



Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Forensic STR profiling based smart barcode, a highly efficient and cost effective human identification system

Andleeb Zahra^a, Bilal Hussain^b, Amer Jamil^c, Z. Ahmed^d, Shahid Mahboob^{d,b,*}^a Department of Bioinformatics, Government College, University Faisalabad, Pakistan^b Department of Zoology, Government College University, Faisalabad, Pakistan^c Department of Biochemistry, University of Agriculture, Jail Road, Faisalabad Pakistan^d Department of Zoology, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

ARTICLE INFO

Article history:

Received 16 August 2018

Revised 30 September 2018

Accepted 1 October 2018

Available online 02 October 2018

Keywords:

DNA profiling

Human identification

Barcode

HIBS

ABSTRACT

In forensic science, the human identification is the major goal in criminal cases and in paternity dispute, based on DNA profiling of individual. Analysis of short tandem repeat (STRs) markers used as a reliable technique for this purpose. In this study the main objective was to develop a human identification based on DNA fingerprinting using Human Identification Barcode System (HIBS). HIBS may be used to provide a unique molecular signature of human in the form of a barcode. DNA was isolated from blood by using PCR technique to detect bands to design human barcode using self-designed system. HIBS is a web based application that can be accessed via www.hibs.com.pk. HIBS can be accessed with internet access and may be introduced on security checkpoints to identify an individual based on his own DNA instead of conventional procedures of identification. The barcode generated through DNA fingerprinting will be stored in a HIBS, and may also be a part of CNIC. It may be successfully used against suicide bombers, target killers, etc., as even a single blood spot, a few skin cells, root of the hair etc., to identify such culprits. It may also be effectively used to relieve the individuals with false accusations.

© 2018 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

From last two decades, DNA profiling has become incumbent technology, which has diverted the molecular scientists towards forensic identification (Walsh, 2005). DNA amplification and DNA typing through multiplexing the STR loci been the subject of ongoing debate (Ip et al., 2014). From paternity testing for criminal investigations, Final curtain of forensics is Human identification. STR analysis is a classical and validated assay for human identification in forensic studies. Continuing Amelioration of technology lead to use of STR in human identification easier and robust (Azari et al., 2007; Collins et al., 2004; Von Wurmb-Schwark et al., 2009). A comprehensive and uniform system of investigation devised by the Federal Bureau of Investigation (FBI) Laboratory in

November 1997. It is based on a core set of STR loci, including CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, and vWA are used by the Combined DNA Index system (Budowle and Chakraborty, 2001; Budowle et al., 1998, 1999). Multiplexing these Standardized STR core set Amelogenin is making the identification more reliable. Some of rapid methods have already been published and utilization of such methods to reduce and time is gaining momentum (Azari et al., 2007; Wang et al., 2009). These efficient techniques were merged with technology to maintain the criminal DNA databases at national and international level (Wang et al., 2009) as well as the establishment of new technological methods for human identification like human barcode and RFID chip as an identifier within limited range. These conventional techniques including a microchip that can transmit a static identifier has many potential socioeconomic and health hazards. It invades the personal privacy, suspected of hacking and injecting inside the body is itself is hazardous (Smith, 2008). This research was based on amplification of STRs through PCR and designing of human barcode using self-designed artificial intelligent system named Human Identification Barcoding system "HIBS". HIBS are web based application that can be accessed by www.HIBS.com.pk. This Barcoding system takes the input in the form of DNA bands data obtained by running the

* Corresponding author.

E-mail address: shahidmahboob60@hotmail.com (S. Mahboob).

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

PCR amplified products on agarose gel electrophoresis or automated electrophoresis system (Experion or bioanalyzer), and has the ability to compare those to the standard DNA marker. On the basis of size and number of bands detected from the picture, it generates a barcode loaded along with the samples. Being STRs, band size and/or number of amplified DNA bands present in each individual differ from each other; therefore, a unique barcode is generated for each individual that serves as Molecular Signature or Molecular ID of an individual. Being web-based application, it can be accessed from anywhere and on acquiring a barcode thorough registration it provides an online authenticated identification (based on molecular data of a person) and profile review of a person via scanning barcode through the barcode scanner. The system is so designed that only “one man, one molecular signature” is possible.

