ASSESSING THE GENETIC INTEGRITY OF TUMOURS BY MICROSATELLITE LOCI OF FORENSIC INTEREST

Hanna Vauhkonen

Department of Forensic Medicine
University of Helsinki
Helsinki, Finland
and
Department of Pathology, Jorvi Hospital
HUSLAB, Helsinki University Central Hospital
Espoo, Finland

Academic dissertation

To be publicly presented with the permission of the Medical Faculty of the University of Helsinki, in the lecture hall of the Department of Forensic Medicine on February 11th, 2005, at 12 noon.

Helsinki 2005

Supervised by Professor Antti Sajantila, M.D., Ph.D.
Department of Forensic Medicine
University of Helsinki
Helsinki, Finland

Professor Pentti Sipponen, M. D., Ph.D.
Department of Pathology, Jorvi Hospital
HUSLAB, Helsinki University Central Hospital
Espoo, Finland

Reviewed by Professor Dr. habil. Manfred Kayser
Department of Forensic Molecular Biology
University Medical Centre Rotterdam
Rotterdam, The Netherlands

Docent Ismo Ulmanen, Ph.D.
NPHI, Molecular Medicine
University of Helsinki
Helsinki, Finland

Official Opponent Professor Ann-Christine Syvänen, Ph.D.
Department of Medical Sciences – Molecular Medicine
Uppsala University
Uppsala, Sweden

ISBN 952-91-8227-9 (paperback)
ISBN 952-10-2282-5 (pdf)
http://ethesis.helsinki.fi

Helsinki 2005
Yliopistopaino

CONTENTS

ABBREVIATIONS

AMEL	amelogenin
APC	adenomatous polyposis coli
BAT-26	"Big A-Tract", a marker of 26 adenine residues
BAX	BCL-2 associated X protein
CRC	colorectal cancer
CSF1PO	c-fms proto-oncogene for CSF-1 receptor
DCC	deleted in colorectal cancer
DGCA	diffuse type gastric cancer
FGA	fibrinopeptide alpha
GI	gastrointestinal
GIST	gastrointestinal stromal tumour
HET	heterozygosity index
HNPCC	hereditary non-polyposis colorectal cancer
IGCA	intestinal type gastric cancer
IGFR2	insulin-like growth factor receptor 2
K-RAS	Kirsten rat sarcoma viral oncogene homolog
LOH	loss of heterozygosity
LOH-H	high-level loss of heterozygosity
LOH-L	low-level loss of heterozygosity
MLH1	mutL homologue 1
MSH2, 6	mutS homologue 2 and 6
MMR	mismatch repair
MSI	microsatellite instability
MSI-H	high-level microsatellite instability
MSI-L	low-level microsatellite instability
MSS	microsatellite stable
NCI	National Cancer Institute (USA)
p	"petit", short arm of the chromosome
PCR	polymerase chain reaction
PIC	polymorphic information content
q	"queue", long arm of the chromosome
SNP	single nucleotide polymorphism
STR	short tandem repeats
TFF1	trefoil factor 1
TGFβ	transforming growth factor beta
TH01	thyroxine hydroxylase 1
TPOX	thyroid peroxidase
TP53	tumour protein 53
vWA	von Willebrandt antigen

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following articles, which will be referred to in the text by their Roman numerals.

I Vauhkonen H, Hedman M, Vauhkonen M, Kataja M, Sipponen P, Sajantila A. Evaluation of gastrointestinal cancer tissues as a source of genetic information for forensic investigations by using STRs. Forensic Sci Int 139 (2004) 159-167.

II Vauhkonen H, Hedman M, Vauhkonen M, Sipponen P, Sajantila A. Typing of XY (male) genotype from malignant neoplastic tissue by the amelogenin-based sex test. J Forensic Sci 49 (2004) 222-226.

III Vauhkonen M, Vauhkonen H, Sajantila A, Sipponen P. Intestinal and diffuse type of gastric cancer represent different genetic pathways as assessed with non-cancer related tetranucleotide microsatellite markers (2004), submitted.

IV Vauhkonen H, Vauhkonen M, Sipponen P, Sajantila A. Comparison of the allelic distribution of X-chromosomal STRs in the normal population and in phenotypically different tumour tissues. A study using four X-STR loci (DXS7423, DXS8377, ARA, DXS101). Ann Hum Gen 68 (2004) 555-562.

Additional unpublished data have also been included in this thesis.

SUMMARY

Microsatellites, or structures of tandemly arranged repeats, are polymorphic loci in the genome which are inherited in a Mendelian pattern. At each locus both the maternal and the paternal counterpart bears an allele with a certain number of repeats. Multiplex analysis of several microsatellite loci creates an individual-specific DNA profile. Therefore, they are widely used for identification purposes, establishment of family relationships, genetic mapping and population studies. However, due to the repetitious structure, microsatellite DNA array is prone to mutations during the replication. This may cause problems in particular applications, e.g. in paternity analyses or disease linkage mapping, where a mutation may lead to false exclusion. Knowing the mutation frequency at a particular microsatellite locus may be essential in assessing the statistical interpretation of the DNA profile, e.g. in parentage testing. Mutation rates have been estimated by family studies, population-based algorithms or pedigree analysis. The rates have been found to be different between various loci and among alleles of a particular locus, depending on the repeat structure of the microsatellite locus.

The aim of this study was to evaluate the stability of a set of microsatellite markers in gastrointestinal cancers from the perspectives of forensic DNA profiling and paternity testing, genetic integrity of cancer tissues, and population genetics. In tumour tissues which are known to harbour genetic instability two types of alterations may be observed. Firstly, the microsatellite instability (MSI) type with emergence of new alleles, and secondly loss of heterozygosity (LOH) type with loss of one allele of a heterozygote. High-level microsatellite instability (MSI-H) has been described as a distinct own entity, the mutator phenotype, in carcinogenesis. In this study a total of 66 gastrointestinal tumours (27 colorectal and 39 gastric) were analysed in parallel with their histologically normal adjacent tissues. The DNA profiles of tumour and matching normal tissue were obtained using fifteen autosomal polymorphic tetranucleotide markers and the gender marker amelogenin. In addition, four X-chromosomal tri- and tetranucleotide markers were studied to determine their allelic distribution in the Finnish population and in the gastrointestinal tumours.

From the forensic and paternity testing perspective, the aim of the present work was to evaluate the appropriateness of tumour tissue as the source of a DNA profile of an individual. The interest was focused on the stability of the DNA in gastrointestinal tumour tissues, and on the occurrence of possible errors in DNA profiling when using such tissue material. Gastric and colorectal cancers were chosen for the study because of their high incidence in the population and because of their inherent genetic instability. The reliability of the analysis is critical in cases where DNA profiling and subsequent identification or paternity testing should be performed from a tumour tissue sample. Only 27 (41%) of the 66 tumour samples tested showed no alterations in their DNA profiles. The rest (59%) of the tumours had at least one alteration, either MSI or LOH. Male samples with frequent LOH also frequently showed loss of the Y-chromosomal gender marker amelogenin. The results show that when testing identity from tumour samples a heterozygote individual may be interpreted as a homozygote, a homozygote as a heterozygote, and a male as a female. An MSI-H sample without the corresponding control tissue could easily be detected, with at least one locus manifesting extra alleles, but the detection of LOH without the corresponding tissue was not straightforward. Thus careful interpretation of the data, and possibly the use of stable surrogate markers, is warranted when using tumour tissues as the source of DNA for forensic identification, especially in paternity analysis.

The applicability of the human identification-targeted markers in revealing MSI and LOH in colorectal and gastric cancers was also tested, with a special emphasis on gastric cancer. In comparison with previous studies using dinucleotide markers located on chromosomal regions known to be deleted in gastric cancer, this study assessed if the forensically widely used tetranucleotide markers would similarly detect aberrations of the cancer genome. The results indicate that MSI-H was efficiently detected by the markers, and LOH and low-level MSI (MSI-L) were found in equal amounts as in previous studies. Importance of detection of MSI-L has been debated, since no markers specific for it has been found. MSI-L has recently been suggested to represent a real phenomenon in e.g. colorectal and

endometrial cancer, however, these studies did not include gastric cancer. In the present work, we stratified gastric cancers according to MSI and LOH and found cancers with genetic instability to associate with the intestinal type of gastric cancer. In contrast, genetic stability, as revealed by these markers, was more prevalent in diffuse type cancer. MSI-L was suggested as a true phenotype also in gastric cancer, and in a novel finding, it was to shown to coexist in tumours with LOH-H.

The third point of interest, relating to population genetics, was to evaluate the possibility that gastrointestinal tumours reflect the mutational properties of microsatellite loci in germ-line. According to our results, as assessed with X-chromosomal markers with homogeneous repeats, the likelihood of a marker to allelic alterations in MSI-H tumours correlated with its allelic diversity in the population. It was proposed that a set of MSI-H tumours might serve as a test model to pre-evaluate the mutational properties of a particular marker in a population. However, as tumours are subject to different selectional and tissue-specific effects than the germ-line cells and human populations, more studies are needed for straightforward comparisons.

REVIEW OF THE LITERATURE

1. ORGANISATION OF THE HUMAN GENOME

The diploid human genome consists of 6400 gigabase pairs of DNA, of which half is inherited from the mother and half from the father. The 23 chromosome pairs have therefore a maternal and a paternal counterpart. In addition to this, a small fraction of DNA, 16.5 kilobase pairs, corresponding to approximately 0.0003% of the total genome, is maternally inherited in mitochondria. The first human genome sequences were reported in 2001 (Venter *et al.* 2001, The International Human Genome Sequencing Consortium 2001) and a more accurate whole-genome assembly was reported three years later (Istrail *et al*. 2004). According to the whole-genome sequence, only 1% of the DNA was found to consist of exons, 24% of introns, while the function of the remaining 75% is not precisely known to date (Venter *et al.* 2001).

The human genome can be divided into single-copy and repetitive DNA, and regions coding for genes are usually placed in the former category. The latter category comprises roughly 45% of the genome. The repetitive elements may be categorised into short interspersed elements (SINE) and long interspersed elements (LINE), and tandemly arranged repeats e.g. mini- and microsatellites (Li *et al*. 2001). The human genome has been reported to contain more than 4.3 million repetitive elements, with *Alu* repeats of the SINE category and L1 repeats of the LINE category being the most frequent (Li *et al.* 2001). Figure 1 illustrates an overview of the organisation of the human genome.

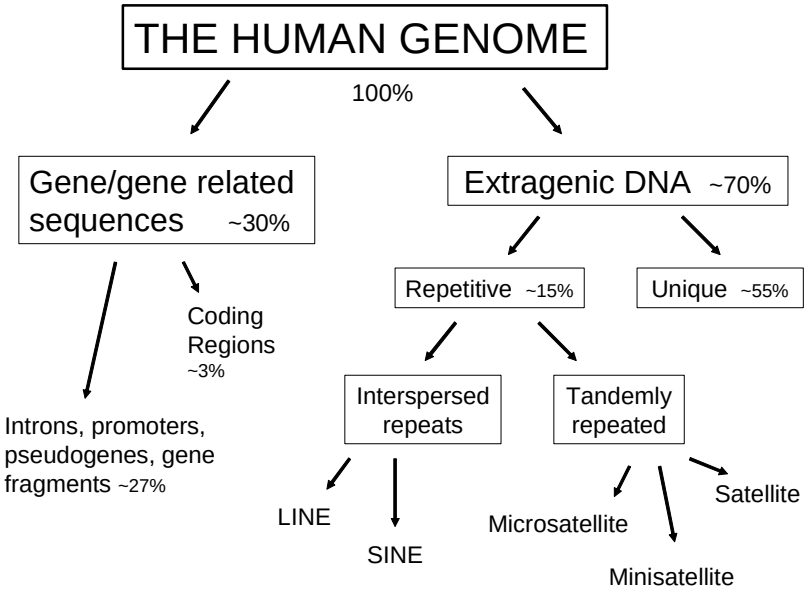

Figure 1. *Organisation of the human genome and division into categories of single-copy and repetitive elements. Figure modified from Bennett (2000).*

2. MICROSATELLITES

Microsatellites, or short tandem repeats (STRs), are regions of repetitive DNA in the genome (Tautz 1989, Weber and May 1989). The length of the repeating unit is 1-6 basepairs (bp) and the size of the repeat ranges around a mean of 100 bp (reviewed by Koreth *et al.* 1996 and Ellegren 2004). Microsatellites show a high degree of polymorphism at the population level and are found in all eukaryotic genomes like yeast, fruit flies, mice, primates or plants, e.g. *Cannabis sativa* (Tautz 1989, Tòth *et al.* 2000, Gilmore *et al.* 2003). The polymorphism is due to a high number of alleles per locus, each bearing a different number of repeats, which are inherited in a Mendelian fashion. Thus every individual carries a unique set of microsatellites in his or her genome and passes half of them to their offspring. Therefore, microsatellite analysis is widely used in mapping of genetic disease and in population genetic and evolutionary studies, and in forensic applications (Koreth *et al.* 1996). The repetitive DNA regions are prone to mutations and thus they can be used in determining the relationships among very closely related species or individuals (Koreth *et al.* 1996, Tòth *et al.* 2000, Forster *et al.* 2000).

The human genome is estimated to have 40 000 microsatellite loci, comprising about 3% of the total DNA (Subramanian *et al.* 2003). Although the overall densities of microsatellite repeats are relatively uniform across human chromosomes, the densities vary according to location (Subramanian *et al.* 2003). Microsa-tellites can be divided into mono-, di-, tri-, tetra-, penta- and hexanucleotides depending on the length of the repeat unit. Mononcleotide repeats are the most copious (Tòth *et al.* 2000, Subramanian *et al.* 2003). The abundance of microsatellites decreases with the length of the repeat unit, i.e. penta- and hexanucleotide repeats are relatively rare. The repeat structure also correlates with the abundance of the repeat, e.g. $(AC)_n$ and $(AT)_n$ occur more frequently than $(AG)_n$ and $(GC)_n$ (Rubinsztein *et al.* 1995, Subramanian *et al.* 2003). Microsatellites are distributed throughout the genome in coding and noncoding regions but their precise role, if any, is not well understood (Tòth *et al.* 2000, Subramanian *et al.* 2003).

Trinucleotide repeats have been found in coding regions with a twofold density compared to other repeats; they can show length variation without interrupting the reading frame (Submaranian *et al.* 2003). The length polymorphism at these repeats may cause susceptibility to certain clinical conditions, e.g. polymorphism at the androgen receptor (ARA) CAG region has been suggested to correlate with male infertility (Wallerand *et al.* 2001). However, for unknown reasons, some of these coding triplet repeats undergo rapid expansion causing severe diseases such as Huntington's disease, Fragile X syndrome, Kennedy disease or Myotonic dystrophy (Biancalana *et al.* 1992, Fu *et al.* 1992, Snell *et al.* 1993, Verheij *et al.* 1993, Cummings and Zoghbi 2000). A recent whole-genome *in silico* survey found 62 such triplet-repeat genes, of which 14 are known to have disease-causing expansions (Collins *et al.* 2003).

