
WHO WROTE THIS SEQUENCE?¹

Babita Singh
Independent Researcher

Genethropic

With
Apart Research

Abstract

Today, any curious mind can open the laptop, design a novel enzyme, order it synthesised, and have it on a bench before any registry knows it exists. This is an extraordinary scientific advancement, but without the right infrastructure, a biosecurity problem waiting to compound. Generative AI is producing novel proteins and genes faster than the field can catalogue, evaluate, or attribute them. No shared infrastructure exists to distinguish AI-designed sequences from naturally occurring ones, screen them for biosafety, or credit their creators.

ArtGene-Archive (artgene-archive.org) is the first dedicated registry for AI-generated biological sequences. Every submission passes an automated three-gate biosafety pipeline, receives a cryptographically signed certificate anchored to a tamper-evident audit log, and is issued a citable Registry ID. Built on experience at the European Genome-phenome Archive and grounded in emerging AI biosafety research, this dedicated archive solves a specific structural gap: provenance and safety certification at the point of design, not the point of discovery. What it needs now is what GenBank needed in 1982 - institutional commitment, knowledge contribution, and collective adoption.

¹ Research conducted at the [AIxBio Hackathon](#), April 2026

1. Motivation

1.1 *A Personal Reckoning*

I have spent the last twenty years witnessing the field of genomics transform from a craft, when sequencing a single gene took weeks, into a data-science problem operating at petabyte scale. From the microarray era, the arrival of next-generation sequencing, the single-cell revolution, CRISPR-CAS systems, along with the parallel-paced evolution of bioinformatics pipelines and computational infrastructure. Each transition was disorienting in the best possible way (*amaze amaze amaze*).

However, the entry of generative AI into the design of new biological sequences feels categorically different. It is the first time in my career that I feel genuinely conflicted. Not because the science is not extraordinary but because for the first time we are designing sequences that evolution never stress-tested. Every therapeutic protein currently in clinical use was, at some point, a naturally occurring molecule that a billion years of selection had already vetted for basic structural stability. That is no longer the baseline. Protein language models like ESMFold and RFdiffusion, and diffusion-based structure generators like AlphaFold3, are now proposing functional proteins that no ribosome has ever translated [1, 2]. In 2025, AI-designed GLP-1 receptor agonists demonstrated plasma half-lives approximately three times longer than semaglutide in animal models [3]. Earlier that same year, AI-designed proteins neutralised lethal snake venom toxins that classical drug discovery had not touched in decades [4]. These are not incremental results.

The same coin has a darker side. The tools that design a therapeutic protein can, with careless prompting, generate sequences that evade the biosecurity screening systems currently used by DNA synthesis companies [5]. A 2025 paper in *npj Biomedical Innovations* noted that automated synthesis platforms could be exploited to optimise the virulence of known pathogens through AI-guided sequence generation [6]. A majority of surveyed biosecurity experts (approximately 76%) expressed concern about AI misuse in biology, and 74% called for new regulatory frameworks [7]. The U.S. White House, in 2025, issued an executive order specifically requiring strengthened nucleic acid screening in response to this threat class [7].

My position is not that we should stop designing artificial sequences. We absolutely should not stop. My position is that we need a system that separates nature's code from human-made codes, records what we did, screens for harm before it propagates, and attributes what was created to who created it. That system does not yet exist. ArtGene-Archive is my attempt to build it.

2. Related Work

2.1 What GenBank Teaches Us, and What It Cannot Do

In 1979, Walter Goad at Los Alamos National Laboratory established the Los Alamos Sequence Database in direct response to a specific technological inflection point: automated Sanger sequencers were generating more nucleotide data than any laboratory, journal, or filing cabinet could coherently manage [8]. The scientific community recognised that without centralised, open, curated infrastructure, the data would fragment into silos or simply be lost. In 1982, the database became GenBank, co-funded by NIH, NSF, the Department of Energy, and the Department of Defense. By 1983 it held 2,000 sequences. By 2024 it held sequences from **557,000** species, **3.7 billion** accession records, and 25 trillion base pairs [9, 10].

