D!DA DNA DATA APPLICATIONS | European Innovation Council | Funded by the European Union

Deliverable 4.1: State of the art and commercial needs for authentication and in-product documentation

Scope

DNA data storage and cryptography can be combined to create innovative security solutions. DNA's unique properties—high data density and natural complexity—enable the development of new cryptographic schemes, such as DNA-based signatures and encrypted sequences. Another direction of related investigation pertains to including data or information in physical objects. A task that can be achieved using data encoded into DNA.

In analyzing and describing authentication schemes, which are the main focus of this document, we use the following nomenclature:

- Manufacturer M: Creates the product and DNA
- Customer C: Checks the authenticity of the product
- Adversary A: Sells the fake product and tries to mimic the behavior of the DNA

Scientific state-of-the-art

We first review the state-of-the-art in the context of using DNA to label physical objects and also describe relevant concepts from related fields such as cryptography. Classical cryptography theoretical frameworks will probably form the basis for designing and analyzing DNA-based cryptographic schemes.

Classical cryptography

Cryptography forms the backbone of secure communication in digital systems, offering mechanisms to ensure confidentiality, integrity, and authenticity of

data. **Confidentiality** refers to the protection of information from unauthorized access, typically achieved through encryption schemes that transform readable data (plaintext) into an unreadable format (ciphertext). **Integrity** ensures that data has not been altered in an unauthorized manner, typically verified through cryptographic hash functions or message authentication codes (MACs). **Authenticity** confirms the identity of the sender and the origin of the data, often provided by digital signatures or authentication protocols.

Encryption schemes primarily focus on confidentiality. Symmetric encryption, where the same key is used for both encryption and decryption, is efficient but requires secure key distribution. Asymmetric encryption, using a pair of public and private keys, facilitates secure communication without sharing a secret key but is computationally intensive.

In contrast, **signature schemes** emphasize authenticity and integrity. A digital signature, generated using a sender's private key, can be verified by anyone with the corresponding public key, ensuring that the message is indeed from the claimed sender and has not been tampered with.

For an in-depth overview of classical cryptography, we refer the reader to "Introduction to Cryptography" by Buchmann et al.[1], "Foundations of Cryptography" by Goldreich[2] , and "Introduction to modern cryptography" by Katz and Lindell[3].

Cryptographic assessment schemes

Cryptographic security is assessed through various attack models. **IND-CPA (Indistinguishability under Chosen Plaintext Attack)**[3] measures the strength of an encryption scheme by testing its resistance to an adversary choosing plaintexts and attempting to distinguish between their ciphertexts. **EUF-CMA (Existential Unforgeability under Chosen Message Attack)**[3] evaluates the robustness of a signature scheme by assessing its ability to withstand attempts to forge signatures on arbitrary messages. There are other approaches to evaluating security. For example, see a review in Lindell[4]. Different use cases can, generally, be analyzed via one or more such frameworks.

In summary, while encryption schemes focus on confidentiality, signature schemes provide authenticity and integrity, both critical for secure communication in a digital landscape. Understanding these distinctions and attack models is essential for designing resilient cryptographic systems.[5]

Cryptography in DNA-based data storage

As digital communication increasingly intersects with biotechnology, DNA-based data storage emerges as a novel field with significant implications for cryptographic security. Just as in

digital communication, the principles of confidentiality, integrity, and authenticity are crucial for safeguarding DNA-stored information. However, the unique properties of DNA introduce new challenges and opportunities in this domain.

Firstly, in the context of **Genetically Modified Organisms (GMOs)**, ensuring the integrity and authenticity of genetic material is paramount. Users must have confidence in the origin of GMOs and assurance that their genetic sequences have not been tampered with. Cryptographic techniques, such as digital signatures, could be adapted to verify the authenticity of genetic sequences, ensuring that they remain unaltered from their original form.[6,7]

Secondly, when DNA is utilized as a **medium for data storage**, it must meet the same stringent standards as traditional digital storage. This includes safeguarding the confidentiality of stored data against unauthorized access, maintaining its integrity to prevent alterations, and ensuring authenticity to verify the source of the information. As DNA data storage becomes more prevalent, applying and adapting existing cryptographic methods to this medium will be essential.[8]

Lastly, the unique properties of DNA open up a new field: **in-product authentication**. DNA can be embedded as a signature within various materials, providing a novel method for ensuring the integrity and authenticity of products and for verifying supply chain adequacy. For example, incorporating DNA sequences into materials could serve as a robust marker to verify the genuineness of a product, making it exceedingly difficult for counterfeiters to replicate. As this is a promising field, the main focus of this report will be on this application.[9,10]

Related works

DNA watermarks and digital signature techniques can be used to ensure the authenticity of synthetic DNA sequences[8]. The target of the protection mechanism could be either completely synthetic DNA, genetically modified organisms, or other systems with defined genetic compositions.

