Germline Mutation Characteristics of MSH6, PMS2 And MLH3: A Statistical Analysis Based on All Reported Data

Jiang Fan[1], Fangqi Liu[2], Huining Gu[3], Binbin Cai[4], Ye Xu[5], Peng Nan[6]

[1, 3, 4, 6]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, School of Life Sciences, Fudan University, Shanghai 200433, China
[2, 5]Colorectal Surgery, Shanghai Cancer Hospital of Fudan University, Shanghai 20032, China
[1]10210700003@fudan.edu.cn; [2]lfqman@hotmail.com; [3]07300700112@fudan.edu.cn; [4]07300700020@fudan.edu.cn; [5]yexu@shmu.edu.cn; [6]nanpeng@fudan.edu.cn

***Abstract*-Loss of mismatch repair (MMR) function predisposes to a mutator cell phenotype, microsatellite instability (MSI) and cancer, especially hereditary non-polyposis colorectal cancer (HNPCC). Plenty of studies of MMR gene MLH1 and MSH2 have been conducted and many mutations have been reported. However, there are few studies of other MMR genes, such as MSH6, PMS2 and MLH3, which are also important especially from statistical aspect. In this study, we collected all the related literatures and extracted the useful information to build a database and performed further analysis. The study focused on the germline mutation characteristics of these three genes and proposed different strategies when dealing with patients of different area, genders and ages. The results showed that (1) each gene has its own hot spots, exons and mutation pattern, which we might pay special attention to when processing the clinical detection and diagnosis; (2) Although Asian people account for the minority of the total patients, there are still some important hot spots and exons that are different from the total sample and require special attention; (3) For the patients with mutations of MSH6, the average age of patients (around 50) is older than usual age (45), and many of them would get more than one attack. And it is worth mentioning that most women (51.3%) got both colorectal cancer and endometrial cancer, which implies that when diagnosing colorectal cancer, it is necessary to check the female reproductive systems such as the uterus and ovary for women.**

Keywords- HNPCC; MSH6; PMS2; MLH3; Data Analysis

I INTRODUCTION

Hereditary nonpolyposis colorectal cancer (HNPCC) (MIM#-120435), also known as Lynch Syndrome, is an autosomal dominant disease characterized clinically by early onset of colorectal cancer (CRC) and other associated tumors [1]. The syndrome accounts for approximately 2-5% of all the colorectal cancer cases and is caused by germline mutations in DNA mismatch repair (MMR) genes [2]. So far, 6 related MMR genes have been identified: MLH1 (3p21-23) (MIM*120436), MSH2 (2p21) (MIM*609309), MSH6 (2p21) (MIM*600678), MLH3 (14q24.3) (MIM*604395), PMS1 (2q31-q33) (MIM*600258) and PMS2 (7p22) (MIM*600259). The mutations of MLH1 and MSH2 underlie 90% of the germline HNPCC mutations, followed by MSH6 (7%), and the mutations of the remaining genes account for less than 5% of HNPCC [3].

In comparison with the mismatch repair genes MLH1 and MSH2, the genes MSH6, PMS2, PMS1 and MLH3 are relatively understudied with respect to cancer risk. However, recent research papers show that other than MSH2 and MLH1, these genes also get involved in HNPCC [4]. There are findings showing that association exists between bi-allelic germline PMS2 mutations and severe childhood-onset gastrointestinal manifestations [5]. In a study with large scale of colorectal cancer (CRC) samples (N=2685), MSH6*c.3984_3987dupGTCA, MSH6*c.3959_3962delCAAG, MSH6*c.3984_3987dupGTCA and MSH6*c.3959_3962delCAAG were found in these CRC cases, consistent with a high risk of CRC [6]. There are studies reported that MSH6 mutations are frequent in HNPCC families with normal MSH6 expression as detected by immunohistochemistry [7]. So far, 18 MLH3 germline mutations/variants have been identified in familial colorectal cancer cases. Sixteen of these variants are amino acid substitutions of which the pathogenic nature is still unclear [8]. According to these reported results above, MSH6, PMS2 and MLH3 genes also play an important role in HNPCC and should not be ignored during the study of HNPCC. However, as for gene PMS1, nearly no reports and studies found its function in HNPCC. And according to the biggest database of HNPCC, the InSiGHT database, there are only one variant above all published data. For that reason, we did not study the gene PMS1 in this paper.

