# SEMINAR

## The Application of Microsatellites in Molecular Pathology

Richard NAIDOO, Runjan CHETTY

Molecular Biology Research Facility and Department of Pathology, University of Natal School of Medicine, Durban, South Africa

Present within the genome are large numbers of seemingly unimportant DNA segments arranged in repetitive units. Furthermore, these stretches of DNA contain variations or polymorphisms that are characteristic for an individual and results in a unique DNA fingerprint. Approximately 30% of the DNA repeat sequences are arranged as short tandem repeat sequences, which are called microsatellites. Microsatellites may consist of 1, 2 or 3 nucleotides; dinucleotides being the commonest. Microsatellites are characterised by being: stably inherited and hence highly conserved from one generation to the next, and unique to an individual and the same in different cells from the same individual. As a result of the above features, microsatellites can be used for personal identification, population genetic analysis and construction of evolutionary trees. In addition, they are located in several important gene loci and this allows microsatellites to be used as markers of disease and to provide information about individual gene status, especially in tumors. This can be accomplished by assessing allelic imbalance or loss of heterozygosity of a particular gene by analysing microsatellites located at specific loci in the gene. Recently, mutations within microsatellites have been described as a result of defective DNA repair mechanisms, resulting in the phenomenon of microsatellite instability. This has been implicated in the aetiopathogenesis of several hereditary and non-hereditary conditions. There are several ways of analysing microsatellites, the popular using radioactively-labelled primers and autoradiography. This method has several drawbacks, especially the use of radioactivity and interpretative/technical problems. The use of fluorescently-labelled primers, automated DNA sequencing coupled with a computer software package obviates these problems. This technique has the added advantage of analysing several microsatellites in large numbers of cases, simultaneously. Thus, microsatellite analysis has become an important investigative tool for the molecular biologist and has provided new information in many diseases. (Pathology Oncology Research Vol 4, No 4, 310-315, 1998)

Key words: microsatellites, microsatellite analysis, microsatellite instability, allelic imbalance, genetic linkage, automated DNA sequencing

#### Introduction

We are in the midst of a biomolecular revolution in which DNA recombinant technology has spawned the domain of molecular medicine. This new discipline is a combination of modern medicine and basic science. It has led to an exponential increase in our understanding of the processes governing cell growth and differentiation and

Received: Oct 2, 1998; accepted: Oct 31, 1998

the defects that occur in this, normally, tightly regulated process. An intricate part of this technological revolution is the application of new techniques to unravel the genetic make-up of disease.

#### Microsatellites

Alterations due to mutations in the simple repeat sequences or microsatellites are a feature in many tumors.<sup>3,11,20</sup> It is envisaged that assaying for microsatellite instability may be important in identifying a substantial fraction of human cancers.<sup>15</sup>

The eukaryotic genome contains not only introns and exons, but also large numbers of copies of other seeming-

© 1998 W. B. Saunders & Company Ltd on behalf of the Arányi Lajos Foundation

*Correspondence:* Dr Richard NAIDOO, Molecular Biology Research Facility, University of Natal School of Medicine, Private Bag 7, Congella 4013, Durban, South Africa. Tel: +27-31-2604498; fax: +27-31-252711. e-mail: naidoor@med.und.ac.za

ly nonessential DNA (over 90%). This is occasionally referred to as "junk DNA" in the literature. These DNA sequences do not code for any protein and in many regions, are "repetitious". The variations within these repetitious stretches of DNA are so great that a "DNA fingerprint", for example, can distinguish each human. This "fingerprint" is based on the variation in the repetitive sequences.

The presence of these repeat sequences in higher eukaryotes was first revealed by hybridisation experiments. In this procedure the genome is mechanically broken into short fragments of double stranded DNA molecules which are about 1000 base pairs long. The fragments are then denatured to produce single-stranded DNA. Under conditions that favour re-annealing, the speed with which these single-stranded DNA re-anneal, is dependent on the number of complementary regions each fragment finds. This is usually a slow process, however, when the DNA from a human cell is analysed in this manner, about 70% of the DNA strands re-anneal slowly, as is expected. The remaining 30% of the DNA strands anneal much more quickly. The reason for this is that these strands contain sequences that are repeated many times in the genome and find their complementary strands relatively rapidly.