Heider and Barnekow propose the DNAcrypt algorithm, a watermarking method that includes error-correcting codes to **protect the watermark against mutations**. In silico studies using the Rab7 gene in Saccharomyces cerevisiae and using real POC with Vam7 gene of Saccharomyces cerevisiae were conducted.[11,12] Jupiter et al. aims to track **infectious or other agents**[13]. Lee et al. improved the watermarking technique. Security to protect copyright is done twofold. 1. There are many degrees of freedom to create the watermark. It can be in different positions and different levels (meaning different substitutions of codons). 2. The security is further enhanced by adding a pseudorandom sequence on the watermark[14].

Kar et al. and Gallegos et al. sign **plasmids** to authenticate their origin. A modified RSA scheme is applied to the sequence of the plasmid and attached to the plasmid. No special characteristics of DNA are used[15–18]. Basically, strings from the standard RSA protocol are encoded into DNA as intermediates.

In one of the first reports of DNA steganography, Clelland et al. described hiding messages within **chopped human DNA**. The PCR primers of the message represent the secret key[6]. An **information-theoretic** analysis of the setting in Clelland et al. was performed by Vippathalla and Kashyap[6,19].

In another example, Leier et al. used DNA binary strands and **graphical subtraction of binary gel images** for steganography. Both methods hide a message in a background of dummies and read out the message via gel electrophoretic patterns. The key is either a set of PCR primers or, alternatively, the dummy pool, which is only known to the actor encrypting the message. They show that this method is secure under the assumption that the Adversary has the same technological capabilities as the Manufacturer and Customer.[20] However, with current sequencing capabilities, this system would be relatively easy to decrypt.

A method with enhanced security was published in 2019 by Cui et al. Their novel approach to DNA steganography incorporates randomness to improve security using secondary secret keys and self-destruction mechanisms, similar to quantum key distribution methods. The proposed method aims to improve the robustness and complexity of data encoding in DNA by leveraging a combination of information-carrying DNA and a partially degenerated DNA library.[21]

In a further method that can be categorized into steganography, Volf et al. use a large genome as a basis. They exemplify the approach by demonstrating how it can be used to **sign cell lines**. They then introduce **substitutions at specific multiple locations of the genome**, using multi-site targeted base editing by adenine and cytosine base editors (ABEs, CBEs). They encode binary messages using 0 as no substitution and 1 as substitution at any given position. The authentication step consists of determining the 0/1 content in a given subset of locations. Security comes from the large possibility space of locations in the genome where substitutions could be introduced[7]and the difficulty of distinguishing random mutations introduced through processing steps versus the deliberately induced ones without knowledge of the key.

In a different approach, Schaudy et al. introduce a photolithographic in situ synthesis technique that enables high-density oligonucleotide patterning and spatially organized surface encoding, extending beyond conventional DNA synthesis. By **incorporating L-DNA**, the method adds an independent, bio-orthogonal information channel that prevents cross-hybridization and enhances data density on microarrays. The formation of patterns and signatures are based on the fact that hybridization exclusively occurs between oligonucleotide

strands of equal chirality. This allows for advanced applications such as QR code generation, counterfeit-resistant watermarks, and hidden messages within D-DNA microarrays.[22]

Moving more towards encryption, Grass et al. discuss using synthetic DNA as a medium for storing digital data and extracting encryption keys from **genetic short tandem repeats** (STRs). It highlights the potential for enhanced data security by leveraging the unique characteristics of individual genomes, which can be analyzed using existing DNA sequencing technologies. The research emphasizes the importance of key entropy and distribution in ensuring the robustness of the encryption against brute force attacks.[23]

Luescher et al.[10] use up to 10^{10} randomly generated sequences in an implementation of a **DNA-based unclonable function**. This concept is built on the principle of **physical unclonable functions**, which use randomly manufactured items that are able to process a physical stimulus into an output that is unique to the respective input[24]. Through the random features, it is impossible to reverse-engineer the input from the output, in analogy to a mathematical one-way function such as used in **cryptographic hashes**. The specific instance of random DNA sequences thus form a substrate for a chemical computing unit, which is able to transform an input into an output. The input corresponds to a set of PCR primers, which amplify a very small subset out of the billions of random sequences originally present in the random DNA pool. The amplified sequences are a specific fingerprint to the input (and the specific instance of random sequences) and can be identified using sequencing. The process is designed in such a way that the input cannot be read back from the output, and the random pool can be operated on but can not copied by PCR. Sequencing and synthesizing the entire pool would be prohibitively expensive.