Among all the results reported, there are no hotspots of these genes being reported in the worldwide scale, and there are no studies focused on specific patients, say Asian patients. However, there are reports showing that Caucasians and Asians have significant differences in the distributions of both individuals and mutation types in MLH1 and MSH2, and each gene has its own mutation pattern [9]. Based on this study, and considering the importance of MSH6, PMS2 and MLH3 genes in HNPCC, it is highly likely that these three genes also have their own germline mutation patterns. It is necessary to focus on the statistical analysis of all the reported data and get the complete information of these genes. It is also worthwhile to analyse the hotspots and frequent mutations at a large scale, for the results can offer guides and suggestions in the clinical diagnosis of

HNPCC. For this reason, we collected the related literature, extracted the important information and excluded the invalid data to perform statistical analysis of the whole database. Also, the special features of Asian population were analysed to illustrate the significant differences between different populations.

II METHODOLOGY

A. Data Collection

To get the overall data, we collected all the existing related publications (till March 2014), by searching "PMS2", "MLH1" and "MSH6" in PubMed (http://www.ncbi.nlm.nih.gov/pubmed). We downloaded all the references from PubMed (http://www.ncbi.nlm.nih.gov/pubmed), Springer (http://beta.springerlink.com), Elsevier (http://www.sciencedirect.com), and Wiley (http://www3.interscience.wiley.com) and so on to extract the information of patients and mutations. Each data item includes all the information of patients, such as Ethnicity, Number of patients, and the information of mutations, such as Gene, Exon, DNA change, Protein, Mutation type, Pathogenic, etc. Part of the data items contain more complete information, including Age, Gender and Disease, etc. For this part of data, we can perform further analysis of different ages and genders. To ensure the accuracy of the analysis, we excluded the experiments in tissues or in vitro, and only collected the mutations confirmed in patients.

B. Definition of Mutations Type

On the transcriptional and translational levels, a single nucleotide alteration (change, deletion or insertion) may result in substitution of an amino acid (missense), a change in the reading frame (frameshift), immediate translation termination caused by the introduction of a stop codon (nonsense), or make no change of the amino acid (silent). In MMR genes, another kind of alteration (large deletion or duplication) may occur when the deletion or duplication of the whole exon happens. We classified all the MSH6, PMS2 and MLH1 alterations into nonsense, frameshift, large deletion/duplication, missense, silent mutation and aberrant splicing. The rest alterations, including variants in intron, promoter site, 3' and 5'-UTR, classified into "others", were not considered in this study.

