BRJ BASIC RESEARCH JOURNALS

Basic Research Journal of Medicine and Clinical Sciences ISSN 2315-6864 Vol. 6(10) pp. 99-103 December 2017
Available online http//www.basicresearchjournals.org

Review

Genome wide association studies; discovering human genome and genetics

Dr. Fahmida Khatoon[1],

Assistant Professor, Faculty of medicine, University Of hail, KSA

Author's email: drfahmida1@gmail.com

Accepted 24 December, 2017

ABSTRACT

Aim: The aim of this study is to explore the application of Genome Wide Association Studies in accordance with Biochemical investigation of human population. Introduction: The achievements in Human Genome Project and subsequent advancements observed in Genotyping techniques have led to an influx of exciting new era of discovering human genome and genetics. These technologies have provided the scientists with a comprehensive data on human genomes as human genome is now capable to incite in depth and precise data information and allows access into processes that extract detailed sequences for genomics in order to analyze theoretical questions but also identifies the practically feasible genetic characteristics. Methodology: The methods of this research for this particular study is to demonstrate a review of relevant literature to examine application of Genome Wide Association Studies in accordance with population based forensic investigations. To conduct a literature review a qualitative research design is the most suitable research design. It provides the rationale for assessing the human behavior and assists to legalize and authenticate the data which is selectively collected from the secondary sources. Conclusion: Genome Wide Association Studies associated approaches are been incorporated into validation remains a standard by which the acceptance of most of the hypothesis and techniques can be appraised.

Keywords: Human Genome Project, Genome Wide Association Studies, Genome

INTRODUCTION

The achievements in Human Genome Project and subsequent advancements observed in Genotyping techniques have led to an influx of exciting new era of discovering human genome and genetics. These technologies have provided the scientists with a comprehensive data on human genomes and have introduced state-of-art methods of investigating the fundamentals of forensic medicine (Moore, 2009). The last five years in the field of genomic studies have been an era of ensurgement of discoveries for epidemiological studies associated with the genetic and genomic technological advancements. High-Throughput genotyping and sequencing centers are containing torrents of data and information related to surveying genetic variants across the genome for association of complex diseases, human phenotypical characteristics and quantitative traits (Park et al., 2010). The Genome Wide Association Studies has been the work force for these efforts and now is considered as a standard technique for disease gene mapping.

Misidentification is an error that can be challenging to detect since it is a fundamental feature for many Genome Wide Association Studies that all the samples are unrelated to each other. Thus, the use of already known Mandelian relationships between pairs of individuals included in the study sample is for purpose to infer the tendency to make mistakes which is not exclusively possible to elicit successful outcomes. Instead, misidentification of samples can be detected by following comparisons:

a). Sex characteristics associated with X and Y chromosomes data for genome with not established identity of sex.

b). The inferred data for constructing ancestry details based on principle component analysis with self proclaimed ancestry information (Laurie et al., 2010; Anderson et al., 2010).

Most of the Genome Wide Association Studies are designed in ways so as to include the single individual from the same population in order to minimize errors and bias caused by population stratification.

The utilization of Genome Wide Association Studies for detecting many diseases of forensic interest and traits for sex distributions are due to sex associated variations in prevalence (Laurie et al., 2010). The sex and ancestry related data and genetic information are important components for Quantity Check (QC) in GWAS and can be implemented in practice to remove misidentifications of samples and stratifications in population samples. The approach to identify the misidentification in processes where strong associations are detected between genotype and phenotype characteristics can be inferred if the observed phenotype is expected to be caused by the observed genotype for every subject within the study (Hindorff et al., 2009). This observed phenotype if detected to be too extreme in accordance with the genotype then the person is flagged as a possible subject for misidentification (Laurie et al., 2010).

The study of genomic and associated genetic markers requires the following steps to be wisely followed with techniques, firstly, a set of information about genotype and phenotype relationship need to be established and identified (Hindorff et al., 2009). Secondly, phenotypes need to be ascertained at additional cost before the initiation of study, possibly for reasons of extracting usual phenotype data during the collection method. Thirdly, the modeling for the mixture sample that is under consideration for estimating phenotype and genotype relationship needs to be optimized to deliver highest set of information. Finally, the combination of all the information assessed from the relationships should be higher enough to give better sensitivity and specificity. It can also be possible to assess Single Nucleotide Polymorphisms (SNPs) in determining genotype for each sample, before the execution of costly Genome Associated Studies or Genome Sequencing into practice (Hindorff et al., 2009).

