Predicting Physical Features and Diseases by DNA Analysis: Current Advances and Future Challenges

Caio Cesar Silva de Cerqueira[1]*, Virginia Ramallo[2], Tábita Hünemeier[3], Soledad de Azevedo[2], Mirsha Quinto-Sanchéz[2], Carolina Paschetta[2], Celia Cintas[2], Marina González[2], Lavínia Schüler-Faccini[4,5], Maria Cátira Bortolini[4] and Rolando González-José[2]

[1]Scientific Police of Sao Paulo State, Ourinhos-SP, Brazil

[2]Patagonian Institute of Social and Human Sciences, National Patagonian Center, CONICET, Puerto Madryn, Chubut, Argentina

[3]Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of Sao Paulo (USP), Sao Paulo - SP, Brazil

[4]Department of Genetics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

[5]INAGEMP - National Institute of Medical Genetics Population, Porto Alegre, Brazil

*Corresponding author: Caio Cesar Silva de Cerqueira, Scientific Police of Sao Paulo State, Street Thirteen December, number 32, Centro, Ourinhos -SP, Brazil, Tel: +5514998995137; E-mail: c_aioc@yahoo.com.br

Abstract

The 'omics' era and its concomitant technological advances have brought great insight into genetics. One of the most promising fields within human genetics is the prediction of physical traits from analysis of genetic material. Besides the predictive potential of DNA, the traceability of pathogenic agents in the human body through molecular analysis is also a field to be further exploited. In this review, we aim to discuss specific aspects of phenotypic prediction by analysing DNA, with special emphasis on normal variation, and the application of a technology known as 'Forensic DNA Phenotyping' (FDP). We also suggest the term 'Phenotype Informative Markers' (PIMs) to designate any molecular markers responsible for normal or pathological human phenotypic variation. In addition, we raise some recommendations related to forensic genetics, the molecular diagnosis of human diseases, and the traceability of pathogens in the human body, giving special emphasis to the need for validation of these tests with strict protocols. Some relevant concerns about privacy, ethics, and legality of such predictions have also been discussed. Finally, we look at perspectives on the use of epigenetic tools, and quote some examples of what has been done in this specific field.

Keywords: Legal concerns on prediction; Phenotype informative markers; Forensic DNA phenotyping; Molecular diagnosis; Pathogen traceability

Introduction

Since the discovery of the DNA double helix [1], the development of innovative sequencing technologies [2-4] and different modalities of the PCR technique [5], among others, have led to substantial improvements in the available information detailing molecular markers for normal human phenotypic constitution/variation, as well as several specific markers that indicate predisposition to certain diseases. These advances have fuelled important areas of research such as forensic genetics and preventive medicine. The number of publications since 1953 (year that the DNA double helix was described by Watson and Crick [1]) containing the key-words "prediction" or "molecular diagnosis" is illustrated in Figure 1. The progressive increase in the research around these concepts clearly indicates the importance and growing interest in genetic variation as a tool to predict both normal and pathological phenotypes. It also denotes the important challenges and difficulties that any established and powerful prediction technique has to face.

Given the increasing importance of predictive methods for several areas of human genetics and anthropology, the purpose of this paper is to review the current cutting edge technologies. We do not intend to do a systematic compilation of all the possible phenotypic information to be predicted by DNA analysis, but rather we wish to give a broad perspective on these areas, as well as some of the most relevant studies on normal and pathological human phenotypes that can be derived from the analysis of DNA. This information is useful mainly for Forensic DNA Phenotyping [6] and molecular diagnosis of human diseases. Hereafter, we will use the term "Phenotype Informative Markers" (PIMs) to refer to genetic or molecular markers of the human genome related with normal or pathological traits.

Understanding the genotype-phenotype relationship

When studying the genotype-phenotype relationship, we must consider a wide view of the biological system being studied including the (usually) unknown particularities of the genetic-phenotypic points of view. To do so, a mandatory first step is detailed characterisation of the phenotype, including proper delimitation of the observable phenotypic states. This kind of approach has to be extensive, covering as many as possible phenotypes in order to uncover possible pleiotropic effects, and intensively, featuring each phenotype with the maximum amount of detail, as suggested by Houle et al. [7]. Some genotype-phenotype pathways are relatively easy to observe (for instance, characteristics considered to be monogenic), however the vast majority of human physical characteristics (normal or pathological) need a more comprehensive and detailed approach to be properly established. For practical purposes, continuous traits are usually categorized as discrete entities. However, it is important to note that many qualitative classifications used in the scientific literature are

just simplifications of continuous, multivariate traits, a phenomenon that has fuelled an intense debate in the field of systematic [8].
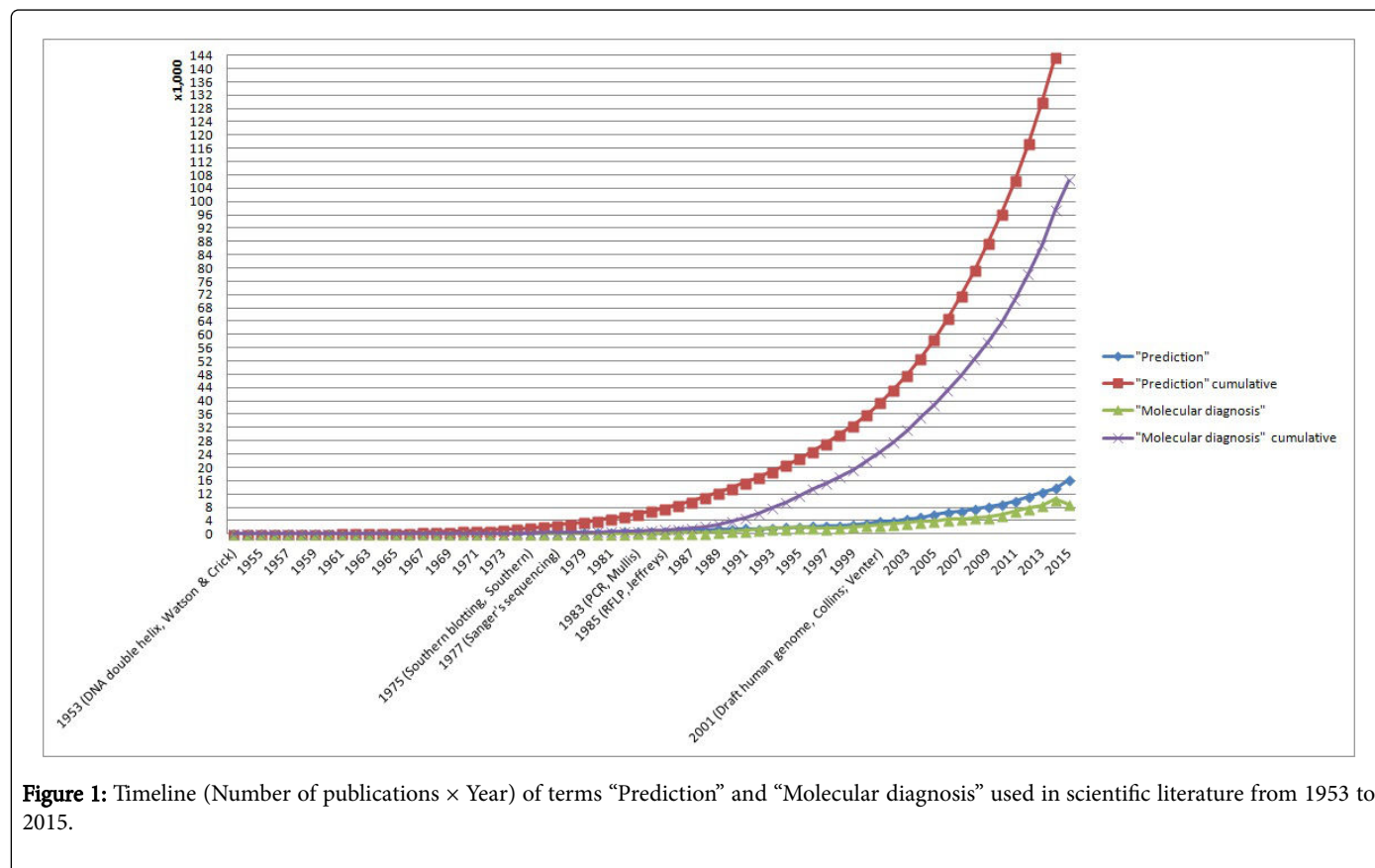