The tools built on top of GenBank - BLAST (1990), Entrez (1991), the infrastructure for the Human Genome Project, collectively transformed every branch of biology. None of that was inevitable. It happened because the right people built the right infrastructure at precisely the right moment [11].

I spent four years inside one of GenBank's European counterparts, at the European Genome-phenome Archive (EGA), operated jointly by EMBL-EBI and the Centre for Genomic Regulation (CRG), it is the primary repository for human genomic and phenotypic data that require controlled access due to data-protection obligations. Working there gave me a detailed, operational understanding of what a production genomic archive actually requires: robust deposition pipelines, long-term data integrity guarantees, policy-aware access controls, cryptographic data handling, and critically, the institutional trust that makes researchers willing to deposit their most sensitive data into a centralised system.

Both archives have something in common: their design assumptions treat the sequence as a *discovered* object, something found in nature or produced by an experimental process. Neither was built to handle sequences that were designed by a machine. Neither watermarks their submissions. Neither runs automated biosafety screening at the point of deposit. Neither issues creator attribution certificates. These are not gaps that the archives can easily patch; they are structural absences that a new class of archive needs to fill.

Table 1: Registry Comparison - GenBank, EGA, ENA, and ArtGene Archive

Feature	GenBank (NCBI)	EGA (EMBL-EBI / CRG)	ENA (EMBL-EBI)	ArtGene-Archive
Primary data type	Natural + synthetic sequences	Human genomic + phenotypic	Natural experimental sequences	AI-generated sequences

AI-origin labelling	No	No	No	Yes - required at deposit
Automated biosafety screening	No	No	No	Yes - four-gate pipeline
Codon watermarking	No	No	No	Yes - TINSEL HMAC-SHA3-256
Tamper-evident audit log	No	Partial	No	Yes - blockchain-style chaining
Post-quantum signing	No	No	No	Yes - WOTS+ (live), LWE (planned)
Embargoed deposit	Yes	Yes (controlled access)	Yes	Yes
Creator attribution certificate	Accession number only	No	Accession number only	Signed HybridCertificate
Citizen science access	Restricted	No	Yes	Yes - free API, open submission

The comparison is intended to show that the problem space has changed. ArtGene-Archive is building for AI-designed sequences what GenBank did for sequencer-derived ones: creating the infrastructure that the moment demands.

3. Methods

ArtGene-Archive is implemented as a Python/TypeScript monorepo. Three backend packages handle distinct concerns: *tinsel-core* : data models, TINSEL encoder/decoder, Reed-Solomon codec, *tinsel-gates* : the four-gate biosafety pipeline with a live ESMFold adapter (Gate α), an offline composition heuristic layer (Gate β + Gate γ), and mock stubs for SecureDNA DOPRF and IBBIS pending credential integration, and *tinsel-api* : FastAPI application, PostgreSQL persistence, Alembic migrations, and WOTS+ signing. A Next.js 16 dashboard provides the submission and certificate-browsing interface. The full stack runs locally with a single docker compose up --build command; no external services are required in development mode. Production adapters replace mocked implementations when SENTINEL_ENV=production is set.

3.1 Registration Pipeline

Important: For the limited time of this hackathon, the pipeline was optimised for amino acid sequences, DNA and RNA sequences are yet to be optimised.

Every submission follows a fixed execution order (figure 1). FASTA input is parsed and normalised, then a fragment-assembly cross-check screens the sequence against an SHA3-256 k-mer index of all archived sequences. *Gate α* runs next and is the only fail-fast gate: a failure terminates the pipeline immediately, skipping Gates β and γ . If Gate α passes, other gates run

concurrently. If all four pass (fourth under development), a SHA3-256 deduplication check prevents re-registration of identical sequences. A stub WOTS+ certificate is then signed and written together with the audit log entry in a single atomic database commit, preventing partial registration states. (real WOTS+/LWE crypto are planned for a later phase.)