Table 1. *Examples and characteristics of simple and complex repeats.*

Locus name	Structure	Repeat[a]	Alleles[b] (n)	Repeat Range[c]	Reference
HPRTB	simple	$(tcta)_x$	8	9-16	Edelmann *et al.* 2001
D7S808	simple	$(cttt)_x$	12	7-11	Akiyama *et al.* 2002
D6S683	complex	$(tcta)_x(ta)_1(tcta)_{0-2}(ta)_{0-1}(tcta)_x$	12	13.2-22	Barrall *et al.* 2000
D22S6823	complex	$(tatatc)_x(tatc)_x(atc)_{0-1}(tatc)_x$	20	12-21	Barrall *et al.* 2000

[a] *The repeating unit is indicated in parentheses, where x is the number of repeats.*
[b] *Number of alleles found in a population.*
[c] *The range of repeat number.*

The structures of microsatellites can be divided into simple and complex repeats. The simple repeats consist of an uninterrupted run of units sharing the same sequence (Urquhart *et al.* 1994, Kayser *et al.* 2004). For nomenclature purposes the observed number of repeats forms the basis of allele calling (Bär *et al*. 1997, Kayser *et al.* 2004). Simple repeat loci may also have variant alleles with an insertion or deletion of one base in one of the repeat units. Their nomenclature is based on the number of perfect repeats followed by the number of bases in the imperfect portion. For instance, the sequence for the TH01 allele 9.3 is $(AATG)_6$-ATG-$(AATG)_3$, whereas for allele 9 it is $(AATG)_9$. Complex microsatellites consist of interrupted repeats or repeats different in sequence or size. Uncertainties in recognising the repeat and its direction may lead to confusion (Szibor *et al.* 2003a, Kayser *et al.* 2004). The nomenclature can be based either on the overall number of repeats or by the length of the amplicon in base pairs (Bär *et al.* 1997, Kayser *et al.* 2004). Examples of simple and complex microsatellites are shown in Table 1.

3. MICROSATELLITE ANALYSIS IN FORENSIC MEDICINE

3.1. Human identification

The invention of the polymerase chain reaction, PCR (Saiki *et al.* 1985, Mullis and Faloona 1987), represented a break-through for the analysis of the human genome in biomedical sciences. For resolving forensic questions, individual variation in a set of polymorphic loci could be revealed e.g. by simultaneously amplifying several microsatellite regions. By using modern techniques the individual-specific DNA profile can be obtained from a small number of cells and analysed through high-throughput electrophoresis (reviewed by Butler *et al.* 2004). Because the DNA sequence in general is uniform in all tissues of an individual almost any tissue can serve as a DNA source. By choosing a set of highly polymorphic microsatellites for the amplification procedure, an efficient and reliable way to discriminate between individuals is achieved. Tri- and tetranucleotide microsatellites are preferred in human identification, because of the accurate allele calling due to lower occurrence of PCR-based artefacts (Urquhart *et al.* 1994). The multiplex amplified PCR products are separated according to their sizes, and the differently migrating fragments are visualised by DNA staining or detection of fluorescent signals ("peaks") generated by labelled PCR primers (Butler *et al*. 2004). Allelic ladders help the allele-calling in samples of unknown identity. Based on the frequencies of different genotypes in a given population, a probability for the incidence of a particular genotype may be determined.

Gender determination is usually included the amplification procedures by analysing of an X-Y homologous gene amelogenin (Sullivan *et al.* 1993). Designing the amelogenin primers so that a region of different length in both chromosomes is amplified, unequivocal gender determination (female XX or male XY) is achieved. However, certain Y-chromosomal haplotypes have been shown to lack the amelogenin gene, hence the individual is typed as XX instead of XY (Santos *et al.* 1998, Roffey *et al*. 2000, Steinlechner *et al*. 2002, Tharangaj *et al.* 2002, Chang *et al*. 2003, Michael and Brauner 2004). Also the X-chromosomal amelogenin signal has been found to be absent as a result of a C-to-G transversion in the region where the commonly used genotyping primers bind (Shadrah *et al*. 2004).

Different commercial genotyping kits for detecting different loci are currently available. In order to facilitate comparison of DNA profiles, a set of core markers has been agreed upon in the forensic community (Butler *et al.* 2004). These markers are always typed in casework samples, for population genetic studies, and for criminal offender DNA-register purposes. For example, the 13 core markers selected by the Federal Bureau of Investigation (FBI) include CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11 (Butler *et al.* 2004). Figure 2 illustrates a DNA profile obtained by AmpF*l*STR® SGM™, which includes ten of the markers used by the FBI.

DNA profiles can also be reliably achieved from minute amounts of DNA or severely degraded human-derived material. A hair shaft or a cigarette butt from a crime scene can serve as a source of DNA (Hochmeister *et al.* 1991, 1995). Decomposed bodies and bone samples can also be used as a source of DNA because bone tissue DNA is well protected against degradational artefacts (Hoff-Olsen *et al.* 2001). For instance, the identification of Joseph Mengele (Jeffreys *et al.* 1992) and the Romanov family (Gill *et al.* 1994) has been based on microsatellite analysis of skeletal remains. Choosing primers which result in shorter amplicon lengths may enrich the chance to obtain results from degraded DNA (Hoff-Olsen *et al.* 2001).

In some cases, cancerous tissue has also served as a DNA source in identification studies. For instance, when tissue blocks are mislabelled in a pathology laboratory and their sources has to be identified, DNA profiles between the tissue specimen and the patient should match (Banaschack *et al.* 2000, Iwamoto *et al.* 2003). In some cases where reference DNA is required for the identification of a deceased person or a disputed paternity has to be confirmed, the only reliable source may be a paraffin block of a malignant tissue from a histology archive. In such cases, the DNA profiles must be interpreted with caution (Rubocki et al. 2000, Pai *et al.* 2002, Poetsch *et al.* 2004, Schwark *et al.* 2004).

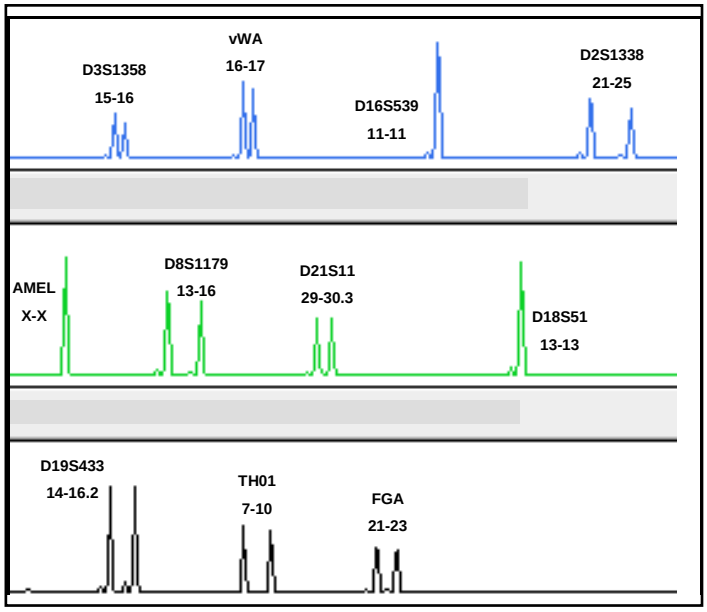

Figure 2. *Electropherograms obtained with AmpFlSTR SGM Plus set of markers. Three different dye channels (blue in the upper, green in the middle and black in the lower graph) and the genotype of individual loci are indicated.*

3.2. Paternity analysis

Paternity and kinship analysis can be effectively performed by analysis of microsatellites (Thomson *et al.* 1999, 2001). The parents pass one of their alleles from each of the loci to their descendants. In cases with normal chromosomal transmission, one allele of a particular autosomal microsatellite locus is inherited from the father and the other from the mother. By comparing the DNA profiles of the child, mother and putative father, statistical interpretation in a form of the paternity index can be calculated in cases where an parenthood is not excluded (Thomson *et al.* 1999). In the future, however, it may be foreseeable that a high number of single nucleotide polymorphisms (SNPs) will be analysed in addition to or instead of microsatellite loci. The advantage of SNPs is the low mutation rate (10^{-8}) in comparison to microsatellites (in the range of 10^{-3}), although the number of SNPs needed for achieving high paternity indexes is much higher that that of microsatellites (Gill 2001, Amorim and Pereira 2004). However, since the detection of multiple SNPs per sample needs high throughput detection and re-generation of criminal offender registers, the present system is to date the most convenient method.

When the descendant in question is a male the Y-chromosomal microsatellites may be used in paternity analyses, especially in deficiency cases where the alleged father is deceased (Kayser *et al.* 1997, Ruitberg *et al.* 2000). The non-pseudoautosomal region of the Y chromosome does not recombine during meiosis and it is inherited as a haplotype. Therefore, all the loci found in the father and the descendant are of one genotype unless mutations occur, and the Y-chromosomal haplotype may be traced as a paternal lineage to distant male relatives or pedigree ancestors (Carvalho-Silva *et al*. 1999, Forster *et al.* 2000). For instance, Y-chromosomal analysis has revealed new insight to the disputed paternity of President Thomas Jefferson of Sally Heming's descendant(s) (Forster *et al.* 1998).

In some particular cases when the child with disputed parentage is a female X-chromosomal microsatellite markers can be used (Edelmann *et al.* 2001, Zarrabeitia *et al.* 2002a, 2002b, Wiegand *et al*. 2003, Szibor *et al.* 2003a, 2003b, Shin *et al.* 2004). The putative father carries a single X-chromosome, and the haplotype is passed on to the daughter. Therefore, the analysis of X-chromosomal microsatellites may be, e.g. in deficiency cases where one of the parents is missing, more informative than conventional autosomal microsatellite analysis (Szibor *et al.* 2003b). Figure 3 and Table 2 illustrates X-chromosomal analysis in a family case.

However, as with all genetic markers, solving disputed paternity or kinship may be problematic when mutations are found in the descendant (Thomson *et al.* 1999, Kayser and Sajantila 2001). Therefore, the estimation of microsatellite mutation rates is of great importance in forensic DNA studies. Markers with high mutability may result in erroneous identity when analysing kinship (Brinkmann *et al.* 2001), e.g. in investigations by reverse paternity for victims of war (Gornik *et al.* 2002). In addition, understanding the mechanisms and factors underlying the mutations may reveal new insights to the evolution of the microsatellites.

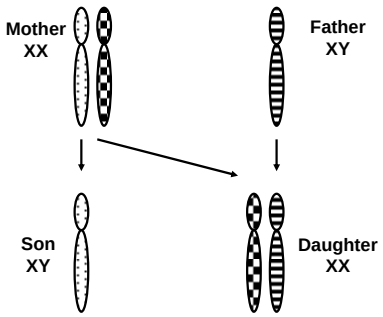

Figure 3. *The pattern of X-chromosomal inheritance.*

	DXS7423	DXS101	DXS8377	ARA
Mother	13-15	24-24	48-49	23-25
Father	13	26	49	30
Son	13	24	49	23
Daughter	13-15	24-26	48-49	23-30

Table 2. *The X-chromosomal genotype in a family case. The son has maternal alleles. The daughter has a combination of maternal and paternal alleles.*

4. MICROSATELLITE MUTATIONS IN THE GERM-LINE

4.1. Mutational mechanism and dynamics

Molecular mechanisms underlying the mutations of microsatellite loci are thought to result from the interplay between replication errors and their post-replication repair. The error-producing mechanism has been suggested to be slippage of the DNA polymerase enzyme (reviewed by Ellegren 2004). Replication slippage or slipped-strand mispairing, as illustrated in Figure 4, refers to the misalignment of the two DNA strands after strand dissociation ("strand breathing") prior to polymerisation (Ellegren 2004). The observed slippage rates vary between organisms and it may explain the observed variation in microsatellite length distribution (Krugylak *et al.* 1998, Ellegren 2004). The "stepwise mutation model", where alleles increase or decrease by one unit, was concluded to originate from replication slippage, observed as mutations occurring in single steps (Levinson and Gutman 1987). The best fit with the allelic distributions seen in populations was obtained with this model (Weber and Wong 1993, Valdes *et al*. 1993, Chakraborty *et al.* 1997, Klintschar and Wiegand 2003). Data obtained from family studies corroborate the single-step model, although longer changes of two or three repeat units have also been noticed (Sajantila *et al.* 1998, Kayser *et al*. 2000, Brinkmann *et al.* 1998, Leopoldino and Pena 2003). However, another study found dinucleotide markers to change mostly by larger steps (Huang *et al.* 2002). This may indicate a different mechanism for dinucleotides compared with longer repeat microsatellite motifs.

The evolution of a microsatellite sequence can be considered as a balance between slippage events and point mutations (Krugylak *et al.* 1998). A microsatellite may contain more than one repetitive tract, but only one repeat region is suggested to be polymorphic and thus harbour all the detectable mutations (Leopoldino and Pena 2003). Depending on the primer locations, however, the amplification product may contain more than one microsatellite tract and thus more polymorphic regions may be observed (Kayser *et al.* 2004). Various factors have been found to affect mutation rates (Schlötterer 2000, Ellegren 2004)

.

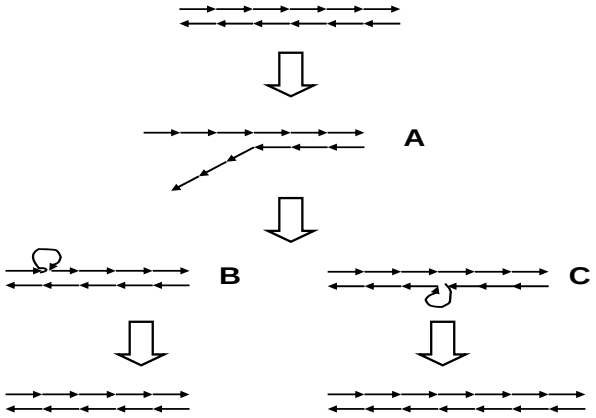

Figure 4. *The strand slippage model for microsatellite mutations. The DNA strands dissociate from each other (A) and re-anneal with a mismatch of one repeat unit. A mismatch in the template strand results in contraction (B), while a mismatch in the nascent stand results in expansion (C). Figure modified from Ellegren (2004).*

For instance, the average repeat number at a locus has been shown to be proportional to its polymorphism (Chakraborty *et al.* 1997, Ellegren 2000). Accordingly, the geometric mean of the longest run of perfect repeats has been found to correlate with germ-line mutations (Brinkmann *et al.* 1998, Leopoldino and Pena 2003). Smaller alleles seem to have a tendency to gain repeat units, whereas long alleles tend to lose repeats (Xu *et al.* 2000, Huang *et al.* 2003, Dupuy *et al.* 2004). However, the existence and nature of upper (and lower) boundaries on allele size are still unknown (Ellegren 2000, Xu *et al*. 2000). Sequence interruptions by imperfect repeats lead to stabilisation of the microsatellite (Brinkmann *et al.* 1998, Bacon *et al.* 2000, Sturzeneker *et al.* 2000). In addition to the differences in mutation rates in different microsatellite loci, the mutation rates have also been observed to vary among the alleles of the same locus (Jin *et al*. 1996, Carvalho-Silva *et al.* 1999, Ellegren 2000).

A male bias in germ-line mutations has been noticed (Ellegren 2004), but the tendency for maternal and paternal mutations may also be locus-specific (Henke and Henke 1999). The sex-specific mutation rates are suggested to originate from a larger number of mitotic cell divisions in spermatogenesis than in oogenesis, with an assumption that most mutations are replication-dependent (Ellegren 2000). Therefore, the mutations should also correlate with paternal age, but as yet no significant difference has been observed (Sajantila *et al.* 1998, Kayser *et al.* 2000, Kayser and Sajantila 2001, Leopoldino and Pena 2003, Dupuy *et al.* 2004).

4.2. Inferring mutation rates

The mutations of microsatellites manifest cell division. Thus, having an estimation of the mutation rates at certain loci is critical e.g. when dating the last recent common ancestor from phylogenetic data (Morral *et al.* 1994). The most straightforward way to obtain mutation frequencies is to analyse family data and directly observe inconsistencies in parent-descendant allelic transfers. However, a large number of samples are needed for reliable mutation rates (Sajantila *et al.* 1998, Brinkmann *et al.* 1998, Xu *et al.* 2000, Leopoldino and Pena 2003, Dupuy *et al*. 2004). One study has been conducted on mutational frequencies observed directly from diluted pools of sperm DNA (Holtkemper *et al*. 2001). However, this study directly observed only the repeat gains from which the total mutation rate was extrapolated, giving a result of higher estimates than in studies with direct pedigrees. Table 3 summarises the observations of six studies with parent-descendant allelic transfers.