In a use case, a refined implementation of DNA-based unclonable functions has been applied as a secure anti-counterfeiting tag for oral pharmaceuticals. The DNA is stabilized in a non-toxic silica matrix, which can be admixed to drug substances and retrieved for authentication. In contrast to other methods (e.g. barcodes, watermarks), DNA allows for in-product authentication on several levels as required (e.g. product, batch, production plant, etc.) while maintaining cryptographic security.

The concept of in-product information is further used in the “DNA-of-things” scheme, where the **information related to a physical product is embedded within the product** itself[25]. The difference to product-integration of unclonable functions lies in the focus on information storage rather than authentication. The term was coined by Koch et al., who stored an .stl CAD instruction file within the polymer matrix of a 3D-printed plastic bunny, which was printed from this particular .stl file. Much like biological DNA, the object then contained its own building instructions within its material. The DNA-of-things concept can thus be used to stably store relevant information within objects during their manufacturing process with the

advantage that the information is inextricably linked with its materiality. Since appropriately encapsulated DNA exceeds the stability of most materials - lasting for thousands of years - this is a promising tool for recording supply-chain information, manufacturing conditions, and more. Through its extremely high storage density in the order of several exabytes per gram, DNA is a very efficient information carrier of which only minute amounts are needed, which makes it compatible with non-destructive authentication of many materials and objects. Furthermore, DNA can be used in a "DNA-of-things" scheme where the information related to a physical product is **embedded within the product** itself[25].

Components of DNA-based authentication systems

In digital authentication schemes, computational power is the primary factor determining both applicability and security. In contrast, DNA-based schemes introduce a wider range of processing steps. The creation, verification, or potential forgery of DNA signatures involves not only computational tools but also chemical processes and specialized lab equipment. Additionally, due to the physical nature of DNA and its chemical properties, new types of attacks may emerge, exploiting these unique characteristics.

Chemical and Biological tools

Beyond computational methods, several biochemical tools can be leveraged toward the goal of developing in-product authentication schemes.

Degenerate bases or **composite DNA** can be used to increase the effective size of the DNA alphabet beyond A, C, G, and T and, therefore, efficiently add large amounts of entropy. Composite alphabets can have different complexities. The fully random alphabet is a simple one that can be used to increase entropy.

It is possible to create **unsequenceable templates** that are biochemically hard to sequence or amplify. For example – one can use (non-DNA) linkers as well as 3p-3p and 5p-5p junctions as part of the synthesis – yielding obstruction to standard amplification, hybridization, and extension processes.

CRISPR-Cas enzymatic approaches provide precise methods to create authentication schemes based on un-sequenceable and un-amplifiable chemically modified DNA and RNA.

Standard techniques like **PCR**, various types of **ligation**, **restriction enzymes,** and other NA-cutting enzymes are also potentially used by any scheme or protocol. Steps involving these techniques can be used by all three types of involved parties – M, A, and C.

Attacks

In-product authentication using DNA introduces three primary attack vectors for signature forgery.

1. **Dilution of the DNA Signature**: In this approach, adversary A obtains samples of the original product containing the authentic DNA signature and blends them with counterfeit material. While this method allows the transfer of the signature to the fake product, its effectiveness is limited, as a substantial quantity of the original DNA would be needed to convincingly mimic authenticity. A potential countermeasure involves specifying the expected concentration of DNA in the genuine product, making it challenging to dilute the sample without detection. This attack is also only relevant in some use cases where dilution is possible but not in others.
2. **Sequencing and Synthesis of DNA**: An Adversary could sequence the DNA of the signature and then synthesize an identical copy to forge it. To mitigate this risk, DNA sequences can be designed to be either exceptionally difficult or impossible to sequence using current technology, thereby hindering attempts to replicate or analyze them.
3. **Chemical Cloning of DNA**: Through techniques like Polymerase Chain Reaction (PCR), an attacker could amplify the DNA directly in a laboratory, exploiting DNA's natural ability to be replicated without exposing its full content. This method allows the creation of counterfeit DNA signatures through direct cloning of the original sequences.