About a third of all these repeat sequences are short tandem repeat sequences referred to as microsatellites which are scattered throughout the human genome. In one of the early studies, huge tracts of dT-dG alternating sequences were identified by Southern blotting and hybridisation analysis using<sup>32</sup> P-labelled poly (dT-dG).(dC-dA) as a probe.<sup>9</sup> These sequences have been shown to adopt a lefthanded DNA conformation, also referred to as Z-DNA. This conformation was thought to influence the expression of certain cellular genes. This hypothesis was tested in a CAT assay system and their results indicate that the poly (dT-dG).(dC-dA) sequence is capable of enhancing gene expression *in vivo*, however, the exact mechanism by which this enhancer activity is brought about is not clear.<sup>10</sup>

The repeating units may be as short as one, two or three nucleotides; the most common being the dinucleotide repeat, for example the CA dinucleotide in a  $(CA)_n$  repeat.<sup>19</sup> These microsatellites are non-transcribable and constitute 10–15 percent of the total mammalian genome. The length of these microsatellites are unique to each individual and vary between individuals, however, there is no variation between different cells in the same individual.<sup>1</sup> In addition, these sequences are highly conserved and are stably inherited.<sup>24</sup>

Novel or new alleles at (CA)/(TG) microsatellites and at tetranucleotide repeats are known to be formed without exchange of flanking markers. This means that they are not generated by unequal crossover. Since the mutant allele has been observed to differ by a single repeat unit from the original parent allele, the most likely mechanism to explain this length polymorphism is due to DNA slipped strand mispairing also referred to as DNA slippage.

This occurs when the normal base pairing between the two complementary strands is altered by staggering of the repeats on the two strands, leading to incorrect pairing of repeats.<sup>23</sup> Microsatellites are located in the heterochromatin near chromosomal centromeres and telomeres. In the eukaryotic genome, microsatellites are identified as stretches of dT-dG sequences with varying length. Microsatellites are estimated to occur between 55 000 and 100 000 copies in the human genome, providing a marker density of one microsatellite every 100 000 base pairs, even by the most conservative estimates.<sup>13</sup> However, although widely distributed, microsatellites are not regularly spaced in the chromosome.

#### Uses (see Table 1)

Weissenbach *et al.* isolated and mapped a large number of microsatellites in order to construct a linkage map of the human genome with an average resolution of 5 centimorgans (cM).<sup>25</sup> The sequence data obtained from this study is widely used for various microsatellite applications. Microsatellites may be used for personal identification, population genetic analysis and in the construction of the human evolutionary tree.<sup>6</sup> Further, they are located and linked to several important gene loci, thus they are associated with human diseases not only as markers but also directly in disease aetiopathogenesis, providing insight into the replication, repair and mutation of eukaryotic DNA.<sup>16</sup>

Microsatellite PCR can be used in genetic mapping studies where multiple loci can be examined. This includes the analysis of polygenic, multifactorial diseases.

*Table 1.* Neoplastic and non-neoplastic conditions where microsatellite PCR has been employed

Neoplastic conditions	References
Testicular germ cell tumors	35
Malignant melanoma	36, 37
Lobular breast carcinoma	38
Prostate cancer	39
Head and neck squamous carcinoma	40
Gastric cancer	21, 41
Uterine leiomyoma	42
Transitional cell carcinoma	43
Hodgkin's disease	44
Non-neoplastic conditions	
Population genetic studies	45
Down's syndrome	46
DNA linkage analysis	47
Inflammatory bowel disease	48

In addition, this technique can be used to examine the status of specific single genes such p53, APC, DCC, etc in any disease entity. Gene mutations and/or deletions in malignancies can be sought by determining allelic imbalance or loss of heterozygosity.