Table 3. *Mutation rates and the proportion of single step events at autosomal and Y-chromosomal loci observed in family studies.*

Study	Allele transfers (n)	Locus Type [a]	Locus (n)	Averaged mutation rate	Single step events	Paternal origin [b]
Sajantila *et al.* (1998)	29.640	AS	5	0.6×10^{-3}	56%	100%
Brinkmann *et al.* (1998)	10.844	AS	9	2.1×10^{-3}	96%	85%
Leopoldino *et al.* (2003)	24.224	AS	10	9.5×10^{-4}	96%	82%
Xu *et al.* (2000)	287.786	AS	122	1.8×10^{-3}	85%	76%
Kayser *et al.* (2000)	4.999	Y-STR	15	2.8×10^{-3}	93%	
Dupuy *et al*. (2004)	15.894	Y-STR	9	2.3×10^{-3}	96%	

[a] *AS, autosomal, Y-STR, Y-chromosomal.*
[b] *Percent of cases where the mutating allele could be deduced.*

Mutation rates may also be inferred by deep rooting pedigrees e.g. in the non-recombining Y-chromosomes or autosomal microsatellites with linkage to a haplotype (Forster *et al.* 2000, Carvalho-Silva *et al.* 1999, Heyer *et al.* 1997, Lin *et al.* 1996, Hästbacka *et al.* 1992, Morral *et al.* 1994). These methods give only estimates of the actual mutational frequencies, because the ancestral haplotypes and thus paternity are not confirmed. In addition, as both additions and deletions of the repeat units are possible, the results are not reliable because an allele can mutate back to its original length (Carvalho-Silva *et al.* 1999). The allelic distribution seen in a population can also help in inferring mutation rates, assuming that the demographic history of the population is known (Chakraborty *et al.* 1997, Di Rienzo *et al.* 1998, Andrés *et al.* 2003). Because of obligatory uncertainties in the demographic history of populations and population effects unknown to date, the mutation rates obtained this way may not be fully reliable. However, loci with the highest levels of polymorphism at the population level have been suggested to have the highest rate of germ-line mutations (Ellegren 2004).

5. MICROSATELLITE ANALYSIS IN CANCER RESEARCH

In the progression of cancer the tumour accumulates genetic changes in genes that control cell birth and death, thus enabling growth and metastasis by the tumour cell population. These changes can affect the whole genome by alterations in chromosome number or chromosomal translocations or amplifications, or affect single genes by introducing point mutations, insertions or deletions. The selective conditions in the tumour environment give rise to clonal expansion, allowing the tumour cell with the mutated allele to outnumber its sister cells (reviewed by Lengauer *et al.* 1998 and Loeb *et al*. 2003). However, the mechanisms which enable these mutations to arise are still unknown. Normal mutation rates have been suggested to be insufficient to produce the observed mutations, despite the high number of cell divisions during the final clonal expansion of the tumour. Thus it has been proposed that an underlying genetic instability, i.e. mutator phenotype, is required to generate the numerous mutations observed in particular tumours (Loeb *et al.* 2003). In order to produce the mutator phenotype a primary loss of function in a key process is needed which enables the increase in mutation rates. These key processes may be controlled by factors which are involved in controlling genetic integrity, i.e. enzymes involved in replication (e.g. helicases), error correction (e.g. mismatch repair) or mitosis (e.g. chromosome segregation) (Lengauer *et al.* 1998, Jallepalli and Lengauer 2001, Levitt and Hickson 2002 and Loeb *et al.* 2003). Furthermore, some of these genes have also been found to predispose to recessive genetic syndromes with increased risks of cancer (reviewed by Hickson 2003 and D'Andrea and Grompe 2003).

Microsatellites have been frequently used in cancer research, because they are distributed throughout the genome, act as mutational hotspots due to their repetitious nature, and allow determination of genomic changes due to high heterozygosity (reviewed by Koreth et al 1996 and Bennett 2000). Thus, microsatellite analysis may be used as an indication of genomic integrity. The detection of single point mutations or small insertions or deletions in non-repetitive areas is laborious because the regions manifesting mutations cannot be predicted. Comparing the DNA profiles obtained by microsatellites of a tumour and its neighbouring non-cancerous tissue (or blood lymphocytes) may reveal two different types of alterations. The emergence of alleles not present in healthy tissue is termed microsatellite instability (MSI) (reviewed by Loeb 1994, Lengauer et al. 1998 and de la Chapelle 2003), whereas loss of one of the heterozygote alleles is termed loss of heterozygosity (LOH) (Lasko et al. 1991, Tomlinson et al. 2002). Figure 5 illustrates MSI and LOH.

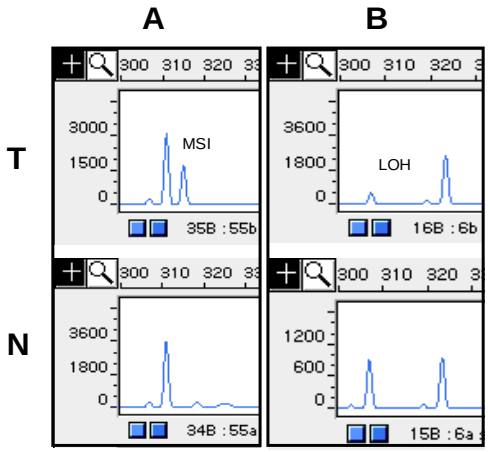

***Figure* 5.** *Examples of MSI (panel A) and LOH (panel B) at the vWA microsatellite locus. T denotes tumour and N normal tissue.*

5.1. Microsatellite instability

The MSI phenomenon, occurring in a subset of colorectal cancers, was found simultaneously by three research groups (Aaltonen *et al.* 1993, Ionov *et al*. 1993, Thibodeau *et al.* 1993). Most of the tumours found in the patients with the hereditary non-polyposis colon carcinoma (HNPCC) syndrome, a dominantly inherited disorder, were found to manifest MSI (Aaltonen *et al.* 1993). The linkage of MSI with a hereditary condition lead to the characterisation of two predisposing chromosomal regions (Peltomäki *et al.* 1993, Lindblom *et al.* 1993). Soon thereafter, similarities between the syndrome phenotype and mismatch repair (MMR) deficiency of *E. coli* was found, and the defective genes in HNPCC were found to be *E. coli* MMR-component homologues *MLH1* and *MSH2* (Fishel *et al*. 1993, Bronner *et al.* 1994). The key event leading to the MSI phenotype in HNPCC was found to be the loss of MMR function by somatic inactivation of the wild-type allele, thus resulting in the mutator phenotype (Loeb 1994). MSI was also found to occur in somatic tumours and in cancers other than colorectal ones (Rhyu *et al*. 1994, Boland *et al.* 1998), summarised by Lawes *et al*. (2003). In 90% of the sporadic gastrointestinal tumours the mutator phenotype has been shown to result from down-regulation of *MLH1* by promoter hypermethylation (Thibodeau *et al.* 1998, Herman *et* al. 1998, Fleisher *et al.* 1999, Yamamoto *et al.* 1999). The MMR system was originally characterised in *E. coli*, where it was found to recognise and correct mismatches in hemimethylated DNA (Lu *et al.* 1983). Bacterial strains with deficient MMR genes were also found to accumulate mutations in their genomes (reviewed by Schofield and Hsieh 2003). Individuals with homozygote MLH1 mutant alleles (MLH1-/- genotype) have recently been found, all of which share phenotypical features of neurofibromatosis 1 and hematological malignancies (Ricciardone *et al.* 1999, Wang *et al.* 1999). Investigation of one MLH1-/- case revealed wide-spread MSI in the apparently healthy somatic tissues of this individual (Vilkki *et al.* 2001).

MSI tumours accumulate point mutations and small insertions and deletions, but the "hot spots" are regions of repetition, e.g. microsatellite loci including also regions of mononucleotide repeats. Because the replication errors are less accurately corrected due to impaired function of the MMR genes, an insertion or deletion at a coding region may lead to a shift in the reading frame and thus to the loss of function of the protein product (Duval and Hamelin 2002). A two-base (AA) deletion in the *TGFβ* receptor II gene and subsequent truncation of the protein is illustrated in Figure 6. The *TGFβ* receptor II gene was the first target gene found, being mutated in 90% of MSI cancers (Parsons *et al.* 1995). Many other genes, such as *BAX*, *hMSH6* and *IGF2R*, have been described as target genes in cancers with the mutator phenotype (Oliveira *et al.* 1998, Ottini *et al.* 1998, Vilkki *et al.* 2001, Mori *et al.* 2001, Duval *et al.* 2002). However, the critical targets for mutation in MSI-

driven carcinogenesis are still poorly characterised, target mutations may also show tissue specificity, as reported for endometrial, gastric and colorectal cancers (Duval *et al.* (2002).

Tumour tissues with somatic MSI can be subdivided into low- and high-frequency MSI phenotypes. For the distinction between these two types, the National Cancer Institute (NCI) has selected a panel of five mono- and dinucleotide markers. High frequency MSI tumours (MSI-H), with novel alleles in at least two of the five markers, display the mutator phenotype (Boland *et al.* 1998). The MSI-H genotype may also be assessed with any set of markers, where the cut-off limit for MSI-H can be set between 33-40% (Boland *et al.* 1998). Therefore, also tri- and tetranucleotide repeats may be used for determination of MSI phenotype (Samowitz *et al.* 1995, Dietmaier *et al.* 1997, Boland *et al.* 1998). For instance, a complex Alu variable poly(A) marker MYCL1 with mono- and tetranucleotide repeats has been favoured in MSI detection (Young *et al.* 1995, Dietmaier *et al.* 1997, Jass *et al.* 2001). MSI-H tumours also show specific contractions at certain mononucleotide repeats. A quasi-monomorphic 26-adenine repeat (BAT-26) located in the fifth intron of MSH2 has been widely used in studies of microsatellite instability. This marker has been reported to reveal the MSI-H phenotype of sporadic colorectal and gastric cancers with 99.4%-100% accuracy (Hoang *et al.* 1997, Halling *et al.* 1999) and it has been suggested to be sufficient for MSI-H detection as reviewed by Duval and Hamelin (2002). The even in tumours without the matching normal tissue. However, since polymorphism at the BAT-26 locus has been detected in individuals of African origin (Pyatt *et al.* 1999), careful interpretation is needed if MSI-H detection is based solely on this marker.

Low-frequency MSI (MSI-L) tumours exhibit sporadic new alleles. In MSI-L, less than one-third of the loci analysed, or one of the loci in the NCI panel, is affected (Boland *et al.* 1998). BAT-26 is never or rarely altered in MSI-L tumours. The mechanism for microsatellite alterations in MSI-L tumours is currently unknown, but they may originate from an increased rate of errors upon DNA replication (Jass *et al.* 2002) or from normal replication errors which become detectable by the clonal expansion of a tumour tissue (Laiho *et al.* 2002, Halford *et al.* 2002). Differentiation of MSI-L from microsatellite stable (MSS) tumours is uncertain and controversial (Jass *et al.* 2001), because when the number of markers is increased more MSS tumours will probably be re-categorised as MSI-L (Laiho *et al.* 2002, Halford *et al.* 2002). No specific markers for MSI-L or clear biological differences between MSI-L and MSS cancers have been found (reviewed by Tomlinson *et al.* 2002). Therefore the MSI-L tumours are usually included in the MSS group (Umar *et al.* 2004). However, the MSI-L phenotype has been recently shown to associate with ovarian, endometrial and colorectal cancers and it was therefore suggested to represent a real phenomenon (Halford *et al*. 2003).

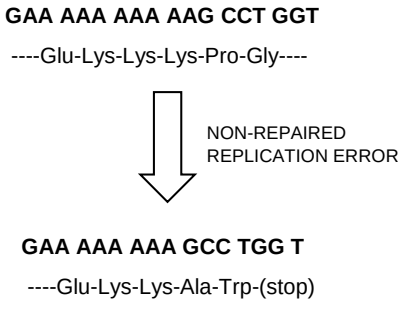

Figure 6. *Microsatellite mutation at the (A)8-repeat of the TGFβ-receptor II gene in colon cancer. The two-base deletion in the coding region interrupts the reading frame, leading to a premature translational stop codon. Figure adopted from Lengauer et al. (1998).*

5.2. Loss of heterozygosity

The concept of LOH is based on the hypothesis that two hits are needed for inactivation of tumour suppressor genes (Knudson 1971). According to the Knudson theory, both alleles of the tumour suppressor gene need to be inactivated for tumorigenesis, either by intragenic mutations, loss of the corresponding chromosomal region or methylation of the promoter (reviewed by Jones and Laird, 1999), of which only chromosomal losses may be detected by microsatellite analysis. Normally, the two alleles of a heterozygote should be detected at near-to-equal intensities depending on the detection methods. In tumours LOH is seen as a loss or reduced intensity of the other allele of a heterozygote genotype (reviewed by Lasko *et al.* 1991, Knudson 2001 and Tomlinson *et al.* 2001, 2002b). Usually heterozygote signal ratios falling below 0.5 in tumour tissues are interpreted as LOH. However, low DNA concentration and fragmented or chemically modified genomic material, e.g. formaldehyde-fixed paraffin-embedded tissue, may result in artefactual LOH (Sieben *et al.* 2000). Apparent LOH does not necessarily mean chromosomal loss; the cells may still have two chromosomes (a duplication of one and a loss of the other) or have severe rearrangements (Thiagalingam *et al.* 2001). Therefore, LOH should not solely be considered as a deletion, as frequently reported, but alternatively as a gain of genetic material resulting from e.g. chromosomal duplication or sub-chromosomal amplification (reviewed by Tomlinson *et al.* 2001 and Thiagalingam *et al.* 2002). However, it must also be kept in mind that an apparent LOH may also be a result of a point mutation in the region for amplification primer binding (Clayton *et al.* 2004). Therefore, the putative loss has to be confirmed by amplification with alternate primers.

Usually LOH is thought to indicate chromosomal instability, which can be seen as chromosomal rearrangements and aneuplody at the karyotype level (Thiagalingam *et al.* 2001). The amount of LOH has been suggested to correlate with tumour progression and growth patterns (Choi *et al.* 1998). It has also been postulated that LOH and MSI-H occur as mutually exclusive (Lengauer *et al*. 1998), but this has recently been questioned (Goel *et al*. 2003, Tang *et a*l. 2004, Hiyama *et al*. 2004). However, distinguishing LOH from MSI in MSI-H tumours may be troublesome, since an apparent LOH may be a homozygous genotype generated by MSI or result from a mutation in the primer binding site. Assessing LOH in allelotyping analysis has been used in the search for tumour suppressor genes, because the loss of one functional gene by LOH and impaired function of the other by some other mechanism might give a growth advantage to the cell clone and enhance tumorigenesis (Vogelstein *et al.* 1989, Shen *et al.* 2000). However, although LOH is frequent in tumours, very few studies have succeeded in identifying new tumour-suppressor genes as targets of LOH (Tomlinson *et al.* 2002).

5.3. MSI and LOH in colorectal and gastric cancer

Colorectal cancer (CRC) is the most common cancer in Western countries. Approximately 15% of sporadic CRC cases show MSI-H phenotype; this is mainly caused by epigenetic silencing of MLH1. In HNPCC, which occurs in 1-5% of all CRC cases, almost all individuals harbour a mutation in one of the three genes (*MLH1, MSH2 and MSH6*) coding for MMR proteins (Fishel *et al.* 1993, Bronner *et al.* 1994, de la Chapelle 2003). Almost all of the HNPCC tumours show MSI-H, hence detection of unstable microsatellite loci is one of the diagnostic methods used (Dietmaier *et al.* 1997, Boland *et al.* 1998). The MSI-H phenotype is known to associate with tumours arising in the right colon, with poor differentiation or mucinous cell type, lymphocyte infiltration, older age and female gender (reviewed by Duval and Hamelin 2002). Recently, the MSI-L phenotype has been reported to occur in 7-14% of sporadic CRCs (Goel *et al.* 2003, Tang *et al.* 2004) of cases.

