The case of the missing lung: how tissue identity is confirmed in histopathology

Daniel T Field

David A Dorward

Abstract

This case describes a situation of mislabeling of specimen pots that led to a malignant diagnosis being attributed to the wrong patient's surgical resection. Short tandem repeat analysis was used to determine which specimen originated from the two patients involved. This report describes a not infrequent scenario for pathologists and how current genetic techniques can be used to resolve issues of patient/tissue identity.

Keywords DIPplex; genotyping; mislabeling; molecular pathology; short tandem repeats

This Short Case is brought to you in association with the Pathological Society of Great Britain and Ireland. Each month we feature a Short Case written by a member of the Trainees' Subcommittee of the Pathological Society of Great Britain and Ireland. This case, including its scanned slides, is published on the Pathological Society's website https://www.pathsoc.org/training_education/education/diagnostic_histopathology_short_cases.aspx?edit=true (earlier cases can be accessed by members via the education portal login). The Short Case includes a series of Test Yourself Questions at the end to check your understanding of the case. We hope you enjoy reading it.

Pathological Society
Understanding Disease — Guiding Therapy

Case report

The pathology department was alerted by the thoracic surgical team that, 2 weeks post-surgery, no pathology report was visible on the clinical portal for a patient who had undergone a right middle lobectomy and lymphadenectomy for suspected lung cancer (patient 1, female). There was, however, no record of the specimen arriving or being logged into the pathology department, despite confirmation that the operation took place, and that the specimen left the theatre immediately after the procedure. Neither the specimens nor request form were found on detailed search of all parts of the pathology department and theatres.

Subsequently, it became clear that two specimens had been received, reported and attributed to a separate patient on the same theatre list (patient 2, female). She had undergone a right upper lobe bullectomy for severe bullous emphysema immediately prior to patient 1. The first pathology report related to the bullectomy, the second a right middle lobectomy for lung malignancy. A review of pathology request forms and specimen pots determined the same patient identification labels were used in both cases. This was compounded by the specimens being trimmed on different days and reported by different pathologists due to their time of arrival.

With this information, it was presumed that patient's 1 specimen pot and form had been mislabeled with patient 2's details in theatre. To confirm this, DIPplex testing (DNA Intelligence Profile Plus) was undertaken (details below). Previous GI biopsies for each patient were used as reference samples to confirm patient identity. Initial testing confirmed that the bullectomy specimen belonged to patient 2. Unfortunately, the DNA from the presumed mislabeled lobectomy was of poor quality but analysis suggested it belonged to an unidentified *male* patient. Upon further investigation, the analysis showed few amplified markers, and it was noted that fixation was delayed with the specimen remaining fresh at room temperature for 16 hours. Repeat DIPplex testing was therefore undertaken on a separately submitted lymph node that had been in formalin from the time of surgery along with a further reference sample (colonic polyp). Genotyping confirmed that the lymph node had the same Short tandem repeat (STR) profile as both reference specimens and therefore belonged to patient 1 (Figure 1).

Pathology records were updated accordingly, the patients informed of the error and their care proceeded based on the correct histological results. Local incident reporting and review procedures were undertaken in line with hospital policy.

Discussion

Mislabeling of pathological specimens is an infrequent occurrence but one in which the consequences can be severe.[1] DNA identification may be required in such cases, where the patient identity of a specimen is uncertain (e.g. due to mislabeling, lack of labelling) or, more commonly, 'carry over' whereby erroneous tissue is present within a slide that may have diagnostic consequences. The same techniques and principles are applied to areas such as forensic science and transplantation medicine.[2,3] Current concepts in DNA identification derive from work conducted in the 1980's characterizing an individual's genetic 'fingerprint'.[4] STR are an example of a type of polymorphism within the genome. In its simplest form, STR are short sequences (∼2–6) of nucleotides, usually present in introns, that are repeated a varying number of times in tandem between individuals (typically up to around 15–20 times) (Figure 2a). The variation in the number of repeats in the population can be two (bialleic – e.g. only 10 or 11 repeats possible) or more (multialleic –10, 11, 12, 13 etc repeats). The number of times a STR is repeated is dictated by parents and the number of repeats can vary on each allele of a chromosome (e.g. 11 repeats on maternal and 12 on paternal chromosomes) (Figure 2b). By using a range of multialleic STR

