

Management summary

Use of DNA databases as part of a root search

In this study, a weighing framework has been drawn up that can be used by the Expertise Centre for Intercountry Adoption to support adoptees in deciding whether to use commercial international DNA databases. This assessment framework includes all relevant technical, legal and social criteria that can or must be taken into consideration when determining whether and which DNA database should be chosen when using DNA technology in a root search - the search for (close) biological relatives by an adoptee.

DNA technology can be used to supplement the administrative trail within a root search. If adoptees' documents turn out to be incomplete or incorrect, the use of DNA matching is often the only option to find relatives.

DNA testing can be done through commercial DNA databases. The result of that test (the DNA profile) can then be used to search one or more databases for biological relatives. This is done by looking at the degree of overlap between one's own DNA profile and that of people already in the DNA bank's customer database. The greater the amount of shared DNA, the higher the degree of kinship.

In a root search, the most important result of a DNA test is the matches found in the database. With each additional degree of kinship, the volume of people also increases. For fourth-degree relatives (second cousins), this number increases exponentially and, moreover, without DNA testing of other relatives or extensive 'paper trails', it is almost impossible to reconstruct a family tree that starts so far away. Therefore, it is often said that anything further away than a second-degree connection (brother/sister, grandparents) is not useful in a root search.

Because most matches are several degrees away from the adoptee's parents, the root search often does not end with finding a match. Therefore, as a result of a match, there may be a desire to take DNA from possible relatives in the country of origin for a DNA test to reconstruct the family tree and find possible relatives who are closer. An important question here is whether these people are willing and/or able to cooperate (voluntarily) (e.g. whether they fully understand what the DNA tests entail, and what the consequences may be). The test can reveal information that the user and/or his or her family members would rather not have known and that can drastically change their lives. It may be that the sibling of the seeking adoptee does not want contact with his or her parents and actually wants to remain anonymous or hidden. If the seeking adoptee manages to make contact with the parents, it may become much more difficult for the sibling to remain anonymous.

Selected DNA databases in the study

Based on the desk research and the interviews, we created a generic framework. This framework includes all relevant technical, legal and social assessment criteria that can be considered when determining if and which DNA database should be selected when using DNA technology in a root search.

The framework was applied and tested on six major international DNA databases: Ancestry (US, 21 million DNA profiles), 23andMe (US, 12 million profiles), MyHeritage (Israel, 6 million profiles), FamilyTreeDNA (US, 2 million profiles), GEDmatch (US, 2 million profiles), Living DNA (UK, 0.3 million profiles). The selection of the databases was based on four criteria: (1) the DNA database must offer the functionality to match DNA profiles; (2) the DNA database must have sufficient critical mass; (3) the DNA database must offer a possibility to match DNA with other profiles; (4) DNA profiles must be able to enter the DNA database from the Netherlands. Based on (1) and (2), the

Indian MapMyGenome was not included, based on (4) the Chinese databases 23Mofang and WeGene were not included.

DNA databases often offer other services besides 'matching'. Ancestry analyses have little or no value from a scientific point of view: the results are simply too unreliable. This is because the quality (representativeness) of the reference populations used for the analyses is generally (very) low. This applies to an even greater extent to people with origins outside Western Europe. Medical analyses are separate from the root search but can, for example, provide the adoptee with information about any genetic anomalies that exist in the biological family.

The framework

The framework relates only to DNA matching. The framework consists of three parts. These relate to three different subjects that play a role in the use of a (commercial) DNA database: technical aspects, legal aspects and social aspects. These aspects in turn consist of a large number of indicators: measurable phenomena that say something about the quality (or lack thereof) of a DNA database in that specific aspect.

Technical aspects

DNA analyses are done on different types of DNA. This results in different types of tests, each with a specific function: Autosomal DNA testing, X-DNA testing, Y-DNA testing and mtDNA testing. All DNA databases offer autosomal DNA testing. Because the other types of testing cannot confirm a specific (recent) family relationship, these tests generally add little to the standard autosomal test.