CRC has also served as a model for genetic changes in tumour progression. Non-MSI-H cancers have been observed to lose on average 25% of randomly chosen alleles, occurring in stepwise alterations from normal epithelium via

adenoma to subsequent carcinoma (Fearon and Vogelstein 1990, Kinzler and Vogelstein 1996). This pathway or "Vogelgram", as presented by Goel *et al.* (2003), of tumour suppressor gene inactivation is based on the sequential accumulation of particular mutations and LOH (e.g. *APC*, *K-RAS*, *TP53* and *DCC* genes) during the adenoma-carcinoma transition (reviewed by Fearon and Vogelstein 1990 and Kinzler and Vogelstein 1996).

Gastric cancer (GCA) is globally the second most common cause of cancer death. GCA is a heterogeneous disease where a variety of subtypes with different prognoses can be seen. Different classifications for GCA are used (Laurén 1965, Oota and Sobin 1990, Ming 1998). Association of *Helicobacter pylori* infection with 80% of gastric cancers has been described. The eradication and decreased incidence of *H pylori* has lowered GCA incidence significantly in Western countries (reviewed by Sipponen 2002). Correa (1992) and Sipponen (2002) have described sequential histological changes (i.e. chromic atrophic gastritis, intestinal metaplasia and dysplasia) in normal gastric epithelium leading to carcinoma. However, the molecular mechanisms beyond these changes are poorly understood, since no single genomic abnormality is known to be specific for sporadic GCA (reviewed by Zheng *et al.* 2004). Based on histology, GCA can be divided into two major morphologic categories, i.e. intestinal (IGCA) and diffuse (DGCA) types, which comprise approximately 90% of the adenocarcinomas (Laurén 1965). The remainng 10% represent unclassified (mixed) type (Laurén 1965, Sipponen 2002). IGCA shows differentiated structures, e.g. gland formation, and its pathogenesis has been linked to pre-existing gastric atrophy, intestinal metaplasia and dysplasia (Correa 1992). DGCA has no known precursor lesions, and the cancer cells grow as scattered single mucous cells in gastric mucosa (Correa 1992, Sipponen 2002).

The MSI-H phenotype is reported to occur in approximately 20-33% of gastric cancers (Fleischer *et al.* 1999, Choi *et al*. 2000), representing the highest level among human sporadic cancers. IGCA tumours show higher prevalence (15-23%) of MSI-H cancers than DGCA tumours (2.5-7%) (Buonsanti *et al*. 1997, Choi *et al*. 2000, Habano *et al*. 2000, Hayden *et al*. 2000, Kim *et al*. 2003). Already at early stages GCA is thought to harbour LOH (Tamura *et al*. 1996, Nishizuka *et al.* 1998). However, unlike in colorectal cancer, no clear sequential accumulation of genetic alterations has been found upon progression of GCA (Zheng *et al.* 2004).

6. TUMOURS AS MODELS FOR MICROSATELLITE MUTATIONS

As described above, microsatellite mutation rates have been inferred from family data, pedigree analysis or from population diversity data. Mutations most probably occur also in somatic tissues, but seldom become visible because of the polyclonality of cells (Mann *et al.* 2003). Tumours, however, represent genetically unstable lineages of cell clones which have undergone clonal selection and considerable numbers of cell division. Thus tumour genomes manifest numerous mutations (Tomlinson *et al.* 2002c). Few reports on the use of tumours for modelling microsatellite mutations have been published (Hoff-Olsen *et al.* 1998, Di Rienzo *et al.* 1998, Sturzeneker *et al*. 1998, 2000, Bacon *et al.* 2000). It has been suggested that the somatic mutation patterns and mechanisms observed in cancer do not differ significantly from those in normal germ-line cells (Di Rienzo *et al.* 1998, Sturzeneker *et al*. 1998, 2000). The amount of MSI seen at a particular locus in tumours has therefore been observed to be associated with its diversity at the population level (Di Rienzo *et al.* 1998, Sturzeneker *et al*. 2000).

Generally the somatic mutation rates at most loci are too low, even in tumours, to be efficiently detected by PCR (Shibata *et al*. 1996). However, the mutation rates in MSI-H tumours or MSI-H cell lines have been estimated to be 100 to 1000 fold higher than in MMR-proficient cells (Bhattacharyya *et al.* 1994, Eshleman *et al.* 1996, Umar *et al.* 2004b). Unrepaired slippage during DNA replication is the likely mechanism for microsatellite mutations in MMR-deficient cells, thus linking mutation with cell division (Blake *et al.* 2001). Accordingly, a mutation frequency of 5-10 $x10^{-3}$ has been observed for dinucleotide microsatellites in MMR-deficient cell lines (Shibata *et al.* 1996). Poly(A) deletions (like at the 26 adenine residues of BAT-26 locus) occur stepwise in MMR-deficient cancers. Larger

poly(A) deletions in tumours suggest for multiple contraction events, and thus counting the number of deletions gives an estimate of the number of cell divisions after the loss of MMR proficiency (Blake *et al.* 2001, Kim *et al.* 2002). Differences between the germ-line genotype and the final founder cell of the tumour reflect the number of mitotic divisions that occur during tumorigenesis and thus can be utilised to reconstruct the history of the tumour (Tsao *et al.* 2000). Comparing the mutational events at different loci in a MSI-H tumour may give an estimate of their mutational frequencies.

AIMS OF THE PRESENT STUDY

The aims of the present study were as follows:

1. To evaluate neoplastic tissues as a source of genetic information in forensic settings.

2. To investigate the applicability of tetranucleotide markers, primarily used in forensic identity testing, to assess the genetic integrity of gastrointestinal tumours with a special emphasis on gastric cancer.

3. To estimate the variability of the alleles generated by MSI tumours and to correlate their distribution in alleles found at corresponding loci in the Finnish population.

MATERIALS AND METHODS

1. TISSUE SPECIMENS AND DNA EXTRACTION (I-IV)

Surgically resected tissue specimens from primary gastrointestinal (GI) cancers and from their completely separated cancer-free adjacent areas were collected and fresh-frozen during the period 1995 to 2003 at the Department of Pathology of Jorvi Hospital, Espoo, Finland. Histopathological diagnosis was obtained from routinely stained tissue sections. Care was taken during preparation on the histology of the tumour samples to achieve adequate tumour cell content. The tissues were disrupted by proteinase K digestion and the DNA was purified in Qiaquick columns (Qiagen, Hilden, Germany).

A total of 66 GI cancer specimens were used in the studies, of which 27 were colorectal and 39 were gastric cancers. In addition, in study II two gastrointestinal stromal tumours (GIST) were included. Some of the samples were used in multiple studies, and study III included GCA samples not included elsewhere, thus the total number of samples is greater than in any of the individual studies. The number and origin of GI tumours analysed in studies I-IV are indicated in Table 4.

In study III the 37 gastric cancers investigated were stratified into intestinal (25) and diffuse (12) type by routine microscopy according to Laurén (1965), and the tumour cell content (50%-100%) was estimated from the tissue samples. The 103 samples representing the Finnish population (IV) consisted of adjacent tissues of surgically removed tumours and lymphocyte samples of non-related healthy individuals. The DNA from lymphocytes was purified with proteinase K digestion followed by phenol/chlorophorm extraction (Sambrook *et al.* 1989).

Table 4. *Number of gastric (GCA) and colorectal (CRC) tumour samples used in each study (I-IV).*

Study	GCA (n)	CRC (n)	Total (n)
I	18	23	41
II	22	22	46
III	37	-	37
IV	30	27	57

2. TUMOUR GENOTYPING (I-IV)

The tumours and their adjacent healthy tissues were genotyped separately with the AmpF*l*STR® SGM Plus™ and AmpF*l*STR® Profiler ™ kits (Applied Biosystems, Foster City, CA, USA), which consist of primers for 15 autosomal tetranucleotide markers and primers for the amelogenin gene for gender typing. The chromosomal locations of the loci are shown in Table 5. The PCR procedures, fragment separation and genotyping analysis were carried out by ABI Prism CE 310 and 3100 capillary electrophoresis (Applied Biosystems) according to the manufacturer's instructions. The resulting DNA profile from each tumour and its adjacent normal tissue were compared and the tumour samples were categorised into phenotypes according to the autosomal allelic alterations.

2.1. Microsatellite unstable phenotypes (I-IV)

The MSI phenotype showed emergence of novel alleles and was divided into two groups according to the NCI recommendations (Boland *et al.* 1998).

The samples with extra alleles at ≥ 5/15 loci were designated as MSI-H whereas samples with extra alleles at < 5/15 loci were designated as MSI-L. The accuracy of the forensic markers to detect the MSI-H phenotype was confirmed with amplification of BAT-26 (Hoang *et al.* 1997, Halling *et al.* 1999).

2.2. Phenotypes with LOH (I-III)

A tumour sample was considered to have LOH at a locus when the ratio of fluorescence intensities (LOH ratio) between the tumour sample (Yt/Xt) and the respective normal tissue (Yn/Xn) fell below 0.5 (D'Adda *et al.* 1999). A locus was considered as informative for LOH when the normal genotype was heterozygous and the locus did not show MSI in the tumour tissue. For studies I and II, samples with LOH at <33% of the informative loci were designated as low-level LOH (LOH-L) and samples with LOH at ≥33% of the loci were designated as high-level LOH (LOH-H). In study III the value of ≥25% for LOH-H was used according to Choi *et al*. (2000).

2.3. Microsatellite stable phenotypes (I-IV)

In genetically stable tumours no alterations were detectable. In the studies I and II both MSI and LOH negative tumours were categorised as microsatellite stable (MSS). In study III the LOH negative phenotype was designated as LOH-non detectable (LOH-N) according to Choi et al. (2000) and the genetically stable phenotype as MSS/LOH-N.

Table 5. *Loci used in studies I-IV. The chromosomal locations are given according to the EMBL database (www.ensembl.org), Human Genome Database (www.gdb.com) or Y-STR database (www.yhrd.org).*

Locus	Multiplex used		Chromosomal location
TPOX		Profiler Plus [a]	2p25.3
D2S1338	SGM [b]		2q35
D3S1358	SGM	Profiler Plus	3p21.31
FGA	SGM	Profiler Plus	4q31.3
CSF1PO		Profiler Plus	5q33.1
D5S818		Profiler Plus	5q23.2
D7S820		Profiler Plus	7q21.11
D8S1179	SGM		8q24.13
TH01	SGM	Profiler Plus	11p15.5
VWA	SGM	Profiler Plus	12p13.31
D13S317		Profiler Plus	13q31.1
D16S539	SGM		16q24.1
D18S51	SGM		18q21.33
D19S433	SGM		19q12
D21S11	SGM		21q21.1
AMEL	SGM	Profiler Plus	Yp11.2, Xp22.2
DXS7423	X-tetraplex [d]		Xq28
DXS8377	X-tetraplex		Xq28
DXS101	X-tetraplex		Xq22.1
ARA	X-tetraplex		Xq12.2
DYS393	9-plex [d]		Yp11.2
DYS389	9-plex		Yq11.21
DYS391	9-plex	10-plex [d]	Yq11.21
DYS437		10-plex	Yq11.21
DYS438		10-plex	Yq11.21
DYS439		10-plex	Yq11.21
DYS390	9-plex		Yq11.221
GATA-H4		10-plex	Yq11.221
DYS385	9-plex		Yq11.222
DYS392	9-plex	10-plex	Yq11.222
GATA-A7.1		10-plex	Yq11.222
DYS19	9-plex	10-plex	Yp11.2
DYS435		10-plex	Yq11
DYS436		10-plex	Yq11
BAT-26			2p21

[a] *AmpFlSTR® Profiler Plus™*
[b] *AmpFlSTR® SGM™*
[d] *multiplexes developed in our laboratory*

3. AMELOGENIN TYPING (I,II)

AmpF*l*STR® SGM Plus™ and AmpF*l*STR® Profiler ™ multiplexes include the determination of gender by amplification of the amelogenin locus in the X and Y chromosomes. Although given the same gene name, the X- and Y-chromosomal amelogenin genes are not identical. The primer set which is commonly used in gender determination flank a region where the Y-chromosomal gene harbours a 6-bp insertion as compared to the X-chromosomal homologue (Sullivan *et al*. 1993). The different sizes of the amplification products, 103 bp and 109 bp for X and Y, respectively, thus allow typing of the gender. Alternate PCR primers (Steinlechner *et al.* 2002) for amelogenin amplification were synthesised with binding sites flanking the amelogenin amplicon of the genotyping kits.

4. MULTIPLEX Y-CHROMOSOMAL ANALYSIS (II)

Y-STR amplification was perfomed in two multiplexed PCR reactions (9-plex and 10-plex, Table 5) for 16 different Y-chromosomal STRs (Kayser *et al.* 1997, Ruitberg and Butler 2000, Hedman *et al.* 2004) and analysed by ABI Prism CE 310 capillary electrophoresis. The number of repeat units was calculated by the apparent fragment size and by comparison with individuals with known haplotype (Hedman *et al.* 2004). To exclude the possibility that the Y-chromosomal haplotype came from a minority of cellular clones while the majority lacked the entire Y chromosome, a stable X-chromosomal marker DXS7423 (Zarrabeitia et al. 2002a) was included in the 9-plex reaction. The sets of nine and ten Y-STRs have four markers in common, which enabled comparison of both sets with respect to the DXS7423 signal. The peak intensity ratios of the X- and Y-chromosomal signals were analysed in both tumours and control tissue.

5. MULTIPLEX X-CHROMOSOMAL ANALYSIS (IV)

The PCR primers for DXS7423, DXS8377, DXS101 and ARA were synthesised as described by Zarrabeitia *et al.* (2002b) with appropriate fluorescent labels (TAGC, Copenhagen, Denmark). The amplifications were performed in a tetraplex reaction with a touch-down protocol described by Kayser *et al.* (2004) and analysed by capillary electrophoresis. The number of repeat units was calculated by the apparent fragment size according to Zarrabeitia *et al.* (2002b) and verified by sequencing of appropriate samples.

6. STATISTICS (I-IV)

SPSS software was used in the statistical analyses. Pearson correlation test, t-test, Mann-Whitney U test, Fisher's exact probability test and Chi-square test were used where appropriate. Polymorphic information content (PIC) and heterozygosity index (HET) for the X-chromosomal population data was calculated using the PowerStats V12 software.

7. ETHICS (I-IV)

The study protocol was evaluated and approved by the institutional ethics committee of the University Hospital Helsinki.

RESULTS

1. MSI AND LOH PHENOTYPES (I-IV)

In studies I-IV a total of 66 gastrointestinal tumours were analysed for MSI, of which 11 (17%), 15 (22%) and 40 (61%) were categorised as MSI-H, MSI-L and MSS phenotypes respectively. Additional phenotyping by the BAT-26 assay showed 100% consistency with the observed MSI-H phenotypes of the autosomal tetranucleotide microsatellite markers. The apparently normal control tissues produced normal-looking genotypes with near-to-equal peak intensities for heterozygote signals. Table 6 indicates the proportions of MSI phenotypes found in gastric and colorectal cancers (I,II,IV) and in the intestinal and diffuse type of gastric cancers (III). The two GIST tumours in study II were of the MSS type (omitted from the summarised results). However, Table 6 indicates separately the distributions of the LOH and MSI phenotypes. These phenotypes were partly overlapping, but 41% of the samples did not manifest either MSI or LOH categories of alterations (i.e. the MSS/LOH-N phenotype). No statistically significant difference in MSI frequency between GCA and CRC was found. In gastric cancer IGCA harboured significantly more MSI (p=0.012, Fisher's exact test) than did DGCA. However, the difference in distribution of MSI-L (p=0.064) or MSI-H (p=0.389) alone between IGCA and DGCA was not significant.

