dr Magdalena Zubańska
head of Forensic Science Department in Faculty of Internal Security of Szczytno Police Academy
Przemysław Knut
student at Faculty of Internal Security of Szczytno Police Academy

Procedure for phenotypic data collection as an important stage of predictive studies

Summary

The article addresses the issues related to phenotypic data collection at the stage of acquiring research sample and indicates the importance of this stage for studies on DNA-based predictive analysis. It presents the assumptions and the execution of a task carried out as part of the project entitled: *"A genetic portrait of the offender and the victim – developing a system to determine human physical appearance and biogeographical origin, based on DNA analysis, involving Next Generation Sequencing (NGS)"* – the acronym NEXT.
In addition, the results of a questionnaire survey administered to a group of people who have taken part in the above project as sample donors, are discussed. The purpose of the study was to determine the motivation for participation in a scientific undertaking related to genetics. Based on the respondents' answers, the general conclusion can be formulated that this type of research enjoys positive social perception.

Keywords: research material, phenotypic data, forensic phenotyping, next generation sequencing

Forensic science constitutes a theoretical basis for practical activities of law enforcement authorities, hence any research in this field must be adapted to the areas of activity and needs of forensic practice (Kołecki, 2008). This relationship between forensic capabilities and the needs of practical crime fighting is one of the criteria for the assessment of the development of forensic science. In professional journals, one can read that the greatest value of forensics lies in applying a holistic approach to particular cases, by providing a contextual contribution both to investigatory proceedings and evidence (Black, Nick Daeid, 2015). Criminal proceedings entail the need for scientific methods of crime detection and identification of criminals. The need for continuous improvement of equipment, forensic techniques and research methods results, however, not only from the development of crime, but also from overall technical progress taking place in our time. Without a doubt, the development of modern technology is an ally of forensics. An example here (one of many) is DNA analysis, frequently used to support forensics and causing it to constantly evolve.

It is commonly accepted that the implementation of genetic identification was, without a doubt, the greatest achievement of modern forensics (Moszczynski, 2015). Indeed, the DNA analysis-based forensic identification method has secured its position within just a few years. The authorities responsible for criminal proceedings (investigators, prosecutors, judges) generally agree on the significance of certain types of forensic expertise, i.e., genetic tests are perceived as the most valuable (Zubanska, 2016). Genetic expertise has made its way into the courtroom – says T. Tomaszewski – and it is currently difficult to imagine the identification of crime perpetrators without it (Kwiatkowska-Wojcikiewicz, 2010). It is one of the most valuable types of evidence (Moszczynski, 2015). In other words, the position of a genetic expertise is well established in criminal proceedings. Is it therefore necessary to further improve this method of forensic identification? Definitely yes, as genetics has to offer forensic practice much more. Therefore, genetic methods of biological material identification have been constantly improved, which in turn opens up new possibilities for both the experts and law enforcement and judicial authorities.

For some time, a new research direction has been developed, the so-called forensic DNA phenotyping. To determine the person's appearance on the basis of DNA analysis is one of the biggest challenges faced by contemporary forensic genetics. According to a criminologist Manfred Kayser (Erasmus University Rotterdam) the testimony of an eyewitness can often lead astray (lets recall that the value of witness testimony was already questioned many years ago by Edmond

Locard). However, in the case of phenotyping, DNA itself becomes a biological witness of a crime. It seems that science has never been so close to achieving the abovementioned objective. The development of modern technologies, for example the introduction of an innovative method of Next Generation Sequencing (NGS), opened up wider possibilities for forensic genetic analysis. According to common opinion, the NGS has become a universal and indispensable tool in molecular biology. Compared with the conventional method, the main advantages of this approach consis in applying a large number of high-resolution genetic markers simultaneously, low cost and improved performance, enabling the sequencing of multiple samples at the same time. Another advantage of the NGS is the ability to distinguish the differences in individual sequences of allelic variants, which increases the potential of markers to be analyzed (Branicki et al., 2014). It can be expected that in the nearest future the solution to such important issues as the analysis of mixtures and trace amounts of degraded material will be found. New technologies enable an analysis of multiple markers, aimed at identifying important biological characteristics, including age, biogeographical origin, tissue type and the external physical characteristics. A number of studies have been performed on the use of the NGS for the purpose of legal proceedings. However, the implementation of this technology for routine use in forensic practice is conditioned upon the fulfillment of certain requirements, hence research must be continued (Sanaa, Dalia, 2015). Enough to mention at this point that the satisfaction of forensic scientists from transforming scientific discoveries into the judicial instrument does not always raise the enthusiasm among lawyers, which should be borne in mind (Kwiatkowska-Wojcikiewicz, 2010).