2. Materials and methods

2.1. Selection of STRs

A set of 14 CODIS primers was chosen which are validated and most reliable for the study of human identification from (Kupfer et al., 2006). The selection was made due to the high level of polymorphism and ability to be used differently in Multiplex system.

2.2. DNA extraction from blood samples and Multiplex PCR amplification

Blood samples were collected from volunteers and their profile was maintained in a database. All individuals were informed about the utilization of their samples and data by taking their consent according to the research policy of Institutional research ethics committee. Genomic DNA from blood was isolated using a PAXgene blood extraction kit (cat#761133) preanalytix Quagen followed by their protocol provided and it was optimized according to requirements. Quantified DNA samples were subjected to multiplex polymerase chain reaction using a set of 14 primers

(Kupfer et al., 2006) at 45 °C annealing temperature. The amplified products were analyzed on 2% agarose gel electrophoresis under UV light on gel documentation system (Syngene UK) and low range DNA Ladder (Fermentas cat# SMO 323) 100 bp DNA ladder as a marker. Soft and hard copy of this image was attained for further processing and generation of barcode.

2.3. Generation of barcode

Scanned or saved soft copies of the images were uploaded to HIBS system. To ensure 100% accuracy of the data we designed a module for sending the email to the admin so that the admin could verify the data provided by users and the pictures if those pictures were according to specified format and protocol. After verification admin could send the activation code by email to the user. “Get barcode module” was designed to enter that activation code to retrieve the barcode. The barcode is decoded on barcode reader that displays the personal profile of the registered user.

For identification of unknown samples, PCR is done and the image is uploaded and a barcode is generated. Soft and hard copy of this image was attained for further processing and generation of barcode as shown in Fig. 1.

2.4. Design and implementation

This system was designed using Visual studio 2010, open CV and database was designed in My SQL. Users were provided with user friendly interface and directed step by step with Help menu. Input file is softcopy of STR based amplification image. It has three main modules. (1) Registration, (2) Unregistered User, (3) Display profile.

2.4.1. Registration

In this module, User profile is retrieved to save it in database for future use. After profile, user is directed to upload the image according to preset standard which is also saved in database. As soon as the process is completed, administrator receives an email

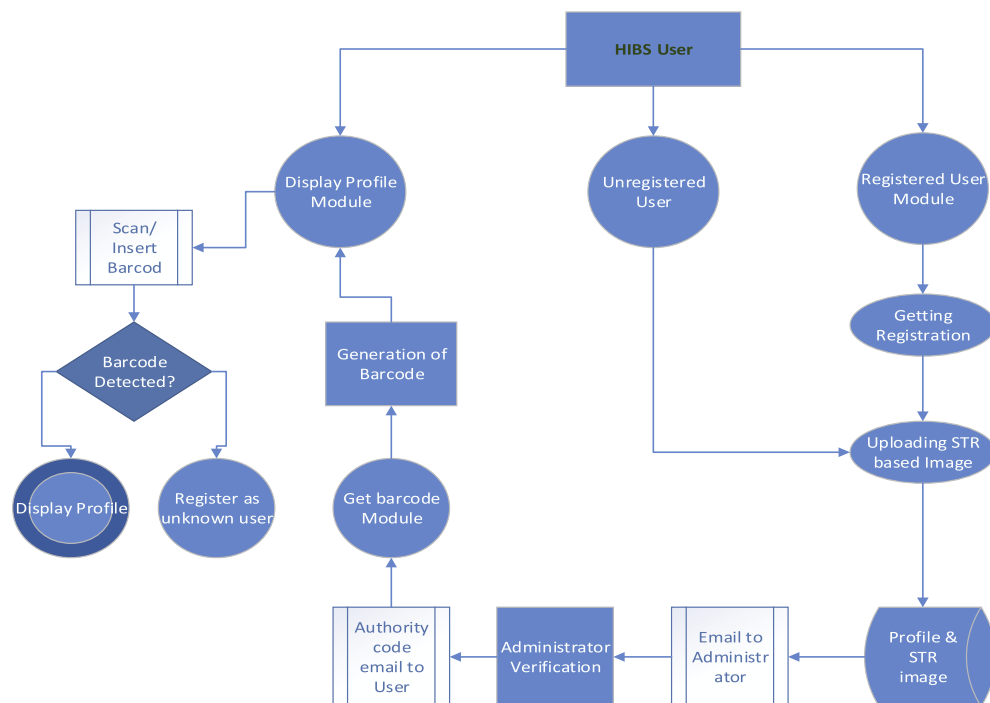


Fig 1. Schematic Work Flow of working of HIBS.