Figure 7 indicates the instability of individual autosomal markers and BAT-26 in the MSI tumours. The marker instability rates correlated with the instability in the two MSI phenotypes, ranging from 0 to 20% and from 9 to 91% in MSI-L and MSI-H, respectively. No marker was stable in all MSI-H tumours. Only three markers (TPOX, D3S1358 and TH01) were unstable in less than 50% of the MSI-H tumours. TH01 was found to be an exceptionally stable marker, which also had the smallest variation at the population level (data not shown). In MSI-L tumours, FGA and D19S433 were the most unstable, showing instability in 20% of the tumours. The tumour samples were also stratified according to their LOH events (I-III, H. Vauhkonen, unpublished results). All of the 15 autosomal loci showed at least one incidence of LOH in the tumour samples. The prevalence of LOH was higher in colorectal cancers (57%) than in the gastric cancers (33%) (Table 6, H. Vauhkonen, unpublished results), but the difference was not statistically significant (p=0.08, Fisher's exact test). The proportions of informative loci for LOH ranged from 7% to 60% in individual tumours (data not shown). The relative frequency of LOH at the individual loci in the 23 LOH-positive cancers is indicated in Figure 8. In study III, where IGCA and DGCA were investigated, a significant (p<0.001) association of LOH with the IGCA type was found compared with DGCA, where LOH was found only at locus CSF1PO in one sample. The amounts of LOH in the non-MSI-H IGCA samples are also shown in Figure 8 on page 28.

Table 6. *Proportions of the MSI and LOH phenotypes in colorectal (CRC) and gastric (GCA) cancer, and in intestinal (IGCA) and diffuse (DGCA) type gastric cancer. A cut-off of 25% was applied to delineate LOH-H and LOH-L*

	CRC n=27	GCA n=39	IGCA n=27	DGCA n=12		CRC n=27	GCA n=39	IGCA n=27	DGCA n=12
MSI-H	15%	18%	24%	8.3%	LOH-H	22%	18%	24%	0%
MSI-L	22%	23%	40%	8.3%	LOH-L	35%	15%	8%	8.3%
MSS	63%	59%	36%	83.4%	LOH-N	43%	67%	68%	91.7%

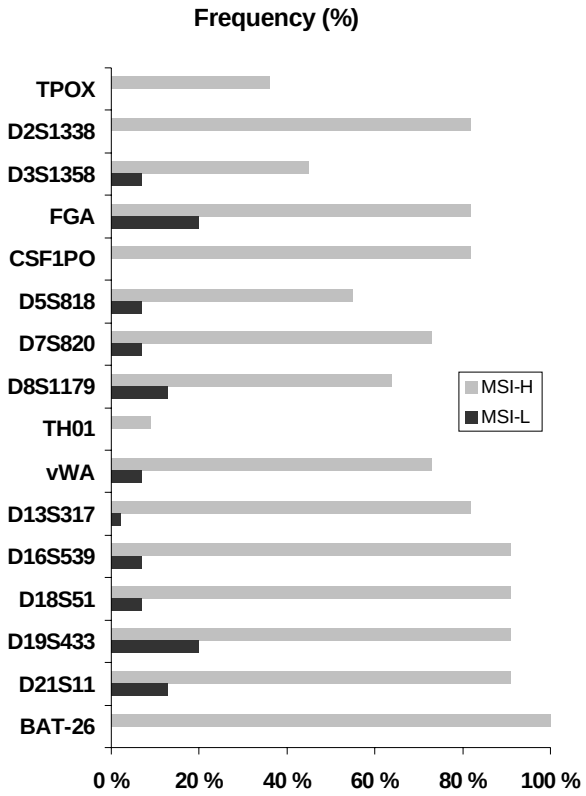

Figure 7. *Instability rates of the autosomal loci and BAT-26 in the 26 gastrointestinal cancers showing MSI-H and MSI-L phenotypes.*

2. LOSS OF THE Y-CHROMOSOMAL AMELOGENIN GENE (II)

Seven of the 46 male tumour samples in study II were found to show a diminished (LOH ratio < 0.5) Y-chromosomal amelogenin signal. A mean ratio of 0.91 (range 0.58-1.42) was calculated for the 46 adjacent healthy tissue samples. The apparent amelogenin deletion was also observed when the samples were assayed with the set of alternative primers flanking the originally amplified region. Results from amplification of the 16 Y-chromosomal microsatellite loci suggested for preservation of the Y chromosome. Four of the tumours with observed loss of the Y-chromosomal amelogenin gene were of LOH-H, one MSI-H and one MSS (a GIST tumour) type, comprising 9% (2/22) of colorectal and 18% (4/22) of gastric tumours from males.

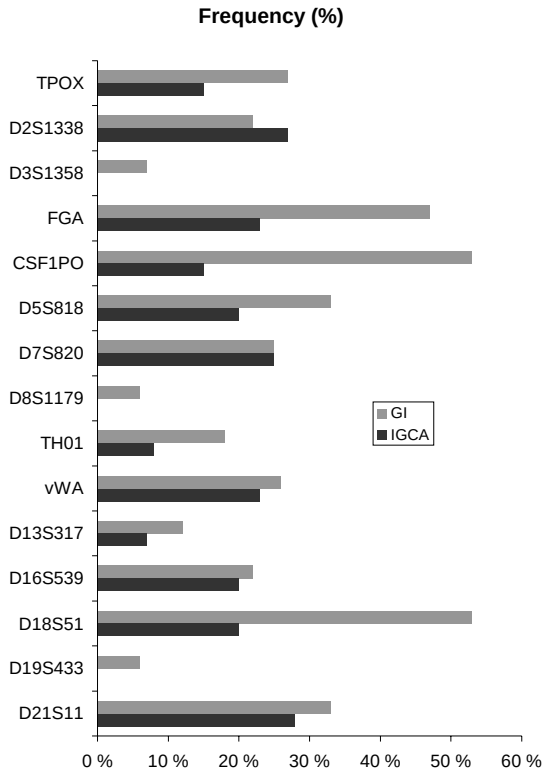

Figure 8. *Frequency of LOH at 15 autosomal loci in the 23 LOH positive gastrointestinal tumours and in the 19 non-MSI-H IGCA samples. Only informative loci are shown.*

3. X-CHROMOSOMAL MICROSATELLITE ANALYSIS (IV)

In the Finnish population sample consisting of 103 unrelated individuals the X-chromosomal loci DXS7423, DXS8377, DXS101 and ARA were found to have 4, 19, 15 and 17 alleles, respectively. The allelic frequency distributions of the X-chromosomal markers in the Finnish population sample correlated well ($p<0.001$) with those reported for German, Spanish and Korean population samples (Edelmann *et al.* 2001, Zarrabeitia *et al.* 2002a, 2002b, Shin *et al.* 2004). The values calculated for PIC and HET describing the polymorphic characteristics of the loci in populations were comparable to those reported earlier for other populations. The relative instability of the X-chromosomal markers analysed in MSI-H, MSI-L and MSS gastrointestinal tumours is indicated in Figure 9. All of the X-chromosomal markers were unstable in MSI-H cancers, and in MSI-L and MSS tumours a rather similar behaviour of the markers could be seen. The marker DXS7423, with the least alleles in the Finnish population, was unstable in only a subset of MSI-H tumours. DXS8377, with the highest number of alleles and highest number of repeats, was unstable in 53% of non-MSI-H tumours. DXS101 was the most unstable in MSI-H tumours, where it most efficiently produced new alleles (20 new alleles in 8 MSI-H tumours). Table 7 shows the allelic distribution of DXS101 in MSI-H tumours, which was comparable by 66-100% to two Caucasian and one Korean population.

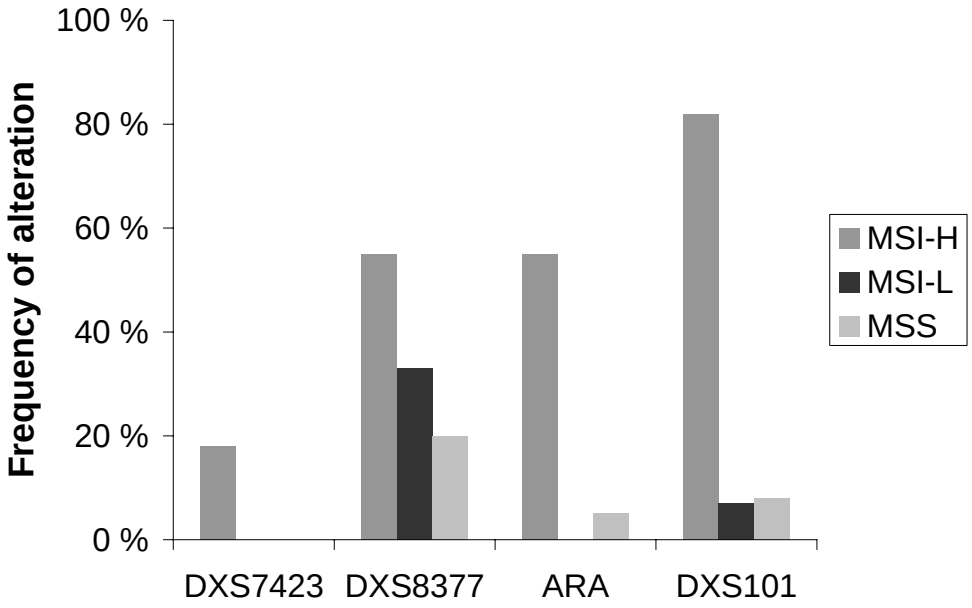

Figure 9. *Frequency of alterations at the X-chromosomal loci in MSI-H, MSI-L and MSS GI tumours.*

Table 7. *Allele distribution of DXS101 in eight MSI-H tumours and in the Finnish (IV), German (Edelmann et al. 2001), Spanish (Zarrabeitia et al. 2002b) and Korean (Shin et al. 2004) populations.*

Allele distribution (%)	DXS101 Allele number																		
	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
MSI-H	-	-	-	-	3.4	-	3.4	3.4	3.4	13.8	17.2	17.2	17.2	10.3	3.4	3.4	3.4	-	-
Finland	-	4.1	0.7	-	9.6	2.7	2.7	3.4	1.4	4.1	30.8	19.2	8.9	5.5	4.1	2.1	0.7	-	-
Germany	0.2	4.4	0.5	0.2	8.4	4.7	1.2	3.2	2.2	6.6	21.3	15.6	11.4	7.9	7.0	2.7	2.4	-	0.1
Spain	-	5.5	-	0.5	8.5	7.5	2.5	1.5	2.5	4.5	15.6	22.1	17.1	3.0	4.0	2.5	2.5	-	-
Korea	-	-	-	-	-	-	-	1.1	3.8	10.9	27.8	21.3	18.0	11.3	4.0	0.7	1.1	-	-

DISCUSSION

1. USE OF TUMOUR DNA IN FORENSIC SETTINGS (I,II)

In some cases, the absence of a sample source from a deceased or missing person may necessitate the use of tumour tissue in DNA-based paternity or identification analysis. One aim of the present work was to evaluate the usefulness of gastrointestinal tumour samples in forensic analyses.

Only 27 (41%) of the 66 gastrointestinal cancers analysed in studies I-IV showed no alterations at their autosomal microsatellite loci. The rest (59%) showed at least one alteration, either MSI, LOH or both. In MSI cases multiple alleles were present, most commonly as a three-allele pattern at a locus. LOH at a certain locus was most often seen as an imbalance of the alleles. A total loss of the heterozygote signal was seldom seen, most probably because of background signals given by normal tissue or due to the presence of different cellular clones. It was also found that male samples with high-level autosomal LOH frequently showed LOH at the Y-chromosomal amelogenin allele resulting in a female-like genotype (II). Consequently, a failure in gender determination may lead to false identity, because samples are often grouped dichotomously by sex. False results in gender typing have been reported to occur also at the population level, e.g. 0.2% of males in Sri Lankan populations have been observed to lack the Y-chromosomal amelogenin peak when analysed with the commonly used primers (Santos *et al.* 1998, Roffey *et al.* 2000, Steinlechner *et al.* 2002, Tharangaj *et al*. 2002, Chang *et al.* 2003, Michael and Brauner 2004).

Recently, analysis of the applicability of tumour DNA in forensic settings has been conducted using oral (Pai *et al.* 2002), pancreatic (Schwark *et al.* 2004), head and neck, renal cell and colon carcinomas, and melanomas (Poetch *et al.* 2004). The results from these studies indicate that use of neoplastic tissue in DNA profiling may lead to erroneous interpretation of identity. Degraded or scarce material, like paraffin-embedded specimens, may also yield artefactual MSI and LOH findings (Sieben *et al.* 2000, Tomlinson *et al.* 2002b). In general, when neoplastic samples are used as a source of DNA one must be aware that an apparent homozygote might result from LOH (Rubocki *et al.* 2000) and an apparent heterozygote from MSI alteration. A simple solution would be a histological analysis of the specimen to obtain a marginal cut from normal-looking areas. However, histologically normal tissues adjacent to a tumour may also harbour DNA alterations (Lindforss *et al.* 2003). In less than one half of the cases DNA from neoplastic tissues could be considered as suitable for forensic casework. An MSI-H sample without the corresponding control tissue could easily be detected by at least one three-or-more-allele locus, but the detection of LOH without control tissue may not be straightforward. Especially in paternity cases, the observed somatic mutations may lead to false exclusions. In conclusion, when inconsistencies are noticed, deductions from these profiles should be performed with extreme care.

2. ASSESSING THE GENETIC INTEGRITY OF TUMOURS BY MARKERS OF FORENSIC INTEREST (I-IV)

The NCI panel of microsatellite markers (Boland et al. 1998) was established to achieve a more uniform approach to evaluate the MSI-H phenotype in colorectal cancer. The five markers chosen for the panel were mono- and dinucleotides, while tetranucleotide markers were considered to be less mutable (Dietmaier *et al.* 1997, Boland *et al*. 1998). The polymorphic autosomal tetranucleotide markers used in studies I-IV are routinely applied in human identification. These markers were frequently unstable in the tumour specimens, and collectively they revealed the same MSI-H cancers than BAT-26. The incidence of the MSI-H phenotype was similar

between gastric (18%) and colorectal (15%) cancers (Table 6), which is in line with previous reports (Boland *et al*. 1998, Halling *et al.* 1999, Choi *et al*. 2000, Halford *et al.* 2003, Lawes *et al.* 2003) showing MSI-H in 12-15% and 9-17% of GCA and CRC, respectively. The distribution of the MSI-L and MSS phenotype in gastric and colorectal tumours was also similar (Table 6). The results indicate that the autosomal tetranucleotide markers used in studies I-IV are valid in MSI detection and that MSI phenotypes occur similarly in GCA and CRC.

