

# Three novel missense germline mutations in different exons of *MSH6* gene in Chinese hereditary non-polyposis colorectal cancer families

Shi-Yan Yan, Xiao-Yan Zhou, Xiang Du, Tai-Ming Zhang, Yong-Ming Lu, San-Jun Cai, Xiao-Li Xu, Bao-Hua Yu, Heng-Hua Zhou, Da-Ren Shi

Shi-Yan Yan, Xiao-Yan Zhou, Xiang Du,Tai-Ming Zhang, Yong-Ming Lu, Xiao-Li Xu, Bao-Hua Yu, Heng-Hua Zhou, Da-Ren Shi, Department of Pathology, Cancer Hospital/Institute, Fudan University; Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

San-Jun Cai, Department of Abdominal Surgery, Cancer Hospital/Institute, Fudan University, Shanghai 200032, China

Supported by Shanghai Medical Development Fund for Major Projects, No. 05111004 and Shanghai Pu Jiang Projects for Talented-Men, 06PJ14019

Correspondence to: Da-Ren Shi, 270 Dong An Road, Department of Pathology, Cancer Hospital, Fudan University, Shanghai 200032, China. shidaren2000@yahoo.com Telephone: +86-21-64046008 Fax: +86-21-64046008

Received: July 4, 2007 Revised: July 31, 2007

## Abstract

**AIM:** To investigate the germline mutations of *MSH6* gene in probands of Chinese hereditary non-polyposis colorectal cancer (HNPCC) families fulfilling different clinical criteria.

**METHODS:** Germline mutations of *MSH6* gene were detected by PCR-based DNA sequencing in 39 unrelated HNPCC probands fulfilling different clinical criteria in which *MSH2* and *MLH1* mutations were excluded. To further investigate the pathological effects of detected missense mutations, we analyzed the above related *MSH6* exons using PCR-based sequencing in 137 healthy persons with no family history. The clinicopathological features were collected from the Archive Library of Cancer Hospital, Fudan University and analyzed.

**RESULTS:** Four germline missense mutations distributed in the 4<sup>th</sup>, 6<sup>th</sup> and 9<sup>th</sup> exons were observed. Of them, three were not found in international HNPCC databases and did not occur in 137 healthy controls, indicating that they were novel missense mutations. The remaining mutation which is consistent with the case H14 at c.3488A>T of exon 6 of *MSH6* gene was also found in the controls, the rate was approximately 3.65% (5/137) and the type of mutation was not found in the international HNPCC mutational and SNP databases, suggesting that this missense mutation was a new SNP unreported up to date.

CONCLUSION: Three novel missense mutations and a new SNP observed in the probands of Chinese HNPCC

families, may play an important role in the development of HNPCC.

© 2007 WJG. All rights reserved.

Key words: Hereditary non-polyposis colorectal cancer; *MSH6*; Missense mutation; Colorectal cancer

Yan SY, Zhou XY, Du X, Zhang TM, Lu YM, Cai SJ, Xu XL, Yu BH, Zhou HH, Shi DR. Three novel missense germline mutations in different exons of *MSH6* gene in Chinese hereditary non-polyposis colorectal cancer families. *World J Gastroenterol* 2007; 13(37): 5021-5024

http://www.wjgnet.com/1007-9327/13/5021.asp

## INTRODUCTION

Hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch syndrome, is an autosomal dominant inherited disease characterized by susceptibility to a wide spectrum of cancers, including cancer of the colon, rectum, endometrium, small bowel, and urinary tract<sup>[1]</sup>. HNPCC syndrome accounts for 5%-10% of all colon-rectum cancer cases<sup>[2]</sup>. Germline mutations in the hMLH1 (MIM#120436; GDB: 249617) and hMSH2 (MIM#120435; GDB: 203983) genes are detected in 30%-70% of HNPCC families<sup>[3-5]</sup>. Recently, mutations in another MMR gene (MSH6), accounting for 10% of HNPCC kindreds, have also been shown to result in HNPCC<sup>[6]</sup>. The number of mutations in the MSH6 gene keeps increasing<sup>[7]</sup> (http://www.nfdht.nl). Germline mutations of MSH6 gene are mainly associated with patients with no MSH2 and MLH1 mutations. This study was to investigate the MSH6 gene germline mutations by DNA sequencing in 39 Chinese HNPCC kindreds fulfilling the criteria in which MSH2 and MLH1 mutations are excluded, in order to identify HNPCC families and provide experimental information for HNPCC database.