***Daniel T Field** MBBS BSc PGCert Trainee Histopathologist, Department of Pathology, NHS Lothian and The University of Edinburgh, Edinburgh, UK. Conflicts of interest: none declared.*

***David A Dorward** MBChB PhD FRCPath Consultant Pathologist, Department of Pathology, NHS Lothian, UK. Conflicts of interest: none declared.*

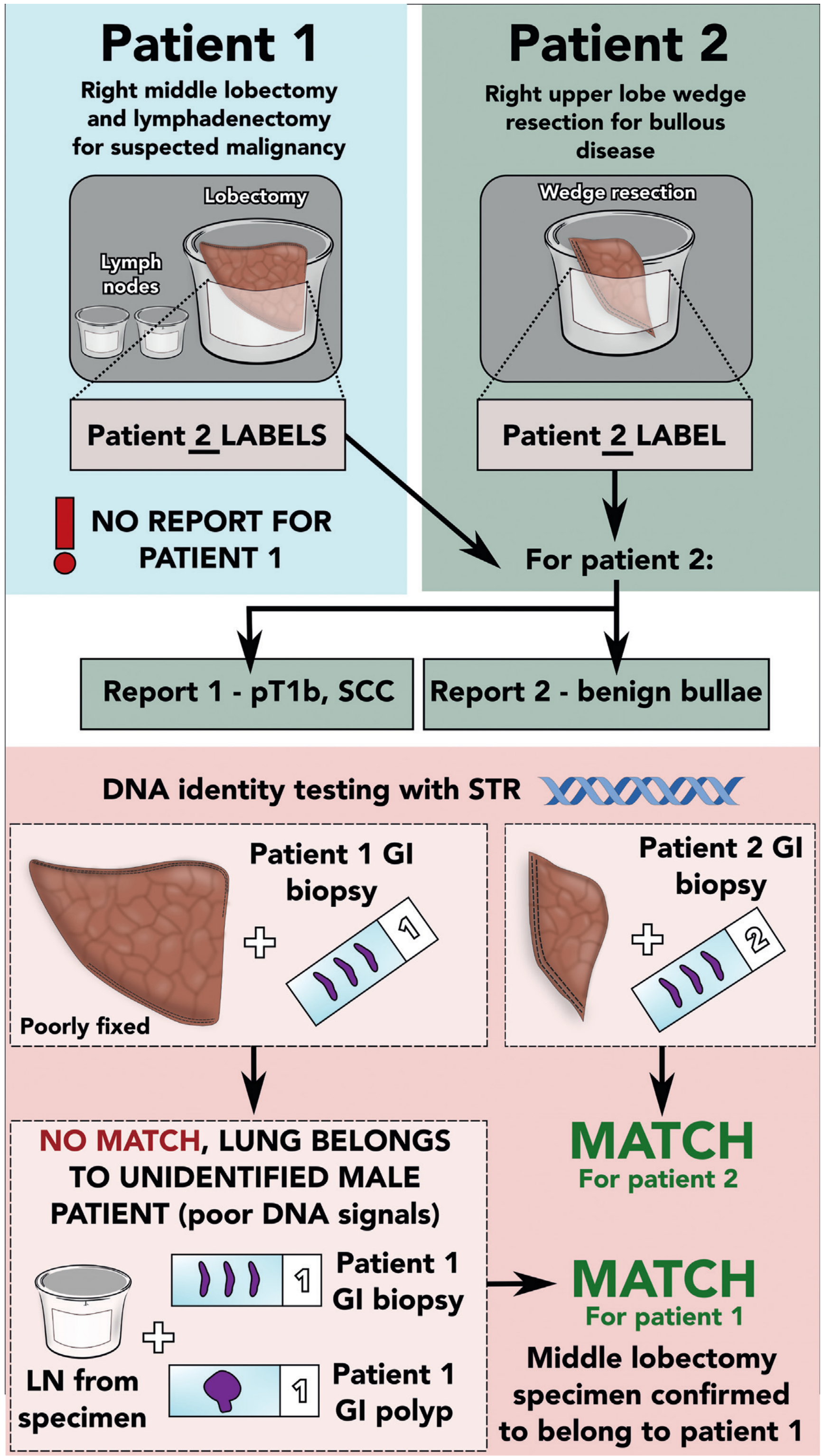

Figure 1 Schematic of the mislabeling problem and how this was resolved. Patients 1 and 2 were operated on the same theatre list, however patient 1's specimen pots were erroneously labelled with patient 2 details. Two reports for both a benign and malignant resection of the right lung were present on patient 2's electronic patient records. Short tandem repeat testing (DIPplex) was undertaken using both the lung resections and previous GI biopsies for each of the patients. The benign wedge resection was confirmed to belong to patient 2 as expected, but the malignant resection was determined to not belong to patient 1, but an unidentified male patient. The DNA quality was however poor due to delayed fixation of tissue. A separate lymph node from this specimen/request form was used to compare against two historical samples and it was determined the malignant resection did belong to patient 1. The poor-quality DNA was the likely reason for the poor signal amplification during DIPplex testing.

