

A New Remedy in Pathology Practice: Molecular Solution to Sample Mix-Up

Patoloji Pratiğinde Yeni Bir Çare Uygulaması: Örnek Karışıklığına Moleküler Çözüm

Hüseyin BALOĞLU¹, Nuri YİĞİT²

Department of Pathology ¹GATA, Haydarpaşa Education Hospital, İSTANBUL, TURKEY and ²GATA, Faculty of Military Medicine, ANKARA, TURKEY
This study was presented as oral presentation at the 20th National Pathology Congress at Eskisehir, held between September 29 and October 3, 2010.

ABSTRACT

Objective: Although very rare, suspicious situations about the identity of diagnostic tissue material have been encountered in pathology practice. Such situations undoubtedly have the potential to create undesirable results. In the present study, an application targeting getting rid of any doubts about the identity of the diagnostic tissue samples is described.

Material and Method: A combination of short tandem repeats (STR) of the human genome consisting of CSF1PO, TH01, TPOX, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539 and Penta E were selected on the basis of ease of application and bioinformatic discrimination power. Possible forms of diagnostic tissue mix up were set in 3 different models with 3 diagnostic tissue samples of 2 different cases. Of the tissue samples selected, A (salivary gland) and B (striated muscle) belonged to the same case and C (uterus wall) belonged to another case. In the first model, there was no problem about tissue identity (M1: A/B). In the second model, two different diagnostic material were mixed up (M2: B/C). In the last model, there were 3 diagnostic material obtained from 2 different cases (M3: A/B/C). DNA was extracted from all tissue samples and all of the selected 10 STR were amplified with specially designed primers by PCR. After chemical denaturation, amplicons were submitted to polyacrylamide gel electrophoresis for discrimination of single DNA strands according to their conformation polymorphism (SSCP). Special patterns of each STR in the gel matrix obtained from M1, M2 and M3 models, were evaluated on the principle of being 'same or different' to determine the diagnostic material identity.

Results: Each of the salivary gland, striated muscle and uterus wall samples were correctly identified (matched with the right source cases) after evaluating 10 different STR SSCP patterns designed under M1, M2 and M3 models.

Conclusion: This application targeting to solve diagnostic tissue identity problems is a simple and cheap application of SSCP and its efficacy was proven on the designed models.

Key Words: Tissue identity determination, Single-Stranded conformational polymorphism, Short tandem repeats

ÖZ

Amaç: Patoloji pratiğinde, tüm önlemlere rağmen, nadir de olsa tanı örneğinin 'aidiyeti' konusunda kuşkular yaşanmaktadır. Bu durumun istenmeyen sonuçlara yol açma potansiyeli taşıdığı kuşkusuzdur. Bu çalışmada, tanı örneklerinin karışıklığı kuşkusunu gidermeye ve örneklerin 'aidiyetini' belirlemeye yönelik bir uygulama tanımlanmıştır.

Gereç ve Yöntem: Genomda tanımlı kısa tekrar bölgelerinden (STR) CSF1PO, TH01, TPOX, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539 ve Penta E'yi içeren biyoinformatik ayırım gücü yüksek ve uygulanabilirliği kolay bir kombinasyon seçilmiştir. Olası tanı materyali karışıklığı, iki farklı kişiye ait 3 doku örneği üzerinden model 1, model 2 ve model 3 olarak 3 farklı durumda kurgulanmıştır. Seçilen dokulardan A (tükürük bezi) ve B (çizgili kas) kodlu olanlar aynı kişiye, C (uterus duvarı) kodlu olan ise farklı bir kişiye aittir. Modellerden ilkinde karışıklık yoktur (M1: A/B). İkincisinde 2 tanı materyali arasında karışıklık vardır (M2: B/C). Sonucuda ise 2 kişiye ait 3 tanı materyali arasında karışıklık vardır (M3: A/B/C). Örneklerden DNA ekstraksiyonu yapılmış ve her bir STR için özel primerlerle hedeflerin ayrı ayrı PCR amplifikasyonu yapılmıştır. Amplikonlar kimyasal denatürasyondan sonra, DNA tek zincirinin konformasyon farklılığına (SSCP) dayalı ayırım için poliakrilamid jel matriste elektroforeze tabi tutulmuştur. Bu matriste her STR için oluşan özgün konformasyon paternlerinin 'aynı/farklı olma' prensibine göre, M1, M2 ve M3 modellerindeki tanı materyalinin aidiyeti irdelenmiştir.

Bulgular: M1, M2 ve M3 modellerinin her birinde 10 ayrı STR SSCP paterni değerlendirilerek, tükürük bezi, çizgili kas ve uterus duvarı örnekleri tek tek ayırt edilebilmiştir.