In 2016, a research project entitled: "*A genetic portrait of the offender and the victim – developing a system to determine human physical appearance and biogeographical origin, based on DNA analysis, involving Next Generation Sequencing (NGS)*" – the acronym NEXT, was started. The primary objective of this scientific undertaking is to develop a system for predicting the selected features of human physical appearance and biogeographical origin, which shall be capable of performing technical analysis of genetic predictors (hundreds of DNA variants) by means of the NGS method as well as mathematical interpretation of the obtained test result by means of a computer software containing mathematical predictive formulas, developed for this purpose. Achieving the primary objective of the project requires, inter alia, the associative analyses to be conducted on a group of subjects with defined, different phenotypic characteristics (related to human physical appearance). The success of the above task is determined by the possession of suitable study material. It must be remembered that, in accordance with the basic principle of statistical inference, in order for the sample-based population inference to be valid, the test sample must be representative for the entire population, for which conclusions are to be formulated From the point of view of studies on phenotype prediction, a carefully organized and well planned data acquisition process is one of the key elements determining the correctness of the further course of research. In the NEXT project, this stage is carried out as part of the task no. I.2. The objective of this task is to collect detailed information on the origin and selected anthropometric parameters and biological (blood) samples from a representative group of 1000 Polish people. Data collection procedure consists of the following actions: filling the questionnaire, anthropometric measurements, spectrophotometric measurements, taking photos and collecting a single blood sample (5 ml). It should be mentioned that the elaboration of the questionnaire was a task of one of the consortium members, i.e. the research team from the Jagiellonian University. A condition *sine qua non* for the execution of the above task was the approval of the Bioethics Committee. A positive opinion on the proposed study was expressed by the Bioethics Committee at the Jagiellonian University, under the obligation, inter alia, to obtain a written consent of each person willing to participate in this experiment. A special leaflet has been elaborated, containing synthetic information about the project. Regardless of getting familiar with this material, persons who express their will to participate in the study are informed about its purpose as well as about the security of sensitive data, course of the study, possibility to ask questions (where necessary) and finally, the possibility to withdraw their voluntary consent for participation, at any stage, without giving any reason. The participation in the study does not involve any risk, which also shall be communicated to the persons concerned. In the organizational model adopted, the first stage of the material collection procedure consist in signing the necessary forms, in which a person declares his/her consent to participate in the project, involving genetic testing and the processing of personal data (to the extend justified by the project assumptions). After the completion of these important formalities, the investigator shall fill out the questionnaire, whereby no personal data of the participant are entered. It was assumed that the participants are assigned codes, according to the following formula: NEXT-0001 (the code of the first participant) to NEXT-1000 (the code of the last participant), respectively. The participants thus remain anonymous. The questionnaire takes into account information regarding gender, age, biogeographical origin, anthropometric parameters such as height, weight, eye color, eye defects, hair color, degree of curling, skin lesions, earlobe (attached, unattached, partially attached). The next stage of the procedure is to take craniofacial measurements, using

Fig. 1. Bow compass, large, for anthropometric measurements Source: author's own collection.

anthropometric compass (fig. 1). These measurements include:

– distance between the eyes,
– the width of the mouth,
– the width of the nose,
– the width of the face,
– distance right eye–nose,
– distance left eye–nose.

Next, the investigator proceeds to measure a person's skin and hair color. For this purpose, a KONICA MINOLTA CM-600d portable spectrophotometer is used (fig. 2). The results of anthropometric spectrophotometric measurements are entered into the questionnaire form. High-definition photographs are taken with a Nikon D5300 camera (with lens Nikkor AF-S Micro 60/2.8 D ED) placed on a tripod, supplemented with a set of wireless flash lamps for macro photography (fig. 3). In further studies, they will allow an assessment of such parameters as: eye color (pixel index), hair color, degree of baldness (androgynous type), degree of earlobe attachment and hair morphology. A set of photographs

Fig. 2. KONICA MINOLTA CM-600d spectrophotometer.