Potential use cases

Implementing DNA-based authentication to ensure the authenticity and integrity of physical products offers immense value across various industries. By embedding DNA markers directly into products, companies can create robust, tamper-proof systems for tracking, verifying, and certifying goods. It also enables accountability, ensuring compliance with regulatory standards and providing a safeguard if issues arise, such as performance failure or adverse effects. Below and in Figure 1: Potential use cases for DNA-based authentication methods are some key use cases where this technology can make a significant impact:

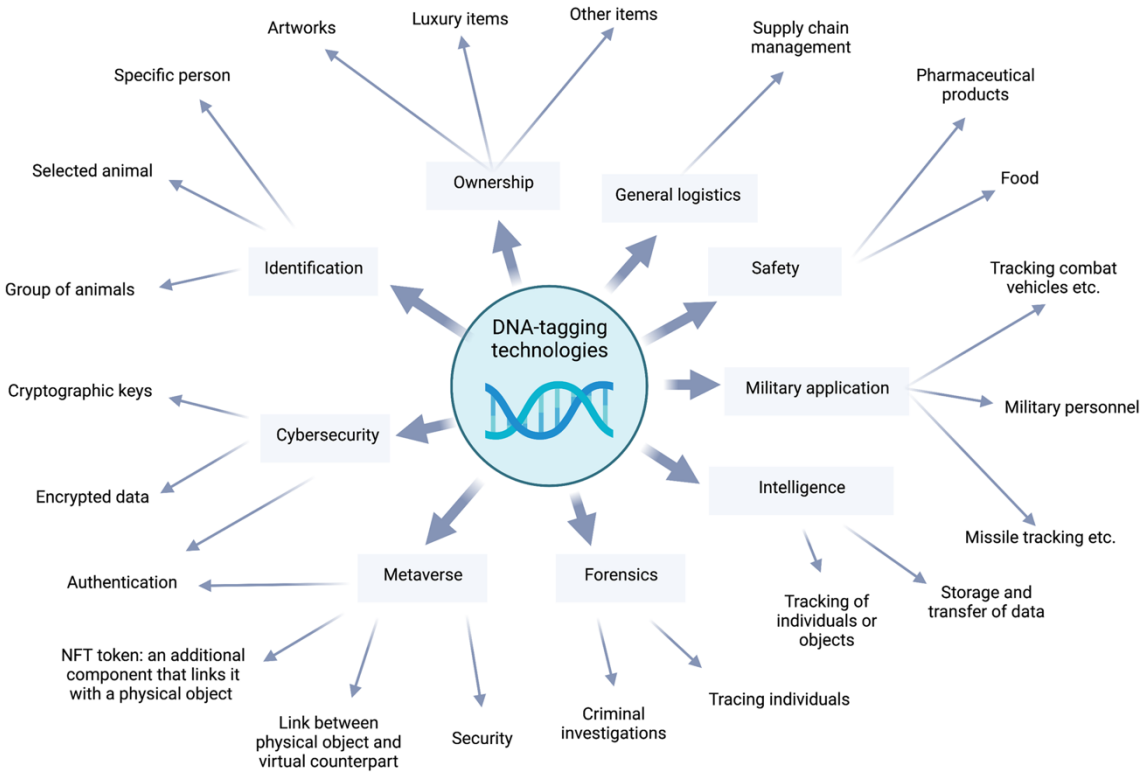

Figure 1: Potential use cases for DNA-based authentication methods[26]

Traceability

- **Gemstones**: DNA markers can be applied to gemstones, allowing for precise tracking of their origin down to the specific mine. This provides transparency for buyers and helps ensure that gemstones are sourced ethically and conflict-free.
- **High-stakes industries**: Critical components can be embedded with DNA markers to store information about the part's origin, manufacturer, and safety certifications. This creates a tamper-proof tracking system that enhances safety and accountability throughout the part's lifecycle. One example of this would be the airplane industry, where all components must be failure-proof.
- **Textiles**: Manufacturers can create a transparent supply chain by marking textiles with DNA at the source. This ensures that the materials used in garments are traceable throughout every stage of production, offering consumers confidence in the authenticity of their products and the ethical sourcing of materials.
- **Food of problematic origin**: For products like cocoa or coffee, which are often sourced from regions with problematic labor practices, DNA markers can help make the supply

chain transparent. This allows buyers to verify the origin of the product and ensure it meets ethical and environmental standards.