Aim and objectives of the study

1. To explore the phenomenon of Epitasis and its relevance with Genome Wide Association Studies Genomic data and privacy factors for public access of population genomic information and access to genomic datasets.
2. To evaluate the challenge of development of statistical techniques and implementing computational efficacy in evaluating genetic variance and their interactions with genomic expression in Logistic Regression (LR), Multifactorial Dimensionality Reduction (MDR), Random Forests (RF) and Evaporative Cooling (EC) Methods.

LITERATURE REVIEW

The section will open up with the review of literature of the published work related to the topic. It will also highlight the theories close to the subject of research along with the general review of the studies selected. The aim of this review is to study the application of Genome Wide Association Studies incorporated into Medicine.

Ancestral sampling in genome wide association studies

The alternative approach used to amplify genetic causes of complex traits is population-based Genome studies. These studies focus on the correlation between genetic markers and relevant traits among unrelated population samples (Cardon et al., 2001). The reasons for ancestral mutation of genes that occurred hundreds of years ago and of which lineages of all descendents carry the casual variants are thus considered susceptible for polygenic mutations. If this casual variant is assayed into Genome Wide studies' perspectives, persons carrying the derived allele will disperse it in different phenotypes as compared to those who carry ancestral alleles (Cardon et al., 2001).

Global studies Jakobsson et al. (2008) reveal that genome wide population structure on various levels can be assessed using statistical approaches. Individuals in different continents are differentiated in genetic framework through genome-wide SNP data but however, some level of over-lapping do exists between continental regions (Jakobsson et al., 2008). This genetic overlapping increases with decreasing distances between continental regions. For example, there is a strong relationship between geographic correlation and genetics similarity of an individual who resides in Europe, but there is a considerable overlap of genetic similarity among

neighboring European subpopulations which determines accurate determinations for genetic variability (Jakobsson et al., 2008).

Scientists can use data from this technique and search for genes that are responsible for predisposing a person at risk of developing disease or a trait important in terms of forensic interest (Lueders et al., 2003). Selective genotyping is an application used to identify different ethnic groups or a mix population and can be beneficial in identifying specific traits in a homogenous ancestry distribution. Similarly, by comparing relative allele frequencies between different phenotypic groups, Genome Wide Studies can detect SNPs that are associated specifically to a disease process or an ethnic trait. GWAS are preferred by many researchers because they provide more accurate localization of casual genes and also they provide unbiased detection of whole genome for genetic association (Lueders et al., 2003).

Phenotypical variance incorporated into genomic studies and SNPs

There are numerous forensic cases in which DNA-based inferences of biogeographic ancestry information is vital to suggest police investigations to find unknown individuals or unidentified victims (Mangold et al., 2010). However, it is important to acknowledge that when and how biogeographic ancestry is applied in answering forensic queries and where DNA testing techniques can be introduced to determine the level of ethnic disparity using useful DNA marker sets in association studies. It is evident to keep in mind that the genetic diversity allows the involvements of various appearance traits likely to be error prone as no phenotype is restricted to a certain geographic region. In similar terms, appearances of the unidentified individuals can be known by utilizing markers from genes incorporated into data sets for genomic studies that are functionally active and strongly associated with person's appearances and skin color (Mangold et al., 2010).

The genetic effect of Single Nucleotide Polymorphisms (SNPs) on the phenotype of an individual depend on the number of contributing SNPs and non-genetic influences like environmental effects in determining accurately an identity of a person. Phenotype studies illustrate that eye color is the most successfully predictable phenotype essential in accurately identifying basic physical factors in an individual (Mangold et al., 2010). The challenges undertaken in applying DNA prediction into categorizing appearance traits and eye color is its expected variability in conceptual understanding of trait information. For example, people assign same eye color to various color categories and therefore look different from others using an eye color provided by DNA predictions. In order to minimize this problem, studies investigated about the genetic basis of variation in eye color utilizing SNP data analysis (Mangold et al., 2010).

