose foundation for sovereign, privacy-respecting genomic data handling.

1. Introduction & Problem Statement

Modern DNA sequencing systems generate highly sensitive raw data that must be processed through laboratory IT environments before encryption, storage, or analysis. This early handling stage creates persistent technical challenges:

- Raw genomic data is exposed during initial processing.
- Dataset provenance and sequencing origin are difficult to verify cryptographically.
- Chain-of-custody is challenging to audit across systems and over time.
- Verification of integrity and biological claims requires access to raw data.

As genomic data becomes integral to healthcare, research, and regulated workflows, these limitations introduce increasing operational, ethical, and compliance risks. A standardized method for converting sequencing reads into verifiable cryptographic objects remains undeveloped.

2. Core Invention – The bDNA Medium

The bDNA Medium is a protocol and device that performs secure conversion at the boundary between sequencing instruments and downstream digital systems. It addresses the problem of early-stage exposure by transforming raw sequencing output into verifiable cryptographic objects before any external system can access the plaintext data.

Key Characteristics

- Deterministic and reproducible conversion.
- Tamper-resistant, isolated processing environment.
- Immediate encryption and data minimization.
- Irreversible deletion of plaintext after conversion.
- Output limited to cryptographic artifacts (encrypted objects, commitments, proofs).

3. System Architecture

The system consists of a hardware-software converter that interfaces directly with sequencing instruments.

Input Stage

- Raw sequencing data (FASTQ, BAM, POD5, or equivalent) is captured at the point of intake from the sequencer.
- Data is streamed or transferred in real-time or batch mode.

Processing Stage

- Data is ingested into a tamper-resistant isolated environment.
- Conversion is performed deterministically.
- Cryptographic operations generate integrity proofs and commitments.
- Plaintext data is encrypted and minimized; keys are managed securely.
- Original raw data is irreversibly deleted after conversion.

Output Stage

- Encrypted genomic object.
- Merkle root representing a deterministic fingerprint commitment.
- Signed sequencing proof package containing metadata and hashes.
- Cryptographic attestation report.

4. Technical Principles

- Protected execution environment ensures isolation and tamper resistance.
- Single-pass streaming processing minimizes data retention.
- Cryptographic provenance is bound to the sequencing process.
- Data minimization and irreversible deletion reduce exposure risk.
- Verifiable outputs support third-party verification without disclosure.

5. High-Level Flow

1. Sequencer generates raw data (FASTQ/BAM).
2. Data is captured at intake by the bDNA Medium device.
3. Device processes data in isolated environment.
4. Conversion generates encrypted object + commitment + signed proof.
5. Device outputs cryptographic artifacts.
6. Artifacts can be anchored on-chain or shared for verification.

6. Use Cases

- Sovereign genomic programs (national data custody with verifiable integrity).
- Laboratory provenance and chain-of-custody auditing.
- Research reproducibility without raw data sharing.
- Blockchain-based verification of genomic claims.
- Cross-institutional sharing with selective disclosure.

7. Conclusion

The bDNA Medium provides a reusable, open infrastructure primitive for trustworthy genomic data handling. By performing secure conversion at the sequencer boundary, it addresses early-stage exposure risks and establishes a foundation for privacy-respecting, verifiable genomic workflows.

References

- FASTQ format specification
- BAM format specification
- Cryptographic commitment schemes
- Merkle trees and proofs of integrity

- Secure enclave technologies

Publication Note

This document is published for defensive purposes to establish prior art. The protocol is intended to be open source under Apache 2.0. No proprietary implementation details are disclosed.

End of Whitepaper