Figure 1: Timeline (Number of publications × Year) of terms "Prediction" and "Molecular diagnosis" used in scientific literature from 1953 to 2015.

All human phenotypes are determined by the interaction of genes with environmental and developmental factors, each having an influence to a variable degree. We can extend this classic definition [9] for both polygenic and multifactorial phenotypes, as well as for monogenic phenotypes. There is a small percentage of so-called monogenic traits that do not directly correspond to the "rules" as predicted by the genotype and/or characteristics that have certain "amplitude" of influence in the expression of phenotypes (incomplete penetrance, variable expressivity, etc.).

In polygenic and multifactorial traits, it is plausible to think that different genetic combinations can lead to the same phenotypic result. Notably, some of these genetic combinations are easier to detect than others are, in a given sample. Thus, a main goal in prediction studies of physical features from DNA is to find the specific PIM profile (that is, a specific combination of genotypes in different loci associated with the phenotype) related to a specific phenotype with the highest probability.

In this way, we can predict phenotypes as having a more frequent genetic profile, that is, one could describe PIM profiles that are found almost exclusively in specific phenotypes (Figure 2), which is, per se, a major breakthrough in human genetics. Nevertheless, we are still far from even knowing a tenth of the gene-gene, gene-RNA, gene-protein connections, and epistatic interactions that occur in our genome. In this context, any minimal amount of useful information to predict a given phenotype is important.

As an example, we define phenotype in the qualitative spectrum as having a hypothetical Gaussian distribution in the population. Each color in the column in this graphic represents the probability in predicting the respective phenotypes. For example, more red color in a column means that the specific PIM profile has greater probability of predicting dark brown/black eyes, and so on.

In this picture, blue represents the probability of blue/grey eyes, green represents the probability of green eyes, orange represents the probability of honey/light-brown eyes, and as mentioned, red represents the probability of dark brown/black eyes. The aim of studies regarding prediction of physical features through DNA is to find the specific PIM profile related with a specific phenotype with highest probability.

Notably, the definition of phenotype initially coined by Johannsen [10] "The phenotype of an individual is the sum total of all his expressed characters does not stop at skin surface" [11], which may even include some human feeding behaviours, for example. A relevant example is the tolerance for certain foods: some people like shrimps, but cannot eat them because they have a severe allergy [12,13] or people who may like milk or derivatives who exhibit a strong reaction to lactose [14,15] or even people who simply "dislike" certain foods due to greater sensitivity to bitter taste [16,17].

These are, perhaps, the more obvious examples of genetic influence underlying certain feeding behaviors. A deep understanding of the genotype-phenotype relationship is essential to discover the highest genome modulation of such complex characteristics, such as responses to foods, liquids, or drugs.
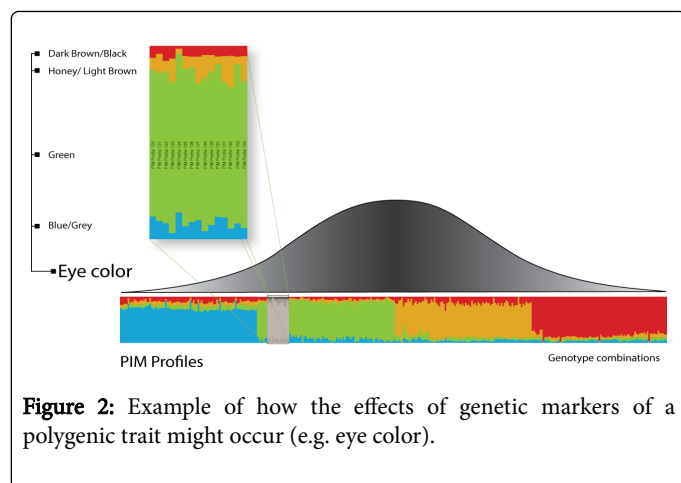
Figure 2: Example of how the effects of genetic markers of a polygenic trait might occur (e.g. eye color).

Despite the complexities in the genotype-phenotype map, it is now possible to predict a handful of traits with a reasonable degree of probability. In the following sections, we present some examples of normal and pathological human phenotypes that are being investigated, which are likely to be predicted in the near future with some accuracy through DNA analysis.

PIMs for prediction of normal human traits

As pointed out by Cho and Sankar [18] and Kayser [19], considerably less attention is given to research on genetic variants for non-medical purposes. However, FDP technology, as well as the characterisation of PIMs, is only possible thanks to the effort and intensification of investments by funding agencies in these types of research. Initiatives such as CANDELA (Consortium for the Analysis of the Diversity and Evolution in Latin America) [20,21] and GIANT (Genetic Investigation of ANthropocentric Traits) [22,23] are examples of basic and applied research in these areas.

Although FDP is still under development and discussion in the legal and administrative scope of most countries, specialists in the UK and Netherlands have already tried to incorporate some of these predictions into forensic routine [6,19,24], as an additional tool of police intelligence in the pursuit of criminal suspects. The main purpose of this tool is to reduce the number of potential suspects through the prediction of physical features from donors of biological samples left at the crime scene. Once the number of suspects are reduced, a smaller 'DNA dragnet' (collection of DNA samples from hundreds or even thousands of 'volunteers' for comparison with a genetic profile found at a crime scene) is performed [18], and a conventional STR (Short Tandem Repeat) profile can be generated to help in solving the crime.