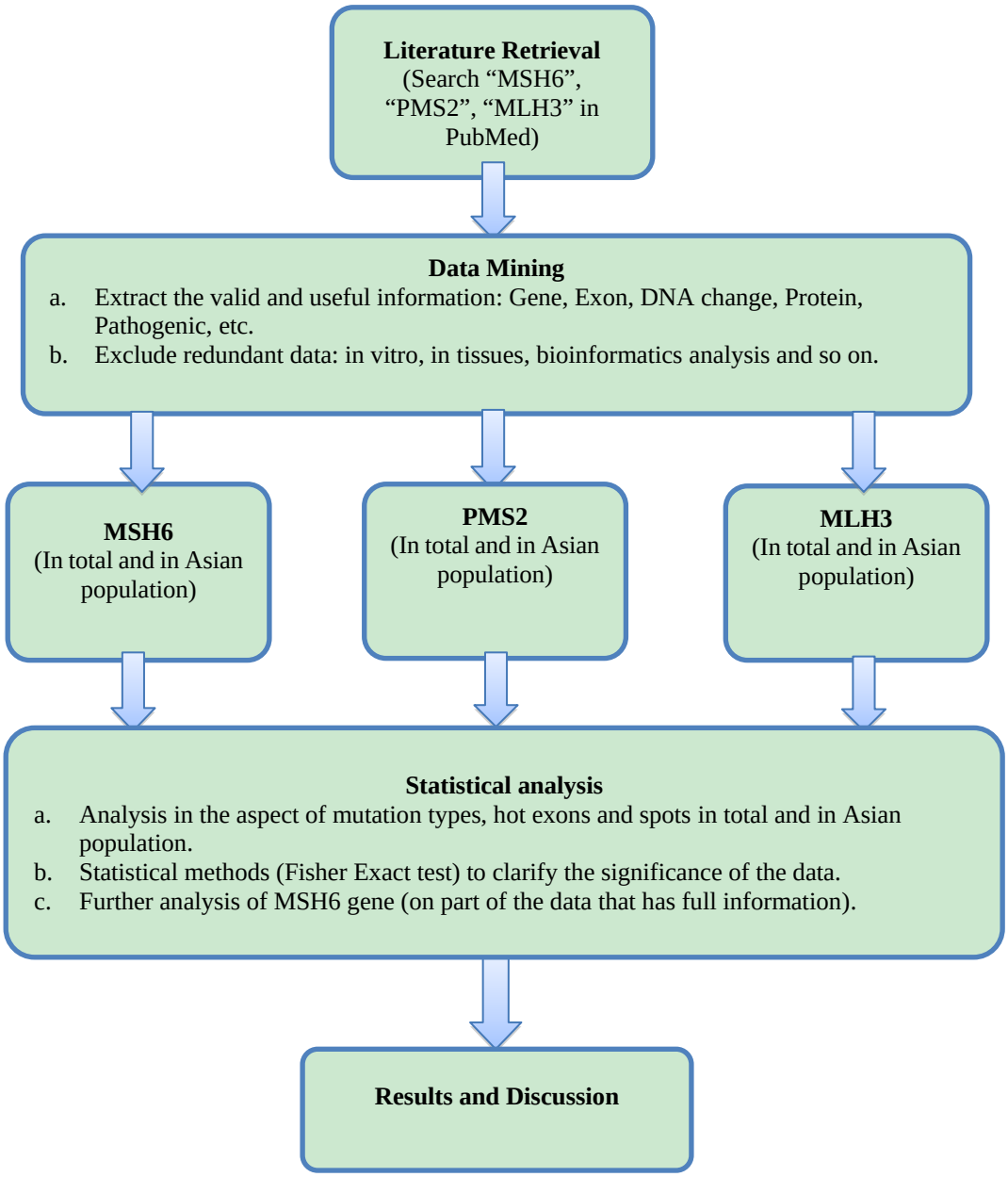
C. Race of Patients

According to all the information we collected, patients can be classified into three groups: Caucasian, Asian and Africa-American. The sample size of Africa-American was too small for analysis. In this study, we mainly focused on the total sample and the Asian population. Patients in some Asian countries, such as China, Japan, and Korea and so on, were grouped into Asian population.

D. Statistical Analysis

Fisher Exact test was used to calculate the significance of the difference between Asian population and the total sample in the distribution of each exon, for the data did not follow normal distribution (the null hypothesis is that the racial difference between two groups is not significant). A P-value of <0.05 was considered statistically significant.

To illustrate our study, a framework of the logic and steps is presented below.



III RESULTS

In this study, we collected the papers published in PubMed that related to MSH6, PMS2 and MLH3, organized and reanalysed the data from these papers to display the complete gene information. We deleted the redundant data, such as in vitro studies and bioinformatics analysis, and only collected the mutations confirmed in patients. In total, we extracted 1519 items of data in MSH6, 1256 in PMS2 and 163 in MLH3 from all the literature publications found in PubMed. In these items, Asian patients account the minority of all patients, with the numbers 45, 2 and 60 in MSH6, PMS2 and MLH3, respectively (as shown in Table 1). In that case, the number of individuals with variant in the Asian population only accounted for the minority of total sample. Since the Africa-Americans were so few in our data, the majority of the samples were Caucasians. In gene PMS2, there was only one unique DNA variant in the Asian population and two individuals carried this variant.

The number and percentage of mutations in each mutation type of MSH6, PMS2 and MLH3 in total and Asian populations are shown in Table 2. Obviously, the most prevalent mutation types are missense and frameshift, accounting for about 70% of the total mutations of each gene. Large deletion/ duplication also take place in MSH6 and PMS2, with the percentages of 1.5% and 10.6% respectively, while the number is 0 in Asians. The number of patients with each mutation type in each exon is also shown in Fig. 1. As the patterns of MSH6 in total sample and Asian population are significantly different, the figure of Asian population is displayed separately. As is shown in the figure, missense and frameshift still take up the majority of all the individuals with variant. In gene MSH6, the mutations take place most frequently in exon 4, while in Asians it is in exon 5, followed by exon 4. Fisher Exact test was used to examine the differences between Asians and the total sample in each exon. According to our statistical analysis of Fisher test, the p values of all exons are larger than 0.05 except for exon 4 and 5 (as

marked in Fig. 1b. The p values of the exons are shown in a table in Fig. 1b). There are significant differences between Asians and the total sample in the distributions of exon 4 and 5. In gene PMS2, most mutations occur in exon 11. And in MLH3 almost all mutations take place in exon 2.