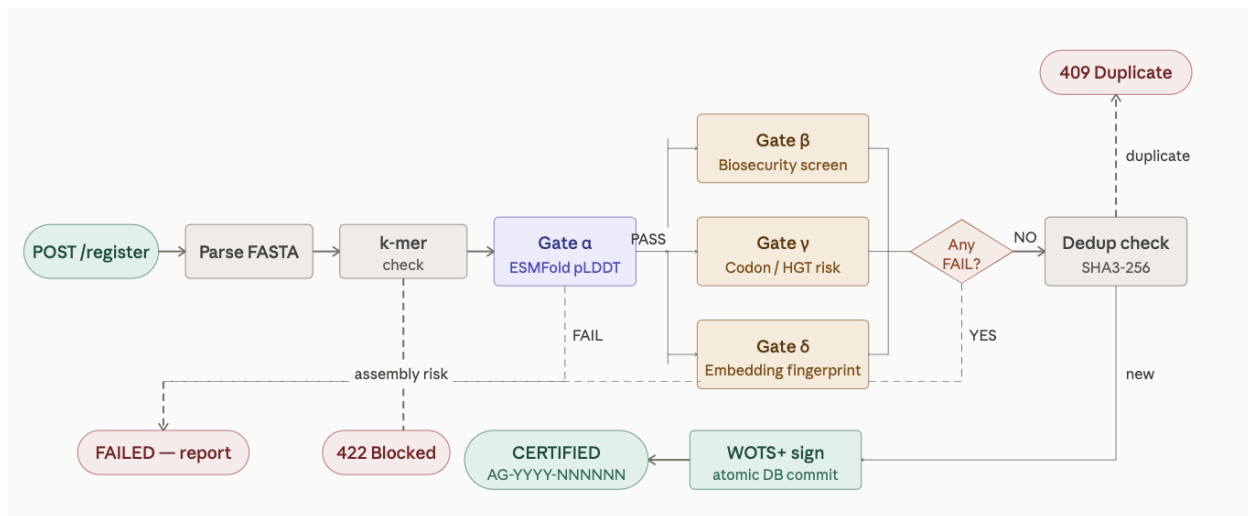


Figure 1. ArtGene Archive registration pipeline. Gate α is fail-fast; gates β and γ run concurrently once α passes. All steps from watermark embedding onward execute in a single atomic transaction.

3.2 Biosafety Gates -- Design Choices and Justification

Gate α - Structural confidence. ESMFold pLDDT scores are the primary filter. Sequences where more than 20% of residues fall below pLDDT 50 fail this gate; the 20% threshold was set conservatively to catch structurally implausible sequences without penalising inherently disordered proteins that remain biologically relevant. RNA minimum free energy (Δ MFE, via LinearFold) is assessed in parallel. Fail-fast design was chosen because structural invalidity is the strongest single predictor of a sequence being computationally generated noise rather than a viable design.

Gate β - Off-target and composition screen. Amino acid composition is evaluated for Kyte-Doolittle hydrophathy (GRAVY score), cationic/amphipathic toxin probability, allergen probability, and k-mer overlap with known antimicrobial peptide scaffolds.

In full-production mode the gate will call SecureDNA's DOPRF (Distributed Oblivious Pseudorandom Function) service and the IBBIS provider screen - the two most widely deployed cryptographic biosecurity screening platforms in the synthesis industry. DOPRF was chosen specifically because it allows sequence screening without exposing the raw sequence to SecureDNA's servers, preserving submitter privacy. Full BLAST screening against pathogen databases is in development for a subsequent release.

Gate γ - Ecological risk. Horizontal gene transfer (HGT) propensity scoring draws on codon usage patterns, GC content, and host-organism compatibility supplied at submission. This gate was included after recognising that sequences optimised aggressively for a microbial expression

host can have elevated HGT propensity independent of their direct toxicity, an ecological risk class not addressed by composition-based screens alone.

When any gate fails, the API returns a structured consequence report naming the gate, the failing metric, the observed value, and the threshold. This design choice of *failing informatively* rather than silently is deliberate, the archive is intended as a tool for improving designs, not merely blocking them.