Sonuç: Tanı örneği karışıklığını çözmeyi amaçlayan bu uygulama basit ve ucuz bir SSCP uygulamasıdır ve ne kadar etkili çalıştığı olası model karışıklıklarında gösterilmiştir.

Anahtar Sözcükler: Doku aidiyeti tayini, DNA tek zinciri konformasyon farklılığı, Kısa tekrar bölgeleri

Received : 07.12.2010

Accepted : 17.01.2011

Correspondence: Nuri YİĞİT

Department of Pathology, GATA, Faculty of Military Medicine, ANKARA, TURKEY

E-mail: nyigitsum@gmail.com Phone: +90 312 304 20 00

INTRODUCTION

One of the most important activities during the practice of the science of pathology is to diagnose disorders. This process starts with the receipt of the biological diagnostic material and ends with providing a diagnosis using evidence-based approaches as guided by the principles of science. The process initiated by the arrival of the diagnostic material at the pathology department requires an algorithm where the “ownership is protected” at all cost. Most sample mix-ups can be solved by using simple and important data such as the person sending the sample, where the sample was obtained, the time the sample was obtained and the patient’s gender. However, there are also moments during routine pathology practice when the abovementioned simple recorded data are inadequate and one needs other solutions to prove ownership and remove a suspicion of a mix-up. It is also obvious that such situations may lead to undesirable results.

We define a procedure aiming to remove any suspicion of specimen mix-up and to determine sample ‘ownership’ in this study. A unique feature of the human genome is used for this procedure. Although the pattern of the four letters that form the sequence for the human genome are generally the same (>99.9%) the remaining part has regions that make every individual unique. One of these special human genome regions have been defined as short tandem repeats (Short Tandem Repeats: STR) (1-4). The human genome is generally a book that can be written with the four letters of A (adenine), T (thymine), G (guanine) and C (cytosine) and has many repeated letter sequences. Those repeats where the number of repeated letters are 2 to 6 are named ‘short’. The letters of this short sequence and the number of consecutive repeats, i.e. the ‘sequence and number of repeats’ (STR) is different for each genome (1-4). The features of the STR regions defined in the human genome are an open source accessible to everyone (5). As can be seen in this source, there are many STR regions that have been defined in the human genome. However, some of these are only useful to ‘differentiate the genome’. It is possible to greatly increase our discrimination ability by increasing the number of STRs with selections from this region. STR combinations with a very high discrimination values can be found using bioinformatic analysis of data obtained from data banks. However, professional studies using the earth’s population and their geographic characteristics together with some other technical details provide the logical STR combinations in data banks.

We defined an easy-to-use and low cost molecular technique to determine “ownership of the sample” and solve possible diagnostic material mix-ups in pathology practice

by creating various experimental models in this study. We created a combination based on bioinformatics that included 10 separate STR regions to create a differentiation power that was valid but did not have to be very sensitive. The easy use of the PCR technique to be used for STR analysis and the differentiating power of the polyacrylamide gel electrophoresis to be performed afterwards were taken into account in creating this combination.

MATERIAL and METHOD

We chose the following 10 STR regions defined in the genome: CSF1PO, TH01, TPOX, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539 and Penta E. There were 3 main factors in using this combination. The first one was high bioinformatic differentiation power (assuming that the frequency of each allele in the population is 5-20%, this would be 1/106-1/1014). The second factor was the low cost and short duration of the PCR amplification procedures necessary at the analysis stage. The final factor was the optimization of amplicon lengths to stay within the known differentiation power for base length of the polyacrylamide gel electrophoresis to be used following the amplification procedures. Table I presents the 10 STR regions chosen accordingly, their localization in the human genome and the PCR primers developed for the PCR amplification of these regions.

Possible diagnostic material mix-up was modelled in 3 ways taking the practice of pathology into account. Model 1: There is no diagnostic material mix-up in this model. Salivary gland (A) and striated muscle tissue (B) from the same person were used. Model 2: The material from two different patients was mixed up. Striated muscle (B) and uterine wall (C) samples from two different persons were used. Model 3: There was mix-up between three pieces of diagnostic material from two different patients. Salivary gland (A) and striated muscle tissue (B) from one person and uterine wall (C) from another person was used.

We made it possible to use an important control parameter (gold standard) for ‘ownership’ by selecting tissues that can be easily differentiated microscopically.