The chromosomal instability of tumours can be assessed by e.g. detecting abnormal DNA content by flow cytometry, detecting aneuploidy or chromosomal rearrangements by cytogenetic studies or detecting chromosomal deletions by LOH studies (reviewed in Lengauer *et al.* 1998 and Jallepalli and Lengauer 2001). However, the mechanism generating apparent LOH is not fully understood, as point mutations in the primer binding sites also result as apparent LOH. Whole-genome LOH analysis, i.e. allelotyping, is used for the detection of chromosomal losses that are common and develop sequentially in the tumour-type of interest. The best known example of a carcinogenetic pathway revealed by allelotyping analysis is colorectal cancer (Vogelstein *et al.* 1989). Most of the studies on LOH have focused on certain chromosomal areas with known or suspected tumour suppressor genes. In studies I-III the LOH events were found to be unevenly distributed among the markers (Figure 8). A high frequency (>40%) of LOH was noticed at the FGA, CSF1PO and D18S51 loci. The latter marker, located in 18q, was mostly deleted in colorectal tumours (8/13, 62%, data not shown), in accordance with previous studies (Fearon *et al.* 1990, Goel *et al.* 2003). The markers CSF1PO, D5S818, D18S51 and D21S11 are located in chromosomal arms containing the tumour suppressor genes APC, DCC and TFF1, which are known to be deleted during gastric carcinogenesis (McKie *et al.* 1993, Fang *et al.* 1998, Carvalho *et al.* 2002). These markers showed LOH in 15%, 20%, 20% and 28% of the non-MSI-H IGCA tumours, respectively, which is in line with previous reports (Tamura *et al.* 1996, Choi *et al.* 1998, Nishizuka *et al.* 1998). We conclude that autosomal markers used in DNA profiling of individuals are sensitive also in detecting LOH events and are comparable to the markers more commonly used in tumour studies. Assessing MSI and LOH in the two types of gastric cancer indicated that they represent for the most part different genetic pathways. A genetically stable phenotype is the major entity in DGCA, whereas severe genetic lesions (MSI-H and LOH) are associated with IGCA.

The males of study II with frequent LOH also showed LOH at the amelogenin region. The possibility of allele drop-out as a result of point mutations, as reported by Boutrand *et al.* (2001), was excluded by using alternative primers. Previous studies report Y-chromosomal deletions, either entire or partial, in male gastric tumours (Castedo *et al.* 1992, Van Dekken *et al.* 2001). However, deletion of the amelogenin gene region has not been previously reported. The X- and Y-chromosomal amelogenin genes have been shown to be active in the ameloblasts of developing tooth buds, where the amelogenin proteins have been found as components of the extracellular matrix (Salido *et al.* 1992). Recently the amelogenins have also been suggested to have a role in cell adhesion and signalling (Veis *et al*. 2000, Hoang *et al.* 2002). So far, there is little evidence for a cancer-related role for the amelogenin proteins. Accordingly, to our knowledge, no reports on cancer predisposition in males lacking the Y-chromosomal amelogenin gene exist to date.

3. TUMOURS WITH MICROSATELLITE INSTABILITY MODELLING GERM-LINE MUTATIONS (I,IV)

A relationship between the diversity of alleles at a particular locus in a population and the relative mutability in MSI-H cancers has been previously observed. Sturzeneker *et al.* (2000) found the level of instability of individual markers in colorectal MSI-H tumours to correlate with variables linked to germ-line mutation frequency, i.e. the population average heterozygosity and variance of repeat number of the loci. Di Rienzo *et al.* (1998) estimated the distributions of the size of gains and losses at various loci in 219 genetically uncharacterised colorectal tumours and related

them to allele distributions in populations. They found a positive correlation between the locus mutability and allelic distribution in different populations with known demographic histories. The intriguing possibility to use tumours as a model to test allelic mutations in relation to a population has not been studied previously.

In study I, we found a positive correlation between the germ-line mutation frequency and the level of MSI at a given locus in tumours. The data for MSI frequencies (I-IV) and germ-line mutation rates (www.cstl.nist.gov/ div831/strbase) at the 15 autosomal loci is summarised in Table 9 and Figure 10. The observed positive correlation in the mutability of loci was more significant between MSI-H and the germ-line mutations (r=0.656, p=0.008) than between MSI-L and germ-line mutations (r=0.547, p=0.035) or MSI-H and MSI-L (r=0.480, p=0.070) (Figure 10). This finding suggests that the microsatellite instability of gastrointestinal tumours reflects the germ-line mutation susceptibility, significantly in MSI-H and approaching significance in MSI-L.

Table 9. *Frequency (%) of MSI at different loci in MSI-H and MSI-L tumours with respect to the observed mutation rates in population.*

Locus	MSI-H[a] n=11	MSI-L[b] n=15	Population[c]
TPOX	36%	0	0.01%
D2S1338	82%	0	0.09%
D3S1358	45%	7%	0.13%
FGA	82%	20%	0.30%
CSF1PO	82%	0	0.16%
D5S818	55%	7%	0.12%
D7S820	73%	7%	0.10%
D8S1179	64%	13%	0.13%
TH01	9%	0	0.01%
vWA	73%	7%	0.16%
D13S317	82%	20%	0.15%
D16S539	91%	7%	0.11%
D18S51	91%	7%	0.25%
D19S433	91%	20%	0.11%
D21S11	91%	13%	0.21%

[a] *Frequency of instability in MSI-H tumours.*
[b] *Frequency of instability in MSI-L tumours*
[c] *Data from www.cstl.nist.gov/div831/strbase.*

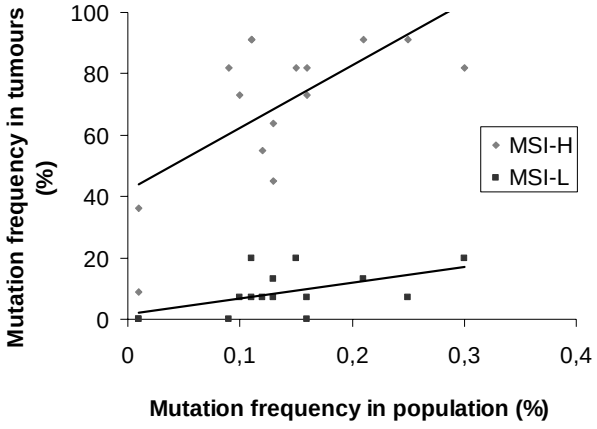

Figure 10. *A graphic presentation of MSI-H and MSI-L mutation frequencies as compared to that of the germ-line. The numerical data is presented in Table 9.*

Assuming that the mutational mechanisms of tumours and the germ-line are similar, tumours may be considered as microscopic populations with the microsatellites evolving under the same constraints than in a human population, but over a more rapid time scale (Di Rienzo *et al.* 1998, Sturzeneker *et al.* 1998, 2000). The different cell clones present in an MSI-H tumour originate from the founder cell, or the ancestral genotype, which has lost its capability to repair replication errors (Shibata *et al.* 1994, Kim *et al.* 2002). The final population of cells that form the tumour can be considered as descendants of the same "pedigree". Because a MSI-H tumour may show multiple alleles at a single locus (Figure 11), different allele combinations represent different lineages of tumour cells.

In study IV the primary question was if gastrointestinal tumours representing different levels of MSI can produce the set of alleles found in a population. It was found that the MSI-H tumours most efficiently produced new alleles, all of which fell well within the allelic distribution found in the Finnish population. The instability of a locus was found to correlate with its allelic diversity in population. Locus DXS7423, with only four alleles in the Finnish population, was the least unstable in tumours, while locus DXS8377, with nineteen alleles, was unstable in non-MSI-H tumours. Locus DXS101 produced the greatest number of alleles in the MSI-H tumours. These alleles correlated well with the allelic distribution observed in the Finnish, German, Spanish and Korean population samples (r=0.822-0.955, p<0.001) (Edelmann *et al*. 2001, Zarrabeitia *et al.* 2002a, 2002b, Shin *et al.* 2004). The results suggest that a series of MSI-H tumours can be used in estimation of the allelic distribution of a candidate marker in a population. However, since gastrointestinal tumours, germ-line cells and populations are subject to different selection, population effects, and tissue-specific effects, more knowledge is needed before straightforward interpretation of the results can be performed. Furthermore, marker-specific factors which affect the number of alleles and thus variance, such as size boundaries and the existence of variant alleles, all complicate direct comparison of microsatellite mutation frequencies between tumour and germ-line cells. It would be interesting to further evaluate this "tumour model" with data obtained from autosomal and Y-chromosomal markers.

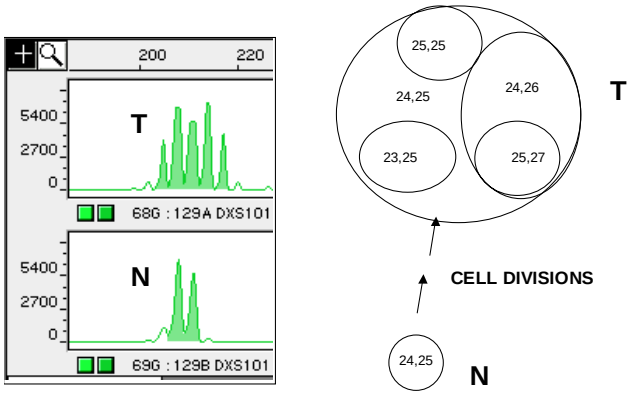

Figure 11. *Formation of an allelic population in a MSI-H tumour as exemplified with DXS101.* The *electropherogram on the left indicates the genotype (alleles 24 and 25) in normal (N) and in tumour (T) tissue (alleles 23,24,25,26 and 27). The circles on the right describe the accumulation of single-step mutations in different cell clones. The presence of the normal genotype in the tumour tissue may be due to back-and-forth stepwise alteration or retaining of an unaltered cell clone, or due to contamination with normal tissue.*

CONCLUSIONS

In forensic settings, if tumour tissues are to be used as a source for DNA profiling, careful interpretation of data and use of stable surrogate markers is warranted. The widely used tetranucleotide microsatellites as well as the gender-specific amelogenin genes are prone to MSI and LOH in the gastric and colorectal cancer studied here (I,II). In 59% of the cases, alterations were found in the DNA from gastrointestinal tumour tissues, thus such samples were concluded to be unsuitable for forensic studies, especially paternity analysis, without strict guidelines.

Previous studies on MSI and LOH in gastric cancer were mainly conducted with mono- and dinucleotide markers. Our studies show that polymorphic tetranucleotide markers also readily detect aberrations in the genetic integrity of gastrointestinal cancers. In addition, different alteration frequencies were found between IGCA and DGCA, where the former harboured a majority of the MSI-H, MSI-L and LOH phenotypes, while the latter was mainly manifested the stable (MSS/LOH-N) phenotype (III). The tumours with high LOH frequency were also found to present the MSI-L phenotype (III). Another novel finding was the loss of the Y-chromosomal amelogenin gene in male samples with frequent LOH (II).

Comparison of allelic distributions between gastrointestinal tumours and a Finnish population sample consisting of unrelated individuals suggest that similar mechanisms of mutation in both tumours and germ-line (IV). Complexity of allelic structure and number of repeats may contribute to the mutational spectrum of a particular locus in the evolution of microsatellites upon tumour development and as well as in the parental transmission of alleles. However, since tumour and germ-line cells are subject to different cell-specific effects, and tumour cell and human populations are subject to different population effects, further investigations are needed for straightforward integration of the data.

ACKNOWLEDGEMENTS

This study was carried out at the Department of Forensic Medicine, Universtity of Helsinki, and Department of Pathology, Jorvi Hospital, HUSLAB Helsinki University Central Hospital, during the years 2002-2004. The former and present head of the Department of Forensic Medicine, Professors Antti Penttilä and Erkki Vuori, are greatly thanked for establishing and providing excellent research facilities at the Laboratory of Forensic Biology.

I am deeply grateful to all the people that I have worked with and I wish to express my sincere gratitude especially to:

Professor Antti Sajantila and Professor Pentti Sipponen, my supervisors, for providing the guidance and supportive attitude which have encouraged me throughout the study. I want to thank Antti for introducing me into the STR-logy of human identification and Pentti for providing the tumour specimens "Pentin Penni Palat" and for the excellent guidance in histology and pathobiology of gastrointestinal cancer.

My dear husband and co-author Matti Vauhkonen, M.D., Ph.D., at the Department of Medicine of Jorvi Hospital, Helsinki University Central Hospital. I want to thank him also for the moments of scientific jargon around the kitchen table.

Professor Dr. habil. Manfred Kayser, and Docent Ismo Ulmanen, Ph.D., the official pre-examiners of this thesis, for careful reviewing and constructive comments. Jodie Painter for the excellent language revision of this thesis.

Eve Karvinen, Kirsti Höök and Teija Vanhala at the Department of Forensic Medicine, and Soili Kaliste and Sue Hyypijev at Department of Pathology, Jorvi Hospital for their excellent technical assistance. Sari Karesvuori at the Endoscopy Unit at Jorvi Hospital for excellent management of the patient records.

Minttu, Ville, Johanna, Anna, Silvia and Yukiko at the Forensic DNA-lab for sharing the moments of happiness and despair, and some bottles of sparkling wine. Also other colleagues, past and present, are acknowledged.

My family, relatives and friends. And last, but not the least, my dear children Aino and Ilmari, to whom I dedicate this work.

REFERENCES

Aaltonen LA, Peltomäki P, Leach FS, Sistonen P, Pylkkänen L, Mecklin JP, et al. Clues to the pathogenesis of familial colorectal cancer, Science 260 (1993) 812-816.

Akiyama K, Youshii T, Nogami M. Analysis of the polymorphic structure of the D7S808 – short tandem repeat (STR) locus. Legal Med 4 (2002) 178-182.

Amorim A and Pereira L. Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs. Forensic Sci Int (2004) in press.

Andrés AM, Lao O, Soldevila M, Calafell F, Bertranpetit J. Dynamics of CAG repeat loci revealed by the analysis of their variability. Hum Mut 21 (2003) 61-70.

Bacon AL, Farrington SM, Dunlop MG. Sequence interruptions confer differential stability at microsatellite alleles in mismatch repair-deficient cells. Human Molecular Genetics 9 (2000) 2707-2713.

Banaschack S, Du Chesna A, Brinkmann B. Multiple interchanging of tissue samples in cases of breast cancer. For Sci Int 113 (2000) 3-7.

Barral S, Lareu MV, Salas A, Carracedo A. Sequence variation of two hypervairiable short tandem repeats at the D22S683 and D6S477 loci. Int J Legal Med 113 (2000) 146-149.

Bennett P. Microsatellites. Mol Pathol 53 (2000) 177-183.

Bhattacharyya NP, Skandalis A, Ganesh A, Groden J, Meuth M. Mutator phenotypes in human colorectal carcinoma cell lines. Proc Natl Acad Sci 91 (1994) 6319-6323.

Biancalana V, Serville F, Pommier J, Julien J, Hanauer A, Mandel JR. Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. Hum Mol Genet 1 (1992) 255-258.

Blake C, Tsao JL, Wu A, Shibata D. Stepwise deletions of poly(A) sequences in mismatch-repair deficient colorectal cancers. Am J Pathol 158 (2001) 1867-1870.

Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Herman JR, Burt RW, et al. A National Cancer Institute workshop on microsatellite instability for cancer detection on familial predisposition; development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 58 (1998) 5248-5257.

Boutrand L, Egyed S, Füredi S, Mommers N, Mertens G, Vanderberghe A. Variations in primer sequence are the origin of allele drop-out at loci D13S317 and CD4. Int J Legal Med 114 (2001) 295-297.

Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62 (1998) 1408-1415.

Brinkmann B, Pfeiffer H, Schurenkamp M, Hohoff C, The evidental value of STRs. Int J Legal Med 114 (2001) 173-177.

Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colorectal cancer. Nature 368 (1994) 258-261.

Buonsanti G, Calistri D, Padovan L, Luinetti O, Fiocca R, Solcia E, Ranzani GN. Microsatellite instability in intestinal- and diffuse-type gastric carcinoma. J Pathol 182 (1997) 167-173.

Butler JM, Buel E, Crivellent F, Mc Cord BR. Forensic DNA typing by capillary electrophoresis: using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis 25 (2004) 1397-1412.

Bär W, Brinkmann B, Carracedo A, Gill P, Lincoln P, Mayr W, et al. Further report of the DNA commission of the ISFG regarding to the use of short tandem repeat systems. Int J Legal Med 110 (1997) 175-176.