The prediction of physical traits useful for FDP consists mainly of 'phenotype informative SNPs' typing [25]. It is worthy to say that with conventional STR typing it is possible to determine a unique phenotype: the sex of the investigated sample through the amelogenin locus. The X chromosome copy has a deletion of 6 bp in relation to the copy of the Y chromosome, which makes it possible to differentiate between a man (XY) and a woman (XX) [26,27]. This is done largely in forensic practices around the world, although it was known for a long time that this prediction test was not error-free [28] and more reliable forensic kits have therefore been developed in more recent years [29,30] to better confirm the gender of the biological sample.

The most promising phenotypes for FDP, in addition to sex, are traits related to pigmentation [31,32]. There are at least ten published studies presenting methods to explore the predictive values of combinations of SNP alleles of pigmentation phenotypes (skin, eye and hair colour) and presence/absence of freckles [33-42], two of which [34,40] have already been validated [43-45]. Additional research is under development regarding these and other phenotypes, such as estimation of age [46-49], height [50,51], hair shape [52-54], facial features [55,56], baldness [54,57,58], and adult stuttering [59,60]. Besides physical phenotypes, additional information of individuals can be obtained with some reliability by analysing DNA, such as biogeographical ancestry [61,62] and surname of the sample [6,19,63-67].

In the context of normal phenotypes, Koops and Schellekens [6] earlier suggested some recommendations aimed to enhance the implementation of FDP technologies in forensic cases. Here, we comment some of these recommendations.

1) Many of the characteristics that can be predicted by DNA are multifactorial, resulting from the interaction of genes with specific environmental factors. Thus, police intelligence professionals should be aware that no method of prediction is deterministic, but probabilistic. In other words, to say that the subject has a genetic profile for a certain trait is not the same as to say that the individual will have a specific phenotype determined with complete certainty; in which case, it is important to note how much of the individual variation may be determined by the specific genotype, taking into account environmental features [68].

2) FDP is actually useful for excluding people from suspicion, thus avoiding the conduction of extensive DNA dragnets (as mentioned earlier); in fact, perhaps this is the most important aim of FDP technology.

3) FDP only considers traces with 'a relatively high likelihood of manifestation, such as 75% or more (i.e. 75% or more of people with this genotype(s) actually develop the phenotype).

4) Despite that, FDP is particularly useful when several features are combined, and the use of a maximum of three or four independent phenotypes is recommended for a more conservative approach. More than that could cause the overall cumulative probability to rarely exceed 50%, which is very low (for instance, using a threshold of 75% in the prediction of two independent phenotypes would give a 56% probability, which is uninformative).

5) Finally, since FDP is an emerging technology applicable in the near future, it is essential to discuss this technology in the context of civil society in order to standardise techniques that can be used in forensic routines, especially with respect to the interpretation of the generated data. More importantly, the development of a legislative policy in each country to regulate and authorize explicitly (or not) the fulfilment of these procedures in criminal justice is of crucial importance.

The practical application of FDP technology requires robust validation protocols, but it is also necessary to account for population differences regarding the association of genetic markers used, as suggested by Cerqueira et al. [69]. For example, although there are highly significant polymorphisms associated with a particular phenotype in some populations, these same PIMs have effects that are apparently too small or null in others, suggesting that they can only be used as markers in the context of specific populations. Furthermore,

analysis with too few markers may not ever be sufficient for the prediction of the various existing complex physical characteristics. Therefore, regarding some phenotypes, we agree with Kayser and Knijff [67] in the sense that it is 'unlikely that FDP technology to be achievable with small sets of genetic markers'. Technological advancement has allowed large scale DNA sequencing and genotyping technologies to be increasingly cheaper and faster [3,4], enabling massive analyses of polymorphisms; therefore, the predictive power of phenotypes is expected to be increasingly reliable. The problem with using large sets of markers is the low quantity and quality of the biological samples recovered in the forensic material.

In Table 1, we point out some of the more frequently studied phenotypes in FDP technology ("Green list of FDP technology"), which are broadly named as externally visible characteristics-EVCs [28]. There is a good amount of information generated on these phenotypes and some of these can be predicted with a reasonable degree of reliability. Most of the phenotypes described in the Green list (Table 1) still need validation for ensuring greater reliability in prediction. Despite that FDP technology is discussed mainly in relation to normal human characteristics, we also present a "Red list of FDP technology", which includes some phenotypes that still need extensive regulation regarding ethical issues in case of to be potentially implemented in the future. Such traits are useful nowadays only for clinical practice (molecular diagnosis and genetic counseling) or are even non-pathological phenotypes (e.g. personality features) that deserve extreme caution due to ethical issues on privacy of phenotypic data of the investigated subjects and the intrinsic difficulty of measuring such phenotypes. Additionally, if eventually these kinds of phenotypic characteristics are included in forensic practice, corresponding protocols should be extensively regulated by the law to ensure the proper access of this information by the general public and police staff.

One of the further advantages of performing FDP is to increase the statistical confidence in further analyses performed with conventional STR profiles [18]. The conclusion about genetic identity or non-identity between two samples is probabilistic. If we have predicted phenotypic information on a suspect prior to making the STR profile, the information gathered with the FDP technology will be useful for increasing the reliability of the claim that a particular person actually left his sample at the crime scene. This increase in reliability of the STR profile has been recommended by the NRC [70] (National Research Council, USA) when conducting tests to characterise the biogeographic group belonging to an unknown sample.

Another FDP potentiality also includes assistance in finding missing persons through phenotypic reconstruction from DNA extracted from bones (for example, in forensic anthropology, the traditional method is the analysis of the femoral size, skull and other bones to estimate height and age) as well as to assist in phenotypic reconstruction of missing children to check how they would look to be in adulthood from DNA analysis, which would assist the method already used in computer forensics through photo analysis.

Green list of FDP technology		
Adult stuttering	Facial features	Handedness
Age	Freckles	Height
Ancestry	Gender	Skin colour
Baldness	Hair colour	Sample's surname
Eye colour	Hair shape	EVCs in general
Red list of FDP technology		
Achondroplasia	Cancers	Obesity
ADHD	Diabetes mellitus type I or II	Personality traits
Alcoholism	Familial hypercholesterolemia	Phenylketonuria
Alzheimer's Disease	G6PD deficiency	Schizophrenia
Autism	Haemophilia	Sickle cell disease
Bipolar disorder	Huntington's Disease	Tay-Sachs disease
Cystic fibrosis	Marfan syndrome	Others diseases or not-EVCs

Table 1: List of traits predicted (currently or in near future) from DNA analyses.