According to the data we collected, the top ten mutations of MSH6, PMS2 and MLH3 in total sample are shown in Fig. 2. In MSH6, c.651dupT takes up the biggest percentage (2.9%), with 48 patients carried this variant, and in PMS2, the most frequent mutation is c.876A>G, with 91 individuals carried this variant and took up 8.4%, and in MLH3, 81 individuals carried the "hot" mutation c.4377G>A, and the percentage is up to 50.3%. The top mutations in Asians are marked with red squares in Fig. 2. And in Fig. 2c, the three mutations in red squares are not only top mutations but also specific mutations in Asian population. Since the sample size of Asian population is much smaller than that of the total population, the top two instead of the top ten mutations of each gene are listed (Table 3). Also, there are mutations that only take place in Asian population. In Table 3, these specific mutations only take place in Asian population are listed, so are the top two mutations in Asian population. The specific mutations of each gene are: c.440T>A, c.3200C>A and c.3246T>A in MSH6, c.691A>C, c.2896T>C and c.2825C>T in MLH3 and c.705G>A in PMS2 (for it only has one specific mutation in Asian population). The mutation c.691A>C of MLH3 in Asian population also has a relatively high percentage (26.7%).

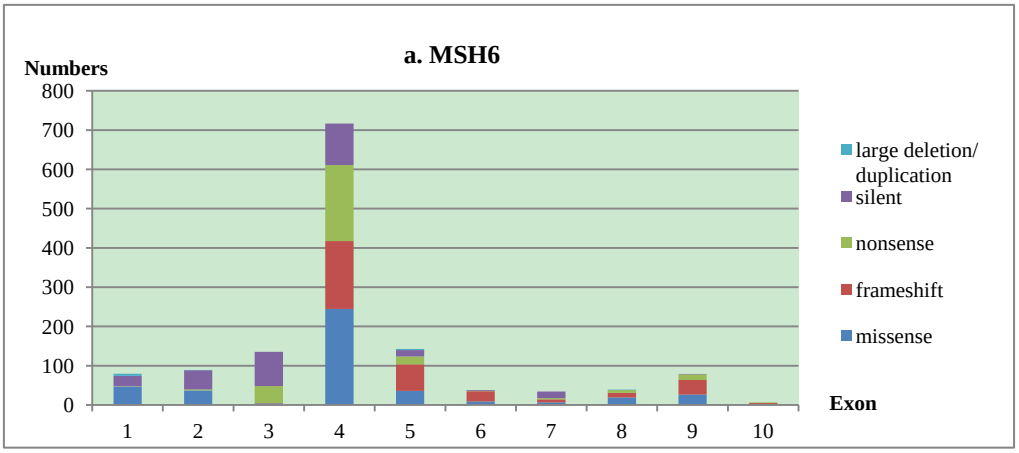
The items related to MSH6 are not only significantly more than those of MLH3 and PMS2, but also mainly contain the complete information, including gender, age, and disease information and so on. In the total number of 1519 items, there are 620 items containing the complete information. And in these 620 items, only 27 items are from Asian area. The Asians with complete information are so few that we did not separate them from the total population. In that case, the 620 items were analysed independently from the aspects of gender, age and disease information (Fig. 3). For all the 620 individuals, the first attack of HNPCC occurs at the age of 50 averagely, which is much later than the ordinary age of HNPCC. And in these 620 patients, 169 patients got 2-5 attacks of HNPCC (Fig. 3a). Among all the diseases related to the mutation of MSH6, the incidence of colorectal cancer is the highest (54%), followed by the incidence of endometrial cancer (21%) and ovarian cancer (4%), as shown in different genders in Fig. 3b and Fig. 3c. Furthermore, the female patients account for 78.9% of all the patients who got more than one attack. We analysed the female patients who got two attacks (in total 115 samples), and found that most of them (51.3%) caught both colorectal cancer and endometrial cancer (Fig. 3d).

