IDENTIFYING INDIVIDUALS BY MEANS OF FORENSIC GENOTYPING

Ph.D. Ioan Aron*

Abstract:

We may say that modern genetics was born in 1900, when Mendel's rules were rediscovered, as presented by their author in a series of scientific communications starting with 1865. The word „gene" was printed for the first time in 1909, while the first demonstration according to which each gene may be associated to a certain chromosome was made known to the scientific world in 1910. The first genetic map – showing the relative location of six genes on one chromosome – was published in 1913. That would be, briefly, the history of the beginning of the greatest scientific adventure lived by the humanity.

Keywords: *individual identification, forensic genotyping, chromosome, genetic imprint, DNA, genetic markers.*

1. Considerations on genetic imprinting

For a long time, it has been believed that the development of DNA-related research is of a peripheral interest for forensics. In 1985, the paper delivered by the English Professor Alec J. Jeffreys on the possibility of individual identification based on the human DNA repetitive hyper variable regions was a breakthrough in forensics. The research made in the project of the human genome, consisting of the identification of the DNA code for each living cell of a body demonstrated that such code is the support of heredity, in other words, the unique genetic imprint of an individual.

The DNA is a polymer, a very big molecule made by the joining of a series of units (nucleotides) which repeat themselves and whose number is ca. three billion. There are 4 types of nucleotides, traditionally designated by the letters ACGT (adenine, cytosine, guanine and thymine), which are grouped around a coil and whose structure is called a „double helix". Those who succeeded to reveal the characteristics of DNA molecules are the researchers James Watson and Francis Crick in 1950, who, in 1962, received the Nobel Prize for having discovered the DNA structure. From the onset, the criminalists asked the scientific world several questions whose answers are of essence in using DNA tests as evidence in forensic research : how big, physically speaking, should the biological sample be so that the laboratory test can reveal the DNA structure of an individual?, how old should such sample be so that it is conclusive?, can or cannot human DNA be confounded with the DNA of other living bodies?.

The researchers involved in the project of the human genome answered these questions, clarifying each of their aspects.

Firstly, it has been scientifically demonstrated that the sample size under analysis for DNA identification may be reduced to the size of a molecule, as the DNA is present in absolutely all the cells of a living body, irrespective of where in such body they come from.

As far as the age of the biological sample is concerned, the dilemma was solved by research made on samples taken from bodies older than several thousands of years. For example, in 1994, DNA research on a Peruvian mummy that was 1000 years old revealed the existence of tuberculosis as an infectious disease caused by Mycobacterium tuberculosis, in the bodies of Native Americans 500 years before the arrival of the Europeans, brought by Christopher Columbus. The same kind of research, made on the oldest mummy discovered in the Atacama desert, whose age was 9,000 years, led to the discovery of the „trypanosoma cruzri" parasite, which causes the Chagas disease, a

* Assistant Proffesor, Faculty of Law – University „Transilvania" from Brasov, e-mail: aron.ioan.mai@gmail.com.

disease that is spread on wide regions in Latin America. To identify this disease, the research was supplemented by samples taken from the skeletons and viscera of 27 mummies of men, women and children, being proven that seven of them were infected with Chagas disease. In order to explain why approximately 20% of the population of Egypt suffer from schistosomiasis, Egyptian mummies have been examined, which led to the conclusion that the disease originates in the use of contaminated water and that such disease is genetically inherited.

A fascinating discovery has led to the conclusion that not only the samples taken from mummified bodies, that is preserved bodies, show results, but also the DNA tests on materials used by Ancient people in artistic or lucrative activities may lead to the identification of the DNA of living bodies. The DNA tests on rock paintings, discovered in the region of Lower Pecas in Texas (between 2,950 and 4,200 years old) have concluded that the dyes used consisted of blood, urine, milk, eggs, vegetal juice and animal fat.

The DNA tests may identify genetic markers that approach or differentiate the ethnicity and customs of the individuals of a community, thus being able to explain significant cultural and social aspects. In the mid 80s, this kind of test revealed that North America was colonised by three distinct migration waves, while the archaeological digging in the town of „Three Hills" in China showed the Tuja minority population was a descendant of the Ba culture bearers (206 B.C.-220 A.D) and even of the Western Zhou culture bearers (1100-771 B.C).

As for the strict delimitation existing among the DNAs of living bodies, by species and sub-species, the research of the latest 25 years has solved the problem. Today we know precisely the DNA structure of several species and sub-species, but, above all, the year 2000 brought a scientific breakthrough : the complete genome mapping and, confounding it with the DNA of other living bodies, either insects or animals or plants, is practically impossible. Thus, one of the main forensic conditions, namely the particularisation and assignment of biological traces, irrespective of their nature, taken on the crime scene, receives a complete answer – as well as complex one– leading to the accurate solving of cases.