Carvalho R, Kaydemir T, Soares P, Canedo P, Sousa C, Oliveira C, et al. Loss of heterozygosity and promoter methylation, but not mutation, may underlie loss of TFF1 in gastric carcinoma. Lab Invest 82 (2002) 1319-1326.

Carvalho-Silva DR, Santos FR, Hutz HM, Salzano FM, Pena S. Divergent human Y-chromosome microsatellite evolution rates. J Mol Evol 49 (1999) 204-214.

Castedo S, Correia C, Gomes P, Seruca R, Soares P, Carneiro F, Sobrinho-Simoes M. Loss of Y chromosome in gastric carcinoma. Cancer Genet Cytogenet 61 (1992) 39-41.

Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R. Relative mutation rates at di-, tri- and tetranucleotide microsatellite loci. Proc Nat Acad Sci 94 (1997) 1041-1046.

Chang YM, Burgoyne LA, Both K. Higher failures at amelogenin sex test in an Indian population. J Forensic Sci 48 (2003) 1309-1313.

de la Chapelle A. Microsatellite instability. New England J Med 349 (2003) 209-210.

Choi SW, Park SW, Lee KY, Kim KM, Chung YJ, Rhyu MG. Fractional allelic loss in gastric carcinoma correlates with growth patterns. Oncogene 17 (1998) 2655-2659.

Choi SW, Choi JR, Chung YJ, Kim KM, Rhyu MG. Prognostic implications of microsatellite genotypes in gastric carcinoma. Int J Cancer 89 (2000) 378-383.

Clayton TM, Hill SM, Dentton LA, Watson SK, Urquhart AJ. Primer binding site mutations affecting the typing of STR loci contained within the AMPF*l*STR® SGM Plus™ kit. Forensic Sci Int 139 (2004) 255-259.

Collins JR, Stephens RM, Gold B, Long B, Dean M, Burt SK. An exhaustive micro-satellite map of the human genome using high-performance computing. Genomics 82 (2003) 10-19.

Correa P. Human gastric carcinogenesis: multistep and multifactorial process. Cancer Res 52 (1992) 6735-40.

Cummings CJ and Zoghbi HY. Fourteen and counting: unravelling trinucleotide repeat diseases. Hum Mol Genet 9 (2000) 909-916.

D'Adda T, Candidus S, Denk H, Bordi C, Höfler H. Gastric neuroendocrine neoplasms: tumour clonality and malignancy-associated large X-chromosomal deletions. J Pathol 189 (1999) 394-401.

D'Andrea AD and Grompe M. The Fanconi anemia/BRCA pathway. Nat Rev Cancer 3 (2003) 23-34.

Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, Rüschoff J. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. Cancer Res 57 (1997) 4749-4756.

Di Rienzo A, Donnelly P, Toomajian C, Sisk B, Hill A, Petzl-Erler, Haines GK, Barch DH. Heterogeneity of microsatellite mutations within and between loci and implications for human dermographic histories. Genetics 148 (1998) 1269-1284.

Dupuy BM, Stenersen M, Egeland T, Olaisen B. Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. Hum Mutat 23 (2004) 117-124.

Duval A and Hamelin R. Genetic instability in human mismatch repair deficient cancers. Ann Genet 45 (2002) 71-75.

Duval A, Reperant M, Compoint A, Seruca R, Ranzani GN, Iacopetta B, Hamelin R. Target gene mutation profile differs gastrointestinal and endometrial tumors with mismatch repair deficiency. Cancer Res 62 (2002) 1609-1612.

Edelmann J, Hering S, Michael M, Lessig R, Deichsel D, Meier-Sundhausen G, et al.. 16 X-chromosome STR loci frequency data from a German population. For Sci Int 124 (2001) 215-218.

Ellegren H. Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5 (2004) 435-445.

Ellegren H. Microsatellite mutations in the germline: implications for evolutionary inference. Trends in Genetics, 16 (2000b) 551-558.

Eshleman J, Markowitz S, Donover P, Lang E, Lutterbaugh J, Li GM, et al. Diverse hypermutability of multiple expressed sequence motifs present in a cancer with microsatellite instability. Oncogene 12 (1996) 1425-1432.

Fang DC, Jass JR, Wang DX. Loss of heterozygosity and loss of expression of the DCC gene in gastric cancer. J Clin Pathol 51 (1998) 593-596.

Fearon ER and Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 61 (1990) 759-767.

Fearon ER, Cho KR, Nigro JM, Kern SE, Simons JM, Ruppert JM, et al. Identification of a chromosomal 18q gene that is altered in colorectal cancer, Science 247 (1990) 49-56.

Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, et al. The human mutator gene homolog MSH2 and its association with hereditary non-polyposis colon cancer. Cell 75 (1993) 1027-1038.

Fleischer AS, Esteller M, Wang S, Tamura G, Suzuki H, Ying J, et al. Hypermethylation of the hMLH1 gene promoter in human gastric cancers with microsatellite instability. Cancer Res 59 (1999) 1090-1985.

Forster EA, Jobling MA, Taylor MG, Donnely P, de Knijff P, Miermet R, et al. Jefferson fathered slave's last child. Nature 396 (1998) 27-28.

Forster P, Röhl A, Lünemann P, Brinkmann C, Zerdal T, Tyler-Smith C, et al. A short tandem repeat – based phylogeny for the human Y chromosome. Am J Hum Genet 67 (2000) 182-196.

Fu YH, Dissuti A, Fenwick Jr RG, King J, Rajnarayan S, Dunne PW, et al. An unstable triplet repeat in a gene related to myotonic dystrophy. Science 255 (1992) 1256-1258.

Gill P. An assessement of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. Int J Legal Med 114 (2001) 204-219.

Gill P, Ivanow PI, Kimpton C, Piercy R, Benson N, Tully G, et al.. Identification of the skeletal remains of the Romanov family by DNA analysis. Nat Genet 6 (1994) 130-135.

Gilmore S, Peakall R, Robertson J. Short tandem repeat (STR) DNA markers are hypervariable and informative in *Cannabis Sativa*: implications for forensic investigations. For Sci Int 131 (2003) 65-74.

Goel A, Arnold CR, Niedwiecki D, Chang DK, Ricciardiello L, Carethers JM, et al. Characterization of sporadic colon cancer by patterns of genomic instability. Cancer Res 63 (2003) 1608-1614.

Gornik I, Marciki M, Kubat M, Primorac D, Lauc G. The identification of war victims by reverse paternity is associated with significant risks of false inclusion. Int J Legal Med 116 (2002) 255-257.

Habano W, Sugai T, Nakamura SI, Uesugi N, Yoshida T, Sasou S. Microsatellite instability and mutation of mitochondrial and nuclear DNA in gastric carcinoma. Gastroenterology 118 (2000) 835-841.

Halford S, Sasieni P, Rowan A, Wasan H, Bodmer W, Talbot I, et al. Low-level instability occurs in most colorectal cancers and is a nonrandomly distributed quantitative trait. Cancer Res 62 (2002) 53-57.

Halford SER, Sawyer EJ, Lambros MB, Gorman P, MacDonald ND, Talbot IC, et al. MSI-low, a real phenomenon which varies in frequency in cancer types. J Pathol 201 (2003) 389-394.

Halling KC, Harper J, Moskaluk C, Thibodeau S, Petroni G, Yustein AS, et al. Origin of microsatellite instability in gastric cancer. Am J Pathol 155 (1999) 205-211.

Hayden JD, Cawkwell L, Dixon MF, Pardal F, Murgatroyd H, Gray S, et al. A comparison of microsatellite instability in early onset gastric carcinomas from relatively low and high incidence European populations. Int J Cancer 85 (2000) 189-191.

Hedman M, Pimenoff V, Lukka M, Sistonen P, Sajantila A. Analysis of 16 Y STR loci in the Finnish population reveals a local reduction in the diversity of male lineages. Forensic Sci Int 142 (2004) 37-43.

Henke J and Henke L. Mutation rate in human microsatellites. Am J Hum Genet 64 (1999) 1473.

Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, Issa JPJ, et al.. Incidence and functional consequence on *hMLH1* promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci 95 (1998) 6870-6875.

Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P. Estimating Y chromosome frequencies using deep rooting pedigrees. Hum Mol Genet 6 (1997) 799-803.

Hickson ID. RecQ helicases: caretakers of the genome. Nat Rev Cancer 3 (2003) 169-178.

Hiyama T, Tanaka S, Yoshihara M, Sasao S, Kose K, Shima H, et al.. Chromosomal and microsatellite instability in sporadic gastric cancer. J Gastroentero Hepatol 19 (2004) 765-760.

Hoang AM, Klebe RJ, Steffensen B, Ryu OH, Simmer JP, Cochran DL. Amelogenin is a cell adhesion protein. J Dent Res 81 (2002) 497-500.

Hoang JM, Cottu PH, Thuille B, Salmon RJ, Thomas G, Hamelin R. BAT-26, an indicator of the replication error phenotype in colorectal cancers and cell lines. Cancer Res 57 (1997) 300-303.

Hochmeister MN, Budowle B, Jong J, Borer UV, Comey CT, Dirnhofer R. PCR-based typing of dNA extracted from cigarette butts. Int J Legal Med 104 (1991) 229-233.

Hochmeister MN, Budowle B, Borer UV, Rudin O, Bohnert M, Dirnhofer R. Confirmation of the identity of human skeletal remains using multiplex PCR amplification and typing kits. J Forensic Sci 40 (1995) 701-705.

Hoff-Olsen P, Jacobsen S, Mevag B, Olaisen B. Microsatellite stability in human post-mortem tissues. For Sci Int 119 (2001) 273-278.

Hoff-Olsen P, Meling G, Olaisen B. Variation in mutation rate and direction between tetranucleotide STR loci in human colorectal carcinomas, Ann Hum Genet 62 (1998) 1-7.

Holtkemper U, Burkhard R, Hohoff C, Forster P, Brinkmann B. Mutation rates at two human Y-chromosomal microsatellite loci using small pool PCR techniques. Hum Mol Genet 10 (2001) 629-633.

Huang QY, Xu FH, Shen H, Deng HY, Liu YJ, Liu YZ, et al.. Mutation patterns at dinucleotide microsatellite loci in humans. Am J Hum Genet 70 (2002) 625-634.

Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E. Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. Nat Genet 2 (1992) 204-211.

International Human Genome Sequencing Consortium. A physical map of the human genome. Nature 409 (2001) 860-921.

Ionov Y, Peinando M, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 363 (1993) 558-561.

Istrail S, Sutton GG, Florea L, Halpern AL, Mobarry CM, Lippert R, et al. Whole-genome shotgun assembly and comparison of human genome assemblies. Proc Natl Acad Sci 101 (2004) 1916-1921.

Iwamoto S, Kamesaki T, Kunada M, Omi T, Okuda H, Hasegawa T, et al. DNA-based identification resolved suspected misdiagnosis due to contaminated cytological specimens. Legal Med 5 (2003) 246-250.

Jallepalli PV and Lengauer C. Chromosome segregation and cancer: cutting through the mystery. Nat Rev Cancer 1 (2001) 109-117.

Jass J, Whitehall VLJ, Young J, Leggett B, Meltzer SJ, Matsubara N, et al. Correspondence re: Laiho P et al., Low-level microsatellite instability in most colorectal carcinomas. Cancer Res 62: 1166-1170, 2002. Cancer Res 62 (2002) 5988-5989.

Jass JR, Young J, Leggett BA. Correspondence: Biological significance of microsatellite instability-low (MSI-L) status in colorectal tumors. Am J Pathol 158 (2001) 779-780.

Jeffreys AJ, Allen MJ, Hagelberg E, Sonnberg A. Identification of the skeletal remains of Joseph Mengele by DNA analysis. Forensic Sci Int 56 (1992) 65-76.

Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E. Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. Proc Nat Acad Sci 93 (1996) 15285-15288.

Jones PA and Laird PW. Cancer epigenetics comes of age. Nat Genet 21 (1999) 163-167.

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al.. Evaluation of Y chromosomal STRs: a multicenter study. Int J Legal Med 110 (1997) 125-133.

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, et al.. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father / son pairs, Am J Hum Genet. 66 (2000) 1580-1588.

Kayser M and Sajantila A. Mutations at Y-STR loci: implications for paternity testing and forensic analysis. Forensic Sci Int 118 (2000) 116-121.

Kayser M, Kittler R, Erler A, Hedman M, Lee A, Mohyuddin A, et al. A comprehensive survey of human Y-chromosomal microsatellites. Am J Hum Genet 71 (2004) 1183-1197.

Kim KM, Kwon MS, Hong SJ, Min KO, Seo EJ, Lee KY, et al. Genetic classification of intestinal –type and diffuse-type gastric cancers based on chromosomal loss and microsatellite instability. Virchows Archiv 443 (2003) 491-500.

Kim KM, Salovaara R, Mecklin JP, Järvinen HJ, Aaltonen LA, Shibata D. Poly(A) deletions in hereditary nonpolyposis colorectal cancer. Mutations before a gatekeeper. Am J Pathol 160 (2002) 1503-1506.

Kinzler KW and Vogelstein B. Lessons from hereditary colorectal cancer. Cell 87 (1996) 159-170.

Klintschar M and Wiegand P. Polymerase slippage in relation to the uniformity of tetrameric repeat stretches. Forensic Sci Int 135 (2003) 163-166.

Knudson AB. Mutation and Cancer: statistical study of retinoblastomas. Proc Natl Acad Sci 68 (1971) 820-823.

Knudson AG. Two genetic hits (more or less) to cancer. Nat Rev Cancer 1 (2001) 157-162.

Koreth J, O'Leary JJ, O'D McGee J. Microsatellites and PCR genomic analysis. J. Pathol 178 (1996) 239- 248.

Kruglyak S, Durrett RT, Schug MD, Aquardo CF. Equilibrium distribution of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc Natl Acad Sci 95 (1998) 10774-10778.

Laiho P, Launonen V, Lahermo P, Esteller M, Guo M, Herman JG, et al. Low-level microsatellite instability on most colorectal carcinomas. Cancer Res 62 (2002) 1166-1170.

Lasko D, Cavenee W and Nordenskjöld M. Loss of constitutional heterozygosity in human cancer. Annu Rev Genet 25 (1991) 281-314.

Laurén P. The two histologic main types of gastric carcinoma: diffuse and so-called intestinal type carcinoma. An attempt at a histological classification. Acta Pathol Microbiol Scand 64 (1965) 31-49.

Lawes DA, SenGupta S, Boulos PB. The clinical importance and prognostic implications of microsatellite instability in sporadic cancer. Eur J Surg Oncol 29 (2003) 201-212.

Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in human cancers. Nature 396 (1998) 643-649.

Leopoldino A and Pena S. The mutational spectrum of human autosomal tetranucleotide microsatellites. Hum Mutat 21 (2003) 71-29.

Levinson G and Gutman GA. Slipped-strand mispairing. A major mechanism for DNA sequence evolution. Mol Biol Evol 4 (1987) 203-221.

Levitt NC and Hickson ID. Caretaker tumour suppressor genes that defend genome integrity. Trends in Molecular Medicine 8 (2002) 179-186.

Li WH, Gu H, Wang H, Nekrutenko A. Evolutionary analyses of the human genome. Nature 409 (2001) 847-849.

Lindblom A, Tannergård P, Werelius B, Nordenskjöld M. Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. Nat Genet 5 (1993) 279-282.

Lindforss U, Papadogiannakis N, Lidberg G, Olivercrona H. Distribution of genetic variants in preneoplastic areas of colorectal tumours. Eur J Surg Oncol 29 (2003) 201-212.

Loeb LA. Microsatellite instability: Marker of a mutator phenotype in cancer. Cancer Res 54 (1994) 5059-5063.