It is almost unheard of for autosomal DNA tests to find a match, if the persons are not actually related (false positive). In a search, however, a person with whom the match is so small is irrelevant, because that putative relative would then be too far away for a search to be initiated. The possibility of missing a family member (false negative) is more common. These false negatives are usually not caused by technical errors, but by the way DNA is passed on. Because the splitting of DNA is not always 50/50, one may not share DNA with people who are indeed biologically related. Again, however, this only applies to relatives who are distant, and who would also likely be irrelevant to starting an international search.

Some databases also offer the ability to see exactly on which chromosomes the match occurred. This can be helpful in discovering the specific relationship. These common matches can help in the construction of a family tree because certain matches can be linked and family structures can emerge. The proper interpretation of these results does require specialized knowledge.

The DNA sample can be collected via spit or via a swab of the cheek mucus. Which method is chosen by the DNA database will affect the availability of the tests in the country of origin. DNA databases that use a swab can ship their kits to almost all countries, while this is not the case for databases that use spit. This has to do with transporting the sample and local laws. As a result, these tests are often only available in Western countries. This indicator is an important consideration in choosing a DNA database, because the chances of finding a relative are lower if a database does not send tests to the (presumed) country of residence of relatives.

When a DNA profile is uploaded to a database, it is compared to the profiles already in the database. When using a DNA database to perform an untargeted search for possible matches, the probability of finding a match depends in part on the numbers of profiles in the database: the more profiles, the greater the probability of a match. With some databases it is also possible to upload raw DNA data from other companies/databases. In this way, with one test, the DNA profile can be included in multiple DNA databases.

To increase the chance of a match, it is important to not only ensure that there are enough profiles in the database used, but also that the right profiles are in the database. Thus, it is important to find

out what the population of a database is. The focus of the profiles in the databases is on individuals with European ancestry and these are often not the relevant profiles for intercountry adoptees. The percentage of the population present from the country of origin is extremely low in most cases. Although the exact numbers are not known, the necessary number is currently far from being reached in any (large) DNA database. For an adoptee from a non-Western country, the chance of a match is therefore extremely small, regardless of which database is chosen.

Legal aspects

In assessing commercial DNA databases, the focus was on how these databases handle the DNA of their users. A total of 60 variables were identified. From the 60 variables, a core set of 24 indicators was finally selected. Most of these variables are an elaboration of the data processing principles from the AVG, the overarching European privacy law.

The AVG also applies to companies outside the EU when they process data of EU citizens. This therefore applies to all selected DNA databases. When data is transferred outside the European Economic Area (EEA), for example when biological material or DNA data is sent to laboratories or repositories in the United Kingdom or the United States, there is no immediate guarantee that the data is as well protected as it is within the EEA. In the case of none of the selected databases, it is currently possible to exclude the possibility that such data may leave the EEA. However, the AVG does provide a number of ways to still ensure a comparable level of protection. One way to do this is through so-called standard clauses. In 2021, the European Commission approved a model agreement for the processing of personal data by a processor on behalf of a controller in the EEA. Currently, four of the six databases use a model agreement.

With regard to privacy, users should be given a full picture of the data being collected about them before they agree and send their data. It must also be clear what security measures are taken and for what purposes the various data are collected. The extent to which the data are actually secure is difficult to verify because the selected databases provide little concrete information about their security measures. However, all six DNA databases have appointed a Data Protection Officer. According to the AVG, this is an essential precaution to ensure responsible data management, which covers the entire data management process.

For all the different ways that DNA data can be used by the database itself, or possibly by a third party, explicit prior consent must be obtained. Users must be able to opt out and opt in at any time by different purposes. The databases must also provide a complete list of the parties with whom the data will be shared, why these parties receive the data, and under what conditions. Investigative agencies (such as the FBI or the NSA in the US) can force DNA databases (through a court order) to surrender DNA data. Some databases publish requests they have received, what types of organizations the requests came from, and how many requests were actually fulfilled.

Personal data should not be kept longer than necessary to achieve the set purposes. As long as the user still wants to participate in matching, it is obviously necessary to keep most of the data, and at least the DNA data. All databases therefore retain DNA data until the user deletes his or her account or requests that the data be deleted. Specifically for the biological material, (the DNA sample), this is data that is sensitive, but in principle not necessary to keep after the material is tested and the DNA data is stored. It is therefore good practice to offer the user to destroy the biological material after the test results have been received. In practice, however, we see that this rarely happens, the user himself must (proactively) submit a request to destroy the material.