- **Fertilizers and pesticides**: DNA-based certification can be applied to fertilizers and pesticides, enabling precise tracking of their use in agriculture. This helps with regulatory compliance, ensuring that only authorized and safe quantities are used while preventing counterfeit or harmful products from entering the market.
- **Hydrogen products**: As the hydrogen economy grows, distinguishing between green hydrogen (produced via electrolysis) and hydrogen from methane-based sources becomes crucial. DNA markers can ensure the origin of hydrogen products, making the tracking and certification of their source transparent.
- **Biological products**: For biological products like genetically modified organisms (GMOs) or cell lines, traceability to the manufacturer is essential to ensure safety and proper function. Knowing the origin verifies that these products behave as intended, reducing risks of contamination, unintended mutations, or unexpected behavior.

Counterfeit protection

- **Precious materials**: DNA markers can be embedded in materials like porcelain, marble, or other valuable resources, protecting against counterfeiting. This ensures that high-quality materials can be distinguished from cheap imitations, providing buyers with confidence in their purchases.
- **Pharmaceuticals**: In the pharmaceutical industry, DNA markers can ensure the authenticity of drugs, protecting against counterfeit products that pose serious health risks. By embedding unique DNA signatures in the products themselves, manufacturers can verify the legitimacy of medications throughout the supply chain. This prevents fake drugs from entering the market, safeguards patient health, and maintains trust in pharmaceutical brands.[27]
- **Cosmetics**: In the cosmetics industry, where counterfeit products pose significant health risks, DNA authentication can verify the authenticity of ingredients and finished products. This not only protects consumers but also preserves the reputation of legitimate brands.
- **Jewelry**: DNA can be embedded into high-value jewelry and watches. Instead of providing a certificate with the product, the authenticity can be ensured by embedding DNA in the product itself.
- **Certification authorities**: Any authorities that certify products or goods, starting from cars up to medical devices, can utilize DNA. Then, the certificate can be an integral temper-resistant part of the product itself.

These use cases demonstrate the versatility of DNA-based authentication across multiple sectors, highlighting the technology's potential to enhance both traceability and counterfeit protection. As the demand for product transparency and security increases, DNA markers can become a critical tool for ensuring integrity across global supply chains.

In-product documentation

The ability to store any digital file in DNA goes far beyond simple markers or tags. This opens up opportunities for enhancing products with in-product documentation, creating an immutable memory that lasts as long as the product itself. By embedding DNA within products, we can store valuable information that remains accessible over the product's entire lifespan. Below are some potential applications of this concept:

- **Building Materials:** Embedding DNA in building materials could store critical information about their origin, composition, and recycling instructions. This would provide traceability throughout the supply chain and offer guidance on sustainable disposal or reuse. Additionally, DNA could retain architectural plans and material specifications directly within the structure, ensuring indefinite access to the blueprint of the building, even decades after its construction.
- **Repairability:** DNA embedded within (high-tech) components could store repair instructions, making it easier to fix damaged parts without needing to determine their composition or manufacturing details first. This would preserve essential data, ensuring that information about how to maintain or repair a component never gets lost, which is crucial for extending product lifecycles and reducing waste.
- **Circular Economy:** To support a truly circular economy, tracking materials throughout their lifecycle is essential. DNA-encoded data can provide a unique method for this, especially for bulk materials or products that are difficult to label. In-product documentation becomes even more relevant with the European Union's introduction of a mandatory Digital Product Passport, offering a streamlined way to trace and manage materials as they move through cycles of use, reuse, and recycling.
- **Cultural Heritage:** Embedding DNA into cultural artifacts, artwork, or historical items could serve as a way to preserve their provenance, creator information, and care instructions. This would create an enduring link between the item and its history, maintaining a chain of custody and detailed records for future generations, ensuring the legacy of culturally significant objects is never lost.
- **Medicine and Foodstuffs:** DNA can store critical information about the production, safety standards, and expiry dates of medicines and food products. This ensures that

even if packaging is lost or labels fade, essential details about the product remain intact, helping both consumers and regulatory bodies maintain trust and safety.

- **Medical Implants:** Medical implants could carry detailed information about their design, materials, immune related characteristics, and even the surgical procedure used for implantation. This in-product documentation would allow healthcare providers to access vital information for future surgeries, maintenance, or replacements, improving patient outcomes and reducing risks in long-term medical care.
- **Automotive and Aerospace Components:** Critical components in vehicles or aerospace manufacturing can carry service histories, technical specs, and repair logs through in-product DNA information. This enhances safety, ensures proper maintenance, and creates a tamper-proof record, especially for high-stakes industries where precision and accountability are essential.