Fig. 3. Nikon D5300 camera on a tripod, supplemented with a set of wireless flash lamps to perform macro photography.

consisting of five shots includes: right facial profile, top of the head, right eye, left eye, and forearm images (fig. 4-8). The last step of the data collection procedure is collecting a blood sample, which is labeled with the same code as that assigned to the participant in the questionnaire (fig. 9). It should be noted that the blood is collected in the treatment room by a qualified person (medical staff).

As already stressed, the stage of gathering study material is crucial as an incorrect sample selection can impact the entire study. Let's recall, that the NEXT project assumed that a group of sample donors will count 1000 people from throughout the country, ensuring the representativeness of the data collected. Meeting this assumption therefore requires reaching 1000 people, who after hearing information about the project will express their willingness to participate in such a scientific undertaking. This is the basic difficulty of the task I.2. The persons concerned must be provided, in a way that is accessible, with comprehensive information about the purposes and effects of the project, as well as the usefulness of the method to be developed. The investigator must first of all explain the role currently played by genetic testing in criminal proceedings, and then underline that genetics has much more to offer, which is to be demonstrated by the study in question. Moreover, the importance of data provided by persons who agree to participate in this project must be emphasized. During a series of interviews, it became apparent that already such terms as: forensic studies, genetic testing or genetic portrait of a person – the latter particularly appealing to the imagination – make a significant impression on the interlocutors. Generally, an acceptance towards genetic testing was noticeable. For comparison, it is worth to note that judicial authorities are equally positive about this type of testing, as indicated by the results of many studies (Achrem, 2012). It is argued

Fig. 4. Photo of the right facial profile.

Fig. 5. Photo of the top of the head.

Fig. 6. Photo of the right eye.

Fig. 7. Photo of the left eye.

that we are currently dealing with a phenomenon that is known as fascination with genetic testing method or enchantment with its scientific nature. In the literature, there are many examples which would absolutize evidence from DNA analysis – says J. Wojcikiewicz (2000) – while R. Jaworski (1999) believes that, both in theory and in practice, the methods that make an impression of absolute certainty on the authorities responsible for criminal proceedings are favored. Obviously, the above refers to genetic testing. This section of the discussion can be summarized as follows: the perception of genetic testing is positive.

The subject of interest for the team collecting research material was the motivation of persons who have agreed to using their data in the NEXT project. Consequently, a subgroup of 160 study participants were asked the following questions:

1. What prompted you to participate in the NEXT scientific project?
2. Would take part in another research project related to forensic genetics?

The composition of a group of respondents is shown in figures 10 and 11.

The first question was answered as follows: (questioned persons had the opportunity to give more than one reason):

– 101 people (63%) answered that it was a will to participate in scientific progress.

It is worth to add a comment to this result. Namely, among the respondents a large group were students of internal security and criminology. Therefore, it can be

Fig. 8. Photo of the inside of the forearm.

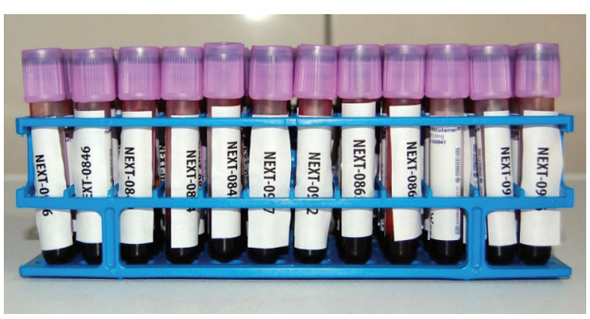

Fig. 9. A batch of blood samples labeled with codes.

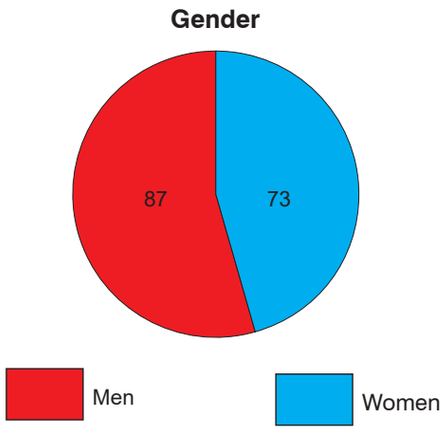

Fig. 10. Distribution of respondents by gender.