## MATERIALS AND METHODS

## Materials

From January 1998 to October 2005, 39 Chinese HNPCC families fulfilling the HNPCC clinical criteria were registered at the Department of Abdominal Surgery

in Shanghai Cancer Hospital/Institute. Eleven families fulfilled Amsterdam criteria (AC I) and II (AC II)<sup>[8,9]</sup>, 11 additional families Japanese criteria<sup>[10]</sup> and the remaining 17 kindreds Bethesda guidelines (BG)<sup>[11]</sup>. Germline mutations of MSH2 and MLH1 were excluded by based-PCR sequencing in the probands of all the Chinese HNPCC families. Each proband was asked to give 10 mL peripheral blood samples and consent for access to archival tumor tissue. A total of 137 control blood samples were taken from healthy persons after obtaining informed consent, none of the individuals in the control group had a family history suggesting HNPCC or development of colon cancer in earlier age. This study was proved by the Medical Ethical Committee of Cancer Hospital, Fudan University, and the procedures of the study were in accordance with the international rules and regulations.

## **DNA** extraction

Genomic DNA was extracted with the QIAGEN (Hilden, Germany) DNA extraction kit following the manufacturer' s instructions. Concentrations of the genomic DNA were determined by an ultraviolet spectrophotometer (Beckman, DU640 type).

## PCR amplification and DNA sequencing

According to the exon/intron boundary sequences of MSH6 (GenBank assession number: NM\_000179.1), 18 sets of primers were designed to amplify the entire coding region, including 10 exons and each splicing site of MSH6 (Table 1). The primer pairs used to amplify the 18 fragments of MSH6 have either M13 forward 5' primers or M13 reverse 3' primers to facilitate sequencing after amplification. PCR was carried out with Taq DNA polymerase (Promega, USA) as described elsewhere<sup>[12]</sup>. The PCR conditions were as follows: preheating at 94°C for 7 min, followed by 30-38 cycles of denaturation at 94°C for 45 s, annealing at 56°C-68°C for 45 s and extension at 72°C for 45 s, and a final elongation at 72°C for 10 min (Table 1). PCR products were subjected to 2% agarose gel electrophoresis. After observation of clear and expected size bands, the products were purified and used as a template for sequencing reactions with BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, USA). The sequencing primers were M13F or M13R. Automated fluorescence analysis was performed on a 3700 DNA sequence system (ABI, USA).

## **Bioinformatic analysis**

Each result of sequencing was analyzed by DNAStar5.08 bioanalysis software. The type of mutations and potential significance were determined by comparing the corresponding amino acids and proteins in the following databases (http://www.ncbi.nlm.nih.gov/, http://www.ensembl.org/homo-sapies and http://www.insight-group.org).

## RESULTS

Four missense mutations were found in 39 probands of different Chinese HNPCC families, at codons 468 (CGT>CAT, Arg>His), 1163 (GAA>GTA, Glu>Val) (Figure 1), 666 (TCT>CCT, Ser>Pro) (Figure 2) and 1284 
 Table 1 Primer sequences and PCR condition of different exons of MSH6 gene