These applications highlight the potential of DNA-based in-product documentation. As industries move toward greater transparency, sustainability, and longevity, using DNA to store detailed, immutable information within products may become an essential tool for achieving these goals.

Regulatory aspects

The regulatory requirements for implementing Digital Product Passports (DPPs) under the EU's Ecodesign for Sustainable Products Regulation (ESPR) are designed to enhance transparency, sustainability, and traceability across the lifecycle of products. As part of the ESPR, companies will be required to provide DPPs for a wide range of products starting in 2024, with a phased rollout that will extend to more products by 2030. To comply, each DPP must include a detailed digital record containing information such as the product's unique identifier, global trade identification number, and relevant regulatory documentation like declarations of conformity and certificates of compliance.

The ESPR sets stringent requirements for documenting materials, substances of concern, and the product's environmental impact, including data on recyclability, repairability, and safe disposal. Companies must also include user manuals and information on how to handle the product at its end-of-life stage, ensuring that it aligns with EU circular economy goals. The DPP must provide this information not only to consumers but also to regulatory bodies, treatment facilities, and other stakeholders along the value chain.

In this regulatory landscape, DNA-based in-product documentation offers a secure and tamper-proof solution to meet these compliance requirements. By incorporating DNA markers that are unique to each product or batch, companies can authenticate the origin,

materials, and compliance data of the product, thereby adhering to the ESPR's mandates for traceability and transparency. This approach also supports the EU's broader objectives of reducing environmental and climate impacts by enabling more sustainable, longer-lasting, and repairable products.

Companies active in the field

Several companies are active in the field of securing products using DNA. The current state-of-the-art for commercial use is DNA tagging, where no additional cryptographic schemes are applied. An overview of relevant companies can be found in Table 1 which is taken from Kuzdraliński et al.[26]

Table 1: Companies that provide or create technological solutions for labeling tangible items through the application of DNA[26]

Company name, country of origin and launch date	Main features of the technology	Selected markets
Applied DNA Sciences, United States, 1983	Botanical DNA fragments, detection by PCR and CE, an encapsulation system	Product authentication, supply chain traceability, brand protection, anti-counterfeiting, textiles, pharmaceuticals, etc.
Haelixa, Switzerland, 2016	Synthetic DNA tags, detection by PCR, DNA enclosed in silica	Product authentication, supply chain traceability, intellectual property protection, etc.
Selectamark Security Systems (SelectaDNA), United Kingdom, 1986	Laboratory analysis of DNA for owner identification if microdots are absent (DNA serves as an alternative authentication solution)	Asset protection and recovery, securing high-value items, art and jewelry authentication, IT equipment and vehicle security, forensic applications, theft prevention and deterrence, etc.
TraceTag (CypherMark), United Kingdom / Norway, 2001	Synthetic DNA with unique primers, authorized access to primer sequences, detection using qPCR	Brand safeguarding, industrial applications, cash security, security of documentation, oil and fuel tracking, anti-counterfeiting measures, etc.
Holoptica, United States, 2012	Synthetics DNA tags (100 nucleotides), integration with inkjet cartridges	Artwork, documents and assets protection, verifying product authenticity, food tracking, etc.
DNA Technology, United States, 1993	DNA-laced ink, combination of DNA synthetic segments and optical taggants	Memorabilia and collectibles, limited edition artwork, pharmaceuticals, apparel and luxury goods, health and beauty industry, etc.

Tagsmart, United Kingdom, 2015	Synthetic DNA tags, secure Certificate of Authenticity	Artwork, securing collectibles, verifying paper documents, book manufacturing, etc.
DNA Guardian, Australia, 2007	UV-detectable stain, detection using pyrosequencing	Asset marking, crime prevention, artwork protection, theft deterrence, etc.
Aanika Biosciences, United States, 2018	Genetically modified *Bacillus subtilis* as an encapsulation system for DNA tag	Agriculture and food production, textiles, etc.

This shows that there is a market for DNA-based tagging that can be embedded into various products. As technology advances, it will be the logical next step to also include security mechanisms in these tags such that the methods become more reliable and further use cases can emerge.

Current shortcomings

Many of the DNA-based authentication solutions currently available in the market rely on embedding plain text DNA sequences within products as a means of ensuring authenticity and traceability. While this approach may initially seem secure, it presents several critical vulnerabilities that sophisticated attackers could exploit.

