Genetic Identification: A Review on Autosomal Single Nucleotide Polymorphism’s as Diagnostic Tool for Identifying Human

Fahmida Kahtoon1*, Ruba Mustafa Elsaid Ahmed2, Naveed Iqbal3, Zahid Balouch4 and Fahad Saleh Alenazi5

1Department of Biochemistry, College of medicine, Hail University, Saudi Arabia.
2Department of Pathology, University of Hail, Saudi Arabia.
3Department of Obstetrics and Gynecology, College of medicine, Hail University, Saudi Arabia.
4United Medical College, Jinnah University, Pakistan.
5Department of Pathology, College of Medicine, Hail University, Saudi Arabia.

Authors’ contributions

This work was carried out in collaboration among all authors. Authors FK and FSA designed the study, wrote the protocol and wrote the first draft of the manuscript. Authors NI, RMEA and ZB contributed in literature search and finalized the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2020/v32i4631099
Editor(s):
(1) Dr. Mohamed Fathy, Assiut University, Egypt.
Reviewers:
(1) Sabyasachi Mukherjee, ICAR-National Dairy Research Institute, India.
(2) Suhad Abdul Azeez Ibrahim, Al-Nahrain University, Iraq.
Complete Peer review History: http://www.sdiarticle4.com/review-history/64820

Received 14 November 2020
Accepted 19 January 2021
Published 03 February 2021

ABSTRACT

Single Nucleotide polymorphisms are biological markers, helping researchers to locate genes that are associated with various diseases. When SNPs occur within a gene or in a regulatory region near a gene, they may play a more direct role in disease by affecting the gene’s function. Most SNPs have no effect on health or development. Advancement in the field of genetics has resulted in the application of several techniques of molecular genetics in Pharmacogenomics. Nucleotide Polymorphisms (SNPs) holds the key in defining the risk of an individual’s susceptibility to various illnesses and response to drugs. The body of human beings is composed of DNA which is a chemical molecule responsible for imparting phenotypic and genotypic characteristics to the individuals. The most recent advancement of molecular genetics, which has found application in

*Corresponding author: E-mail: drfahmida24@gmail.com, f.khatoon@uoh.edu.sa
forensic science, is the use of autosomal SNPs because they can provide information about the ancestral genetics of human beings. The primary aim of this research is to explore the significance of autosomal SNPs in forensic science through the identification of humans at a crime scene. A secondary qualitative research design has been selected for conducting this study. This secondary research is based on a systematic review of the studies which have provided an insight in the significance of autosomal SNPs in forensic sciences by using various Data search Engine. SNPs can be used in the forensic investigation for the identification of individuals present at the crime scene.

Keywords: Single nucleotide polymorphism (SNPs); deoxyribonucleic acid (DNA); short tandem repeats (STRs); biogeographic ancestry; signal joint T cell receptor excision circles (sjTRECs).

ABBREVIATIONS

SNPs: Short Nucleotide Polymorphism; DNA: Deoxyribonucleic acid; STRs: Short tandem; sjTRECs: signal joint T cell receptor excision circles; GWAS: Genome Wide Association Study.

1. INTRODUCTION

Recent advancements in the field of molecular genetics have contributed to the development of techniques that have provided benefits to various other fields. Molecular genetics deals with the analysis of the genome of organisms at molecular level. There are numerous genes that constitute a genome of the humans. There are approximately 3 billion base pairs in the human genome [1]. These genes play a crucial role in the development of the DNA which is the core element for the formation of human body [2]. Each cell of the body of an individual contains the DNA which is a chemical code. There is a similarity in nearly 99.9% of the sequences of DNA of humans in every individual. Therefore, the remaining 0.1% of the DNA sequences that are different forms the basis of the differentiation of an individual from the other [3]. This variation enables the identification of a person and imparts the different characters to the human beings that vary from person to person.

Sequencing of the human gene has been performed over the past few years which has enabled the scientists to analyze the sequence of genes in the DNA. It has served as the foundation of research in the modern biomedical research. The technology of DNA sequencing has enabled the amelioration of this process leading to the application of this technique in other fields for the betterment of mankind [4]. It not only enables to determine the sequence of the genes in the DNA but also facilitates the identification of any mutations and variations in the gene sequences [5]. Forensic science is the field that involves the investigation of a crime through the application of various methods and techniques of science and technology. The significance of molecular genetic techniques in forensic cannot be denied as it has been benefitted enormously through molecular genetics. The most recent advancement of molecular genetics which has found application in forensic science is the use of autosomal SNPs as they can provide information about the ancestral genetics of human beings. This has been perceived to serve as a means for the identification of humans and to explore the accomplishment of autosomal SNPs towards the identification of human beings at crime scene.