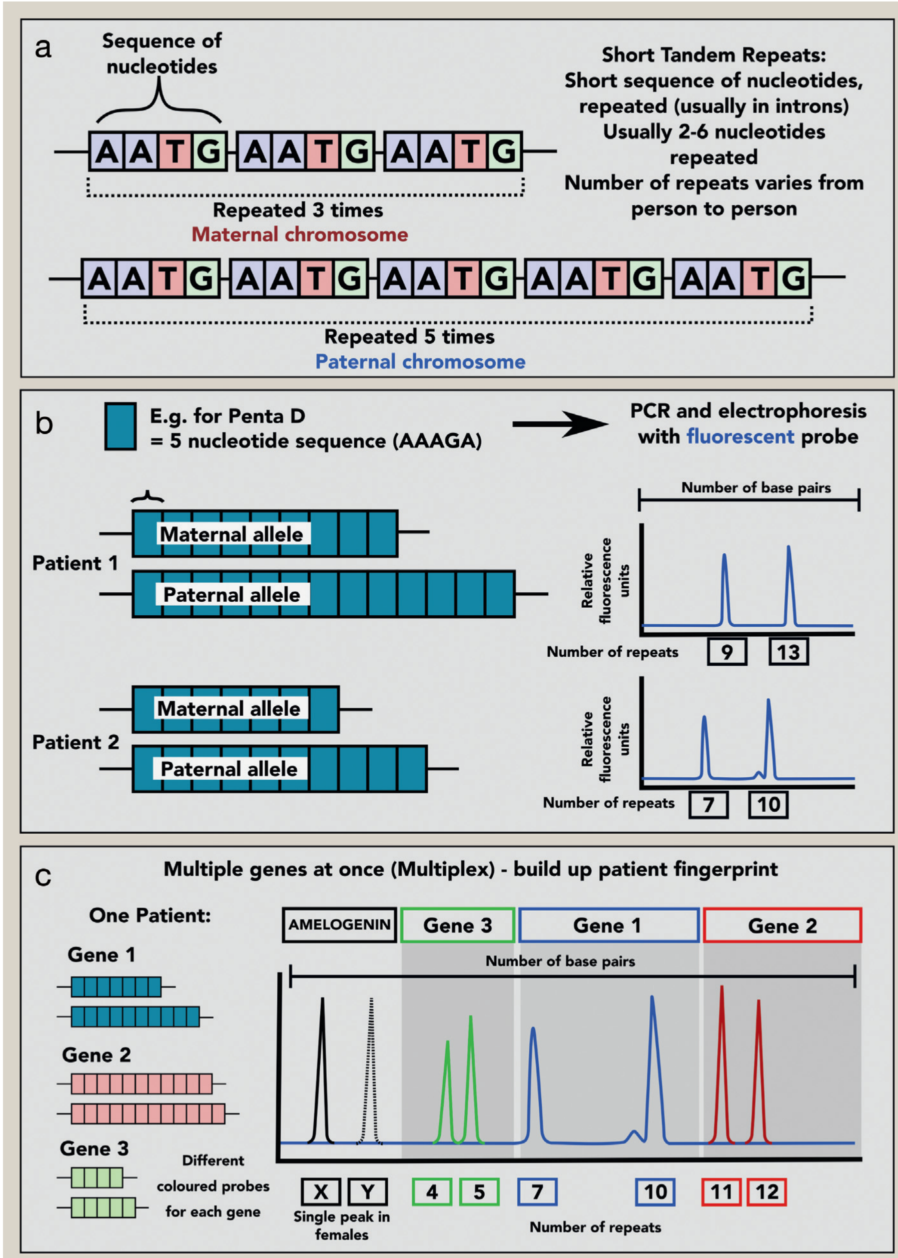

Figure 2 (**a**) Short tandem repeats. Represented is a simple STR for one patient whereby a tetranucleotide sequence is repeated a varying number of times in tandem. The STR is present on both the maternal and paternal chromosome and the number of repeats may be different on each chromosome (alleles). In this case, the maternal chromosome has 3 repeats, and the paternal chromosome has 5 repeats. (**b**) STR of interest are amplified using PCR and tagged using fluorescent probes. Electrophoresis is undertaken that measures the length of the base pairs. The length is determined by the distance genetic material moves in a gel with an applied electric field. The larger the gene (i.e. more nucleotides), the slower it will move. This can be compared to known data and the number of repeats can be deduced for a known STR. Per STR there will be two peaks that represent the repeats for the maternal and paternal alleles. The repeats from person to person varies in the population so can be compared from one tissue sample to another and confirm if there is a match. In the example here, patient one has 9 repeats and 13 repeats in their maternal and paternal alleles respectively. In contrast, patient two has 7 and 10 repeats respectively. (**c**) Using a mix of STR and different coloured fluorescent dyes can be used to build up a unique genetic fingerprint of a person that is almost impossible to be repeated in another person (although monozygotic twins may cause issues). This can then be compared between specimens/people to see if there is a genetic match. Amelogenin is commonly included in STR testing and exploits the shorter amelogenin gene on the X chromosome to determine the XY makeup of a patient. A single peak indicates only X chromosome is present and two peaks indicates an X and Y are present.

(in our case 16), the number of times each STR is repeated can be built up for a person. Combined, this creates a unique signature for a person that is almost impossible to be shared by another individual (Figure 2c).[5]