3.3 TINSEL Watermark

The TINSEL (Traceable INtegrated Sequence ELeMent) watermark is used specifically for **provenance tracing**, providing an unique option to submitters to be able to issue “per-recipient distribution copies” for traceability and avoiding misuse of their sequences. TINSEL uses spread-spectrum codon steganography to embed a cryptographic signature using only synonymous codon substitutions, leaving the protein sequence and function biologically unchanged. The approach is grounded in Liss et al. [12], who first demonstrated permanent information embedding in synthetic genes via synonymous codon choice, and BioCode [13], which extended the concept to both coding and non-coding regions. TINSEL adds HMAC-SHA3-256 spread-spectrum encoding and Reed-Solomon forward error correction to provide robustness against single-base errors introduced by routine re-synthesis or sequencing.

When a researcher issues a distribution copy, a unique HMAC-SHA3-256-derived seed generates a recipient-specific codon pattern. If a copy appears somewhere unexpected, the Verify Source tool identifies the originating recipient by replaying the fingerprint encoding for every issued copy -- without requiring any metadata in the submitted sequence. Watermark tier is assigned automatically by the number of synonymous carrier codons available in the sequence (Table 2).

Table 2. TINSEL Watermark Tiers

Tier is assigned automatically from the number of synonymous carrier codons available in the submitted sequence. Shorter sequences receive weaker watermarks; sequences below REJECTED cannot carry a detectable watermark.

Tier	Min Carrier Codons	Watermark Bits	RS Codec	Error Tolerance
FULL	1,792	128-bit	(32, 16)	8 bytes
STANDARD	896	64-bit	(16, 8)	4 bytes
REDUCED	320	32-bit	(8, 4)	2 bytes
MINIMAL	96	16-bit	(4, 2)	1 byte
DEMO	24	8-bit	none	0
REJECTED	< 24	--	--	too short

3.4 Certificate Issuance and Tamper-Evident Audit Log

Sequences that pass all gates are issued a HybridCertificate: a signed JSON document containing the registry ID, SHA3-512 sequence hash, biosafety gate outcomes, watermark parameters, submitter identity, and timestamp. Certificates are appended to `registry_audit_log`, a PostgreSQL table where each row is chained to the previous entry by:

```
entry_hash = SHA3-256( seq_num || prev_entry_hash || certificate_hash )
```

Immutability is enforced at two independent layers: a PostgreSQL trigger (migration 003) blocks any UPDATE or DELETE at the database level, and an AppendOnlyMixin in the ORM raises a RuntimeError if any field is mutated after the first commit. This dual-layer design means that even a database administrator with full credentials cannot silently alter a historical record -- a requirement borrowed directly from append-only ledger design patterns used in financial audit systems.

Certificates are signed with WOTS+ (Winternitz One-Time Signature Scheme). The keypair is derived deterministically from (`spreading_key`, `registry_id`), so the private key is never stored and is discarded after use. LWE lattice commitments are planned as a Phase 4 post-quantum upgrade, particularly relevant for certificates that will be cited in regulatory submissions years after registration. The stub is wired in the codebase; API responses flag it with "not_implemented": true.

4. Results

- The full ArtGene Archive registration pipeline is live and publicly accessible at artgene-archive.org.
- The demo at artgene-archive.org/showcase allows any visitor to submit a protein FASTA and observe the full biosafety screening and watermark embedding pipeline.

4.1 What Is Live

Table 3 summarises the current implementation state across all major features. The core pipeline - FASTA parsing, deduplication, three-gate biosafety screening, TINSEL watermark embedding, WOTS+ signing, and tamper-evident audit logging - is fully operational. Per-recipient distribution copies and leak attribution via Verify Source are live. Embargoed deposits, compliance manifests (US DURC and EU Dual-Use frameworks), and synthesiser auth document generation are all functional.