PIMs for molecular diagnosis

In the following two sections, we discuss the opportunities regarding the use of DNA analysis within the medical field (molecular diagnosis), and tracking pathogens in the human body. Furthermore, we explain how it connects with FDP, and how the forensic field can benefit from the practice and research avenues already experienced by medical genetics. In the other words, we try to discuss which analysis can be performed through DNA analysis on both forensic and medical points of view, without attempting to exhaust such a vast theme. It is important to clarify the different goals between the technology of FDP and the technology of prediction of human diseases or pathogens traceability.

A large amount of data is available on molecular methods for the diagnosis of a handful of diseases. For example, the GeneTests website

has information about these tests, as well as a list of laboratories worldwide that perform them. There are over 3,000 monogenic diseases that are possible to be diagnosed by molecular methods (OMIM). It is important to note here that these diagnostic methods are extremely useful to detect some diseases in pre- and postnatal testing in order to outline preventive medicine and to avoid or mitigate future damage to the newborn or prepare the family for such a child with an incurable genetic disease [71]. In addition to cytogenetic prenatal tests performed to detect structural or numerical chromosomal changes, which cause certain syndromes (for example: Down's – Trisomy 21; Turner – Monosomy X without another sex chromosome; Klinefelter – Disomy X with the presence of Y; Patau – Trisomy 13; Edwards – Trisomy 18), there are other tests that can be done for detecting smaller genetic mutations (e.g. SNPs) or abnormal gene products, which are also directly associated with various diseases. The March of Dimes global report [71] suggests conducting preventive tests for 29 conditions in the United States, including cystic fibrosis, citrullinemia, sickle cell disease, phenylketonuria, galactosaemia.

Besides enzymatic assay or the analysis of a particular gene product routinely performed for diagnosis of many genetic diseases, tests based on direct analysis of DNA are performed throughout the world for diagnosing these disorders, albeit less frequently. As outlined by Nussbaum et al. [72], this is due to the fact that when a genetic disorder that appears to be a single entity is studied in more detail, it often turns out that it is genetically heterogeneous, making it difficult to perform and interpret these tests. Therefore, it is noteworthy that the resultant phenotype of many monogenic disorders is not caused by just a single mutation. In contrast, there are many cases described in the scientific reports of individuals with the same disease, who do not have the same genetic mutation. Furthermore, there are cases reported where the individual does not manifest the disease but has the mutation (incomplete penetrance), which has caused for revision of the genetic etiology of some disorders in the scientific community [72]. Thus, for many diseases, the dosage of the corresponding gene product is determined instead of analyzing the DNA itself. Therefore, healthcare professionals should be aware of the genetic heterogeneity (allelic or locus) underlying the disorder being diagnosed or managed clinically and if possible, they should do a thorough analysis of all gene(s) (e.g. using sequencing technologies), when appropriate, to check for possible mutations that may be responsible for the disorder in question.

Some examples of human genetic diseases that can be diagnosed early by direct DNA analysis include achondroplasia (Gly380Arg detection of mutation in the receptor 3 of fibroblast growth factor coded by FGFR3 gene, which is present in virtually all cases [73]; cystic fibrosis (the most common mutation detected is phenylalanine amino acid deletion at position 508 in the CFTR gene product), reviewed in Zielenski [74]; Sickle cell anaemia (one form present a substitution of the sixth amino acid of the beta-globin chain – Glu6Val), described in Linus Pauling et al. [75] and reviewed by Bender and Hobbs [76]; Huntington's disease (detection of >35 copies of the trinucleotide CAG sequence in exon 1 of the HTT gene, which encodes a polyglutamine chain in the protein huntingtin), which can also be diagnosed in patients and carriers by analysing a set of 22 non-redundant tagging SNPs, described by Warby et al. [77]; Tay-Sachs disease (one way of detecting is the insertion of four bases – TATC – in the coding sequence of the HEXA/hexosaminidase A gene [78]; among other diseases. More details can be found in Gene Reviews, Online Mendelian Inheritance in Man (OMIM), Human gene mutation

database or in other sites and references previously suggested in this section.

Beyond monogenic diseases with complete or incomplete penetrance, some commercial tests are now available to diagnose some multifactorial diseases (cancers, obesity, hypertriglyceridemia, schizophrenia, autism, ADHD, Alzheimer's disease). For some of these disorders, it is important to mention that there are some isolated mutations already described in less common cases that have great effects on the phenotype, therefore triggering an associated pathology, although the predominant framework is the interaction of multiple genetic and environmental factors in the aetiology of the disorder [68]. Some practical considerations are very relevant to this topic. As it happens with any complex genetic trait, multifactorial diseases are caused by interaction between genes and environment. It is important to raise awareness that in general, just being a carrier of a particular mutation does not mean that the subject will necessarily develop the disease. Detailed analysis of each case and the associated family history will be the most appropriate procedure. In such cases, genetic counselling with a multidisciplinary approach is thus mandatory. Interesting recommendations for personalized medicine and the use of genetic information can be found in Burke and Psaty [79], and Chen and Snyder [80].

In general, characterization of the genetic profile of various mutations that affect a multifactorial trait can be useful for the affected individual to learn to preventatively deal with their possible future condition (e.g. through changes in lifestyle habits, among other factors). However, proper monitoring by multidisciplinary professionals is necessary, and genetic-medical counselling remains extremely useful to interpret data from families affected by these conditions, as done in prenatal diagnoses and tests for monogenic diseases. Thus, patients need to be properly informed that even when they carry a mutation associated with complex diseases, it is also true that there are many mechanisms involved in the development of the disease. This can help to prevent/stimulate the exposure to particular crucial environmental factors and to explain that the patient's health and the genotype-phenotype relationship are not so direct, in most cases [7,68]. With respect to EVCs studied in the forensic field, it is important to mention that such traits are also complex, and are mostly conditional on the genetic background and environmental interference. However, there are both normal and pathological phenotypes that can possibly be detected by PIMs, therefore, the likelihood of developing the trait or phenotype is quite high, often exceeding 90% probability.

Nevertheless, for many multifactorial conditions, knowledge on the underlying phenotypic components of a macro disorder is essential in understanding the genetic basis of a given condition. For example, studies of predictability of human psychiatric diseases through DNA can be benefited from studies like Cerqueira et al. [81], who associated markers in ADRA2A with ADHD endophenotypes.

Minimising the phenotypic spectrum of a macro-disease is an alternative to estimating traces of the disease. We emphasize here that careful evaluation and validation should be performed before practical application of such molecular diagnostic methods.

According to Wienroth et al. [82], the genetic advance, reflected both in the possibility of using innovative technologies for phenotype prediction as well as in next generation sequencing technologies, among others, has led to the diffusion of a boundary between medical and forensic genetics. The use of modern techniques in the forensic

field could be guided by a rationale that the "proportionality" (serious crimes) imposes the alternative uses of innovations in DNA profiling. The possibility to use such modern methods of analysis in forensics, depending on the case investigated, has also triggered ethical, social, and legal aspects of the debates, which have been discussed later in this review.