EN	Primer sequence (5'-3')		AT (℃)	
Exon1	M13F-AGCTCCGTCCGACAGAAC	381	68	38
	M13R-CTGTGCGAGCCTCCCCT			
Exon2	M13F-TGCCAGAAGACTTGGAATTG	325	63	32
	M13R-CAAACACACACACATGGCAG			
Exon3	M13F-GATGGGGTTTGCTATGTTGC	341	67	36
	M13R-TACACCCTCCCCTTTCTTC			
Exon4.1	M13F-GGCTGCACGGGTACCATTAT	390	60	34
	M13R-CATTCTCTTCCGCTTTCGAG			
Exon4.2	M13F-GCCAGACACTAAGGAGGAAGG	386	59	35
	M13R-TAGATGCATCAAAATCGGGG			
Exon4.3	M13F-TGGCTTAAGGAGGAAAAGAGA	378	60	34
	M13R-TCTACATCGTGCCTCCATCA			
Exon4.4	M13F-TTCTGGCTTTCCTGAAATTG	374	60	34
	M13R-TAAATCTCGAACAATGGCGA			
Exon4.5	M13F-TCTGGCCATACTCGTGCATA	329	60	34
	M13R-AGCACCTGGGGTAACATCAC			
Exon4.6	M13F-TCAGGAAGGTCTGATACCCG	353	62	33
	M13R-GCACCATTCGTTGATAGGCT			
Exon4.7	M13F-AAGTGAATTGGCCCTCTCTG	483	62	30
	M13R-TGGTTCTGACTCTTCAGGGG			
Exon4.8	M13F-TTTTGGTAAGCGGCTCCTAA	465	62	35
	M13R-TTTCGAGCCTTTTCATGGTC			
Exon4.9	M13F-TTTCTGCTCTGGAAGGATTC	440	62	30
	M13R-TCGTTTACAGCCCTTCTTGG			
Exon4.10	M13F-TGAACAGAGCCTCCTGGAAT	390	62	32
	M13R-CAGCTGGCAAACAGCACTAC			
Exon5	M13F-CTGATAAAACCCCCAAACGA	403	62	30
	M13R-CTGTGTTTGGAAAATGATCACC			
Exon6	M13F-CCAGTCATAAAAGACCTTTTCC	192	56	34
	M13R-GACTGAATGAGAACTTAAGTGGG			
Exon7	M13F-AAGGTGAAAGTACATTT	119	61	33
	M13R-TTCAAATGAGAAGTTTAATG			
Exon8/9	M13F-CCTTTGAGTTACTTCCTT	573	61	32
	M13R-TCATAGTGCATCATCCCTTCC			
Exon10	M13F-GGAAGGGATGATGCACTATG	254	61	35
	M13R-AAGAAAATGGAAAAATGGTCA			

Abbreviation: EN, Exon name. AT, Anneal temperature. CN, Cycle number. The sequence of M13F was 5'-GTAAAACGACGGCCAGT-3'. The sequence of M13R was 5'-AACAGCTATGACCATG-3'.

(ACG>ATG, Thr>Met), respectively (Table 2). To further investigate the pathological effects of these four missense mutations, we analyzed the four related MSH6 exons using PCR-based sequencing in 137 healthy persons with no family history, showing that the mutation of codon 1163 which is consistent with the case H14 at c.3488A>T of exon 6 of MSH6 gene was also found in the control persons, the rate was approximately 3.65% (5/137).The remaining three missense mutations did not occur in the 137 healthy controls. None of these four muations was found in the MSH6-SNP database (www.ensembl. org/homo-sapies). Thus, except that the mutation at c.3488A>T of MSH6 gene in the H14 HNPCC case was an unreported new single nucleotide polymorphism (SNP), the remaining ones were novel missense mutations.

## DISCUSSION

Hereditary non-polyposis colorectal cancer is an autosomal dominant inherited syndrome<sup>[13]</sup>. Although its clinical diagnostic criteria were established in 1990, known as

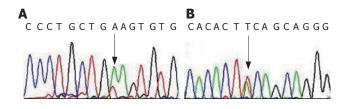


Figure 1 Missense germline mutation of exon 6 of *MSH6* gene in the proband of H14 HNPCC kindreds. Arrow indicates the mutation site, The single basyl substitution was transversed from A to T (A>T) at the codon 1163, the codon from GAA to GTA, causing the amiod acid changes from glutamine to valine, the change was identified as a new SNP. A and B represent the forward sequence and reverse, respectively.