2. AUTOSOMAL SINGLE NUCLEOTIDE POLYMORPHISMS

A Single nucleotide Polymorphism is a change in 1 nucleotide or base-pair within a codon in the DNA. Autosomal single nucleotide polymorphisms are the most common type of the genetic polymorphisms which have alleles that are associated with particular physical characteristics and specific populations. Polymorphic genetic markers of the proteins have been used during the past twenty five years for performing the differentiation of the individuals. There were several factors that limited the utilization of protein based genetic systems There is an ongoing process of identifying the common, biologically relevant SNPs, in particular those that are associated with the risk of disease. The identification and characterization of large numbers of these SNPs are necessary before we can begin to use them extensively as genetic tools [6]. This difficulty has been overcome by the application of techniques which can amplify the DNA available in reduced
quantity or through concentrating on the mitochondrial genome coding region [7]. The most suitable and widely used method for the analysis of the DNA from the degraded sample is the use of Autosomal SNPs [8].

3. IDENTIFICATION OF VICTIMS

It has been expected that the application of identification of biogeographic ancestry of an individual based on the DNA dependent processes can be useful for the identification of victims who are unrecognizable due to brutal murdering and various other reasons. This will also find application in the identification of the victims of terrorist attacks, as it is difficult to identify the victims of such cases. The period of the applicability of biogeographical ancestry of the individuals in the forensic investigation is difficult to predict. It could be determined at the point of detection of the geographic substructure through the utilization of sets of DNA markers. The knowledge present at the current level provides information about the large geographic regions including continents, although differentiation at the level of sub-regions can also be performed. The restriction of the genetic diversity is not followed by the political borders [9]. It is not recommended to apply the ancestry based on genetic testing to a diverse population since the traits responsible for the appearance of a person are not restricted to a specific geographical region and are widely distributed throughout the various regions of the world. Therefore, this can be applied or deriving a general idea about the appearance of a person but with particular considerations regarding the applicability of these processes to the cases in which the individuals ancestry belongs to only a single continental region. There is no information to date available regarding the extent of presentation of the traits of the appearance of an individual that determines the combination of the genes from different continents and regions [9].

4. APPEARANCE OF A PERSON

Considerable amount of information is not present regarding the extent to which the genetic factors determine the appearance of an individual. An insight into the genes which are responsible for the provision of traits to provide information regarding the characteristics such as color of skin, color of hair, traits related to pigmentation such as stature, morphology of hair, and freckles have been recently provided by the genome-wide association studies. The characteristics involved in the incidence of various diseases that are genetically determined such as baldness in male, autosomal dominant woolly hair, and non-syndromic cleft lip can provide information about the genetic fields that are responsible for the variations in appearance of the individuals despite the fact that these traits are not significant for application in forensic investigation. The prediction of a phenotype from a genotype is determined by the extent of genetic effect of SNP on a phenotype and the number of the SNPs that are contributing independently along with the influences that are not related to the genetics. Eye color is the most accurately determined trait from all the traits that have been studied so far [10].

5. PREDICTION OF EYE COLOR

Another limitation with the applicability of DNA for the prediction of various traits that particularly includes the color of eye is dependent on the conceptual understanding of the individuals. For example, there could be assigning of same eye to the categories of different eye color. This could mislead the investigators during the investigation as they could begin searching for the wrong person based on this misleading categorization of the eye color. In order to overcome this issue a study has investigated the quantitative variations in the color of eye that are based on genetics. This study has observed that the there are three eye color genes which are new and the quantitative variation in about 50% of the genes can also be demonstrated through the use data provide by the SNPs [11]. The development of color charts from these data can be of significant value for avoiding the uncertainties that are caused in the prediction of eye color [12].

6. PREDICTION OF HAIR COLOR

Single Nucleotide polymorphism have also been used for the prediction of hair color with particular emphasis on the prediction of red hair color. This hair color is strongly determined by a single gene which provides significant information regarding red hair color when determined through SNPs. The studies that investigated the application of these types of the predictions in forensics have been performed several years ago. In the investigation of 45 SNPs which were obtained from 12 genes has observed an accuracy of 0.87, 0.93, 0.81, and 0.82 for the prediction of black, red, blond, and brown hairs respectively. It has also enabled to distinguish between the categories of the same hair colors. The
dependency of age related to the hair color change might be responsible for the lowest accuracy of prediction of the SNPs for blond hair color [13]. The complete variation in the melanin of hair has been observed to be governed by three SNPs. These SNPs were responsible for 76% of these variations. Recently, two of these SNPs have been confirmed to be highly predictive of the hair color. The third gene, among these observed genes, has been recently examined to be related with the biogeographical ancestry instead to be responsible for the prediction of hair color [13].