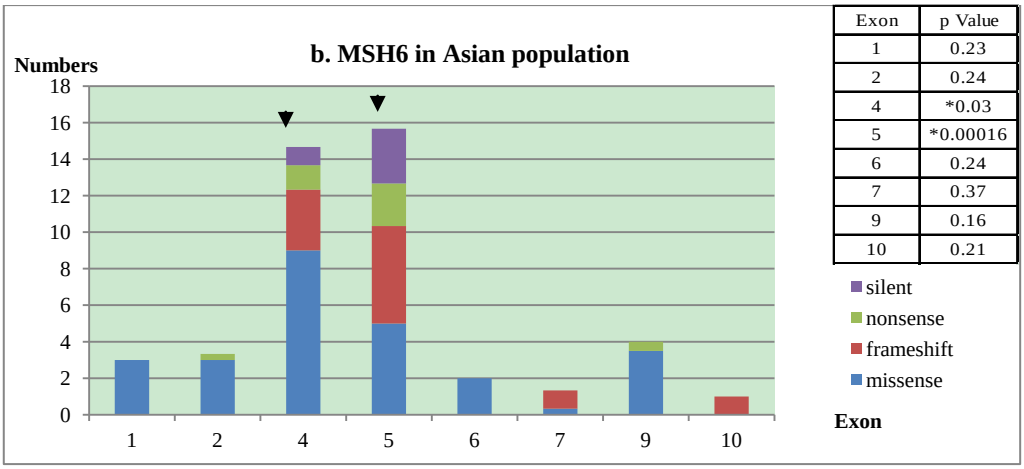
TABLE 1 DNA VARIANTS OF MSH6, PMS, MLH3IN IN TOTAL AND ASIAN GROUPS

	MSH6		**PMS2**		**MLH3**	
	Total	**Asian**	**Total**	**Asian**	**Total**	**Asian**
Total number of unique DNA variants	521	36	217	1	27	12
Total number of individuals with variant (s)	1519	45	1256	2	163	60

TABLE 2 MUTATION TYPES OF MSH6, PMS2 AND MLH3 IN TOTAL AND ASIAN GROUPS

	MSH6		PMS2		MLH3	
	n (%) in Total	n (%) in Asians	n (%) in Total	n (%) in Asians	n (%) in Total	n (%) in Asians
missense	199(38.1%)	19(52.8%)	109(50.2%)	1(100%)	21(77.8%)	10(83.3%)
frameshift	167(32.1%)	7(19.4%)	33(15.2%)	0	4(14.8%)	1(8.3%)
nonsense	88(16.9%)	6(16.7%)	9(4.1%)	0	1(3.7%)	0
large deletion/ duplication	8(1.5%)	0	23(10.6%)	0	0	0
silent	39(7.5%)	3(8.3%)	15(6.9%)	0	0	0
aberrant splicing	7(1.3%)	0	18(8.3%)	0	0	0
others	13(2.5%)	1(2.8%)	10(4.6%)	0	1(3.7%)	1(8.3%)
total	521	36	217	1	27	12