Pathogen traceability in the human body using molecular markers

Existing infectious agents (bacteria, protozoa, viruses, fungi, or worms) in the human body can be detected in various tissues or biological fluids (blood, skin, saliva, etc.) with currently available molecular biology techniques [5]. Traditionally, pathogens in the human body are tracked by using immunological methods, microscopy, or direct detection by isolation of the causative organism.

Many molecular detection techniques are still in the validation phase and not yet widely used in laboratories, however, a few are now available and in the process of being established in several laboratories worldwide. Here, we mention some of these techniques to trace pathogens responsible for some well-known diseases and indicate references that can serve as an initial point of reading on this issue. The list of pathogens that can be detected by DNA/RNA-based technology is large, which includes: analysis for the detection of tuberculosis (e.g. Mycobacterium tuberculosis [83]), human papillomavirus (HPV), Chlamydia trachomatis and Neisseria gonorrhoeae [84-87], avian influenza virus and influenza A (H1N1) virus [88-89], human immunodeficiency virus (HIV) [90,91], measles [92,93], hepatitis A, B, and C [94-97], among other pathogenic agents.

In the forensic field, this technology could be useful to analyse infected materials (e.g., biological weapons), or in the context of human identification, by linking biological materials found in the crime scene to specific suspects exhibiting a specific profile of infectious disease (e.g., linking sample donors with criminal acts) [67].

In the molecular diagnosis of pathogens and several other molecular techniques, it is necessary to follow strict protocols for analysis and carefully validate in-house methods aimed at avoiding false-negative/positive results, cross-contamination, as well as inhibitory substances in the reaction chosen for analysis, among other factors. Notably, the choice of biological material used in the analysis also has fundamental importance, taking into account the tissue in which the pathogenic agents have higher affinity, and the clinical characteristics present in the patient.

Ethical and legal perspectives

The technology of predicting phenotypes with DNA analysis is not new, since it has already been used in the medical field for some years and for a multitude of diseases, as seen earlier and as described in the cited references. Medical genetic experience can teach many things to the forensic area, and knowledge and perspectives on disease diagnosis and ways to trace pathogens in biological material may be useful at some point to such areas. The aspect that is new in this technology is the discussion about the prediction of human phenotypes in the police context, and as the technology of FDP is still recent, we have spoken only in predicting externally visible phenotypes within criminal prosecution, i.e. predicting information that one can visually see in the suspect. We emphasise that this information is not confidential and is present in civil identification documents such as driving licenses and professional cards. Thus, the general argument is that the prediction of these EVCs should not violate any aspect of the suspect's privacy.

The issue of privacy and the right of not knowing has been very well discussed by Koops and Schellekens [6] and the origin of the controversy around this topic may have been with the predictor polymorphisms of disease or stigmatizing characteristics (genetic predisposition to homosexuality, violence, etc.), which could psychologically affect not only the investigation, but also third parties not affiliated with the crime. Additionally, to find mutations in DNA that are predictive of disease and other features indicate possible predisposition of direct relatives not involved in the crime. However, the aforementioned study discusses and questions whether the right of not knowing should be respected in all cases, or if public interest in the criminal investigation should have more value in some cases. One solution to this issue would be to regulate the inclusion of diseases within the scope of FDP technology that have an available cure or diseases with relatively low impact, particularly as a last resource to search suspects in cases of violent crimes, such as murder or rape. In these cases, the public interest in the criminal investigation would have greater value than the right to not disclose a genetic condition. Alternatively, the predicted information in the course of the investigation through DNA analysis could take place in complete secrecy, and would be compatible with the right of not knowing if the information from the DNA is available only at request of a suspect (or defender). In this case, the only information that could be disclosed would be EVCs (or neutral traits like manual dexterity), such as a composite sketch of the suspect.

Although the first article on FDP was published in 2008 [6], very few studies have been published in the developing countries that discuss the legal and ethical perspectives on forensic DNA phenotyping and/or the potential use of information in the molecular diagnosis of diseases in the forensic field. Despite the scarce bibliography and discussion on this issue in some nations, guiding aspects on FDP have been very well established in Europe and United States, such as Koops and Schellekens [6], Kayser [19], MacLean and Lamparello [24], and Wienroth et al. [82], or the documents on forensic genetics in the Euroforgen website.

For example, Wienroth et al. [82] raised some issues about anticipatory governance in relation to the necessity to observe the public/user perspectives, propose legislative changes, validate processes for new technologies, and ensure potential safety mechanisms, emphasizing the emerging need for a wide debate on these aspects, in both, the forensic and the medical field with prediction of phenotypes. Maclean and Lamparello [24] discussed some promises and dilemmas of FDP technology (e.g. DNA phenotypic descriptions versus eyewitness descriptions; externally visible characteristics versus unseen characteristics; phenotyping of stored samples versus crime scene samples; physical characteristics versus behavioural characteristics; among others), which are very relevant to the legal treatment of the topic, as well as to raise awareness by technological operators.

As pointed out by Maclean and Lamparello [24], the concerns that have been raised against the development and implementation of FDP do not justify the elimination or prohibition of this technology. It is also important to note that the discussion of many topics in the medical field can either fuel or help resolve similar concerns in forensic regulations. It is also worth mentioning that FDP should be used only in the investigative procedure and thus would not be used in court. The court is involved only with conventional DNA profiling (i.e. analysis of

STRs), which has more probative value. We also mention that FDP is useful only for tactical purposes intrinsic to police investigation, and should only be used as evidence, since this technology does not have incriminating or absolute power, such as that contained in an STR profile.

There is broad consensus about a lack of regulations against inappropriate use and legislation aimed to limit the phenotypes that might be explored using this technology. Moreover, legislation is needed to allow the use of FDP for "cold" cases or cases that have no progress in the investigation, or lacking eyewitness, for example. Another serious concern is that many countries pose silent legislation, thus opening up the possibility of improper and inadvertent use of the technology.

Regarding the confidentiality of genetic information investigated within both, a medical scientific research and a forensic context, the non-stigmatization and non-discrimination of people based on genetic information, and other sensitive topics, the use of DNA is something guaranteed to all civil societies, in line with recommendations made by international organizations and laws, such as the International Declaration on Human Genetic Data, Universal Declaration on the Human Genome and Human Rights, and the Universal Declaration on Bioethics and Human Rights. However, a question already discussed [6] with respect to genetic data, is who would have access to information through DNA analysis? Would this information be available in a database? To answer these questions, proper regulations are urgently needed.