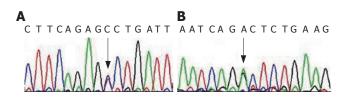


Figure 2 Missense germline mutation of exon 4.6 of *MSH6* gene in the proband of H40 HNPCC kindreds. Arrow indicates the mutation site, The single basyl substitution was transvered from T to C (T>C) at the codon 666, the codon from TCT to CCT, causing the amiod acid changes from serine to proline. **A** and **B** represent the forward and reverse sequence, respectively.

Amsterdam criteria, Japanese Criteria and Bethesda guidelines, a certain number of HNPCC families may be neglected for lacking characteristic clinical manifestations and family history. Molecular genetic screening is been regarded as a standard method for its diagnosis. Mutations of MMR genes are considered golden criteria for molecular diagnosis and monitoring family members. These mutations can be used to select persons who would benefit from genetic counseling and clinical surveillance programs for their relatives to reduce morbidity and mortality due to HNPCC-related tumors<sup>[14]</sup>. Germline mutations in the coding regions of MSH2 and MLH1 are known to be responsible for up to 45%-64% of all HNPCC families<sup>[15]</sup>. We have previously detected germline mutations of the entire coding regions of MSH2 and MLH1 genes in 24 AC probands, 15 JC probands and 19 BG patients using PCR-gene-sequencing with 17 germline mutations detected including two mutations occurred in a same patient<sup>[16]</sup>. Three new mutations have been found by mRNA-based PCR sequencing<sup>[17]</sup>. The remaining 39 probands without MSH2 and MLH1 might be associated with the other abnormal MMR genes such as MSH6.

*MSH6* mutations are involved in the development of colorectal cancer<sup>[18]</sup>. Germline mutations of *MSH6* have been reported in two atypical HNPCC Japanese families lacking mutations in *MSH2* and *MLH1*<sup>[19,20]</sup>. Some researchers believe that *MSH6* gene might be the first candidate gene for detecting germline mutations in HNPCC families in which *MSH2* and *MLH1* mutations are excluded<sup>[21]</sup>. It was reported that *MSH6* mutations account for 10% of kindreds in which *MSH2* and *MLH1* mutations are excluded<sup>[6]</sup>. Our study has demonstrated four missense mutations of *MSH6* gene in the probands Table 2 Germline mutations of *MSH6* gene in 4 probands of HNPCC families

Family No	Exon	Position of mutation	Base change	Result	Mutation type
H3	4.3	Codon 468	c.1403G>A	Arg>His	Missense mutation
H14	6	Codon 1163	c.3488A>T	Glu>Val	New SNPs
H40	4.6	Codon 666	c.1996T>C	Ser>Pro	Missense mutation
H61	9	Codon 1284	c.3851C>T	Thr>Met	Missense mutation

Table 3	Clinical	characteristics	of	4	mutational	probands	of
HNPCC f	families						

No	Sex	Age (yr)	Criteria	Site of cancer	Age at diagnosis of first CRC (yr)	Metachronous tumor
H3	Μ	50	AC	Sigmoid	46	VA
H14	F	26	BG	Rectum	26	NMT
H40	F	39	BG	Rectum	39	NMT
H61	F	51	BG	Descending	33	EC

Abbreviation: AC, Amsterdam criteria. JC, Japanese criteria. BG, Bethesda guidelines. CRC, Colorectal cancer. VA, Villiform adenoma. NMT: No Metachronous tumor. EC, Endometrial cancer.

of 39 Chinese HNPCC families, which have not been reported (http://www.nfdht.nl). However, missense mutations in MMR genes are common and often pose a formidable problem of interpretation, because these changes do not necessarily affect the function of the protein. Further functional studies are required in order to determine whether the missense mutations are neutral polymorphisms or clinically relevant mutations. We detected the exons of these four missense mutations by direct squencing of the MSH6 gene using genomic DNA from blood samples of 137 healthy persons, the mutational rate was approximately 3.65% (5/137). Since "polymorphism" is a term that is usually used for genetic variants with a minor allele frequency  $\ge 1\%$  in a given population<sup>[22]</sup>, single basyl transversion at c.3488A>T of exon 6 might represent a new SNP, although the changes are not found in SNP (http://www.ensembl.org). The remaining three missense mutations were not detected in genomic DNA from 137 healthy persons, suggesting that they are three novel missense muations of MSH6 gene in Chinese HNPCC families. The clinical features of the probands of these three novel missense mutations and one new SNP are shown in Table 3. In brief, the above mutation carriers occur more frequently in the left colon than in MLH1 or MSH2 mutation carriers. However, the relationship between the above typical features and germline mutations of MSH6 gene in the probands of Chinese HNPCC families still remains unclear. To date, the International Collaborative Group on HNPCC (ICG-HNPCC) has found over 30 potentially pathogenic MSH6 mutations. A significant proportion (35%) of them results in a single amino acid substitution, which is difficult to interpret. Since the pathogenicity of HNPCC mutations is linked to malfunction of MMR, whether the above three novel missense mutations are involved in human MMR needs to be further investigated.

