Automating DNA Biometric Recognition for Real-Time Person Identification

* Ogechukwu Iloanusi, Ijeoma Okeke
Department of Electronic Engineering
Faculty of Engineering
University of Nigeria
Nsukka Campus, Enugu State 410001, Nigeria
Email: Ogechukwu.Iloanusi@unn.edu.ng, Ijeoma.Okeke@unn.edu.ng

* Corresponding Author

Abstract

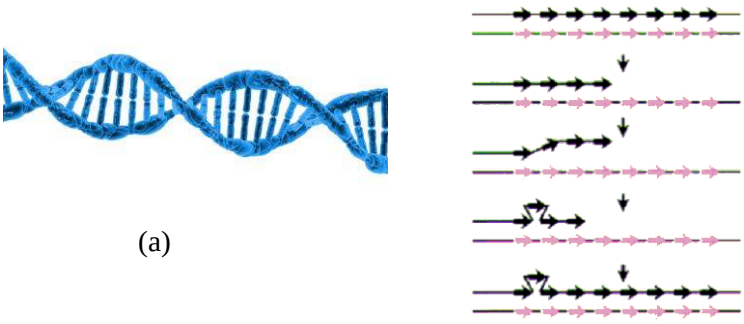
This paper presents possible automation of DNA (DeoxyriboNucleic Acid) profiles using simple sequence repeats (SSR) inherent in a DNA structure. DNA is the most accurate form of biometric and will be useful in environmental and border security, immigration, civil applications and access control though its greatest challenge remains the very significant time it takes to process a DNA profile.

Keywords: Biometrics, DNA, recognition, simple sequence repeats, identification

Introduction

DNA (DeoxyriboNucleic Acid) recognition is identification of persons based on their DNA compositions. DNA, shown in Figure 1 (a), is a hereditary material found all over the human body. It comprises the genetic information for the development and functioning of living organisms. A DNA is a nano-structure measuring between 2.2 to 2.6 nanometers (nm) wide and 0.34nm long [1-3] and it has about 90% of information that is repeatable amongst living beings and about 10% information that is unique to every living being. The 10% comprises information like the Simple Sequence Repeats (SSRs) shown in Figure 1 (b), which are repeating sequences of 2-6 base pairs of DNA [2]. It is this unrepeatable region that is the basis for DNA recognition. The probability of two individuals having the same DNA is one in trillion from research [4]. Some sources of DNA are hair, clothing, glass, plastic, blood, dried skin, used tissue [4, 5]; in fact, our DNA is left behind in many things we come in contact with without realizing it. Every human has a unique composition of DNA except for monozygotic twins. DNA is considered a biometric identifier because of the following characteristics: uniqueness; universality - DNA trait is found in all individuals; permanence over a lifetime – not affected by aging or accident; ease of collection; and most importantly its highest accuracy. DNA identification is currently employed in Forensics for tracing acquired DNAs to their owners, as well as identifying the rightful parents of a child, and it is our hope that its automation would lead to the speedy processing of DNA profiles. Automated DNA profile authentication will be useful in environmental and border security, immigration, civil applications and access control.

The greatest challenge in DNA biometrics is the very significant extraction and process time for a DNA sample [6] due to its lack of automation. DNA is mostly acquired biologically in laboratories and takes about two to five hours to process a DNA sample.



(a) (b)

Figure 1(a): DNA structure adapted from [7]. (b) Simple Sequence Repeats (SSR) adapted from [8].

DNA Recognition

The automation of DNA profiles made up of Simple Sequence Repeats (SSRs) information inherent in DNA structure is presented in Figure 2. Simple Sequence Repeats have repeating structures, which are unique for every individual, and this forms the basis for choosing this sequence as a form for developing unique features for every individual. Ordinarily, DNA extraction, amplification and sequencing take at least 2 hours. DNA amplification requires polymerase chain reaction, which takes at least an hour. A suitable diagram for the biometric recognition of DNA profiles in an identification scenario is also presented in Figure 2. DNA as a biometric can be acquired, enrolled, stored and retrieved for feature matching in an identification or verification system.