Microsatellites were initially thought to play a functional role in the genome, either directly in gene regulation, or indirectly as hot spots for recombination, however, their exact function still remains elusive. Recent studies have shown that mutations occur in these microsatellites, due to mismatch repair errors termed RERs. These mutations are due to an increase or decrease in the number of repeats. As a result, different lengths of DNA are produced which arises directly from the defective repair process. The mutations that occur are referred to as microsatellite instability (MSI), which have been implicated in a host of human disorders, both hereditary and nonhereditary, including tumors, e.g. colorectal carcinomas, breast cancer and prostate cancer.

Over the past five years there has been a dramatic increase in the number of studies which have focused on changes in the molecular mechanisms surrounding tumor biology and tumorigenesis. However, more recent reports seem to indicate a greater involvement of microsatellites in the initiation and development of cancers.<sup>2,12,14</sup>

#### Analysis of microsatellites

The most common way to analyse DNA polymorphisms is by examining RFLPs, in which base pair substitution between the two alleles alter a restriction enzyme consensus sequence, however with microsatellite markers the difference is not a base pair substitution, but rather an alteration in the length of the sequence. Such differences in sequence length can be detected by using PCR. Currently, there are over 2000 polymorphic microsatellite markers identified in the human genome, some of which have been used in genetic linkage studies.<sup>25</sup>

The analysis of microsatellites is conventionally done using radioactively-labelled primers in the PCR reaction.<sup>7,8,12,22</sup> The microsatellite PCR products are resolved on sequencing gels which are fixed and dried. The dried gels are then autoradiographed. However, this method has a number of disadvantages, most of all the hazard of working with radioactivity. The major technical problem in the analysis of microsatellites is the resolution and detection of the amplified DNA. This is due to the presence of additional bands that appear in addition to the microsatellite band, creating ambiguity in the analysis of the results. These bands are referred to as "stutter bands". Another mechanism for microsatellite instability is said to be due to DNA slippage. This occurs when there are multiple repeat units in the template, which gives rise to the formation of secondary structures. The DNA polymerase fails to read

Figure 1. This electropherogram shows the normal DNA (lane 2) exhibiting 2 peaks, which represent both alleles of the gene being analysed (in this particular case it is a DCC locus). Furthermore, these 2 peaks are greater than 2 base pairs apart which means that they are not artefactual or "stutter bands". This case is therefore said to be heterozygous and informative for this particular marker. In the lower trace (lane 3) of tumor DNA, there are also 2 peaks. This trace is in fact, identical to the upper one thus indicating that there is no change between normal and tumor DNA. In other words, there is no loss of heterozygosity nor microsatellite instability. The size of the products (position of the peaks) can be determined from the base pair sizes which are on the x-axis of the electropherogram. Lane 1 represents an external CY5 labelled 50-500 base pair standard which serves as a control to guage the size of the PCR products.

Lane# Peak# Run Time Peak Area Size (BP) Quantity 1 1 21:32 176.11 #50.0 0.30151 2 24:53 152 #100.0 0.26024 З 172.57 0.29544 29:17 #150.0 4 35.03 584.09 #200.0 1 5 40:39 309.4 #250.0 0.52971 6 46:12 396.18 #300.0 0.67828 7 51:57 404.75 #350.0 0.69296 8 57:55 462.38 #400.0 0.79162 9 63:55 473.05 #450.0 0.80988 10 69:51 372.16 #500.0 0.63716 2 1 34:05 1140.8 191.6 1 2 35:34 622.53 204.6 0.5457 1 3 1533.4 34:09 192.2 1 2 35:38 1066.3 205.2 0.6954



**Figure 2.** The upper trace of normal DNA (lane 10) shows 2 peaks representing both alleles. However, the tumor DNA (lower, lane 11) shows only 1 peak representing only 1 allele. This implies that in the tumor DNA the second allele is lost, thus this case demonstrates loss of heterozygosity in the tumor.

through the repeating units within the secondary structure, thus producing a shorter product. Since polymorphisms between alleles occur as multiples of two base pairs, high resolution of acrylamide gels are required for the separation of these fragments. These fragments are visualised as bands by ethidium bromide staining or silver staining.