We performed DNA extraction (QIAamp DNA mini kit, QIAGEN, Hiden, Germany) after obtaining 4 sections 10 µm thick from the paraffin blocks of the selected tissue samples. The 10 selected STR regions were amplified using a PCR reaction mixture and thermal cycle with the obtained DNA as template (Table I). The amplicons were observed in 2% agarose gel and cleaned with the spin column technique. Each STR region that was cleaned and purified was converted to single strand form by chemical denaturation. Electrophoresis for 2 hours at 80 volt constant current in a 7% polyacrylamide gel matrix was used for differentiation on

Table I: The 10 STR regions selected, their localization in the human genome and the PCR primers used for the PCR amplification of these regions

No	STR code and feature	Genomic localization	PCR primers: (5' - 3')	Thermal cycle		PCR tube content
					AT	
1	TPOX 4-16 X AATG	2p25.3	F - AGAACAGGCACTTAGGGAA R - AGCGTTTATTTGCCCAA	{1X 72°C/7'} {1X 95°C/5'} {45 X (95°C/30"; AT°C/30"; 72°C/30")}	53°C	35µL, total volume: 17.5 µL MM* + 50 ng DNA+0.2mM each primer+ddH ₂ O
2	D3S1358 8-20X TCTA/TCTG)	3p21.31	F - CCAACTGGGTGACAGAGCA R - ACTCATGAAATCAACAGAGGCTT		60°C	
3	CSF1PO 5-16XAGAT	5q33.1	F - AGATATTAACAGTAACTGCCTTCA R - CAGATACTATCTCCTGGTGCA		53°C	
4	D5S818 6-18XAGAT	5q23.2	F - CATTTGTATCTTTATCTGTATCCTTA R - CCTCTTGGTATCCTTCTGTAATA		53°C	
5	D7S820 5-16XGATA	7q21.11	F - GATAGAACACTTGTTCATAGTTTAGAA R - GCACCAAATATTGGTAATTAATA		53°C	
6	D8S1179 4-16XTCTA/TCTG	8q24.13	F - GATAGAACACTTGTTCATAGTTTAGAA R - GCACCAAATATTGGTAATTAATA		53°C	
7	D13S317 5-17XTATC	13q31.1	F - CATCTAACGCCTATCTGTATTTACAA R - GCCCAAAAAGACAGACAGAA		53°C	
8	D16S539 4-16XGATA	16q24.1	F - GGAGCAAACAAAGGCAGA R - CCATCTCTGTTTTGTCTTTCAAT		53°C	
9	TH01 3-14XTCAT	11p15.5	F - CATTGGCCTGTTCCCTCCCTTA R - GCAGGTCACAGGGAACACAGA		60°C	
10	Penta E 5-26XAAAGA	15q26.2	F - GGCGACTGAGCAAGACTCA R - GGTTATTAATTGAGAAAACCTTACAA		53°C	

Explanations: AT: Annealing Temperature, *: Master mix: dNTP+Buffer+DNA Polymerase+2.5mM MgCl₂ in X2 concentrated commercial solution (Qiagen MasterMix Plus), F: Forward primer, R: Reverse primer.

the basis of the nucleic acid sequence specific conformation of each strand (Single Strand Conformation Polymorphism: SSCP). The gel was evaluated in a UV imaging system after staining with ethidium bromide. The ownership of the diagnostic material in the M1, M2 and M3 models were evaluated using the 'same or different' principle for the specific conformation patterns created as separate columns for each STR region in the gel matrix. The same pattern in the compared STR regions showed that the two materials were from the same person while a different pattern in at least one STR region indicated that the samples were from different persons.

RESULTS

The comparison of the 10 STR regions for the developed sample mix-up models revealed that the salivary gland (A) and striated muscle tissue (B) samples were from the same person and the uterine wall (C) sample was from another person, using comparisons with the control samples known to belong to the persons involved (Figure 1-3).

DISCUSSION

A series of critical processes start once samples are received for investigation within the practice of pathology, where 'diagnosis of disease' has a major place within the discipline. Some of these processes can be automated to minimize human error. Despite the constantly improving automation processes and careful follow-up procedures, there can rarely be doubts about the ownership of diagnostic material. Such a situation can lead to very serious problems and definite measures that can be used in addition to constant care, monitoring and checks seem to be an important requirement. It is known that such ownership issues are also important for forensic medicine and that the STR regions of the human genome have long been used for this purpose (6-8). Commercial kits have long been available and most have internationally accepted quality certificates (e.g., Biotype, Dresden- Germany; Promega Corporation, Madison-WI-USA; Applied Biosystems, Foster City- CA-USA). It is obvious that these professional solutions can

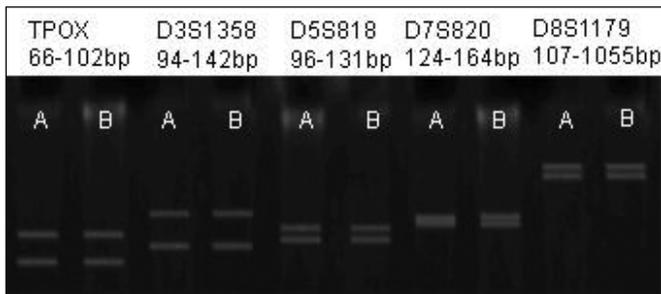


Figure 1: Gel patterns are seen for the first 5 STR regions of the A and B tissues of the same person. The band lengths are seen to be the same for each STR region in sample A and B.

