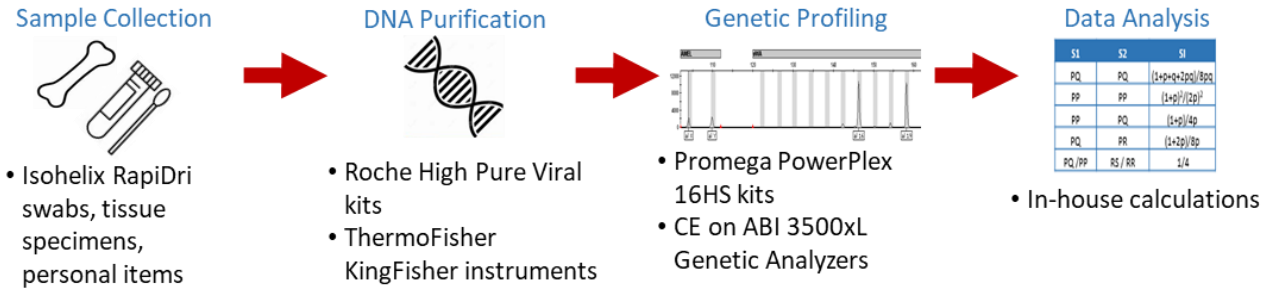
Micropathology Ltd
The Venture Centre, University of Warwick Science Park, Sir William Lyons Road, Coventry, United Kingdom, CV4 7EZ
www.micropathology.com ✉ info@micropathology.com ☎ +44(0)2476 323222

DNA Technology in the identification of Human remains

Summary
We work with organisations, such as HM Coroner's offices, for the identification or re-unification of human remains. The technique we employ for this assay is based upon obtaining a unique DNA profile from up to 15 highly variable, independently segregating loci from across the human genome. Once a DNA profile has been obtained from a set of human remains, the identity of the deceased may be confirmed or excluded by comparison with profiles generated from known reference samples or from known parents/children of the deceased. By analysing 10–15 unlinked genetic markers, each with high variability, the probability of a coincidental match can be reduced to as low as 1 in billions. For more detailed information, including a more in depth discussion of the techniques and statistical interpretation, please see the explanations below. The general process is as follows:



Sample types- UKAS accredited sample types for this assay are serum, blood samples (treated with EDTA or citrate anticoagulant), tissue, mouth-swabs, personal items, and FFPE specimens. We can process almost any biological specimen type that will yield human genomic DNA (except where there is a suspicion of a sample containing a category 4 organism), however non-accredited sample types will be reported with a caveat stating as such.

Reference samples- The preferred samples are from parents/children of the deceased. Siblings can provide useful information however identity can only be confirmed with more modest confidence and in some instances do not provide conclusive results. We do not recommend sending samples from more distant relatives such as cousins. Testing profiles retrieved from personal items known to belong to the deceased can provide a useful comparative reference.

Turnaround time - The stated turn-around time for this assay is 10 days (15 days for bone samples). However cases may take longer than this if samples are received at different times. Please see our Laboratory User Handbook for a full list of turnaround times. Please note, it may be necessary for us to request further specimens for genetic profile analysis in circumstances where our investigations are inconclusive.

Cost - There is a basic cost per sample for generating and analysing a forensic DNA profile - please register on our Access Portal (https://www.micropathology.com/Contact) to refer to our price list. However, additional surcharges may be levied depending upon the nature of the sample provided and the analytical report format required. Skeletal remains and teeth, for example, require considerable extra processing.

Consent - It is the responsibility of the clinician requesting a genetic test to obtain informed consent for testing from individuals providing DNA for reference purposes (or an individual with parental/legal responsibility for this person).

Remote mouth swab collection request - To request a remote sample collection, please email us a completed HID request form (see details below). Including the following details in the email ensures we can match the mortuary sample with the relative's mouth swab:

- Name of the deceased
- Name and address of the next of kin
- Your case number

We will send a mouth swab kit and covering letter to the next of kin. They will collect their own sample, complete the request form with the date of collection, and return everything to us using the prepaid envelope provided. We do not charge for this.

Contact us if you have any questions, please don't hesitate to email us at info@micropathology.com or phone us on +44 (0) 24 7632 3222 and ask for a member of the genetics team.

A downloadable Human Identification (HID) Request Form is available on our website (see: https://www.micropathology.com/Home/DownloadPDF?fileName=HID_Request_Form.pdf)

Please note we do not undertake work related to ongoing criminal investigations.