Two of the SNPs are believed to be highly specific to hair color while the third relates to reflect biogeographic ancestry rather than emphasizing on the hair color (Mangold et al., 2010). Another trait that expresses a physically visible characteristic can be useful in successfully predicting appearance of an individual specially age. Two DNA- based approaches for age predictions are based on mt DNA deletions and age-dependent telomere shrinkage have been suggested for forensic assistance but further research is needed to diversify their effects (Mangold et al., 2010). Genome wide studies, on age related features for analyzing DNA methylation pattern may provide extensive benefits for establishing suitable age-predictive biomarkers (Mangold et al., 2010). By targeting the Genome data that is relevant and informative can be obtained and is generally applicable. This is because of the fact that the specific information contained within the Genome is generally correlated with all the normal cells in the human body as all cells contain similar DNA sequences, same mutations and polymorphisms.

Genetic ancestry and association study analysis

The genomic studies has also been carried within the European population for achieving better inferences in observing a common group of individuals for genomic variance in order to reduce the expenses and efforts for collecting samples and genotyping analysis (Rosenberg et al., 2010). Another reason for European population based studies was the availability of large and homogeneous sample for populations, such as in cohorts for Finland and Sardina derived population samples (Rosenberg et al., 2010). These cohorts are believed to contain extensive collaborations and the prior genetic and phenotypical information to help in analysis and interpretation of Genome Wide Association Studies (Rosenberg et al., 2010).

Structured association and genetic markers in GWAS

The technique of Structured Association works on the principle of Hardy and Weinberg's Equilibrium within populations with linkages to genetic markers that are also responsible for establishing equilibrium to detect and correct stratification in a homogenous sub group cluster based on genotypic characteristics. Test for detecting familial or hereditary pathological traits in the generations can be achieved by studying association of each genetic marker to disease expression phenotype (Alexander, 2009).

METHODOLOGY

For establishing any search of literature, it is important to comprehend any research and its role in informing the clinical practice as well as questioning (Parahoo, 2006). This systemic process was defined by the researcher as utilizing the definite methods for cracking the problem and answering the research question (Parahoo, 2006). Parahoo suggest and argues that the findings of research were not the problem solutions however, the study was more likely offering the current information which assist in the decision making (Parahoo, 2006). Therefore, the eventual purpose of the study is to expand, develop and refine the knowledge of the body (Parahoo 2006).

Qualitative design

The main argument for this particular research study is to carry out a review to examine the molecular genetics and the application of genome and DNA-amplification into forensic investigations. To conduct a literature review a qualitative research design is the most suitable research design. The qualitative research design is the most appropriate research design to examine the human behavior and it assists to legalize and authenticate the data which is selectively collected from the secondary sources. It helps in refining the research and additionally it adds a primary hand worth to it. For many of the researchers the secondary data is crucial in the health care and medicine field as it suggest the previous researches. Creswell (2009), stated that a qualitative research comprise of using exclusive steps of strategies and analysis of inquiry with researchers interpreting what they hear, see, and understand. The method used in this study is qualitative. As compared to quantitative research, Qualitative research is more subjective, and it is based on unlikely methods of gathering information. This research is more or less based on the literature review and the conclusions are drawn on the basis of actual resources (mainly primary researches).

Critical analysis

The critical analysis is central process of any research study. It involves critical thinking which applies logical and rational thinking while deconstructing the text. It is a complex, intellectual activity involving analysis and critical comment on the material formerly gathered. Browne and Keeley (2001) defined the critical thinking as a set of awareness which interrelates with the research question under study which is expected to be critically analyzed. It is an ability to answer and ask the critical question at a specific time and it desires to use the critical questions actively). Consequently, it is not just a descriptive list of the set of accessible summary and material. Without a critical appraisal and systemic literature any academic study cannot be applied to a methodology or in any way put in to the knowledge (Hart, 2005). Critical appraisal of the literature also offers a room for the comparison of the results of diverse researches.

Search strategy

For research evaluation on the molecular DNA and forensics, the inclusive review of most of the studies available was important. The databases used and accessed for this purpose included: ProQuest, PubMed, Cochrane Library, Science Direct and CINAHL.

Inclusion criteria

The search of literature only found out the material published in English and for this reason it created no issues to reject any possible study for the reason of the translating complexities of the papers.

Exclusion Criteria

Exclusion criteria included the articles not published as full manuscripts or not in the peer review literature. The studies including controversial ethical or legislative materials are not included. The studies which were conference papers or studies in progress or government reports will be excluded. The searches will be restricted to retrieve the literature available in English.