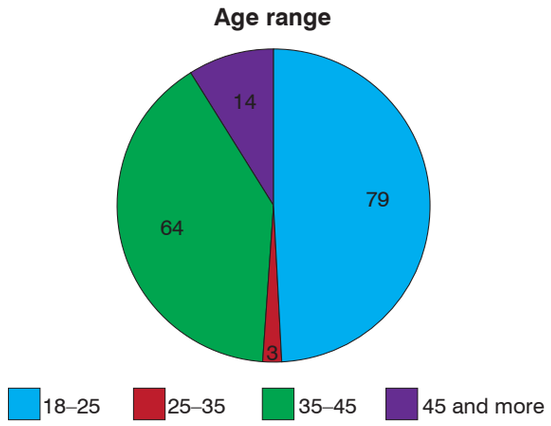

Fig. 11. Distribution of respondents by age.

concluded that by participating they wanted to support research that will contribute to the development of forensics.

– 62 people (about 39%) have become interested in this research project after reading the leaflet;
– 61 people (38%) stated that the decisive factor were their expectations related the advent of a new tool for personal identification;

Such a motivation was given by the group of respondents ages 35–45 and 45 and more. Based on that, it can be concluded that these individuals have a high level of awareness and knowledge of the factual needs of the entities responsible for fighting crime.

– 47 people (29%) reported that their reason was the opportunity to identify genes responsible for appearance characteristics;

It should be noted that such information was contained in the leaflet that was elaborated for study participants.

– 30 people (about 19%) responded that they were encouraged by the person, who had already participated in the study;
– 22 people (about 14%) were guided by the desire to expand their knowledge;
– 15 people (9%) – gave their interest in genetics as the decisive factor.

The respondents' answers to the second question, which concerned the possible participation in future research projects related to forensic genetics, are summarized in figure 12.

It can be easily seen that the vast majority of respondents, a total of 156 persons (97.5%), answered in affirmative (yes and rather yes). It can be therefore concluded that the NEXT project is positively perceived by study participants, who have become convinced of the usefulness of the elaborated method. Only 4 respondents (2.5%) answered that in the future they would not take part in such studies.

In conclusion, it can be once again reiterated that genetic testing is met with a positive public perception. A certain awareness and social responsibility of persons who choose to participate in such studies, contribute to the development of a tool for supporting law enforcement activities. In view of the general acceptance towards such undertakings, a question can be raised about the implementation of regulations into Polish legislation, concerning the scope of interference by authorized entities in the phenotypic data derived from the coding regions of DNA. The discussions about the accessibility of phenotypic characteristics to the criminal justice system are currently in progress in most of the developed countries of the world. It is stressed that, on the one hand, such knowledge may lead to the restriction of civil liberties, but on the other, it may increase the investigational capability of law enforcement authorities (Branicki et al., 2008). This problem, due to its multifaceted nature, is extremely difficult to solve. However, based on the answers received (even though the size of the group was limited), it can be concluded that nowadays, it is important for the community to be provided with a sense of security, and forensic genetics can undoubtedly contribute to this.

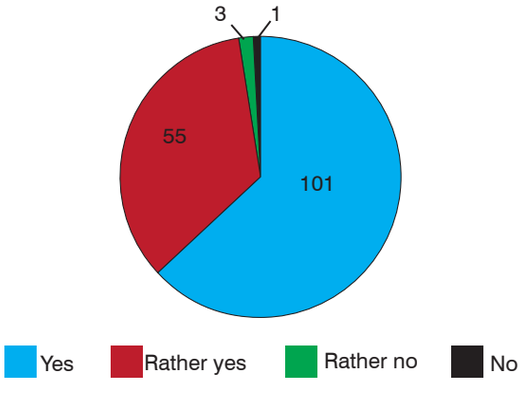

Fig. 12. Distribution of answers to the question: "Would you take part in another research project related to genetics"?

One cannot disagree with Tadeusz Widla, who claims that research conducted in the field of forensics is more applicable than theoretical. These activities consist in developing appropriate tools that can be used by the competent bodies and instances in order to seek truth in an efficient and objective manner (2005). This statement corresponds in full with the objectives of the NEXT project. Therefore, one can actually expect that the tool to be developed will be useful and it that will expand the existing selection of forensic methods. It may allow obtaining data on physical appearance and biogeographical origin of an unknown person, based on his/her biological sample (biological traces, bone material). Such data will significantly step up investigatory proceedings (for example, it will be possible to narrow down the circle of persons who may be related to a certain crime). Currently, DNA analysis does not always provide comprehensive or useful information, for example in cases where there is no investigative hypothesis formulated, regarding the identity of persons who may be related to a certain crime. In such a case, DNA extracted from biological traces or human remains becomes useless as an identification tool. The method being developed can also be used to test samples of unknown origin (unidentified biological traces from the crime scene) or to identify missing persons or disaster victims. Without a doubt, the outcomes of the NEXT project will significantly contribute to the development of forensic studies. Figuratively speaking, it can be assumed that these results will become the so called milestone for forensic practice. In addition, it is worth mentioning once again that already at the beginning of the project, it has been observed that social perception of such studies is positive, which is very important from the perspective of subsequent research projects of this type.