STR analysis is performed by DNA extraction from FFPE tissue and the STRs of interest are amplified by PCR. Each STR of interest is labelled with differing fluorescent dyes and electrophoresis undertaken (Figure 2c). The size (i.e. number of base pairs) and charge of the DNA will dictate the distance travelled in gel with an applied electric field which can be measured (relative fluorescence units) and compared to standard known gene sizes. In turn, the number of repeats can be calculated from the number of base pairs. Because differing-coloured dyes are used for each gene, the number of repeats can be ascribed to each individual gene. Poor quality degraded DNA will not sufficiently amplify and therefore signals are weak or erroneous. In histopathology, historical specimens are often used as a comparator but if these do not exist (i.e. the patient has not had a biopsy before) then contemporary blood sampling is required.

The term microsatellite is often used to describe these STRs although microsatellites can also encompass variable number tandem repeats (VNTR) which are much larger chains of nucleotides (10–100) that are repeated 100–1000 times. Although simple STR (i.e. 3-4 nucleotides repeated in a sequential fashion) are the most utilized, STR can also be non-consensus, compound and complex. Non-consensus STR deviate away from the normal simple repeats with runs of e.g a tetranucleotide sequence (AGAT) but then occasional alternative tetranucleotides (AGAG) followed by repeats of the original sequence. Compound STR describes multiple types of repeat units present in the sequence (e.g. repeats of AGAT and AGAC alternatively), these too can have non-consensus alleles which further add to the variability. Complex STR exhibit intricate patterns of repeats outside of the standard linear fashion.[2] Utilizing a mix of simple, compound, non-consensus and complex STR only bolsters the unique fingerprint created during this testing.