Another debatable concern the necessity of storing already investigated/prosecuted phenotypic information into a database. Preliminarily, one could argue that there is no necessity in creating a database of physical characteristics of crime suspects. Furthermore, it may begin a dangerous debate of enabling research avenues aimed to associate specific phenotypes with criminal acts. Unfortunately, these researches are already in development, and we recently reported a meta-analysis of data demonstrating the lack of association among facial traits and aggressive behaviors in a large and composite database [98]. Due to the risk of stigmatization of a particular phenotypic profile (through the public release of a "molecular portrait" of the suspect, in analogy to what is known as a "sketch"), police searches should always be made considering the good sense and decide between publicly releasing the picture and stealthy tracking. Even when it comes to EVCs, this weighting is necessary because "four individuals with red hair in a small town can be more stigmatized than 1000 left-handed individuals in a medium-sized city" [6]. Finally, the police intelligence team requires a case-by-case basis.

Final considerations

Theoretically, molecular prediction techniques based on DNA, RNA, or proteins can be applied to most phenotypes, whether normal or pathological, or even to detect the many pathogens to which we are exposed throughout our lives. Although not all of the available technologies are used in laboratory routines yet, it is expected that many will be consolidated and implemented in the near future. This is because the power of current diagnostic technologies has expanded our knowledge of human (and other species) genetics in an unprecedented way.

Notably, a very promising avenue of research is the study of DNA methylation patterns. Despite relatively recent knowledge of epigenetics, there are already several possible applications. For example, in forensics, the study of cytosine methylation patterns has been discussed in literature as an opportunity to differentiate identical twins based on DNA [99-101]. To our knowledge, there are still very few protocols to accomplish such tests with a reasonable degree of certainty. Another method that has been proposed to make such a differentiation is based on small differences in the DNA sequence [102,103], particularly using SNPs and other genetic markers (not based on an epigenetic analysis). Other possible perspectives related to the epigenetic field include the diagnoses of cancers and other diseases through the dynamics of DNA methylation [104,105], among other technologies.

Finally, a broad range of techniques is currently available or in quick development to be used in many fields of science. Therefore, it is essential to realize that most issues are not regarding the technology itself but the use made of it. Discussions such as these should be encouraged at conferences, meetings, and multidisciplinary seminars, in order to define proper usage for researchers and specialized communities, and to clarify the benefits of such technologies for civil society as a whole.

Conflict of Interest

Acknowledgments

References

1. Watson JD, Crick FH (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature 171: 737-738.

2. Gupta PK (2008) Single-molecule DNA sequencing technologies for future genomics research. Trends Biotechnol 26: 602-611.

3. Dunn PP (2011) Human leucocyte antigen typing: techniques and technology, a critical appraisal. Int J Immunogenet 38: 463-473.

4. Bohlander SK (2013) ABCs of genomics. Hematology Am Soc Hematol Educ Program 2013: 316-323.

5. Rossetti ML, Silva CMD, Rodrigues JJS (2009) Infectious diseases: molecular diagnosis. Guanabara Koogan, Rio De Janeiro, Brazil.

6. Koops BJ, Schellekens M (2008) Forensic DNA phenotyping: regulatory issues. Columbia Sci Technol Law Rev 9: 158-202.

7. Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. Nat Rev Genet 11: 855-866.

8. González-José R, Escapa I, Neves WA, Cúneo R, Pucciarelli HM (2008) Cladistic analysis of continuous modularized traits provides phylogenetic signals in Homo evolution. Nature 453: 775-778.

9. Mahner M, Kary M (1997) What exactly are genomes, genotypes and phenotypes? And what about phenomes? J Theor Biol 186: 55-63.

10. Johannsen W (1909) Elements of the exact heredity. University of Michigan Library, Michigan, USA.

11. Nachtomy O, Shavit A, Yakhini Z (2007) Gene expression and the concept of the phenotype. Stud Hist Philos Biol Biomed Sci 38: 238-254.

12. Gámez C, Sánchez-García S, Ibáñez MD, López R, Aguado E, et al. (2011) Tropomyosin IgE-positive results are a good predictor of shrimp allergy. Allergy 66: 1375-1383.

13. Leung NY, Wai CY, Shu S, Wang J, Kenny TP, et al. (2014) Current immunological and molecular biological perspectives on seafood allergy: a comprehensive review. Clin Rev Allergy Immunol 46: 180-197.

14. Ingram CJ, Mulcare CA, Itan Y, Thomas MG, Swallow DM (2009) Lactose digestion and the evolutionary genetics of lactase persistence. Hum Genet 124: 579-591.

15. Misselwitz B, Pohl D, Frühauf H, Fried M, Vavricka SR, et al. (2013) Lactose malabsorption and intolerance: pathogenesis, diagnosis and treatment. United European Gastroenterol J 1: 151-159.

16. Ooi SX, Lee PL, Law HY, Say YH (2010) Bitter receptor gene (TAS2R38) P49A genotypes and their associations with aversion to vegetables and sweet/fat foods in Malaysian subjects. Asia Pac J Clin Nutr 19: 491-498.

17. Sharma K, Kaur GK (2014) PTC bitter taste genetic polymorphism, food choices, physical growth in body height and body fat related traits among adolescent girls from Kangra Valley, Himachal Pradesh (India). Ann Hum Biol 41: 29-39.

18. Cho MK, Sankar P (2004) Forensic genetics and ethical, legal and social implications beyond the clinic. Nat Genet 36: S8-S12.

19. Kayser M (2015) Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. Forensic Sci Int Genet 18: 33-48.

20. Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, Quinto-Sanchez M, Jaramillo C, et al. (2014) Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 342 individuals. PLoS Genet 10: e1004572.

21. Adhikari K, Reales G, Smith AJ, Konka E, Palmen J, et al. (2015) A genome-wide association study identifies multiple loci for variation in human ear morphology. Nat Commun 6: 7500.

22. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46: 1173-1186.

23. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, et al. (2015) Genetic studies of body mass index yield new insights for obesity biology. Nature 518: 197-206.

24. MacLean CE, Lamparello A (2014) Forensic DNA phenotyping in criminal investigations and criminal courts: assessing and mitigating the dilemmas inherent in the science. Recent Adv DNA Gene Seq 8: 104-112.

25. Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44: 603-610.

26. Butler JM (2009) Fundamentals of forensic DNA typing. Elsevier, USA.

27. Goodwin W, Linacre A, Hadi S (2010) An introduction to forensic genetics. Wiley-Blackwell, New Jersey, USA.

28. Kayser M, Schneider PM (2009) DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. Forensic Sci Int Genet 3: 154-161.

29. Tozzo P, Giuliodori A, Corato S, Ponzano E, Rodriguez D, et al. (2013) Deletion of amelogenin Y-locus in forensics: literature revision and description of a novel method for sex confirmation. J Forensic Leg Med 20: 387-391.

30. Hennessy LK, Mehendale N, Chear K, Jovanovich S, Williams S, et al. (2014) Developmental validation of the GlobalFiler(®) express kit, a 24-marker STR assay, on the RapidHIT(®) System. Forensic Sci Int Genet 13: 247-258.

31. Draus-Barini J, Walsh S, Pospiech E, Kupiec T, Glab H, et al. (2013) Bona fide colour: DNA prediction of human eye and hair colour from ancient and contemporary skeletal remains. Investig Genet 4: 3.