The answers provided by science for the legitimate questions of criminalists open unlimited perspectives for the interpretation of evidence taken from the crime scene, for whose interpretation and completion, at the time of sampling, there were neither technical means nor sufficient scientific knowledge.

The fact that DNA tests can be performed on minute quantities of evidence coming from formerly living bodies, that its age is no drawback in its identification, as well as the complete mapping of the human genome in 2000, which eliminates possible confusions of human DNA with DNAs dispersed in the genomes of eukaryote species (non-human), less known, make the DNA unique, as the paramount evidence. Under such circumstances, the DNA-related discoveries equal the existence of a universal key for the decryption of any code.

It comes without saying that the mere analysis of a biological sample and the decryption of the DNA code characterising such sample do not automatically lead to the solving of a case, as long as there is not the indispensable comparison element, of personalising the analysis, by designating with certainty the individual to whom the sample belongs. In case there are suspects, the DNA identified in the sample taken on the crime scene may lead, by comparison to the suspects' DNA samples, to establishing either their guilt or their innocence.

One of the methods proposed by the criminalists to ease the investigation is the establishment of some data banks that should include the genetic information of a population sample as broad as possible. In North America, the FBI, the government, military secret services and the insurance companies have began to collect and store genetic information in data banks, the so-called information banks and identification banks. In their turn, the information banks are divided into three general types:

2. Information banks storing genetic data

Information banks that store names, payment receipts, credit reports and medical data – for example, Medical Information Bureau Inc. in Westood Massachusetts (MIB) holds medical

data for ca. 20 million people from all the USA states. At the beginning, this bank was created to prevent insurance fraud, but currently it is used by the insurance companies to decide on the provision of insurance or for the determination of the insurance premium for an applicant, based on strictly confidential genetic information, whereby individuals tested positive for certain diseases may be refused an insurance contract.

In 1994, only the United Kingdom of Great Britain and Northern Ireland and the Netherlands held a national genetic database. In the UK, the genetic database, called the DNA Index, was created in 1989 and became operational as of 01.01.1990.

Such database is limited to the people whose criminal record is connected to the samples analysed, to convicted offenders and to biological samples coming from cases with unidentified offenders. Its main functions are to designate the main suspects based on the evidence taken from the crime scene, to assess the suspects designated by the police, to identify potential serial murders and to develop statistics concerning the DNA profile.

The FBI has piloted a system of DNA profile storage and comparison, which operates on three levels : locally, federally and nationally. This pilot system, called The Combined ADN Index System (CODIS) has two types of folders : The Forensic Index and The Convicted Offender Index. Austria has a database that reunites the genetic imprints of sexual offenders and murderers.

The success of a potential European genetic database closely depends on the way in which national databases are organised in each country. In every state, the decision of creating such a database depends on several different factors, such as :

- the majority of people accepting the use of DNA tests for repressive purposes, but only when social co-living rules are violated;
- request by the bodies in charge with law enforcement;
- using DNA profiling as means of evidence before national courts;
- capacity of national laboratories;
- legal provisions in force on the sampling of alleged offenders and preservation of such samples.

The genetic database, consisting of a collection of DNA profiles stored in computers as Alpha numerical codes, generally includes three types of files :

- profiles that have been obtained following sampling of a crime scene with unidentified offender (profiles belonging to so-called „unsolved samples”);

1. DNA profiles of convicted offenders;
2. DNA profiles of missing people or of the parents of missing people.

The identification banks contain electronic records of physical characteristics (height, weight, hair and eyes colour, skin markers, teeth and papillary imprints) and are meant for business use and for the use of governmental agencies wanting to identify individuals.

All these banks can offer significantly valuable information for forensic identification. For example, the FBI, along with the Defence Department, has developed in over 18 American states the system of the identification banks for the storage of genetic information. The stated purpose of such programmes was to achieve electronic data storage, as well as biological sample storage, for sample mainly coming from people who committed or are suspected to have committed offences that left biological traces at the crime scene. In the state of Virginia blood and saliva were collected from over 40,000 convicted offenders (until 1992), which led to the creation of the largest data bank in the USA. The FBI has established a centralised identification of databases containing genetic information, based on the DNA tests from violent criminals, and, at the same time, has launched a N.D.I.I. (National DNA Identification Index) pilot that should allow, by means of a relevant software, the circulation of data in all the state of the USA. The Defence Department of the USA collected biological samples from ca. two million people in the military to compare their DNA profiles to those of some soldiers that may acquire the status „unidentified” following a future war.