The problem with the former technique is that it is often not sensitive enough to detect the quantity of amplified DNA products, whereas the latter technique is more sensitive. There is however, conflicting reports on the use of silver staining. Koreth *et al.* found that the band intensity on silver staining does not correlate well with radioactive RFLP quantitation, while Schwengel *et al.* found this technique to be adequate.<sup>13,17</sup>

Recently, the use of fluorescence based technology was shown to be a sensitive and useful tool for the detection and analysis of microsatellite PCR products. Skolnick and Wallace in 1988 described the advantages of using the automated sequencer for the analysis of microsatellites.<sup>18</sup> Ziegle et al., reported the use of automated DNA sizing technology for genotyping microsatellites.<sup>26</sup> Fluorescently labelled primers were used in the PCR assay. The labels contained FAM (blue), JOE (green), TAMRA (yellow) and ROX (red). The fluorescently labelled PCR products were separated on polyacrylamide gels and the fluorescent products were detected by laser and analysed by computer software. The software package produces so-called electropherograms, after analysing the fluorescently-labelled PCR products. Typical examples of normal and tumor DNA patterns from cases examined with this technique are demonstrated in Figures 1-3. The computer package calculates areas under the peaks and ratios comparing normal and tumor DNA can be computed. In addition, the running of base pair size controls also allows for interpretation of new or novel peaks as compared to artefactual, stutter hands.

Since then, this technology has been used more often and the results clearly show an improvement in the assessment and analysis of the data.<sup>27-34</sup>

A fluorescent based assay for the detection of MSI in sporadic colorectal carcinomas was demonstrated by Cawkwell et al.<sup>5</sup> Fifty four cases were used in their study, and of these 22% showed MSI for at least one marker. Six of the tumors showed MSI at high frequency, which meant that at least 63% of the markers were affected.

Multiplex polymerase chain reaction is another important technique which has become increasingly popular in the analysis of microsatellites in cancers. In this technique, more than one set of primers are used in the same PCR tube, allowing for co-amplification of multiple products. This method has been used by Cawkwell et al., in which multiple markers were analysed using fluorescent based DNA technology to analyse 20 cases of colorectal carcinomas.<sup>4</sup> The findings from this study indicate allelic loss to be 29% for DCC, 66% for p53 and 50% for the APC/MCC region. A slightly different technique was used to analyse microsatellite instability in 21 gastric cancers and their corresponding normal tissue, using the automated DNA sequencer.<sup>21</sup> The procedure adopted in this study was different in that the two amplified products from both tumor and normal tissue were co-loaded into a single lane of the DNA sequencer and analysed. The results of this study are in keeping with other studies in that microsatellite instability was found to occur in 22,7% of the cases, however, the technique employed was unique,



accurate and proved to be an efficient method for detection of microsatellite instability.

The advantages of using fluorescent-based DNA technology for the analysis of microsatellites are very clear. It allows for easy interpretation of the data, which can be captured, and stored using the DNA software linked to the DNA sequencer. Accurate and impartial scoring of alleles, including interpretation of heterozygous and homozygous copies, has significantly increased the accuracy of the data



**Figure 3.** The upper trace of normal DNA (lane 2) once again shows 2 peaks. However, the lower trace of the tumor DNA (lane 3) now shows 3 peaks. There is an extra peak that is greater than 2 base pairs in size compared to the other 2 peaks. This represents a novel or new allele and classifies the case as demonstrating microsatellite instability.

produced. Further, a high throughput of samples is an added advantage, especially for large-scale linkage studies. The technology also allows for assaying for multiple loci as in multiplex PCR. The use of internal and external standards eliminates lane to lane variation arising as a result of uneven mobility of the PCR products through the gel. The use of the fluorescent tag for detection is highly sensitive, so only  $1-3 \ \mu$ l of PCR product needs to be run on the gel. The cost of the reagents required in the PCR reaction is reduced and therefore there is tremendous cost saving.

Thus, microsatellites and their analysis have revolutionised the assessment of genetic abnormalities. It has provided a reproducible and convenient method for the screening of genetic mutations that are implicated in the causation of both cancers and non-cancer disease states. The advent of automation and computer-based software packages with fluorescently-labelled primers, has enhanced the value and use of microsatellite analysis in disease.

### Acknowledgement

The authors are recipients of grants from the Medical Research Council of South Africa, Kennedy-Potts Cancer Research Fund and the University of Natal Research Fund.