Sources of figures:

Figures 1–3: authors
Figures 4–8: phenotypic data collection protocol – NEXT. The team of Prof. Wojciech Branicki, Jagiellonian University, 2016
Figures 9–12: authors

Bibliography

1. Achrem, W. (2012). Opinia biegłego z zakresu badań genetycznych świetle analizy rezultatów badania ankietowego. Moc dowodowa, wiarygodność i przydatność naukowego środka dowodowego do dowodzenia wybranych rodzajów przestępstw. *Problemy Kryminalistyki*, *282*, 9.
2. Black, S., Nic Daeid, N. (2015). Time to Think Differently: Catalysing a Paradigm Shift in Forensic Science. *Philosophical Transcations Royal Society B*, *370: 20140251*. http://dx.doi.org/10.1098/rstb.2014.0251 [accessed: December 2016]
3. Branicki, W., Kupiec, T., Wolańska-Nowak, P., (ed.). (2008). *Badania DNA dla celów sądowych*. Cracow: Institute of Forensic Research Publishers.
4. Branicki, W., Pośpiech, E., Kupiec, T., Styrna, J. (2014). Nowy wymiar ekspertyzy DNA – potrzeba szkoleń ekspertów i odbiorców ekspertyz. *Archiwum Medycyny Sądowej i Kryminologii, 64* (3). http://www.termedia.pl/Nowy-wymiar-ekspertyzy-DNA-potrzeba-szkolen-ekspertow-i-odbiorcow-ekspertyz,82,24543,1,0.html [accessed: December 2016].
5. Jaworski, R. (ed.). (1999). *Opinia z ekspertyzy poligraficznej jako dowód odciążający*. Wrocław: Kolonia Limited.
6. Kołecki, H. (2008). Zakres i sposób uprawiania kryminalistyki w Polsce. In: H. Kołecki, (ed.), *Kryminalistyka i nauki penalne wobec przestępczości*. Poznan: Publishing House of Poznań.
7. Kwiatkowska-Wójcikiewicz, V. (ed.). (2010). *Kryminalistyka dla prawa. Prawo dla kryminalistyki*. Torun: „Dom Organizatora" Publishing House
8. Moszczyński, J. (2015). Daktyloskopijna i genetyczna identyfikacja osób – „zderzenie" dwóch złotych standardów. In: B. Sygit, (ed.), *Golden Anniversary Book presented to Prof. Brunon Hołyst, awarded Honoris Causa doctorate*. Lodz: University of Lodz Publishing House.
9. Sanaa, M. A., Dalia, M. S. (2015). Sekwencjonowanie nowej generacji (NGS): „złote narzędzie" w arsenale metod kryminalistycznych. *Archiwum Medycyny Sądowej i Kryminologii, 65*(4), 260.
10. Wójcikiewicz, J. (ed.). (2000). *Dowód naukowy w procesie sądowym*. Krakow: Institute of Forensic Research Publishers.
11. Z profesorem Tadeuszem Widłą spacer po kryminalistyce. (2005). *Gazeta Uniwersytecka UŚ, 10*(130). http:// gazeta.us.edu, [accessed: June 2016].
12. Zubańska, M. (2016). Badania daktyloskopijne i genetyczne – TAK, badania poligraficzne i osmologiczne – NIE, czyli o postrzeganiu ekspertyz kryminalistycznych przez decydentów procesowych. In: J. Moszczyński, D. Solodov, I. Sołtyszewski, (ed.), *Przestępczość. Dowody. Prawo. Anniversary book presented to Professor Bronisław Młodziejowski*. Olsztyn: Publishing House of the University of Warmia and Mazury.

Translation Rafał Wierzchosławski

Project No. DOB/BIO7/17/01/2105 financed by the National Centre for Research and Development for security and defence.

Narodowe Centrum Badań i Rozwoju