Amelogenin, a protein involved in tooth enamel production, usually makes up one of the components of STR analysis. Strictly speaking, the gene encoding this protein is not a STR, but its base pair length varies on the X and Y chromosome (being shorter on the X chromosome by 6 base pairs). Therefore, quantifying the length of the amelogenin can determine the XY chromosome makeup of an individual.[6] If there are two peaks, this indicates both an X and Y chromosome exist, whereas a single peak denotes just X chromosomes are present (Figure 2c). Of note in our case, due to the degraded nature of the DNA, it was erroneously determined that the specimen belonged to someone with XY chromosomes. Genetic quality is therefore of paramount importance.

Conclusion

Confirming the identity of tissue is sometimes required in histopathological practice. Knowing how STR analysis is conducted, and the possible limitations and impact of pre-analytical factors, allows pathologists to interpret the results carefully to ensure patient care is maintained. ◆

REFERENCES

1 Nakhleh RE, Idowu MO, Souers RJ, et al. Mislabeling of cases, specimens, blocks, and slides: a college of American pathologists study of 136 institutions. *Arch Pathol Lab Med* 2011; **135:** 969–74.

2 Urquhart A, Kimpton CP, Downes TJ, et al. Variation in Short Tandem Repeat sequences – a survey of twelve microsatellite loci for use as forensic identification markers. *Int J Leg Med* 1994; **107:** 13–20.

3 Schichman S, Suess P, Vertino A, et al. Comparison of short tandem repeat and variable number tandem repeat genetic markers for quantitative determination of allogeneic bone marrow transplant engraftment. *Bone Marrow Transplant* 2002; **29:** 243–8.

4 Jeffreys AJ, Wilson V, Thein SL. Hypervariable “minisatellite” regions in human DNA. *Nature* 1985; **314:** 67–73.

5 Chakraborty R, Stivers DN, Su B, et al. The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis* 1999; **20:** 1682–96.

6 Mannucci A, Sullivan KM, Ivanov PL, et al. Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. *Int J Leg Med* 1994; **106:** 190–3.

Practice points

- The most common and widely used method of DNA identification is using short tandem repeat (STR) analysis. Confirming the identity of a patient may be warranted in cases of mislabeling and carryover
- STR analysis uses the concept of population polymorphism in repeated nucleotide sequences across several well documented genes to build up a unique genetic ‘fingerprint’ for a person
- Amelogenin is a commonly used gene of which length varies between the X and Y chromosome, determining the length of this gene in analysis can help determine the X and Y chromosome pattern in a person
- Poor fixation can degrade the quality of DNA and thus cause spurious results when undertaking genotyping
- If historical pathology specimens do not exist for patients and genotyping needs to be confirmed, blood samples are usually required from patients

Multiple choice questions

1. Which of the following genes is commonly used to characterize the X and Y chromosome makeup of a person?

- **A) Myogenin**
- **B) p53**
- **C) Amelogenin**
- **D) TPOX**
- **E) vWF**

Answer: C. Amelogenin, which encodes a protein involved with enamel production, is of different nucleotide lengths on the X and Y chromosome. Measuring the length of these genes within tissue allows you to determine the X and Y makeup of a patient. The remaining genes are not used for this purpose. TPOX and vWF are examples of commonly used STRs.

2. Which of the following types of polymorphisms is commonly used to confirm the patient identity of tissue?

- **A) Next generation sequencing**
- **B) Sanger sequencing**
- **C) Fluorescent in-situ hydridisation**
- **D) Short tandem repeat analysis**
- **E) mRNA sequencing**

Answer: D. Short tandem repeat analysis is the most widely used method for DNA identification worldwide. Whole genome sequences with NGS can technically be used for person identification, although the current cost makes this a very uncommon modality at present. The other techniques above are not used for this purpose.

3. Which of the following is not a type of short tandem repeat (STR)?

- **A) Simple STR**
- **B) Complex STR**
- **C) Hybrid STR**
- **D) Simple repeats with non-consensus alleles**
- **E) Compound repeats with non-consensus alleles**

Answer: C. Hybrid STR is not a recognized type of STR. Simple STR are linear repeats of a short sequence of nucleotides. Non-consensus alleles describe occasional alternations to the pattern (either simple or compound). Compound repeats show multiple types of repeat units in the sequence. Complex STR follows non-linear pattern rules.