7. PREDICTION OF SKIN COLOR

There is limited availability of the GWA studies for the investigation of the color of human skin. Several genes have been highlighted by these studies but the information related to the determination of skin color based on genes is limited so far. It has been demonstrated by a study performed recently that three SNPs obtained from three genes which are responsible for pigmentation, are accountable for 46% of the variations that are related to color of skin. There is a significant difference in the level of variation caused by the SNPs for the hair color and color of eye with that of the skin color [14]. This is related with a lowered level of variation of the color of skin in the same individuals. It is not possible to recover the variations of full traits by utilizing the GWA studies for the mapping of genes responsible for imparting skin color to the individuals. This limitation can be determined through GWA studies which utilize the samples from the entire world including various continents to predict the variations in the color of skin based on the differences in the SNPs with regards to different continents [15]. Another study has observed 82% variation in the color of skin by identifying five SNPs responsible for the variation in skin color among the population of the world. It has been considered that these markers would be useful for the prediction of skin colors of various individuals [16].

8. PREDICTION OF HEIGHT AND AGE

Height of the human body has also been predicted through the application of SNPs. However, the accuracy of prediction from these SNPs is significantly low which has limited the application of most of the SNPs that are responsible for the prediction of height in forensics. In a recent GWA study that consisted of 180,000 people provided results which identified 180 loci that are responsible for the height of an individual. These studies have provided several SNPs which can be utilized in future studies for identifying their predictive value [17].

Age of an individual can also be predicted through DNA testing. The shrinkage of telomeres and the deletion of mitochondrial DNA related with the process of ageing have found applicability in forensics for the prediction of age of an individual. However, there are certain limitations in their value of practicality. The predictive markers for age related identification of the individuals can be provided through the GWA studies which examine the gene expression related to ageing and patterns of DNA methylation [18].

9. RESEARCH DESIGN

Secondary qualitative research design has been selected by the researcher for conducting this study. This particular research design has been selected as it is appropriate for enabling the researcher to address the aims and objectives of the study. A systematic review of the literature was performed as it provides an analysis of the studies which have been conducted previously on the topic which is being studied. The researcher has been able to collect the findings of several studies and infer them to develop a comprehensive synopsis of the topic of Barbour, published in 2010 [19]. An analysis of the recent literature has been performed for gathering the information from the recent literature by systematically reviewing the studies [20]. Generation of several themes has occurred as a consequence of analysis of the data provided through review of the literature [21]. This secondary research is based on a systematic review of the studies from 2020 to 2004 which have provided an insight in the significance of Autosomal SNPs in forensic sciences through the identification of humans at crime scenes [20].

A logical search of the relevant terms is performed through the use of Boolean operators as shown in Table 1. It also results in the exclusion of several items which are irrelevant. Boolean operators are used for eliminating, joining, or expanding the keywords during a search. A combination of diverse key words and the use of capitalize “AND” to restrict the search and “OR” to expand the search when combined is utilized for the search. In order to refine the search, subject heading with map terms were
used for the identification of terms from the database thesaurus. Boolean operators were utilized in the search that enabled the researcher to perform the search in accordance with the study topic and narrowing the search. It also saved the time to search the useful research materials and enabled the exclusion of irrelevant items. A Systematic review research design was not selected because it allows summarizing the findings of numerous studies. It is aimed at provision of effectiveness of a specific intervention or service. It enables the researcher to select the studies that are relevant to the topic being studied. Assimilation of huge amounts of data can also be performed through a systematic review. It reduces the possibility of biasness since it does not include personal opinions of the reviewer. Thus, the conclusions derived from a systematic review are reliable [22].

### 10. RESULTS

Single Nucleotide Polymorphism (SNPs) have been utilized in cases where the analysis of STRs failed resulted in failure of identification of the humans at the crime scene. In a study by Sun et al in 2012, observed six cases of forensics in which the identification failed through the application of STRs utilized SNPs for analysis [23]. The failure of identification was due to the failure of analysis for STRs in the sample. The analysis of the sample was possible through the application of SNPs by using the SNaPshot technique. The samples included in the study were extracted according to the standard protocol [24]. Y-SNPs can also be used in forensic investigation for the identification of humans at crime scene. It has been demonstrated that phylogenetic tree of Y-chromosomes can be applied for in forensic which requires to be extensively enhanced. As a consequence of presentation of numerous phylogenetic trees of Y-chromosomes by the researchers, it has become difficult to perform accurate reporting of the results of investigation as it is impossible to relate the results with a phylogenetic tree of Y-chromosomes, which has updated information [25]. Therefore, it has become difficult for the forensic investigators to interpret the results of investigation based on recent nomenclature. Optimization of the currently available haploid markers of the phylogenetic tree is possible through the data which has been provided by the WGS studies. Only a limited genome set of the Y-chromosome haplogroups is available. Thus, the newly developed tree can be used for the identification of Y-chromosomal trees based on STPs for ensuring their application in forensic investigation [25]. Considering the increased requirement of application of SNPs in forensic sciences for identifying the humans at a crime scene, it is important to develop the techniques and systems which are efficient for providing results that are free from error. A system for the efficient analysis of Y-SNPs has been developed recently. This system was developed by obtaining information from blood sample of six individuals who are not related to each other. Multiplex system of Y-SNPs analysis was designed which ensured effective and efficient analysis of the samples which were artificially degraded. This suggested that such a system can found application in forensic investigation for performing the analysis of SNPs so that it becomes easier to identify a person at the crime scene. Samples that contain mixtures of the DNA can also be analyzed through this system [26].