In the early days of databases based on genetic information, they won public sympathy both among the North-American population and the population of the European Community, due to the advantages the DNA tests showed when solving special cases, in which mainly criminalists were

involved, such as : ensuring an accurate identification of people, either sought by the police, or under suspicion or declared missing as a result of catastrophes (air, shipwrecking, fire etc.), reuniting members of the same family who were separated under various circumstances (armed conflicts or child replacement after birth), achieving firm containment of criminality.

The use of genetic data by government agents or private companies for purposes other than those initially stated was nonetheless unavoidable. After a short while, after the first wave of enthusiasm faded away, the use of the data stored by the military, police, secret services, insurance companies, and other holders of information and identification banks for purposes other than those stated initially was considered a violation of human rights by the general public and specialist non-governmental organisations.

It was noticed that the biological samples and genetic information stored may lead not only to positive social results, such as containment of criminality, identification of missing people etc., but also to actions prohibiting civil rights, by using the genetic information to limit the access of some people, genetically suspect, to the insurance system and even to professional positions; or for certain negative interpretations in terms of people belonging to certain social and cultural groups or to an ethnic minority.

For example, there have been numerous debates on genetic diagnosis before implantation due to the high risk of its use in eugenic practices, practices prohibited by the European Convention on Human Rights. This kind of procedures aimed at early diagnosis of genetic diseases, couples using it when the risk of serious diseases transmission is present, procedure forbidden although in some legislations and, where allowed, it is limited to situations in which the risk of transmission of hereditary diseases is very high and there is no treatment[1].

3. The structure of the human genome

Every individual is different from their peers by genotypic characteristics (genetic hypervariability) that have a phenotype manifestation, giving birth to the numberless individual characters. Such differences allow for the recognition of some individuals, who, having the same characteristics as the tested samples, are beyond doubt the source of the biological material. These characters are genetically-based, which ensure their constancy during the entire life of an individual. They are hereditary, which also allows for their use in expertise of filiation. These characteristics, called polymorphous, are contained by the inheritance molecule, the DNA.

The human genome, which exists in its entirety in all the cells, except for the erythrocytes, is made of molecules of deoxyribonucleic acid – DNA. In its turn, this macromolecule is made of two polynucleotide „strands" organised as a „double helix" (double coiled band) supported by hydrogen links, resulting from the joining of nitrogenous bases. This results in the complementarity of each „strand". The nucleotides sequence along the two strands defines the genetic code.

A genome consists of :

- Encrypted sequences
- Non-encrypted sequences, distributed along the genome.

Among them there are repetitive regions that have a common central pattern. They were discovered by Alec Jeffreys from the University of Leicester (England) in 1985 and called minisatellites (due to their peripheral position when the genetic material is ultra-centrifuged in the presence of the caesium chloride).

The minisatellites, also called VNTR (variable number tandem repeat) are highly polymorphous regions, in tandem, and made of short base sequences of different lengths (from 6 to 64 nucleotides), which repeat themselves from 1 to several hundred times. The number of sequence repetition is different from one individual to another and is inherited according to Mendel's laws. Consequently, an individual holds half of their minisatellites from their mother and the other half from their father.

[1] Nasty Marian Vlădoiu – Protecția constituțională a vieții, integrității fizice şi a integrității psihice, Hamangiu Publishing House, Bucharest, 2006, p. 137.

Character segregation according to Mendel, as well as the minisatellites is balanced among the population, as their distribution is according to the law of Hardy Weinberg.

The analysis is based on two principles :

- the complementarity of the two DNA strands

- the DNA is made of two nucleotide strands, each representing the genome assemble of an individual. These nucleotides-adenine, guanine, cytosine, thymine – are joined two by two (A-T and G-C). This complementarity allows the „binding" of the two strands to each other, so that the final shape of the DNA is the „double helix". By separating the two strands, a particular sequence may be revealed via a reference DNA (probe) which will join its complementary.

The existence of repeated sequences in the genome, ranging in number from one individual to another.

Such sequences can exist on different chromosomes (multilocus probes) or on the same chromosome. These genome regions are revealed by means of the specific probes of the base unit. The variable number of base units accounts for the diversity noticed in population (over 70 billion possible combinations for some probes).

In the actual context, the knowledge of human genome structure is particulary important, interventions on the human genome being able to affect the integrity of the individual and even threaten the human species[2].

[2] Nasty Marian Vlădoiu, Biocrime – A New Concept, Criminal Law Review, Vol. I, Issue 1, Jan.-Jun. 2013