Table 3. ArtGene-Archive Feature Implementation Status (26 April 2026)

Core pipeline features are fully operational. Post-quantum and governance features are in planned phases. Demo at artgene-archive.org/showcase.

Feature	Status
CORE PIPELINE	
FASTA parsing (protein, RNA detection)	Live*
SHA3-256 sequence deduplication	Live
Three-gate biosafety pipeline (mock in dev, live adapters in prod)	Live
TINSEL codon watermark encoder / decoder + Reed-Solomon	Live
WOTS+ post-quantum certificate signing	Live
Tamper-evident SHA3-256 chained audit log	Live
Fragment 20-mer k-mer cross-check (privacy-preserving hashed index)	Live
Per-recipient codon fingerprint + leak attribution (Verify Source)	Live
REGISTRY & WORKFLOW	
Embargoed / public visibility control + publish workflow	Live
REST API with rate limiting + API key auth	Live
Next.js dashboard (submit, browse, verify, certificate detail)	Live
Compliance manifest (US DURC, EU Dual-Use framework)	Live
Synthesiser auth document generation	Live
IN PROGRESS	
ESMFold live integration (non-mock, production mode)	In progress
FASTA parsing (DNA sequence optimisation)	In progress
PLANNED	
LWE lattice commitments (post-quantum Phase 4)	Planned

CDK / Terraform infrastructure-as-code (Phase 5)	Planned
Merkle inclusion proofs for pathway bundles (Phase 6)	Planned
Institutional API key management portal	Planned
ORCID / academic identity integration	Planned
FUTURE	
Journal submission workflow integration	Future
Benchside synthesiser protocol interface	Future

Table 3. ArtGene-Archive Feature : check all implemented features demo at artgene-archive.org/showcase

*Live = code implemented on Artgene-archive's Github repo.

The **benchside synthesiser** interface deserves particular attention. It would allow ArtGene-certified sequences to be passed directly to desktop or benchtop DNA synthesisers, with the archive acting as an intermediary safety gate. This would make biosafety screening not merely advisory but a practical requirement embedded in the synthesis workflow, analogous to how safety interlocks operate in regulated manufacturing contexts. The regulatory pathway for this feature will depend heavily on the policy environment that emerges in the next two to three years.

4.2 Honest Disclaimers

ArtGene Archive cannot succeed as a solo project, it requires professional engineers to architect the base-infrastructure and biosafety experts to review the pipelines. The biosafety pipeline currently runs in mock mode in the development environment. Production adapters for ESMFold (live structure prediction) and SecureDNA DOPRF are implemented but require external API access and are therefore not active in the open demo. Reviewers testing the system via artgene-archive.org/showcase will observe mock gate outcomes; the pipeline architecture, watermarking, and certificate issuance are fully functional in both modes.

Gate β 's toxin screening is only as good as the databases it queries. Gate γ 's HGT model involves assumptions about gene transfer frequencies that are imperfectly characterised in the literature. These are known limitations: ArtGene-Archive will evolve together with the AI biosafety community. As the new safeguards and screening methods get developed, they will be added as necessary components of a biosafety ecosystem.

The TINSEL watermark, while robust against routine single-base errors (Reed-Solomon correctable), could theoretically be defeated by an actor with full knowledge of the encoding scheme and the spreading key. This is why the spreading key is stored in a secrets vault and never exposed via the API, and why setting `SENTINEL_ENV=production` without a custom `SPREADING_KEY` causes the API to refuse startup.

5. Discussion and Limitations

5.1 Why This Moment Is the Right Moment

The parallel to GenBank in 1979 is not rhetorical decoration; it is the central argument. GenBank was built because a specific technological development was generating data faster than the existing scientific infrastructure could handle, and because a small group of people recognised that without deliberate intervention the opportunity would be lost and the risk would compound.

Generative biology models are being released continuously, each more capable than the last. The research community is generating novel sequences daily. Synthesis costs are falling toward the threshold at which physical realisation of computational designs becomes routine even for small labs and community biology spaces. Regulatory frameworks are scrambling to keep pace. There is currently no standard archive for AI-generated sequences, no standard watermark, no standard safety gate, and no standard attribution mechanism for any of it.