Phenotypic characteristics of the individuals are being predicted through the application of SNPs in forensic investigations. These are valuable markers for the prediction of color of eye, pigmentation of skin, biogeographic origin, and color of hair. However, the implementation of SNPs for investigation purposes in crime scene human identification is still controversial as there are considerations regarding the legal and ethical implication of SNPs testing in forensic investigation for identifying humans at crime scene. There are also several limitations that have been observed regarding the application of SNPs for identification of humans at crime scene. These limitations reside in the limited availability of the technologies that can be used for the typing of SNPs and increasing the ability of systems that are used for typing of SNPs to analyze the biological samples which are highly degraded. Bioethical issues remain a major
hindrance in the applicability of SNPs analysis in forensic genetics. Knowledge and ability of the forensic investigators and researchers is essentially required for this purpose to overcome these hurdles in the excessive application of SNPs for identifying humans at crime scene [27]. Prediction of eye color of the individuals is also possible through the analysis of SNPs which can serve as a beneficial tool for the identification of criminals and victims at the crime scene. SNPs that were responsible for imparting the color of skin to the individuals were also used. The applicability of these SNPs for the prediction of color of skin and eye were observed to be significant with very low rate of error. This suggested that these SNPs can be applied in forensic analysis for identifying the individuals at crime scene through the prediction of their physical characteristics [11].

11. DISCUSSION

There are several studies which have demonstrated the applicability of SNPs in forensic investigations for identifying the humans at crime scene [28]. Several effective and efficient technological systems have been developed by the researchers which are capable of performing analysis of biological samples containing degraded DNA because SNPs can be obtained from these samples. Physical characteristics of the individuals can be predicted through the analysis of SNPs. This can provide significant information about the color of eye, hair and skin of the individuals involved in crime.

There are certain challenges that are presented by autosomal SNPs which results in their resistance of the application in the field of forensics. SNPs do not provide significant information for the analysis in cases where they are collected from the samples that are present in the form of mixture from samples of various individuals. This limitation can be compensated by the use of SNPs that are higher in number and relative to the STRs. It can also be overcome by the utilization of tri-allelic SNPs in combination with the technologies of multiplex genotyping [29]. The identification of male individuals can also be performed through the utilization of Y-SNPs. A phylogenetic tree has been developed for this purpose which has also founded implication in forensic for identifying the criminals and victims through the analysis of the samples which have been obtained from the crime scene. The Y-SNPs could be useful for the identification of male criminals [30]. They can be used for the identification of paternal ancestry which can provide significant information to the forensic investigators regarding the paternal lineage of the criminal or the victim present at the crime scene. This would also be beneficial to reach the criminal by identifying the paternal ancestry of the suspect [31]. SNPs are being used extensively for the determination of biogeographical ancestry of the individuals. Recently, it has also found application in the identification of the humans present at crime scene by predicting their phenotypic characteristics and ancestral lineage. Several proficient systems have been designed for performing the analysis of SNPs. This has aided in the identification of criminals and victims at the crime scene.

12. CONCLUSION

At present implication of SNPs are being explored in various aspects of forensic sciences for performing the identification of the humans at the crime scene. SNPs have proven to be efficient for predicting the color of eye and hair of the individuals along with the color of skin [32]. This can provide enormous information about the criminal or victim present at the crime scene. This information can be used for identifying the person.

Despite the undeniable significance of SNPs in identification of humans, there are certain limitations which serve as barrier in the implication of SNPs in forensics [33]. These challenges are present in the ethical considerations related with the application of SNPs and the unavailability of the databases with information about them [27]. However, these limitations can be overcome and SNPs can be utilized for the identification of humans at crime scene.

13. RECOMMENDATIONS

Further studies are required to explore the significance of SNPs in the identification of humans at crime scene. These studies should be based on control groups to validate the findings. These studies should focus on the prediction of phenotypic characteristics of the individuals for determining the humans in forensic investigation. However, based on the results of review study, it can be recommended that SNPs should be employed in forensic genetics for identifying the individuals,
either victim or criminal, present at the crime scene.

**CONSENT**

It’s not applicable.

**ETHICAL APPROVAL**

It’s not applicable.

**COMPETING INTERESTS**

Authors have declared that no competing interests exist.

**REFERENCES**


Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/64820

© 2020 Kahtoon et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.