32. Maroñas O, Söchtig J, Ruiz Y, Phillips C, Carracedo Á, et al. (2015) The genetics of skin, hair, and eye color variation and its relevance to forensic pigmentation predictive tests. Forensic Sci Rev 27: 13-40.

33. Mengel-From J, Børsting C, Sanchez JJ, Eiberg H, Morling N (2010) Human eye colour and HERC, OCA2 and MATP. Forensic Sci Int Genet 4: 323-328.

34. Spichenok O, Budimlija ZM, Mitchell AA, Jenny A, Kovacevic L, et al. (2011) Prediction of eye and skin color in diverse populations using seven SNPs. Forensic Sci Int Genet 5: 472-478.

35. Cerqueira CC, Paixão-Côrtes VR, Zambra FM, Salzano FM, Hünemeier T, et al. (2012) Predicting Homo pigmentation phenotype through genomic data: from Neanderthal to James Watson. Am J Hum Biol 24: 705-709.

36. Pospiech E, Draus-Barini J, Kupiec T, Wojas-Pelc A, Branicki W (2012) Prediction of eye color from genetic data using Bayesian approach. J Forensic Sci 57: 880-886.

37. Allwood JS, Harbison S (2013) SNP model development for the prediction of eye colour in New Zealand. Forensic Sci Int Genet 7: 444-452.

38. Hart KL, Kimura SL, Mushailov V, Budimlija ZM, Prinz M, et al. (2013) Improved eye- and skin-color prediction based on 8 SNPs. Croat Med J 54: 248-256.

39. Ruiz Y, Phillips C, Gomez-Tato A, Alvarez-Dios J, Casares de Cal M, et al. (2013) Further development of forensic eye color predictive tests. Forensic Sci Int Genet 7: 28-40.

40. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, et al. (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7: 98-115.

41. Gettings KB, Lai R, Johnson JL, Peck MA, Hart JA, et al. (2014) A 50-SNP assay for biogeographic ancestry and phenotype prediction in the U.S. population. Forensic Sci Int Genet 8: 101-108.

42. Sochtig J, Phillips C, Maroñas O, Gómez-Tato A, Cruz R, et al. (2015) Exploration of SNP variants affecting hair colour prediction in Europeans. Int J Legal Med 129: 963-975.

43. Walsh S, Lindenbergh A, Zuniga SB, Sijen T, de Knijff P, et al. (2011) Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. Forensic Sci Int Genet 5: 464-471.

44. Pneuman A, Budimlija ZM, Caragine T, Prinz M, Wurmbach E (2012) Verification of eye and skin color predictors in various populations. Leg Med (Tokyo) 14: 78-83.

45. Walsh S, Chaitanya L, Clarisse L, Wirken L, Draus-Barini J, et al. (2014) Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. Forensic Sci Int Genet 9: 150-161.

46. Meissner C, Ritz-Timme S (2010) Molecular pathology and age estimation. Forensic Sci Int 203: 34-43.

47. Zubakov D, Liu F, van Zelm MC, Vermeulen J, Oostra BA, et al. (2010) Estimating human age from T-cell DNA rearrangements. Curr Biol 20: R970-R971.

48. Yi SH, Xu LC, Mei K, Yang RZ, Huang DX (2014) Isolation and identification of age-related DNA methylation markers for forensic age-prediction. Forensic Sci Int Genet 11: 117-125.

49. Zbiec-Piekarska R, Spolnicka M, Kupiec T, Parys-Proszek A, Makowska Z, et al. (2015) Development of a forensically useful age prediction method based on DNA methylation analysis. Forensic Sci Int Genet 17: 173-179.

50. Allen HL, Estrada K, Lettre G, Berndt SI, Weedon MN, et al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467: 832-838.

51. Liu F, Hendriks AE, Ralf A, Boot AM, Benyi E, et al. (2014) Common DNA variants predict tall stature in Europeans. Hum Genet 133: 587-597.

52. Fujimoto A, Kimura R, Ohashi J, Omi K, Yuliwulandari R, et al. (2008) A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. Hum Mol Genet 17: 835-843.

53. Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, et al. (2009) Common variants in the trichohyalin gene are associated with straight hair in Europeans. Am J Hum Genet 85: 750-755.

54. Adhikari K, Fontanil T, Cal S, Mendoza-Revilla J, Fuentes-Guajardo M, et al. (2016) A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. Nat Commun 7: 10815.

55. Liu F, van der Lijn F, Schurmann C, Zhu G, Chakravarty MM, et al. (2012) A genome-wide association study identifies five loci influencing facial morphology in Europeans. PLoS Genet 8: e1002932.

56. Claes P, Liberton DK, Daniels K, Rosana KM, Quillen EE, et al. (2014) Modeling 3D facial shape from DNA. PLoS Genet 10: e1004224.

57. Hillmer AM, Brockschmidt FF, Hanneken S, Eigelshoven S, Steffens M, et al. (2008) Susceptibility variants for male-pattern baldness on chromosome 20p11. Nat Genet 40: 1279-1281.

58. Richards JB, Yuan X, Geller F, Waterworth D, Bataille V, et al. (2008) Male-pattern baldness susceptibility locus at 20p11. Nat Genet 40: 1282-1284.

59. Raza MH, Riazuddin S, Drayna D (2010) Identification of an autosomal recessive stuttering locus on chromosome 3q13.2-3q13.33. Hum Genet 128: 461-463.

60. Raza MH, Gertz EM, Mundorff J, Lukong J, Kuster J, et al. (2013) Linkage analysis of a large African family segregating stuttering suggests polygenic inheritance. Hum Genet 132: 385-396.

61. Frudakis TN (2008) Molecular Photofitting – Predicting Ancestry and Phenotype Using DNA. Elsevier, USA.

62. Paschou P, Lewis J, Javed A, Drineas P (2010) Ancestry informative markers for fine-scale individual assignment to worldwide populations. J Med Genet 47: 835-847.

63. Sykes B, Irven C (2000) Surnames and the Y chromosome. Am J Hum Genet 66: 1417-1419.

64. King TE, Jobling MA (2009) What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. Trends Genet 25: 351-360.

65. Redmonds G, King T, Hey D (2011) Surnames, DNA, and Family History. Oxford University Press, Oxford, UK.

66. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y (2013) Identifying personal genomes by surname inference. Science 339: 321-324.

67. Kayser M, de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet 12: 179-192.

68. Caspi A, Moffitt TE (2006) Gene-environment interactions in psychiatry: joining forces with neuroscience. Nat Rev Neurosci 7: 583-590.

69. Cerqueira CC, Hünemeier T, Gomez-Valdés J, Ramallo V, Volasko-Krause CD, et al. (2014) Implications of the admixture process in skin color molecular assessment. PLoS One 9: e96886.