The cost of building ArtGene-Archive is small relative to the cost of not building one. A future in which AI-generated sequences are indistinguishable from naturally occurring ones, for example, in forensic samples, in environmental surveillance, in clinical isolates, is a future in which biosecurity investigations become vastly more difficult and the scientific record cannot be trusted. Building the infrastructure now, while the volume of sequences is still manageable, is orders of magnitude cheaper than retroactively auditing a decade of undocumented AI sequence generation.

5.2 Pending Development and Open Questions

Several significant features are stubbed but not yet live:

LWE lattice commitments (Phase 4). Post-quantum signatures with WOTS+ provide forward-looking non-repudiation for certificates that will be referenced in regulatory submissions decades from now. The LWE commitment scheme, which would provide zero-knowledge proofs of biosafety pathway inclusion without exposing sequence data, is architecturally the most ambitious feature on the roadmap. The zero-filled stub is already wired in the API; responses flag it with `"not_implemented": true`.

Merkle inclusion proofs for pathways (Phase 7). Multi-gene pathway bundles are supported via `POST /api/v1/pathways`, but inclusion proofs -- the cryptographic mechanism that allows a third party to verify that a specific sequence was part of a certified pathway without accessing the full bundle -- return `{"not_implemented": true}`. This is critical for regulatory use cases where a synthesis company needs to verify pathway membership without receiving the complete sequence set.

Journal submission workflow integration. The highest-leverage policy outcome for ArtGene Archive would be a requirement analogous to the GenBank deposition requirement for sequencing papers: journals publishing work that uses AI-generated sequences would require an ArtGene Registry ID in the submission. This is a policy decision, not a technical one, and it requires engagement with journal editors and publishers that has not yet begun.

Institutional governance. The ArtGene Consortium referenced on the landing page is, at present, aspirational. Formalising it -- with a charter, a technical advisory board, an independent biosafety review panel, and a data-access policy framework -- is the most important non-technical task ahead.

Benchside synthesiser interface. This feature, which would make ArtGene certification a practical prerequisite for synthesis rather than an optional step, is where the biosafety argument becomes operationally consequential. It requires hardware partnerships, regulatory engagement, and synthesis industry buy-in that are outside the scope of a hackathon build.

5.3 Limitations: What the Pipeline Cannot Guarantee

No biosafety screening system, including ArtGene's four-gate pipeline can guarantee that a sequence is safe. Biosafety is probabilistic, context-dependent, and subject to emergent properties that even sophisticated computational tools cannot anticipate. These are not reasons to not build ArtGene-Archive. They are reasons to position it correctly: as a necessary component of a biosafety ecosystem that also includes synthesis company screening, institutional biosafety committees, and emerging regulatory frameworks - not a replacement for any of them.

As mentioned earlier, ArtGene-Archive will evolve together with the AI biosafety community. As the new safeguards and screening methods get developed, they will be added as necessary components of a biosafety ecosystem.

6. Future Work

What this project needs at this stage is not just funding, though it needs that too. It needs people. Specifically: engineers and biosafety researchers who can stress-test the screening pipeline against sequence classes I have not yet thought to include. Policy specialists who understand how a registry like this gets written into journal submission requirements or synthesis company

compliance frameworks. And most importantly: early adopters, the sequence submitters. If any of these descriptions fit you, the repository is open, the issues tracker is live, and my email is on the site.

7. Conclusion

I would like to conclude this paper with something more personal because the name deserves an explanation.

The obvious reading of ArtGene is *artificial* gene, which is accurate. But the name was also a deliberate signal about what kind of future I think we are moving towards. In the age of genAI, the boundary between science and art is dissolving. The researchers who will produce the most interesting sequences in the next twenty years may not be the ones working in well-funded institutional labs with access to wet-lab infrastructure. Some of them will be curious individuals who have sat with a problem *long enough* to find an angle that no established group thought to try. What they need is not a grant committee; they need a canvas, a way to showcase their work, and a mechanism to share it on their own terms.