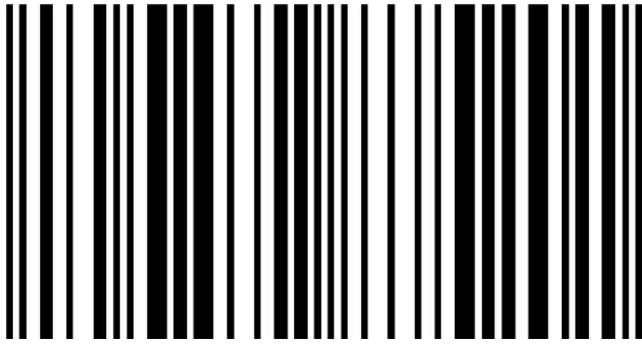


Fig. 2. Barcodes designed by the Human Identification Barcoding system.

for verification of user. This check point ensures the validity of data as well as the user of system. After validation, Administrator sends a code to user for generation of Barcode from get barcode tab as shown in Fig. 2.

2.4.2. Unregistered user

This module is designed for forensic casework specifically. Here we have just STR amplification results image based on DNA obtained from scene investigation. It checks the availability of profile in database by recognizing the pattern of given image. If same pattern already existing in database, it will show the barcode so that profile can be seen.

2.4.3. Display profile

This module displays the personal profile by scanning the barcode. It can be utilized at security check points. Culprit can deceive by any way but cannot change his or her STR Genotype. So the generated barcode is not only an ID but it is Molecular ID which is real time need of this era.

3. Results and discussion

The goal of this study was to develop a unique system for human identification in the form of DNA based barcode. STR amplification of the extracted DNA using simple instruments and procedure followed by electrophoresis and elucidation of obtaining data are the key points of forensic genetic profiling (Liu et al., 2008). Hence, the system was established using PCR of well-characterized, short, repetitive genomic regions followed by agarose gel electrophoresis system.

The HIBS is designed to register, the personal profile along with a STR profiling image which is further decoded and bar code against each STR image is assigned. Registration is used for creation of personal profile of the user for future identification. Barcode after registration may be used in public places such as

airports, security checks or in other places where verification and authentication of a person is required. This barcode could be added in CNIC for quick and highly accurate identification which is based directly on molecular data of a person. To reduce the data duplication, checks were applied that could detect the picture if it is used against any previous record and for high accuracy.

DNA was isolated from whole Human blood using a PAXgene blood extraction kit (cat#761133) preanalytix Quagen followed by their protocol provided. The extracted DNA was used for PCR amplification and 2D electrophoresis image is obtained as shown in Fig. 3.

The 2D electrophoresis image is processed by the HIBS for Bands detection. Automated band reading in STR profiling is complicated due to many challenges such as highly degraded and contaminated sample collection from the scene (Ballantyne et al., 2011). Pattern recognition is one of the technique that can be used to read bands from uploaded images. Pattern recognition of 2D PCR-RFLP images had been used for lane intensity profile by removing background intensity (Maramis and Delopoulos, 2010). Pattern recognition of 2D Images is being extensively used discrimination of Plants, 2D fingerprints, Spot segmentation of 2D histograms etc., (Rodriguez et al., 2014; Ye et al., 2011; Zhai et al., 2010; Zacharia et al., 2010). HIBS also uses Pattern recognition for band detection, but this software has ability to compare Low range DNA ladder 100 bp (Fermentas cat# SMO 323) and PCR product and auto recognize the size of every primer amplification and number of amplifications. Unidentified User module deals with any culprits, including suicide bombers, target killers, other suspects etc. using STR. Multiplex amplification by the polymerase chain reaction (PCR) and typing of short tandem repeat (STR) loci provide a sensitive, potentially highly discriminating and rapid means of characterizing forensic specimens (Moretti et al., 2001). These STRs being varied from human to human provide the unique pattern on the gel, pictures of these gels taken from gel documentation system are analyzed by HIBS. HIBS have the ability to detect the unique pattern of the gel using its image processing, pattern recognition and pixels detection ability of HIBS makes it smarter to detect this unique pattern on 2D gel. It detects that pattern setting specific threshold adjusted by user from giving menu as shown in Fig. 3.