Loeb LA, Loed KR, Anderson JP. Multiple mutations and cancer. Proc Natl Acad Sci 100 (2003) 776-781.

Lu AL, Clarck S, Modrich P. methyl-directed repair of DNA base-pair mismatches *in vitro*. Proc Natl Acad Sci 80 (1983) 4369-4636.

Mann K, Donaghue C, Oglivie CM. In vivo somatic microsatellite mutations identified in non-malignanat human tissue. Hum Genet 114 (2003) 110-114.

McKie AB, Filipe MI, Lemoine NR. Abnormalities affecting the APC and MCC tumour suppressor gene loci on chromosome 5q occur frequently in gastric cancer but not in pancreatic cancer. Int J Cancer 55 (1993) 598-603.

Michael A and Brauner P. Erroneous gender identification by the amelogenin sex test. J Forensic Sci 49 (2994) 258-259.

Ming SC. Cellular and molecular pathology of gastric carcinoma and precursor lesions: A critical review. Gastric Cancer 1 (1998) 31-50.

Mori Y, Yin J, Leggett BA, Young J, Simms L, Kuehl PM, et al. Instabilotyping: comparative identification of frameshift mutations caused by coding region microsatellite instability. Cancer Res 61 (2001) 6046-6049.

Morral N, Bertranpetit J, Estivilli T, Casals T, Giménez J, Reis A, et al. The origin of the major cystic fibrosis mutation (ΔF508) in European populations. Nat Genet 7 (1994) 169-175.

Mullis KB and Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol 155 (1987) 335-340.

Nishizuka S, Tamura G, Terashima M, Satodate R. Loss of heterozygosity during the development and progression of differentiated adenocarcinoma of the stomach. J Pathol 185 (1998) 38-43.

Oliveira C, Seruca R, Seixas M, Sobrinho-Simões M. The clinicopathological features of gastric carcinomas with microsatellite instability may be mediated by mutations of different "target genes". A study of the TGFβRII, IGFII R, and BAX genes. Am J Pathol 153 (1998) 1211-1219.

Oota K and Sobin LH. Histological typing of gastric and esophageal tumors. International histologic classification of tumors. 18. World Health Organisation, Geneva 1990.

Ottini L, Falchetti M, D'Amico C, Amorosi A, Saieva C, Masala G, et al. Mutations at coding mononucleotide repeats in gastric cancer with the microsatellite mutator phenotype. Oncogene 16 (1998) 2767-2772.

Pai CY, Hsieh LL, Tsai CW, Chiou FS, Yang CH, Hsu BD. Allelic alteration at the STR markers in the buccal tissue cells of oral cancer patients and the oral epitheliual cells of healthy betel quid-chewers: an evaluation of forensic applicability. Forensic Sci Int 129 (2002) 158-167.

Parsons R, Myeroff LL, Liu B, Willson JK, MarkowitzSD, Kinzler K, et al. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. Cancer Res 55 (1995) 5548-5550.

Peltomäki P, Aaltonen LA, Sistonen P, Pylkkänen L, Mecklin JP, Järvinen H, *et al*. Genetic mapping of a locus predisposing to human colorectal cancer. Science 260 (1993) 810-812.

Poetsch M, Petersmann A, Woenckhaus C, Protzl C, Dittbeiner T, Lignitz E, et al. Evaluation of allelic alterations in short tandem repeats an different kinds of solid tumors – possible pitfalls in forensic casework. Forensic Sci Int (2004) in press.

Pyatt R, Chadwick RB, Johnson CK, Adebamowo C, de la Chapelle A, Prior TW. Polymorphic variation at the BAT-25 and BAT-26 loci in individuals of African origin. Am J Pathol 155 (1999) 349-353.

Rhyu MG, Park WS, Meltzer SJ. Microsatellite instability occurs frequently in human gastric carcinoma. Oncogene 9 (1994) 29-32.

Ricciardone MD, Özcelick T, Cevher B, Özdag H, Tuncer M, Gürgey A, et al. Human *MLH1* deficiency predisposes to hematological malignancy and neurofibraomatosis type 1. Cancer Res 59 (1999) 290-293.

Roffey P, Eckhoff C, Kuhl J. A rare mutation in the amelogenin gene and its potential investigative ramifications. J Forensic Sci 45(2000) 1016-1019.

Rubinsztein DC, Amos W, Leggo J, Goodburn S, Jain S, Li SH, et al. Microsatellite evolution – evidence for directionality and variation between species. Nat Gen 10 (1995) 337-343.

Rubocki RJ, Duffy KJ, Shepard KL, McCue BJ, Shepard SJ, Wisecarver JL. Loss of heterozygosity detected in a short tandem repeat (STR) locus commonly used for human DNA identification. J Forensic Sci 45 (2000) 1087-1089.

Ruitberg C, Butler J New primer sets for Y chromosome and CODIS STR loci. Poster presented at Int Sym Human Identification 2000, Biloxi, MS, US.

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sicle cell anemia. Science 230 (1985) 1350-1354.

Sajantila A, Lukka M, Syvänen AC. Experimentally observed germline mutations at human micro- and minisatellite loci. Eur J Hum Genet 7 (1998) 263-266.

Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro L. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosome. Am J Hum Genet 50 (1992) 303-306.

Sambrook J, Fritsch EF, Maniatis T. In: Molecular cloning. A laboratory manual. Second edition. Cold Spring Harbor laboratory press, 1989.

Samowitz WS, Slattery ML, Kerber RA. Microsatellite instability in human colorectal cancer is not a useful clinical indicator of familial colorectal cancer. Gastroenterology 109 (1995) 1765-1771.

Santos F, Pandya A, Tyler-Smith C. Reliability of DNA-based sex tests. Nat Genet 118 (1998) 103.

Schlötterer C. Evolutionary dynamics of microsatellite DNA. Chromosoma 109 (2000) 365-371.

Schofield MJ and Hsieh P. DNA mismatch repair: molecular mechanisms and biological function. Annu Rev Microbiol 57 (2003) 579-608.

Schwark T, Bachmann C, von Wurmb-Schwark N. STR typing of ductal adenocarcinomas of the pancreas and healthy control tissue in 18 individuals, Legal Med 6 (2004) 170-173.

Shibata D, Navidi W, Salovaara R, Li ZH, Aaltonen LA. Somatic microsatellite mutations as molecular tumor clocks. Nat Med 2 (1996) 676-681.

Shibata D, Peinando MA, Ionov Y, Malkhosyan S, Perucho M. Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation. Nat Gen 6 (1994) 273-281.

Shadrah B, Commane M, Hren C, Warshawsky I. A rare mutation in the primer binding region of the amelogenin gene can interfere with gender identification. J Mol Diagn 6 (2004) 401-405.

Shin KJ, Kwon BK, Lee SS, Yoo JE, Park MJ, Chung U, et al. Five highly informative X-chromosomal STRs in Koreans. Int J Legal Med 118 (2004) 37-40.

Shen CY, Yu JC, Lo YL Kuo CH, Ye CT, Jou YS, et al. Genome-wide search for loss of heterozygosity using laser capture microdissected

tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. Cancer Res 60 (2000) 3884-3892.
Sieben NL, ter Haar, NT, Cornelisse CJ, Fleuren GJ, Cleton-Jansen AM. PCR artifacts in LOH and MSI analysis of microdissected tumr cells. Hum Pathol 31 (2000) 1414-1419.
Sipponen P. Gastric cancer: Pathogenesis, risks and prevention, J Gastroenterol 37 [suppl XIII] (2002) 39-44.
Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, et al. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nat Genet 4 (1993) 393-397.
Steinlechner M, Berger B, Niederstätter H, Parson W. Rare failures in the amelogenin sex test. Int J Legal Med 116 (2002) 117-120.
Sturzeneker R, Bevilaqua R, Haddad L, Simpson A, Pena S, Microsatellite instability in tumors as a model to study the process of microsatellite mutations. Human Molecular Genetics 9 (2000) 347-352.
Sturzeneker R, Haddad L, Bevilacqua R, Simpson A, Pena P. Polarity of mutations in tumor-assiciated microsatellite instability. Hum Genet 102 (1998) 231-235.
Subramanian S, Mishra RK, Singh L. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Biol 4 (2003) R13.
Sullivan K, Mannucci A, Kimpton C, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 15 (1993) 636-641.
Szibor R, Edelmann J, Zarrabeitia MT, Riancho JA. Sequence structure and population data of the X-linked markers DXS7423 and DXS8377 – clarification of conflicting statements published by two working groups. Forensic Sci Int 134 (2003a) 72-73.
Szibor R, Krawczak M, Heiring S, Edelmann J, Kuhlisch E, Krause D. Use of X-linked markers for forensic purposes. Int J Legal Med 117 (2003b) 67-74.
Tamura G, Sakata K, Nishizuka S, Maesawa C, Suzuki Y, Terashima M, et al. Allelotype of adenoma and differentiated adenocarcinoma of the stomach. J Patol 180 (1996) 371-377.
Tang R, Chanchien CR, Wu MC, Fan CW, Liu KW, Chen HT, et al. Colorectal cancer without high microsatellite instability and chromosomal instability – an alternative genetic pathway to human colorectal cancer. Carcinogenesis 25 (2004) 841-846.
Tautz D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucl Ac Res 17 (1989) 6463-6471.
Thangaraj K, Reddy A, Singh L. Is the amelogenin gene reliable for gender identification in forensic cesework and prenatal diagnosis ? Int J Legal Med 116 (2002) 121-123.
Thiagalingam S, Laken S, Willson JKV, Markowitz SD, Kinzler KW, Vogelstein B, et al. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci 98 (2001) 2698-2702.
Thiagalingam S, Foy R, Cheng KH, Lee HJ, Thiagalingam A, Ponte JF. Loss of heterozygosity as a predictor to map tumour suppressor genes in cancer: molecular basis of its occurrence. Curr Opin Oncol 14 (2002) 65-72.
Thibodeau SN, Bren G, Scaid D. Microsatellite instability in cancer of the proximal colon. Science 260 (1993) 816-819.
Thibodeau SN, French AJ, Cunningham JM, Tester D, Burgart LJ, Roche PC, et al. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. Cancer Res 58 (1998) 1713-1718.
Thomson JA, Ayres KL, Pilotti V, Barrett MN, Walker JIH, Debenham PG. Analysis of disputed single-parent/child and sibling relationships using 16 STR loci. Int J Legal Med 115 (2001) 128-134.
Thomson JA, Pilotti V, Stevens P, Ayres KL, Debenham PG. Validation of short tandem repeat analysis for the investigation of cases of disputed paternity. Forensic Sci Int 100 (1999) 1-16.
Tomlinson I, Halford S, Aaltonen L, Hawkins N, Ward R. Does MSI-L exist? J Pathol 167 (2002a) 6-13.
Tomlinson I, Lambros M. Roylance R. Loss of heterozygosity analysis: Practically and conceptually flawed? Genes, Chromosomes. Cancer 34 (2002b) 349-353.
Tomlinson IP, Roylance R, Houston RS. Two hits revisited again. J Med Genet 38 (2001) 81-85.
Tomlinson I, Sasieni P, Bodmer W. How many mutations in a cancer? Am J Patol 160 (2002c) 755-758.
Tòth G, Gàspàri Z, Jurka J. Microsatellites in different eucaryotic genomes. Survey and analysis. Genome Res 10 (2000) 967-981.
Tsao JL, Yatabe Y, Salovaara R, Järvinen HJ, Mecklin JP, Aaltonen LA, et al. Genetic reconstitution of individual colorectal tumor histories. Proc Nat Acad Sci 97 (2000) 1236-1241.
Umar A, Boland CR, Terdiman JP, Sygnal S, de la Chapelle A, Rüschoff J, et al. Revised Bethesda guidleines for hereditary nonpolyposis colorectal cancer (Lynch Syndrome) and microsatellite instability. J Natl Cancer Inst 96 (2004a) 261-268.
Umar A, Risinger JI, Hawk ET, Barrett JC. Testing guidelines for hereditary non-polyposis colorectal cancer. Nat Rev Cancer 4 (2004b) 153-158.
Urquhart A, Kimpton CP, Downes TJ, Gill P. Variation in short tandem repeat sequences. Int J Legal Med 107 (1994) 13-20.
Valdes AM, Slatkin M, Freimer NB. Alleles frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics 133 (1993) 737-749.
Van Dekken H, Alers J, Riegman P, Rosenberg C, Tilanus H, Vissers K. Molecular cytogenetic evaluation of gastric cardia adenocarcinoma and precursor lesions. Am J Pathol 158 (2001) 1961-1967.
Veis A, Tompkins K, Alvares K, Wei K, Wang L, Wang XS, et al. Specific amelogenin splice products have signaling effects on cells in culture and implants in vivo. J Biol Chem 275 (2000) 41263-41272.
Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al.. The sequence of the human genome. Science 291 (2001) 1304-1351.
Verhej C, Bakker CE, de Graaff E, Keulemans J, Williemsen R, Verkerk AJMH, et al. Characterisation and localization of the FMR-1 gene product associated with fragile X syndrome. Nature 363 (1993) 722-724.
Vilkki S, Launonen V, Karhu A, Sistonen P, Västrik I, Aaltonen LA. Screening for microsatellite instability target genes in colorectal cancers. J Med Genet 39 (2002) 785-789.
Vilkki S, Tsao JL, Loukola A, Pöyhönen M, Vierimaa O, Herva R, et al. Extensive somatic microsatellite mutations in normal human tissue. Cancer res 61 (2001) 4541-4544).
Vogelstein B, Fearon ER, Kern SE, Hamilton SR, Preisinger AC, Nakamura Y, et al. Allelotype of colorectal carcinomas. Science 244 (1989) 207-211.
Wang Q, Lasset C, Desseigne F, Frappaz D, Bergeron C,Navarro C, et al. Neurofibromatosis and early onset of cancers in h*MLH1*-deficient children. Cancer Res 59 (1999) 294-297.
Wallerand H, Rémy-Martin A, Chabannes E, Bermont L, Adessi GL, Bittard H. Relationship between expansion of the CAG repeat in exon 1 of the androgen receptor gene and idiopathic male infertility, Fertility and Sterility 76 (2001) 769-774.
Weber J and Wong C. Mutation of the human short tandem repeat. Human Mol Genet 2 (1993) 1123-1128.
Weber JL and May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44 (1989) 388-396.
Wiegand P, Berger B, Edelmann J, Parson W. Population genetic comparisons of three X-chromosomal STRs. Int J Legal med 117 (2003) 62-65.
Xu X, Pen M, Fang Z, Xu X. The direction of microsatellite mutation is dependent upon allele length, Nat Genet 24 (2000) 396-399.
Yamamoto H, Perez-Piteira J, Yoshida T, Terada M, Itoh F, Imai K, et al. Gastric cancers of the microsatellite mutator phenotype display characteristic genetic and clinical features. Gastroenterology 116 (1999) 1348-1357.
Young J, Searle J, Buttenshaw R, Thomas L, Wand M, Chenevix-Trench G, et al. An Alu VpA marker on chromosome 1 demonstrates that replication errors manifest at the adenoma-carcinoma transition in sporadic colorectal tumors. Genes, Chromosomes & Cancer 12 (1995) 251-254.
Zarrabeitia M, Amigo T, Sañudo C, de Pancorbo M, Riancho J, Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377. Int J Legal Med 116 (2002a) 368-371.
Zarrabeitia M, Amigo T, Sañudo C, Zarrabeitia A, Gonzáles-Lamuño D, Riancho J. A new pentaplex system to study short tandem repeat markers of forensic intrest on X chromosome. Forensic Sci Int 129 (2002b) 85-89.
Zheng L, Wang L, Ajanis J, Xie K. Molecular basis of gastric cancer development and progression. Gastric Cancer 7 (2004) 61-77.