70. National Research Council (USA) (1996) The evaluation of forensic DNA evidence. National Academy Press, USA.

71. Christianson A, Howson CP, Modell B (2006) March of Dimes – Global report on birth defects: the hidden toll of dying and disabled children full report.

72. Nussbaum RL, McInnes RR, Willard HF (2007) Thompson and Thompson Genetics in Medicine. Elsevier, USA

73. Hecht JT, Bodensteiner JB, Butler IJ (2014) Neurologic manifestations of achondroplasia. Handb Clin Neurol 119: 551-563.

74. Zielenski J1 (2000) Genotype and phenotype in cystic fibrosis. Respiration 67: 117-133.

75. Pauling L, Itano HA, Singer SJ, Wells IC (1949) Sickle cell anemia a molecular disease. Science 110: 543-548.

76. Bender MA, Hobbs W (2003) Sickle Cell Disease. In: Pagon RA, Adam MP, Ardinger HH, Bird TD, Dolan CR, et al. eds. Gene Reviews. University of Washington, Seattle, USA.

77. Warby SC, Montpetit A, Hayden AR, Carroll JB, Butland SL, et al. (2009) CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup. Am J Hum Genet 84: 351-366.

78. Bach G, Tomczak J, Risch N, Ekstein J (2001) Tay-Sachs screening in the Jewish Ashkenazi population: DNA testing is the preferred procedure. Am J Med Genet 99: 70-75.

79. Burke W, Psaty BM (2007) Personalized medicine in the era of genomics. JAMA 298: 1682-1684.

80. Chen R, Snyder M (2013) Promise of personalized omics to precision medicine. Wiley Interdiscip Rev Syst Biol Med 5: 73-82.

81. de Cerqueira CC, Polina ER, Contini V, Marques FZ, Grevet EH, et al. (2011) ADRA2A polymorphisms and ADHD in adults: possible mediating effect of personality. Psychiatry Res 186: 345-350.

82. Wienroth M, Morling N, Williams R (2014) Technological innovations in forensic genetics: social, legal and ethical aspects. Recent Adv DNA Gene Seq 8: 98-103.

83. Zakham F, Lahlou O, Akrim M, Bouklata N, Jaouhari S, et al. (2012) Comparison of a DNA Based PCR Approach with Conventional Methods for the Detection of Mycobacterium tuberculosis in Morocco. Mediterr J Hematol Infect Dis 4: e2012049.

84. Yang JL, Schachter J, Moncada J, Habte D, Zerihun M, et al. (2007) Comparison of an rRNA-based and DNA-based nucleic acid amplification test for the detection of Chlamydia trachomatis in trachoma. Br J Ophthalmol 91: 293-295.

85. Lee SH, Vigliotti VS, Pappu S (2009) Molecular tests for human papillomavirus (HPV), Chlamydia trachomatis and Neisseria gonorrhoeae in liquid-based cytology specimen. BMC Womens Health 9: 8.

86. Unemo M, Dillon JA (2011) Review and international recommendation of methods for typing neisseria gonorrhoeae isolates and their implications for improved knowledge of gonococcal epidemiology, treatment, and biology. Clin Microbiol Rev 24: 447-458.

87. Cuschieri K, Hardie A, Hovland S, Hoaas B, Karlsen F, et al. (2013) Comparison of the sensitivities of three commercial assays for detection of the high risk HPV types 1, 18 and 45. J Virol Methods 193: 147-150.

88. Sakurai A, Shibasaki F (2012) Updated values for molecular diagnosis for highly pathogenic avian influenza virus. Viruses 4: 1235-1257.

89. Lim S, Nan H, Lee MJ, Kang SH (2014) Fast on-site diagnosis of influenza A virus by Palm PCR and portable capillary electrophoresis. J Chromatogr B Analyt Technol Biomed Life Sci 963: 134-139.

90. Bourlet T, Memmi M, Saoudin H, Pozzetto B (2013) Molecular HIV screening. Expert Rev Mol Diagn 13: 693-705.

91. McCollum ED, Preidis GA, Maliwichi M, Olson D, McCrary LM, et al. (2014) Clinical versus rapid molecular HIV diagnosis in hospitalized African infants: a randomized controlled trial simulating point-of-care infant testing. J Acquir Immune Defic Syndr 66: e23-e30.

92. Bankamp B, Byrd-Leotis LA, Lopareva EN, Woo GK, Liu C, et al. (2013) Improving molecular tools for global surveillance of measles virus. J Clin Virol 58: 176-182.

93. Michel Y, Saloum K, Tournier C, Quinet B, Lassel L, et al. (2013) Rapid molecular diagnosis of measles virus infection in an epidemic setting. J Med Virol 85: 723-730.

94. Martins PP, Lampe E, Lewis-Ximenez LL, de Souza PS, Fernandes CA, et al. (2013) Performance of molecular methods for hepatitis C virus diagnosis: Usefulness among chronic cases and during the course of infection. Clin Lab 59: 1031-1039.

95. Poovorawan K, Chattakul P, Chattakul S, Thongmee T, Theamboonlers A, et al. (2013) The important role of early diagnosis and preventive management during a large-scale outbreak of hepatitis A in Thailand. Pathog Glob Health 107: 367-372.

96. Heiat M, Ranjbar R, Alavian SM (2014) Classical and modern approaches used for viral hepatitis diagnosis. Hepat Mon 14: e17632.

97. Kania D, Ottomani L, Meda N, Peries M, Dujols P, et al. (2014) Performance of two real-time PCR assays for hepatitis B virus DNA detection and quantitation. J Virol Methods 201: 24-30.

98. Gómez-Valdés J, Hünemeier T, Quinto-Sánchez M, Paschetta C, de Azevedo S, et al. (2013) Lack of support for the association between facial shape and aggression: a reappraisal based on a worldwide population genetics perspective. PLoS One 8: e52317.

99. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, et al. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102: 10604-10609.

100. Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GH, et al. (2009) DNA methylation profiles in monozygotic and dizygotic twins. Nat Genet 41: 240-245.

101. Li C, Zhao S, Zhang N, Zhang S, Hou Y (2013) Differences of DNA methylation profiles between monozygotic twins' blood samples. Mol Biol Rep 40: 5275-5280.

102. Krawczak M, Cooper DN, Fändrich F, Engel W, Schmidtke J (2012) How to distinguish genetically between an alleged father and his monozygotic twin: a thought experiment. Forensic Sci Int Genet 6: e129-e130.

103. Weber-Lehmann J, Schilling E, Gradl G, Richter DC, Wiehler J, et al. (2014) Finding the needle in the haystack: differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing. Forensic Sci Int Genet 9: 42-46.

104. Robertson KD (2005) DNA methylation and human disease. Nat Rev Genet 6: 597-610.

105. Razvi H (2014) DNA methylation and histone modification molecular associations with specific diseases. GENReports: Market & Tech analysis.