This is a UKAS accredited service (ISO15189).

Introduction

DNA consists of a long string of chemical units called nucleotides. There are four different nucleotides often just referred to as 'A', 'C', 'G', and 'T', which may be thought of as four different coloured beads threaded on a long string. In humans, the 'piece of string' has about 3 billion beads in total (known as the genome), and is divided or packaged into smaller pieces called chromosomes (of which there are 23 unique forms). Every individual has two copies of each of these 23 chromosomes, one inherited from their mother and one from their father. Although on a grand scale one person's genome is almost identical to another's, there are many much smaller regions (known as loci) scattered throughout it that can exist in a number of different forms (said to be polymorphic). It is these polymorphic loci that make us all different and allow a unique DNA profile to be generated for everyone (except identical twins). Furthermore, at every locus, this profile must contain one of the two forms (called alleles) present in each of that individual's parents. These are the basic principles of DNA based identification, and are illustrated in the following example:

At a single polymorphic locus, ten forms (each possible form is known as an allele) called A1 through to A10 are known to exist in the population. If an individual's mother has two copies of allele A1 (known as homozygous A1, A1/A1), and the father is homozygous A6 (A6/A6), any child of theirs must have the genotype (combination of alleles) A1/A6 at this locus. If DNA isolated from the remains of an individual believed to be a child of this couple does not show this genotype then their identity may be excluded. If the DNA does match this genotype, then identity as a child of this couple cannot be excluded. This does not however mean that identity is confirmed, since clearly there will be some proportion of the general population that will also possess the genotype A1/A6 by chance. This could be 20%, meaning that we have actually only excluded 80% of the entire population from being the deceased. If we are lucky, and both the alleles A1 and A6 are rare in the population (say 10% or 0.1 each), meaning that the genotype A1/A6 is even rarer (0.1 x 0.1 = 0.01 or 1%), then this will greatly increase our confidence that the deceased is whom we believe them to be (as we will have excluded 99% of the population).

The above example begins to introduce how a degree of confidence or likelihood ratio can be attached to genetic data. However, it should be obvious from this that a single polymorphic locus (or genetic marker as it may also be called) is never really going to be enough to 'confirm' identity, even if we are lucky with rare alleles. Excluding 99% of a population of 10 million still leaves 100,000 people! The solution is to use multiple genetic markers. These markers are selected from different chromosomes so that their inheritance is random (markers on the same chromosome are more likely to be inherited together). This means that likelihood ratios (expressed as an index, see later) attached to each marker may be multiplied together. Again this is best illustrated by example:

DNA is extracted from a toothbrush confirmed by a woman as having belonged to her husband. DNA is also extracted from burned human remains found in a car wreck. One DNA marker alone provides a genotype that matches between the two samples, but that also matches to ¼ of the general population (i.e. this genotype is not at all rare). A second marker also provides exactly the same match, taken together they will only match 1/16th (¼ x ¼) of the population, thus excluding 15/16^{ths} of the population. A third similar marker will result in exclusion of 63/64^{ths}, and so on until by the time you reach 10 similar markers, the chance (based on the DNA evidence alone) of the individual in the car being anyone other than the woman's husband is only 1 in 1,048,576. In real life around 12-15 markers are used and each one is so variable that it normally excludes far more than ¾'s of population as used in the example above. This means that the likelihood of two genetic profiles matching by chance may be only one in many billions.

Samples to collect

There are two major issues to consider when collecting samples for DNA based identification purposes. The first is purely practical and relates simply to what samples give the most and best quality DNA for analysis. From living individuals, blood samples (treated with EDTA anticoagulant) or

mouth swabs are the easiest to collect, store (may be kept for several weeks in the fridge) and process. From deceased individuals, a small (no more than a few grams) deep muscle tissue biopsy is the sample of choice. Muscle tissue is rich in nucleated cells, which contain the most DNA, and samples from deep within the body will be preserved longest and exposed to the least external contamination. Failing this, any flesh, clotted blood or even teeth and bone can often be used successfully. Tissue that is purely fat, along with small bones/front teeth should be avoided. Molars and bones from the lower limbs, especially large bones of the feet, such as the heel bone, are samples of choice if no soft tissue remains are in existence. With regards to reference samples i.e. those known to belong to the deceased; toothbrushes, dental floss and hairbrushes are often used very successfully. It should be noted that hair may not be suitable unless the roots are still attached. Finally, wherever possible samples should be obtained fresh; pathologists tend to preserve key samples in formalin, and whilst this may be very good at preventing obvious decomposition, it can make subsequent DNA analysis very difficult. If samples may be required for DNA analysis at a later date, try to ensure that a small portion is stored frozen at –20°C or below.