The TINSEL watermark was designed with exactly this in mind. A sequence deposited in ArtGene-Archive is not just archived; it is signed. Its creator can issue distribution copies, each fingerprinted to a specific recipient, for a specific purpose, until a specific date. A pharmaceutical company that wants to licence the sequence for wet-lab validation can be issued a copy that is cryptographically tied to that agreement. A community biology lab that wants to test it can receive their own. The creator retains provenance. The work travels, but it carries its creator with it.

ArtGene-Archive demonstrates that we can own the responsibility to provide a bio-safe playground for new bio-artists to emerge. That the cryptographic provenance, automated biosafety screening, and creator attribution for AI-generated biological sequences can be delivered as a single, atomic operation at the point of deposit. The implications will be practical and immediate.

An artGene archiver, to me, will be someone who designs a biological sequence with the same creative ownership that an artist brings to a canvas and the same scientific rigour that a researcher brings to an experiment. We have to build the infrastructure for that. I hope this paper is a start.

Code and Data

- **Code repository:** <https://github.com/babisingh/artgene-archive>
- **Demo:** <https://artgene-archive.org/showcase>
- **Main website:** <https://artgene-archive.org/>

References

- [1] Dauparas, J. et al. (2022). Robust deep learning-based protein sequence design using ProteinMPNN. *Science*, 378, 49-56.
- [2] Watson, J. L. et al. (2023). De novo design of protein structure and function with RFDiffusion. *Nature*, 620, 1089-1100.
- [3] Chen, S. et al. (2025). AI-Driven Efficient De Novo Design of GLP-1RAs with Extended Half-Life and Enhanced Efficacy. *bioRxiv*, doi:10.1101/2025.03.26.645438.
- [4] Torres, S. et al. (2025). De novo designed proteins neutralize lethal snake venom toxins. *Nature*, 639, 225–231.
- [5] Urbina, F. et al. (2022). Dual use of artificial-intelligence-powered drug discovery. *Nature Machine Intelligence*, 4, 189-191.
- [6] Groff-Vindman, C.S. et al. (2025). The convergence of AI and synthetic biology: The looming deluge. *npj Biomedical Innovations*.
- [7] Generative AI for Biosciences: Emerging Threats and Roadmap to Biosecurity. (2025). *arXiv*, 2510.15975.
- [8] Goad, W. (1979). Los Alamos Sequence Database. LANL / NIH.
- [9] NCBI GenBank Statistics. <https://www.ncbi.nlm.nih.gov/genbank/statistics/>
- [10] Pettit, D. et al. (2025). History of Biological Databases, Their Importance, and Existence in Modern Scientific and Policy Context. *Database*, doi:10.1093/database/baae131.
- [11] NCBI. A Brief History of NCBI's Formation and Growth. *The NCBI Handbook*, NBK148949.
- [12] Liss, M. et al. (2012). Embedding permanent watermarks in synthetic genes. *PLOS ONE*, 7(8), e42465.
- [13] Haughton, D. & Balado, F. (2013). BioCode: Two biologically compatible algorithms for embedding data in non-coding and coding regions of DNA. *BMC Bioinformatics*, 14, 121.
- [14] Rudolph, C. et al. (2024). Cryptographic approaches to authenticating synthetic DNA sequences. *Trends in Biotechnology*.
- [15] Gibson, D. G. et al. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329, 52-56.
- [16] De Haro, L. P. (2024). Biosecurity Risk Assessment for the Use of Artificial Intelligence in Synthetic Biology. *Applied Biosafety*, PMC11313549.
- [17] Carnegie Endowment for International Peace. (2024). Mitigating Risks from Gene Editing and Synthetic Biology: Global Governance Priorities.

LLM Usage Statement

I have used Anthropic's Claude-code extensively to develop the application's codes and help me write the Method section. All codes and claims were independently verified and consulted with human experts. The webpage is designed by me and then coding is implemented by using Claude Design.