Digital image processing and pattern recognition have key importance in digital forensics and well known modern method of investigation. Image analysis techniques of auto-montage images provide objective data to complement the hair analysts' microscopic observations, which is non-destructive, would be of obvious benefit in the forensic examination of hairs (Brooks et al., 2011) hence the use of Image analysis, pattern recognition and pixel detection techniques make the HIBS more reliable in digital investigations. Moreover, Bands recognition as these amplifications are the unique character of human, on the basis of these characters' barcode is being generated is also a unique as shown

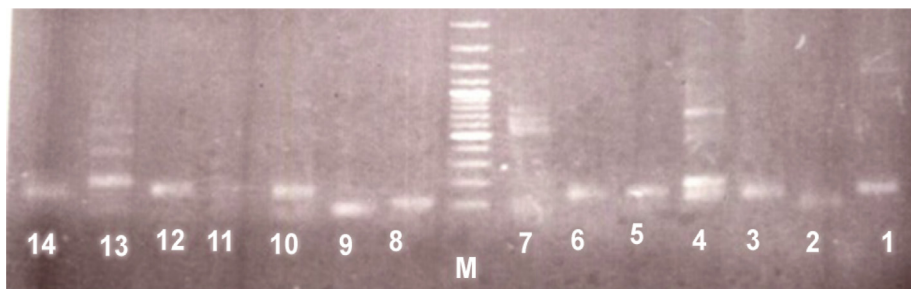


Fig. 3. CSF1PO (1), D3S1358 (2), D5S818 (3), D7S820 (4), D8S1179 (5), D13S317 (6), D16S539 (7), D18S51 (8), D21S11 (9), FGA (10), TH01 (11), TPOX (12), FGA (13) and amelagon (14) along with Low range DNA marker (M) used for comparisons.

in Fig. 2. This barcode is associated with user profile present in the database which was generated at the time of registration of users.

In case of unregistered (unknown) user the barcodes generated and stored in a database, on scanning via barcode reader, display information about the suspects. Hence this barcoding system could be a foolproof security and identification system. This is user friendly and online access system provides a comprehensive security solution. One can get its identical, which is highly secure and reliable. It is an innovative approach that promises the new level of physical security with a comprehensive database and easily scan-able cards or implants, by which one's vital statistics is readily observable by the authorities. Fundamentally, HIBS are a novel and useful idea of providing a unique molecular signature to every individual. Furthermore, it is a low-cost technology for easy implementation without the need of expensive equipment.

4. Conclusion

HIBS is a novel idea that is generated by amalgamation of three cutting edge technologies i.e. DNA amplification, image processing and decoding, which are combined for the first time to get promising unique identification of individuals using Bioinformatics analyses. In this innovative technique PCR amplification using specific STRs ensures the DNA based identification, image processing of the gel documentation system increases reliability of results by reading pixel one by one and barcoding locks the data integrity that provides ultimate and unique molecular signature. It may be successfully used against suicide bombers, target killers, etc., as even a single blood spot, a few skin cells, root of the hair etc., to identify such culprits. It may also be effectively used to relieve the individual s with false accusations.

Acknowledgement

We want to acknowledge Molecular Biochemistry Lab, University of Agriculture Faisalabad Pakistan for providing experimental facilities. The authors (SM & ZA) would like to express their sincere appreciation to the Deanship of Scientific at King Saud University for its funding of this research through the Research Group Project no. RGP-1435-012.

References:

Azari, S., Ahmadi, N., Tehrani, M.J., Shookri, F., 2007. Profiling and authentication of human cell lines using short tandem repeat (STR) loci: report from the National Cell Bank of Iran. *Biologicals* 35, 195–202.