The primary security mechanism for this approach rests on the complexity of DNA synthesis. In theory, the technical expertise and equipment required to produce synthetic DNA create a barrier for counterfeiters. However, this protection is rapidly diminishing as advancements in biotechnology make DNA sequencing and synthesis more accessible. A well-equipped laboratory with readily available technology can now sequence the embedded DNA, reverse-engineer it, and synthesize identical or nearly identical markers. These synthetic DNA markers could then be incorporated into counterfeit products, rendering the original authentication method ineffective.

As the reliance on the inherent difficulty of DNA synthesis as a security measure is becoming increasingly outdated, there is a need for more advanced verification and authentication methods.

In order to achieve this, a cryptographic framework is necessary to assess the schemes that are developed. Similar to IND-CPA and EUD-CPA in classical cryptography, one needs to find a standard to make DNA-based schemes comparable to one another.

The analysis of DNA-based authentication schemes and approaches needs to address, like in computational security analysis approaches, the complexity of reading and reproducing

signatures. In the context of DNA, this complexity has components that mostly relate to the cost of the physical/biological steps required for adversaries to attack – that is, to successfully reproduce signatures. The current literature does not address security analysis under any formal framework.

Randomness is one means that can be used in secure signature schemes. However – some of the current approaches require some overhead from both M and C. Understanding and improving schemes, from this perspective, is a direction where technology can improve. Combining watermarking with information storage is also an interesting potential practical improvement.

Some of the state-of-the-art methods rely on either complex synthesis or on specific live genome characteristics. Either one of these can be a limitation for certain use cases.

In some use cases, dilution attacks can be very effective. A framework for understanding dilution attacks and their limitations can help in this context. Furthermore – through analysis and experimental results, the community will develop, analyze and characterize approaches by which to protect against dilution attacks.

Conclusion and Outlook

The demand for securing physical products with DNA-based authentication methods is already evident in the market. Industries are using in-product DNA markers to protect the authenticity and integrity of goods. Currently, these solutions rely on embedding plain text DNA sequences, offering a basic level of security rooted in the complexity of DNA synthesis. While this approach may suffice for low-risk applications, it falls short when it comes to securing more valuable and critical products.

As the field of DNA-based authentication continues to grow, the limitations of plain text DNA markers are becoming more apparent. Sophisticated attackers with access to modern biotechnology can easily sequence and replicate these markers, undermining their effectiveness. To address these vulnerabilities, there is a clear and urgent need for cryptographic methods that make DNA markers tamper-proof. By integrating encryption and other cryptographic protections, we can ensure the authenticity of DNA markers, making it far more difficult for counterfeiters to replicate or alter them.

In the future, these enhanced cryptographic solutions will be essential for protecting high-value products and ensuring the long-term reliability of DNA-based authentication technologies.

A crucial step toward advancing DNA-based authentication will be the development of comprehensive frameworks for assessing the security and reliability of these schemes. Given the wide variety of approaches within this emerging field, establishing a consistent basis for comparison presents a significant challenge. The complexity arises from the diverse processes involved in creating, verifying, or forging DNA signatures, which differ across methodologies.

To address this, one promising solution is to break down the entire workflow— from signature creation to verification—into standardized lab procedures. By defining and codifying these steps within a structured assessment framework, it will be possible to evaluate the security and robustness of different schemes on a common ground. This standardized approach could also help identify vulnerabilities, establish best practices, and create benchmarks that would drive further innovation. Over time, such frameworks could lead to universally accepted standards for DNA-based authentication, ensuring its reliability across industries and making it a viable solution for protecting high-value products and data.

Bibliography

1. Buchmann, J. A. *Introduction to Cryptography*. (Springer New York, New York, NY, 2004). doi:10.1007/978-1-4419-9003-7.
2. Goldreich, O. *Foundations of Cryptography*. (Cambridge University Press, Cambridge, 2001).
3. Katz, J. & Lindell, Y. *Introduction to Modern Cryptography*. (Chapman and Hall/CRC, 2007). doi:10.1201/9781420010756.
4. Lindell, Y. How to Simulate It – A Tutorial on the Simulation Proof Technique. in *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich* (ed. Lindell, Y.) 277–346 (Springer International Publishing, Cham, 2017). doi:10.1007/978-3-319-57048-8_6.
5. Alagic, G. *et al. Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process*. NIST IR 8413-upd1 https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8413-upd1.pdf (2022) doi:10.6028/NIST.IR.8413-upd1.
6. Clelland, C. T., Risca, V. & Bancroft, C. Hiding messages in DNA microdots. *Nature* **399**, 533–534 (1999).
7. Volf, V. *et al.* Cryptography in the DNA of living cells enabled by multi-site base editing. 2023.11.15.567131 Preprint at https://doi.org/10.1101/2023.11.15.567131 (2023).