#### References

- 1. Beckman S, Weber J: Survey of human and rat microsatellites. Genomics 12:627-631,1992.
- Bedi G, Westra W, Farzadegan H, et al: Microsatellite instability in primary neoplasms from HIV+ patients. Nature Med 1:65-68, 1995.
- Brentnall T: Microsatellite instability Shifting concepts in tumorigenesis. Am J Path 147:561-563, 1995.
- Cawkwell L, Lewis FA, Quirke P: Frequency of allelic loss of DCC, p53, RB1, NF1, NM23, and APC/MCC in colorectal cancer assayed by fluorescent multiplex polymerase chain reaction. Br J Cancer 70:813-818, 1994.
- Cawkwell L, Li D, Lewis F, et al: Microsatellite instability in colorectal cancers: Improved assessment using fluorescent polymerase chain reaction. Gastroenterology 109:465-471, 1995.
- Di Rienzo A, Peterson A, Garza J, et al: Mutational processes of simple-repeat loci in human populations. Proc Natl Acad Sci USA 91:3166-3170, 1994.
- Gruis N, Abeln E, Bardoel A, et al: PCR-based microsatellite polymorphisms in the detection of loss of heterozygosity in fresh and archival tissue. Br J Cancer 68:308-313, 1993.
- Halling K, Honchel R, Pittelkow M, et al: Microsatellite instability in keratoacanthoma. Cancer 76:1765-1771, 1995.
- Hamada H, Petrino M, Kakunaga T: A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc Natl Acad Sci USA 79:6465-6469, 1982.
- Hamada H, Seidman M, Howard B, et al: Enhanced gene expression by poly(dT-dG). poly (dC-dA) sequence. Mol Cell Biol 4:2622-2630, 1984.
- 11. Horii A, Han H, Shimada M, et al: Frequent replication errors at microsatellite loci in tumors of patients with multiple primary cancers. Cancer Res 54:3373-3375, 1994.

- Keller G, Rotter M, Vogelsang H, et al: Microsatellite instability in adenocarcinomas of the upper gastrointestinal tract. Am J Path 1:902-909, 1995.
- Koreth J, O'Leary J, McGee J: Microsatellite and PCR genomic analysis. J Pathol 178:239-248, 1996.
- 14. *Loeb L:* Microsatellite instability: Marker of a mutator phenotype in cancer. Cancer Res 54:5059-5063, 1994.
- 15. McCarthy K: Microsatellites: heroes and villians. Lancet 346:1177, 1995.
- Meltzer S: Dr. Strange DNA, or how I learned to stop cloning and love the computer. Gastroenterology 109:611-614, 1995.
- Schwengel D, Jedlicka A, Nathakumar E, et al: Comparison of fluorescence-based semi-automated genotyping of multiple microsatellite loci with autoradiographic techniques. Genomics 22:46-54, 1994.
- Skolnick M, Wallace R: Simultaneous analysis of multiple polymorphic loci using amplified sequence polymorphisms (ASPs). Genomics 2:273-279, 1988.
- 19. *Thibodeau S, Bren G, Schaid D:* Microsatellite instability in cancer of the proximal colon. Science 75:1027-1038, 1993.
- 20. *Todd J*: La carte des microsatellites est arrivee! Hum Mol Genet 1:663-666, 1992.
- 21. Toh Y, Oki E, Oda S, et al: An integrated microsatellite length analysis using an automated fluorescent DNA sequencer. Cancer Res 56:2688-2691, 1996.
- Tomlinson I, Stickland J, Lee A, et al: Loss of heterozygosity on chromosome 11q in breast cancer. J Clin Pathol 48:424-428, 1995.
- Umar A, Kunkel: DNA-replication fidelity, mismatch repair and genome instability in cancer cells. Eur J Biochem 238:297-307, 1996.
- 24. *Weber J, May P:* Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388-396, 1989.
- Weissenbach J, Gyapay G, Dib C, et al: A second generation linkage map of the human genome. Nature 359:794-801, 1992.
- 26. Ziegle J, Su Y, Corcoran K, et al: Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14:1026-1031, 1992.
- 27. Chetty R, Naidoo R, Schneider J: Allelic imbalance and microsatellite instability of the DCC gene in colorectal cancer in patients under the age of 35 using fluorescent DNA technology. Mol Pathol 51:35-38, 1998.
- Oki E, Oda S, Tokunaga E: A new microsatellite instability using fluorescent primers and laser scanning. Jap J Cancer Chemotherapy 3:436-442, 1998.
- Hayden JD, Cawkwell L, Sue-Ling H, et al: Assessment of microsatellite alterations in young patients with gastric adenocarcinoma. Cancer 79:684-687, 1997.
- Tajiri T, Suita S, Shono K, et al: A microsatellite instability analysis in neuroblastoma based on a high resolution fluorescent microsatellite analysis. Cancer Letters 124:59-63, 1998.
- 31. Oda S, Oki E, Maehara Y, et al: Precise assessment of microsatellite instability using high resolution fluorescent microsatellite analysis. Nucl Acid Res 25:3415-3420, 1997.