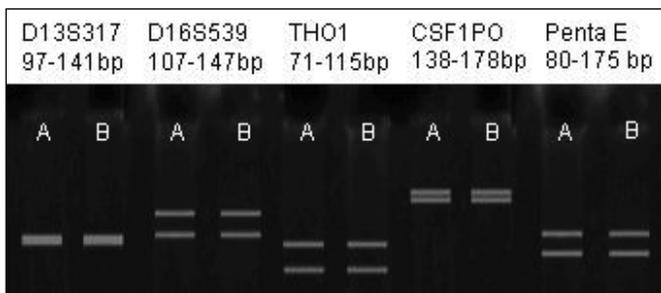


Figure 2: The gel patterns for the second 5 STR regions of the specimen A and B tissues. The band lengths are seen to be the same for each STR region.

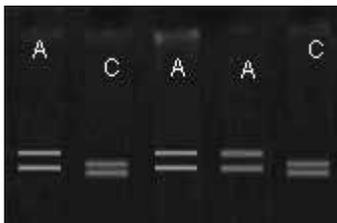


Figure 3: The gel patterns for the TPOX STR region of specimen A and C tissues from different persons are seen. From the left, column 1 of the gel (A) and column 5 (C) are positive comparison columns showing the band length of the TPOX region from the diagnostic material of two separate patients. Columns 2, 3 and 4 of the gel show the TPOX region band lengths of 3 different diagnostic materials from 2 different persons as designed with model 3. Analysis of the patterns shows that the band lengths in column 1, 3 and 4 are the same (tissue samples coded A from the first person) and band lengths in column 2 and 5 are the same (tissue samples coded C from the second person).

also be used to solve problems related to specimen mix-up during pathology practice. However, these procedures require very expensive reagents and equipment. Our aim was to define an alternative method which was similarly based on STR pattern analysis but would be inexpensive and easy to use and at least as effective as its commercial counterparts. The method defined in this article does not provide the very high discriminatory power of the

commercial kits mentioned above that use 13 STR regions. These commercial kits have mostly been designed to provide the discriminatory power needed for forensic medicine procedures (7). However, when one looks at the possible mix-ups in pathology practice, it is clear that there is no need for such high discrimination power. Most mix-ups are between diagnostic samples obtained on the same day or on successive days. This number is not over 500 in even the busiest pathology center in our country (< 125.000 per year, assuming an annual workday number of 250). A procedure that could discriminate one of 1000 samples (the number of materials for two days) would be statistically adequate. Increasing this sensitivity would require increasing the number of selected STR regions and this would lead to the use of more reagents and a longer test. Our aim was to define a procedure that could be easily used and provided quick results with acceptable discrimination power. We saw in our optimization studies that it was possible to greatly increase discrimination power and we used a combination where 10 STR regions were taken into account (assuming the allele frequency for each STR region to be 5-20% in the population, the result would be 1/106-1/1014).

In conclusion, this procedure that aims to eliminate diagnostic sample mix-up is a simple and inexpensive SSCP application and mix-up models have demonstrated its effectiveness. The rationale may seem complicated but pathologists will find it easy to understand and it is simple enough to be used in any laboratory with the technical capacity to perform PCR and electrophoresis.

REFERENCES

1. **Edwards A, Civitello A, Hammond HA, Caskey CT:** DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* 1991, 49:746-756
2. **Gill P, Urquhart A, Millican E, Oldroyd N, Watson S, Sparkes R, Kimpton CP:** A new method of STR interpretation using inferential logic--development of a criminal intelligence database. *Int J Legal Med* 1996, 109:14-22
3. <http://www.forensicdnacenter.com/dna-str.html>
4. http://www.biology.arizona.edu/human_bio/activities/blackett2/str_description
5. <http://www.cstl.nist.gov/strbase/>
6. http://www.ornl.gov/sci/techresources/Human_Genome/elsi/forensics.shtml
7. **Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B:** Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J Forensic Sci* 2001, 46:647-660
8. **Coomber N, David VA, O'Brien SJ, Menotti-Raymond M:** Validation of a short tandem repeat multiplex typing system for genetic individualization of domestic cat samples. *Croat Med J* 2007, 48:547-555