The second consideration regarding samples is more of a statistical matter. Quite simply based on DNA evidence alone, direct comparison to a known reference sample will give the most statistically significant confirmation of identity. Of course the accuracy of this relies entirely upon your confidence that the reference sample is exactly as you believe it to be. Obviously in the example detailed above, DNA evidence would have been no use at all, if the woman, despite categorically stating that the toothbrush she handed over belonged to her husband, did in fact mistakenly belong to the lodger! Unless the lodger was female (which we are able to tell from the DNA profile), this mistake would not be noticeable and the DNA evidence would suggest (incorrectly) that the burned remains were not those of her husband.

If reference samples are not available or confidence in their authenticity is low, then relatives of the alleged deceased are the next best samples to obtain. In these circumstances, ‘completing the trio’ is the phrase to remember. If the deceased is a child, try to get both parents, if they are a father, then try to get a child and the mother. Failing this, try to get one parent and/or all the children. Remember that samples from deceased relatives may still be available in the form of archived samples in hospital laboratories. Only as a last resort should you collect just the siblings of an individual. The reason for avoiding siblings is that in all other cases, identity may either be confirmed, usually with a high degree of confidence, or excluded absolutely. With siblings, identity can only be confirmed with more modest confidence and it can never be entirely excluded. This is because although siblings may reasonably be expected to share DNA more often than you would expect by chance, the simple rules of genetics do not actually require them to share any at all! See the following example:

A couple have two children. At a given genetic marker the mother has the genotype A1/A2, the father A3/A4. Their children can only have the following genotypes, A1/A3, A1/A4, A2/A3 or A2/A4. On one in four occasions the two children will share no alleles i.e. one will be A1/A3, the other A2/A4. In this scenario, and without any knowledge of the parent’s genotypes these siblings would appear as unrelated as two random members of the general population. As the number of markers increases, the chance of this scenario occurring at every one falls, and with 15 markers it is usually possible to say that two genuine siblings are such with about a 90% or greater degree of confidence, yet it remains impossible to prove beyond all doubt that any two individuals pulled from a population at random are not siblings!

Shipping & Documentation

When sending samples for analysis, it is a legal requirement to ensure that they are packaged safely and securely. In simple terms, there must be no chance of anything puncturing through or leaking out of the packaging. This may involve packaging in a rigid box and/or including sufficient absorbent material if any of the samples are liquid. The post office and many courier services produce specialised packaging, which can be requested at a charge. If in doubt, please do not hesitate to contact us for advice.

Samples must always be very clearly/unambiguously labelled and accompanied by some form of paperwork. A downloadable Human Identification (HID) Request Form is available on our website (see: https://www.micropathology.com/Home/DownloadPDF?fileName=HID_Request_Form.pdf) which you may find helpful. The paperwork should make it clear exactly what the samples are, which ones are which (best to include your own reference number with each), where they have come from, exactly what tests or comparisons you would like performed and where you would like the results/invoice sending. It is also very useful to include any hazard information, such as 'subject known to be HIV positive' and relevant background information regarding the sample condition, such as 'underwater for 3 months' or 'preserved in formalin' etc. This is because the history of the sample may affect how it is processed and hence the subsequent success/failure of DNA extraction. It is also helpful to indicate the urgency with which results are required. We routinely send our reports to you as soon as they are approved by email, so please provide a secure email address that can be accessed by anyone from your office authorised to do so. Finally, you should refrain from providing too much other information as this could theoretically compromise confidentiality or the impartiality of the laboratory staff.

Reports & Interpretation

Once we have completed the requested analysis, you will be sent a DNA based report (the exact format will depend on the type of analysis you have requested). This will cite a summary of the sample details including some or all of the following details: Names, DOB's, your sample reference number (if supplied), our own unique laboratory sample reference number (which you should have to hand if you phone us with any queries) and the date the samples were received/tested. You should check that all these details are exactly as you believe they should be.