8. Berezin, C.-T., Peccoud, S., Kar, D. M. & Peccoud, J. Cryptographic approaches to authenticating synthetic DNA sequences. *Trends Biotechnol.* (2024) doi:10.1016/j.tibtech.2024.02.002.

9. Li, Y. *et al.* Genetic physical unclonable functions in human cells. *Sci. Adv.* **8**, eabm4106 (2022).

10. Luescher, A. M., Gimpel, A. L., Stark, W. J., Heckel, R. & Grass, R. N. Chemical unclonable functions based on operable random DNA pools. *Nat. Commun.* **15**, 2955 (2024).

11. Heider, D. & Barnekow, A. DNA-based watermarks using the DNA-Crypt algorithm. *BMC Bioinformatics* **8**, 176 (2007).

12. Heider, D. & Barnekow, A. DNA watermarks: A proof of concept. *BMC Mol. Biol.* **9**, 40 (2008).

13. Jupiter, D. C., Ficht, T. A., Samuel, J., Qin, Q.-M. & Figueiredo, P. de. DNA Watermarking of Infectious Agents: Progress and Prospects. *PLOS Pathog.* **6**, e1000950 (2010).

14. Lee, S.-H. DWT based coding DNA watermarking for DNA copyright protection. *Inf. Sci.* **273**, 263–286 (2014).

15. Kar, D. M., Ray, I., Gallegos, J. & Peccoud, J. Digital Signatures to Ensure the Authenticity and Integrity of Synthetic DNA Molecules. in *Proceedings of the New Security Paradigms Workshop* 110–122 (Association for Computing Machinery, New York, NY, USA, 2018). doi:10.1145/3285002.3285007.

16. Kar, D. M. & Ray, I. That's My DNA: Detecting Malicious Tampering of Synthesized DNA. in *Data and Applications Security and Privacy XXXIII* (ed. Foley, S. N.) 61–80 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-030-22479-0_4.

17. Kar, D. M., Ray, I., Gallegos, J., Peccoud, J. & Ray, I. Synthesizing DNA molecules with identity-based digital signatures to prevent malicious tampering and enabling source attribution. *J. Comput. Secur.* **28**, 437–467 (2020).

18. Gallegos, J. E., Kar, D. M., Ray, I., Ray, I. & Peccoud, J. Securing the Exchange of Synthetic Genetic Constructs Using Digital Signatures. *ACS Synth. Biol.* **9**, 2656–2664 (2020).

19. Vippathalla, P. K. & Kashyap, N. The Secure Storage Capacity of a DNA Wiretap Channel Model. *IEEE Trans. Inf. Theory* **69**, 5550–5569 (2023).

20. Leier, A., Richter, C., Banzhaf, W. & Rauhe, H. Cryptography with DNA binary strands. *Biosystems* **57**, 13–22 (2000).

21. Cui, M. & Zhang, Y. Incorporating Randomness into DNA Steganography to Realize Secondary Secret key, Self-destruction, and Quantum Key Distribution-like Function. Preprint at https://doi.org/10.1101/725499 (2019).

22. Schaudy, E., Somoza, M. M. & Lietard, J. l-DNA Duplex Formation as a Bioorthogonal Information Channel in Nucleic Acid-Based Surface Patterning. *Chem. – Eur. J.* **26**, 14310–14314 (2020).

23. Grass, R. N., Heckel, R., Dessimoz, C. & Stark, W. J. Genomic Encryption of Digital Data Stored in Synthetic DNA. *Angew. Chem. Int. Ed.* **59**, 8476–8480 (2020).

24. Pappu, R., Recht, B., Taylor, J. & Gershenfeld, N. Physical One-Way Functions. *Science* **297**, 2026–2030 (2002).

25. Koch, J. *et al.* A DNA-of-things storage architecture to create materials with embedded memory. *Nat. Biotechnol.* **38**, 39–43 (2020).

26. Kuzdraliński, A. *et al.* Unlocking the potential of DNA-based tagging: current market solutions and expanding horizons. *Nat. Commun.* **14**, 6052 (2023).

27. Altamimi, M. J. *et al.* Anti-counterfeiting DNA molecular tagging of pharmaceutical excipients: An evaluation of lactose containing tablets. *Int. J. Pharm.* **571**, 118656 (2019).