Limitation

In this review there was a likelihood of publication bias. By rejecting those studies with negative outcomes and were unpublished it was probable to overrate the effects of techniques used, nevertheless, for the published literature the comprehensive search for potentially appropriate articles was undertaken by means of a strategy for systemic review for the purpose of avoiding bias.

Ethical considerations

This study lies primarily on the original and primary sources as it is known that the unethical, fraudulent and dishonest researchers can circumvent the scientific method. It is important to properly cite the person whose work is being used. This is because the person who has conducted an actual research may have served a lot of

effort and time to extract the outcomes of the study and hence it is one's responsibility to give credit legally to the individual who has conducted that particular work.

ANALYSIS AND DISCUSSION

This section will include the thematic analysis and critique of the selected articles. This chapter will also discuss the selected researches in detail and their findings, based on the findings of the reviewed literature.

CONCLUSION

Over the span of the last 10 years, the fields of genomics have under taken interesting and dramatic shifts on how to comprehend forensic detailing with respect to investigations about unknown pedigree, unidentified complex diseases and ethnogeographic ancestry of subjects (Scott, 2007). The wealth of information which can be extracted from immense map of human genome along with advancements in genotyping and related technologies has presented many opportunities for the scientists to test assumptions about the complex genomics for human. This has led a new era of genetic investigation like Genomic Association studies that incorporates to solve analytical questions about heritage and population based studies (Scott, 2007).

In 2005, Genome Wide Association Studies has evoked several other successful studies searching for answers for genetic variants related to complexed hereditary information (Scott, 2007; Hunter, 2007). However, other studies also illustrated that results cannot be replicated in other independent sampling and these scenarios demand deeper investigations in order to cater the replication failures. The popularity and the utilization of Genome Wide Association Studies have increased over the years, along with the challenges like identification of stratifications in a population has been made feasible for forensic investigations.

The road ahead in genomic association studies for research and practice

Looking towards the future, it seems evident that the current enhancements in Genome Wide Association Studies is going to transform through changes for finding genetic susceptibilities that are responsible for causing characteristics unique to a familial trait (Bird, 2005). Durbin (2010) provided a series of projects in the '1000 Genomic Project' that has provided an avenue for testing assumptions by utilizing sequencing techniques to identify rare variations in human genome. The use of sequential data for detecting rare genes for variant determination has also been observed possible through association analysis.

REFERENCES

Al Safar HS, Abidi FH, Khazanehdari KA, Dadour IR, Tay GK (2011). Evaluation of different sources of DNA for use in genome wide studies and forensic application. Applied Microbiology & Biotechnology, 89(3), 807-815. doi:10.1007/s00253-010-2926-3

Alexander DH, Novembre J, Lange K (2009). Fast Model-Based Estimation of Ancestry in Unrelated Individuals. Genome Res. 19 (9) 1655-1664.

Anderson C, Pettersson F, Clarke G, Cardon L, Morris A, Zondervan K (2010). "Data quality control in genetic case-control association studies," Nature Protocols, 5, 1564–1573.

Bird TD (2005). Genetic Factors in Alzheimer's Disease. N.Engl.J.Med. 352 (9) 862-864.

Braun R, Rowe W, Schaefer C, Zhang J. Buetow K (2009). Needles in the haystack: identifying individuals present in pooled genomic data. PLoS Genet 5(9): e1000668. doi:10.1371/journal.pgen.1000668

Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, Gabriel SB, Topol EJ, Smoller JW, Pato CN, Pato MT, Petryshen TL, Kolonel LN, Lander ES, Sklar P, Henderson B, Hirschhorn JN, Altshuler D (2004). Assessing the Impact of Population Stratification on Genetic Association Studies. Nat.Genet. 36 (4) 388-393

Genome-Wide Association Scans. Genet.Epidemiol. 30 (4) 356-368.

Gibson G (2010). "Hints of hidden heritability in GWAS." Nature Genetics, 42,558–60.Park, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, Chatterjee N (2010): "Estimation of effect size distribution from genome-wide association studies and implications for future discoveries." Nature Genetics, 42.

Gill P, Puch-Solis R, Curran J (2009). The low template-DNA (stochastic) threshold – its determination relative to risk analysis for national DNA databases. Forensic Sci Int Genet 3:104–111.