## ACKNOWLEDGMENTS

The authors are grateful to the patients who took part in this study and to Departments of Cancer Hospital for sending blood and tumor specimens. The authors also appreciate the help from Professor Sun MH for her detection of germline mutations of *MSH2* and *MLH1* gene in the probands of certain Chinese HNPCC cases and Professor Mo SJ for the supply of certain Chinese HNPCC cases.

## COMMENTS

### Background

Germline mutations in mismatched repair genes, such as *MLH1*, *MSH2* and *MSH6*, lead to hereditary nonpolyposis colorectal cancer (HNPCC) syndrome. Germline mutations of *MLH1* and *MSH2* gene have been reported in Chinese HNPCC families. However, the germline mutation of *MSH6* has not yet been reported.

### **Research frontiers**

Now many researchers are engaged in studies of HNPCC, especially in germline mutations of MMR genes such as *MSH2*, *MLH1* and *MSH6*. These studies can contribute to the early diagnosis of HNPCC and screening of HNPCC families. Few studies on germline mutations of *MSH6* gene are available.

### Innovations and breakthroughs

Three novel germline mutations of *MSH6* gene have been found in 39 probands of Chinese HNPCC families by PCR-based sequencing, and a new SNP has been found by screening the missense mutations of genomic DNA in 137 healthy persons.

### Applications

Germline mutations in genes can be used to diagnose early HNPCC and enrich international HNPCC mutation and SNP databases.

#### Terminology

HNPCC is an abbreviation of hereditary nonpolyposis colorectal cancer. Germline mutations are the mutations in genomic DNA.

### Peer review

This paper is an interesting manuscript, the authors detected new HNPCC-related mutations and discovered three novel mutations and an additional SNP in 39 unrelated HNPCC probands. Data on patient characteristics, such as gender and age at diagnosis of colorectal cancer in the patient (or in the family) should be provided in detail.

## REFERENCES

- 1 Lynch HT, de la Chapelle A. Genetic susceptibility to nonpolyposis colorectal cancer. J Med Genet 1999; 36: 801-818
- 2 Banno K, Susumu N, Hirao T, Yanokura M, Hirasawa A, Aoki D, Udagawa Y, Sugano K, Nozawa S. Identification of germline MSH2 gene mutations in endometrial cancer not fulfilling the new clinical criteria for hereditary nonpolyposis colorectal cancer. *Cancer Genet Cytogenet* 2003; **146**: 58-65
- 3 **Peltomäki P**. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* 2003; **21**: 1174-1179
- 4 Viel A, Genuardi M, Capozzi E, Leonardi F, Bellacosa A, Paravatou-Petsotas M, Pomponi MG, Fornasarig M, Percesepe A, Roncucci L, Tamassia MG, Benatti P, Ponz de Leon M, Valenti A, Covino M, Anti M, Foletto M, Boiocchi M, Neri G. Characterization of MSH2 and MLH1 mutations in Italian families with hereditary nonpolyposis colorectal cancer. *Genes Chromosomes Cancer* 1997; **18**: 8-18
- 5 Wijnen J, Khan PM, Vasen H, van der Klift H, Mulder A, van Leeuwen-Cornelisse I, Bakker B, Losekoot M, Møller P, Fodde R. Hereditary nonpolyposis colorectal cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations. *Am J Hum Genet* 1997; **61**: 329-335