- 32. Yasuda Y, Kashiwabara H, Kawakami K, et al: Detection of microsatellire instability in cancers by arbitrarily primed-PCR fingerprinting using a fluorescently labeled primer (FAP-PCR). Biol Chem 377:563-570, 1996.
- Randerson J, Cawkwell L, Jack A, et al: Microsatellite instability in follicle centre cell lymphoma. Br J Haematol 93:160-162, 1996.
- 34. Larson AA, Kern S, Sommers RL, et al: Analysis of replication error (RER+) phenotypes in cervical carcinoma. Cancer Res 56:1426-1431, 1996.
- Huddart RA, Wooster R, Horwich A, et al: Microsatellite instability in human testicular germ cell tumors. Br J Cancer 72:642-645, 1995.
- 36. *Quinn AG, Healy E, Rehman I, et al:* Microsatellite instability in human non-melanoma and melanoma skin cancer. J Invest Dermatol 104:309-312, 1995.
- Talwalkar VR, Scheiner M, Hedges LK, et al: Microsatellite instability in malignant melanoma. Cancer Genet Cytogenet 104:111-114, 1998.
- Aldez CM, Chen T, Sahin A, et al: Comparative allelotype of in situ and invasive human breast cancer: High frequency of microsatellite instability in lobular breast carcinomas. Cancer Res 55:3976-3981, 1995.
- Watanabe M, Imai H, Shiraishi T, et al: Microsatellite instability in human prostate cancer. Br J Cancer 72:562-564, 1995.
- 40. El-Naggar A, Hurr K, Huff V, et al: Microsatellite instability in preinvasive and invasive head and neck squamous carcinoma. Am J Path 148:2067-2072,1996.
- Nakashima H, Inoue H, Mori M, et al: Microsatellite instability in Japanese gastric cancer. Cancer 75:1503-1507, 1995.
- French D, Cermele C, Lombardi AM: Microsatellite alterations in uterine leiomyomas. Anticancer Res 18:349-351, 1998.
- 43. Christensen M, Jensen MA, Wolf H: Pronounced microsatellite instability in transitional cell carcinomas from young patients with bladder cancer. Int J Cancer 79:396-401, 1998.
- 44. Mark Z, Toren A, Amariglio N, et al: Instability of dinucleotide repeats in Hodgkin's disease. Am J Hematol 57:148-152, 1998.
- Bruford MW, Wayne RK: Microsatellites and their application to population genetic studies. Curr Opinion Genet 3:939-943, 1993.
- 46. Ko TM, Hwa HL, Tseng LH et al: Fluorescence microsatellite analysis to study the parental origin of the supernumerary chromosome in Down's syndrome. Int J Gynec Obstet 61:149-153, 1998.
- Duerr RH, Barmada MM, Zhang L et al: Linkage and association between inflammatory bowel disease and a locus on chromosome 12. Am J Hum Genet 63: 95-100, 1998.
- Mirza MM, Lee J, Teare D, et al: Evidence of linkage of the inflammatory bowel disease susceptibility locus on chromosome 16 (IBDI) to ulcerative colitis. J Med Genet 35:218-221, 1998.