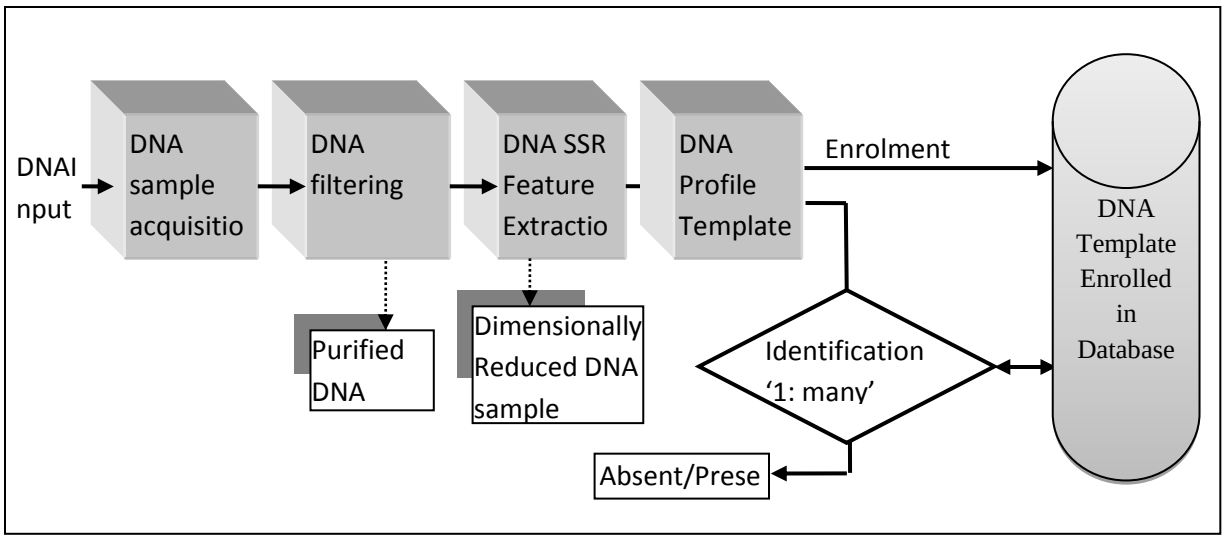


Figure 2: Biometric Recognition of DNA profiles in an Identification Scenario

The following stages are required for the biometric recognition of DNA profiles as shown in Figure 2:

- DNA sample acquisition: DNA sample is easily acquired from the body of a person

or any object s(he) has come in contact with.

- DNA filtering: The acquired DNA is purified if contaminated.
- Extraction of SSR features: SSR features are extracted from the purified sample this stage. This stage includes DNA amplification and sequencing and takes a significant time and hence needs automation.
- DNA template creation and Enrollment: A template is created from the feature extraction stage. The template (not the DNA sample), unique to any individual, is enrolled in a database for later use in person identification.

Enrolment is the process of registering a user's DNA template in a database. It could be an identification database or a memory chip in a card. Identification is the process of verifying a user's real-time DNA template against a database of many DNA biometric templates to seek a match or non-match. The output of this state is a present or absent decision based on the similarity of the input DNA and database DNA templates matched.

Conclusion

Automated recognition of DNA profiles using simple sequence repeats (SSR) is presented in this paper. DNA is the most accurate biometric and is currently employed in forensics and it is our hope that its automation would lead to the speedy processing of DNA profiles. Further advances in DNA biometric research may lead to its deployment in very high security systems.

References

1. Watson, J. and F. Crick, Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature, 1953. 4356: p. 737–738.
2. Alberts, B., et al., (2002). Molecular Biology of the Cell; Fourth Edition. : .2002, New York and London: Garland Science.
3. Butler, J.M., Forensic DNA Typing. Elsevier, 2002: p. 14–15.
4. Butler, J.M., Forensic DNA Typing2005, Burlington, MA: Elsevier Academic Press.
5. Budowle, B., et al., Source attribution of a forensic DNA. Forensic Science Communications, 2000. 2(3).
6. P., V., V. V., and Z. G., DNA Computing based Elliptic Curve Cryptography. International Journal of Computer Applications, December 2011. 36(4): p. 18-21.
7. Public Domain Pictures, D.; Available from: http://www.publicdomainpictures.net/view-image.php?image=42718&picture=dna.
8. Simple Sequences in Proteins and DNA. Available from: http://big.crg.cat/computational_biology_of_rna_processing/simple_sequences_in_proteins_and_dna.