The biological material always remains the property of the user but the information derived from that material (the DNA profile) can become the property of a DNA database. Databases are required to clearly inform their users about the right to inspect, the right to rectify data, the right to delete data and the right to complain to a supervisor. It is desirable that users can download their own DNA

profile after their biological material has been tested. All databases that offer testing currently also give their users this option.

All six DNA databases described in this report meet the minimum requirements for responsible and safe handling of DNA material and biological material. This does not apply to the two Chinese databases (which were not included in the selection).

Social aspects

It varies from one database to another how much attention is paid to the topic of adoption. Most DNA databases do not have a primary focus on supporting root searches. While this is primarily a marketing choice of the DNA database and it does not affect the quality of the test, it does potentially affect the number of adoptees using the service with the result that the chances of a match may vary.

The main source of information for users is the websites of the DNA databases and the reports of DNA results. The websites provide technical and practical information as well as content designed to persuade people to order a DNA test. With the exception of MyHeritage, all DNA databases in the selection communicate entirely in English.

All selected DNA databases provide very comprehensive information about DNA technology, but the information about the (complex) technical matter is not always easy to understand. All major DNA databases are commercial companies and they have an interest in presenting the results in a positive way. This is done, for example, by defining the term "relative" broadly. The probability of finding a fourth- or fifth-degree relative (second cousin/nephew) is fairly high. However, this 'good news' is of little use to intercountry adoptees (see above). Conversely, it is not always desirable to be approached by distant relatives. It contributes little to the root search and can take a toll emotionally. However, none of the selected DNA databases currently offers the option of indicating up to what degree (e.g., only second or first degree) one wants to be approached.

Once a "match" is found, contacting that unknown (alleged) family member can be an exciting affair. Two of the six DNA databases themselves provide (limited) advice on how best to go about this process. Most databases refer to professionals for psychological help. Two other databases offer indirect support by facilitating online communities in which users can share experiences.

Reconstructing a family tree can be part of a root search. Many DNA databases offer explanations and/or services for this, such as lists of DNA matches and the display of DNA test results in a family tree, support by professional genealogists, or digital archives of administrative documents (such as population registers).

Key conclusions & recommendations

- The framework we developed in this study is generic in nature. It can be used to assess any DNA database in the past, present or future.
- Using the framework, an informed choice can be made about using international databases, weighing the benefits (a very small chance of finding a close relative) and costs (possible privacy and security risks, open-endedness, third party interests).
- With the exception of the databases located in China, all six databases described in this report meet the minimum requirements for a responsible and secure way of handling DNA material and biological material. However, there will always be risks associated with the use of DNA databases.
- In technical and legal terms, the databases differ little, but in social terms they do.

- The number of DNA profiles in the database is the most important factor determining the chance of finding a match with a (close) relative. For each database, the number of profiles from non-Western countries is currently so low that the chance of finding a match with someone from the country of origin for adoptees from non-Western countries is extremely small.
- However, the country of origin can have an impact on the chance of finding a match. In fact, there are organizations that actively go to countries of origin to collect DNA profiles there for the purpose of uniting families and adopted children.
- Since the databases are little different from each other in technical and legal terms, the obvious solution is to upload profiles to all of them. After all, the chance of finding a match is greatest when comparing with as many different databases as possible.
- The rights and risks differ per situation and per country of origin. Advice will therefore always remain tailor-made. Close cooperation with relevant interest groups is therefore obvious.
- The aftercare currently provided by DNA databases is very limited. This is a gap which the Expertise Centre - again in close cooperation with the interest groups - could fill.
- All major DNA databases are commercial companies and they have an interest in ensuring that the results are presented positively. There is an important task for the Centre of Expertise to provide professional (counter) expertise in an independent and thorough analysis of the results. For the legal aspects, explaining the terms of use in understandable language is a particularly important task. These terms of use are also subject to regular changes. The legal experts of the Expertise Center should be aware of the most recent changes.