- 6 Kariola R, Raevaara TE, Lönnqvist KE, Nyström-Lahti M. Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. *Hum Mol Genet* 2002; 11: 1303-1310
- 7 Peltomäki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum Mol Genet* 2001; 10: 735-740
- 8 Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum* 1991; 34: 424-425
- 9 Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* 1999; 116: 1453-1456
- 10 Fujita S, Moriya Y, Sugihara K, Akasu T, Ushio K. Prognosis of hereditary nonpolyposis colorectal cancer (HNPCC) and the role of Japanese criteria for HNPCC. *Jpn J Clin Oncol* 1996; 26: 351-355
- 11 Rodriguez-Bigas MA, Boland CR, Hamilton SR, Henson DE, Jass JR, Khan PM, Lynch H, Perucho M, Smyrk T, Sobin L, Srivastava S. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. J Natl Cancer Inst 1997; 89: 1758-1762
- 12 Chen HP, Wang TR, Xu XY, Zhang M, Xiang WP, Jiang RZ, Ma TY. Comparison of three methods for the gene analysis of fetal cells from maternal peripheral blood. *Chin Med J* (Engl) 2004; 117: 507-510
- 13 Yuan Y, Huang YQ, Cai SR, Song YM, Zheng S, Zhang SZ. Genetic characterization of Chinese hereditary non-polyposis colorectal cancer by DHPLC and multiplex PCR. *Jpn J Clin Oncol* 2004; 34: 660-666
- 14 Ripa RS, Katballe N, Wikman FP, Jäger AC, Bernstein I, Orntoft T, Schwartz M, Nielsen FC, Bisgaard ML. Presymptomatic diagnosis using a deletion of a single codon in families with hereditary non-polyposis colorectal cancer. *Mutat Res* 2005; 570: 89-96
- 15 Shin KH, Shin JH, Kim JH, Park JG. Mutational analysis of promoters of mismatch repair genes hMSH2 and hMLH1 in hereditary nonpolyposis colorectal cancer and early onset colorectal cancer patients: identification of three novel germline mutations in promoter of the hMSH2 gene. *Cancer Res* 2002; 62: 38-42
- 16 Cai Q, Sun MH, Fu G, Ding CW, Mo SJ, Cai SJ, Ren SX, Min DL, Xu XL, Zhu WP, Zhang TM, Shi DR. Mutation analysis of hMSH2 and hMLH1 genes in Chinese hereditary nonpolyposis colorectal cancer families. *Zhonghua Binglixue Zazhi* 2003; 32: 323-328
- 17 Wang CF, Zhou XY, Zhang TM, Sun MH, Xu Y, Shi DR. The analysis for mRNA mutation of MLH1, MSH2 genes and the gene diagnosis for hereditary nonpolyposis colorectal cancer. *Zhonghua Yixue Yichuanxue Zazhi* 2006; **23**: 32-36
- 18 Papadopoulos N, Nicolaides NC, Liu B, Parsons R, Lengauer C, Palombo F, D'Arrigo A, Markowitz S, Willson JK, Kinzler KW. Mutations of GTBP in genetically unstable cells. *Science* 1995; 268: 1915-1917
- 19 Akiyama Y, Sato H, Yamada T, Nagasaki H, Tsuchiya A, Abe R, Yuasa Y. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer Res* 1997; 57: 3920-3923
- 20 **Miyaki M**, Konishi M, Tanaka K, Kikuchi-Yanoshita R, Muraoka M, Yasuno M, Igari T, Koike M, Chiba M, Mori T. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. *Nat Genet* 1997; **17**: 271-272
- 21 Wu Y, Berends MJ, Post JG, Mensink RG, Verlind E, Van Der Sluis T, Kempinga C, Sijmons RH, van der Zee AG, Hollema H, Kleibeuker JH, Buys CH, Hofstra RM. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. *Gastroenterology* 2001; **120**: 1580-1587
- 22 **Freeman JL**, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. Copy number variation: new insights in genome diversity. *Genome Res* 2006; **16**: 949-961