- Ballantyne, K.N., van Oorschot, R.H., Mitchell, R.J., 2011. Increased amplification success from forensic samples with locked nucleic acids. *Forensic Sci. Int. Genet.* 5, 276–280.
- Budowle, B., Chakraborty, R., 2001. Population variation at the CODIS core short tandem repeat loci in Europeans. *Leg. Med.* 3, 29–33.
- Brooks, E., Comber, B., McNaught, I., Robertson, J., 2011. Digital imaging and image analysis applied to numerical applications in forensic hair examination. *Sci. Justice* 51, 28–37.
- Budowle, B., Moretti, T.R., Baumstark, L., Defenbaugh, D., Keys, K.M., 1999. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. *J. Forensic Sci.* 44, 1277–1286.
- Budowle, B., Moretti, T.R., Niezgoda, S.J., Brown, B.L., 1998. CODIS and PCR-based short tandem repeat loci: law enforcement tools. In: *Proceedings of the Second European Symposium on Human Identification*; Innsbruck, Austria June 1998. Promega Corporation Madison, WI, pp. 73–88.
- Collins, P.J., Hennessy, L.K., Leibel, C.S., Roby, R.K., 2004. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR identifier PCR amplification Kit. *J. Forensic Sci.* 49 (6), 1265–1277.
- Kupfer, B., Vehreschild, J., Cornely, O., Kaiser, R., Plum, G., Viazov, S., Franzen, C., Tillmann, R.L., Simon, A., Muller, A., Schildgen, O., 2006. Severe pneumonia and human bocavirus in an adult. *Emerg. Infect. Dis.* 12, 1614–1616.
- Ip, S.C.Y., Lin, S.W., Li, C., Lai, K.M., 2014. Forensic DNA typing strategy of degraded DNA on discarded cigarette ends using the AmpFISTR® Identifier®, Identifier® Plus and MiniFiler™ PCR amplification kits. *Sci. Justice* 54, 311–315.
- Liu, P., Yeung, S.H.L., Crenshaw, K., Crouse, C., Scherer, J.R., Mathies, R., 2008. Real-time forensic DNA analysis at a crime scene using a portable microchip analyzer. *Forensic Sci. Int. Genet.* 2, 301–309.
- Maramis, C., Delopoulos, A., 2010. Efficient quantitative information extraction from PCR-RFLP gel electrophoresis images. In: *20th Int. Conf. Pattern Recognit.*, pp. 2560–2563.
- Moretti, T.R., Baumstark, L., Defenbaugh, D., Keys, K.M., Smerick, J.B., Budowle, B., 2001. Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J. Forensic Sci.* 6, 647–660.
- Rodriguez, A., Fernandez-Lozano, C., Dorado, J., Rabuñal, J.R., 2014. Two-dimensional gel electrophoresis image registration using block-matching techniques and deformation models. *Anal. Biochem.* 454, 53–59.
- Smith, C., 2008. Human microchip implantation. *J. Technol. Manag. Innov.* 3, 151–160.
- Von Wurmb-Schwark, N., Preusse-Prange, A., Heinrich, A., Simeoni, E., Bosch, T., Schwark, T., 2009. A new multiplex-PCR comprising autosomal and y-specific STRs and mitochondrial DNA to analyze highly degraded material. *Forensic Sci. Int. Genet.* 3, 96–103.
- Walsh, S.J., 2005. Legal perceptions of forensic DNA profiling part I: a review of the legal literature. *Forensic Sci. Int.* 155, 51–60.
- Wang, D.Y., Chang, C.W., Hennessy, L.K., 2009. Rapid STR analysis of single source DNA samples in 2h. *Forensic Sci. Int. Genet. Suppl. Ser.* 2, 115–116.
- Ye, N., Zhang, L., Gu, X., 2011. Discrimination of green teas from different geographical origins by using HS-SPME/GC-MS and pattern recognition methods. *Food Anal. Methods* 5, 856–860.
- Zacharia, E., Kostopoulou, E., Maroulis, D., Kossida, S., 2010. A spot segmentation approach for 2D gel electrophoresis images based on 2D histograms. In: *20th Int. Conf. Pattern. Recognit.*, pp. 2540–2543.
- Zhai, H.L., Hu, F.D., Huang, X.Y., Chen, J.H., 2010. The application of digital image recognition to the analysis of two-dimensional fingerprints. *Anal. Chim. Acta.* 657, 131–135.