The main body of the report will list all the genetic markers that have been used. These will have names like D8S1179, CSF1PO, Tpox etc. Next to the markers the alleles present in each of the samples will be listed. These are usually just whole numbers e.g. '8', '11' etc. although occasionally they will have decimal points in, such as '31.2'. It is not important to understand what these numbers mean, simply that they should be the same if two profiles match or follow the simple rules of inheritance. For example if two parents are 8/11 and 9/12, a child should be 8/9, 8/12, 11/9 or 11/12, anything else is incompatible with their being a child of both parents. Finally some form of 'Index' figure will be quoted for each marker, this will depend on the hypothesis being tested, PI would be a Paternity Index, MI would be a Maternity Index, SI would be a Sibling Index etc. In all cases, the index figure is a product of the alleged relationship between the samples/individuals, the combination of alleles present in each sample/individual, and the frequency of these alleles in the general population. A bigger index figure indicates that the hypothesis being tested is more likely to be correct; a smaller one means it is less likely. An index of zero means that the data is incompatible with the hypothesis being tested. This would be generated in a scenario such as two parents who are 8/11 and 9/12 at a given marker seemingly having a child who is 7/14!

At the bottom of the report, a Combined Index figure will be quoted. This is simply the product of all the individual marker indexes multiplied together, and is ultimately the figure on the report that means the most. However, to those who are not geneticists/statisticians, this combined index can often mean very little! Consequently, it is frequently 'converted' into a percentage probability, which is also quoted. In all cases, this 'conversion' to a percentage probability relies on what is known as an assumption of prior probability of 50%. What this means is that the quoted percentage probability is based on the DNA evidence alone, and as such only tells you about the samples. This is not necessarily the same as the real life situation. The true probability is actually the probability based on the DNA evidence combined with the prior probability. This is best explained if we consider the case of the woman providing her husband's toothbrush again:

If we assume that the police have managed to definitely identify the wreck as her husband's car, furthermore the pathologist confirms that burned body is a male of the same height, build and

approximate age as the woman's husband, plus of course the husband himself is nowhere to be found. We are now in the situation of already having a very high prior probability that the body is that of her husband. Indeed, this prior probability may be so high that DNA analysis may be considered unnecessary. However, let us say that the body is missing an expected wedding ring (draw your own conclusions as to why!) and that this introduces sufficient doubt to warrant DNA confirmation. The analysis is done and comes back as absolutely no chance of a match (because unknown to everyone the woman has mistakenly handed over the lodger's toothbrush). The DNA evidence is not wrong, because it has correctly told you that the body and reference sample are not from the same individual. Because the probability quoted (as 0%) on the report does not consider the very high prior probability, this is not the same as saying the body in the car is not that of the missing husband. In such a case you should clearly not accept the DNA evidence at face value and continue your investigation for an explanation. The authenticity of the reference sample must be considered doubtful and you should consider collecting a sample yourself from the man's parents (if they are still alive). Note that sampling the man's parents rather than his children is suggested preferentially. This is because the possibility that the lodger may have donated more than just his toothbrush to the man's wife must also be considered!

Generally, if the results of genetic analyses support the hypothesis that the deceased is whom you believe them to be, you will not need to consider prior probabilities. This is because the DNA based probability will be so high that you would need hugely contradictory circumstantial evidence to dispute it. Sometimes however, such as when comparing DNA from bones of an individual with a sample believed to be from the individual's father, the DNA based analysis may actually only provide a probability of 90%. This is because the DNA from the bones may be so degraded that it is only possible to get results for half of the 15 genetic markers used, furthermore the missing mother also reduces statistical power by not 'completing the trio'. In such a case, prior probability considerations are essential. If the body were to be found with the deceased's wallet (which could of course have been stolen by someone else, and hence may not be enough evidence alone), this combined with the 90% from the DNA evidence may be considered enough for a definitive identification. However, if the body has absolutely nothing else to indicate identity, a 90% probability (which may sound good) from the DNA evidence actually means that it could be the body of anyone of 10% of the population, about 6 million individuals in the UK!

To summarise, simply remember that DNA based analysis tells you only about the samples you submit for analysis. If you have a high confidence in these samples, then you may have a high confidence in answers you obtain from any DNA based analyses that you commission. If you have any doubts, or feel that you don't know enough to adequately consider what conflicting factors may influence the DNA based results, then please feel free to contact us using the email or phone addresses at the top of this document.