Exon	p Value
1	0.23
2	0.24
4	*0.03
5	*0.00016
6	0.24
7	0.37
9	0.16
10	0.21

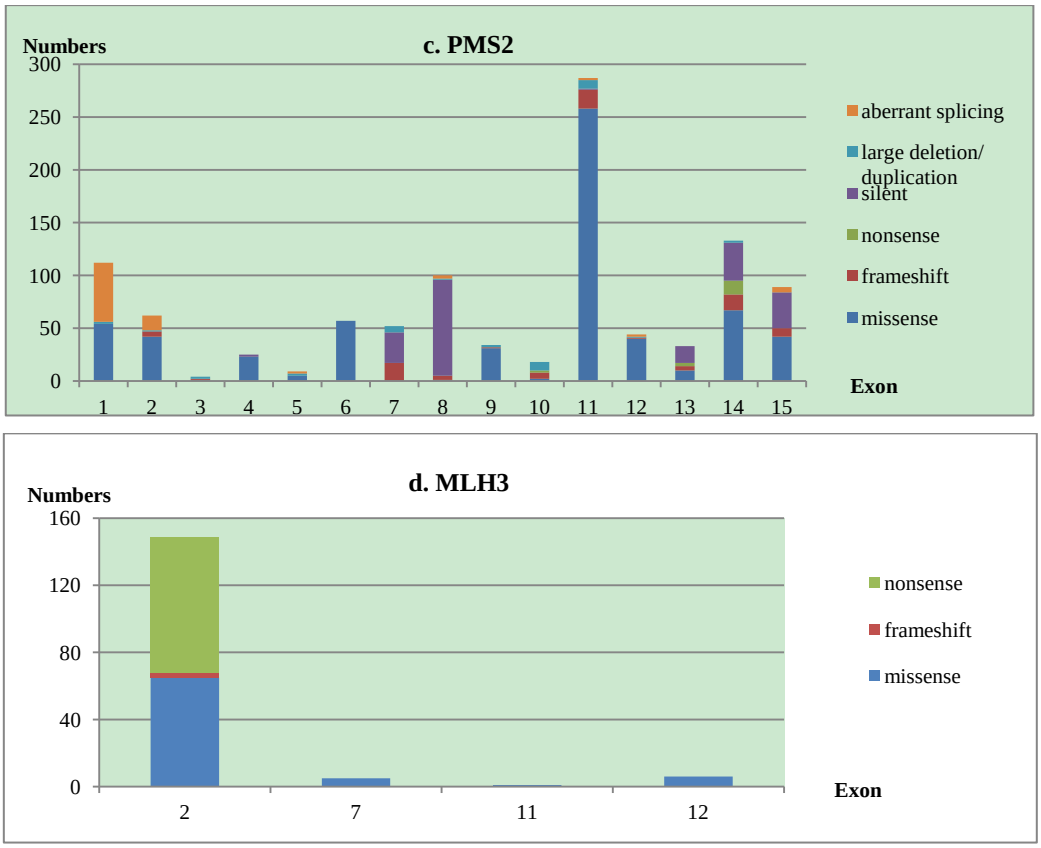
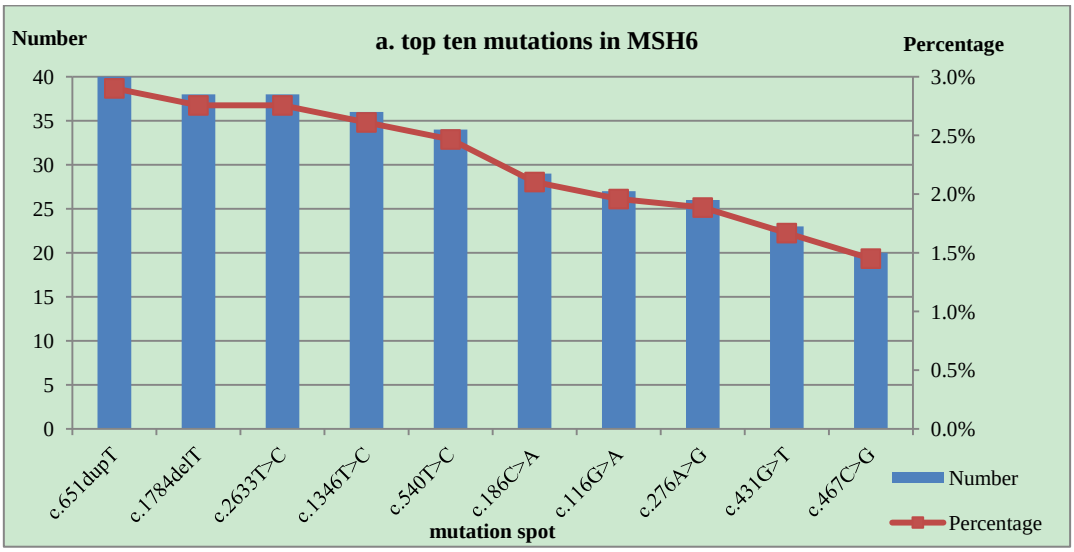


Fig. 1 Distribution of mutation types in each exon of MSH6 (a), MSH6 in Asian population (b), PMS2(c) and MLH3 (d). In this figure, we performed the Fisher Exact test on the differences between Asian population of MSH6 and the total population in each exon. Each p value is shown in the table in Fig. 1b, and two exons were examined to have significant differences, which are marked with "▼" in Fig. 1b



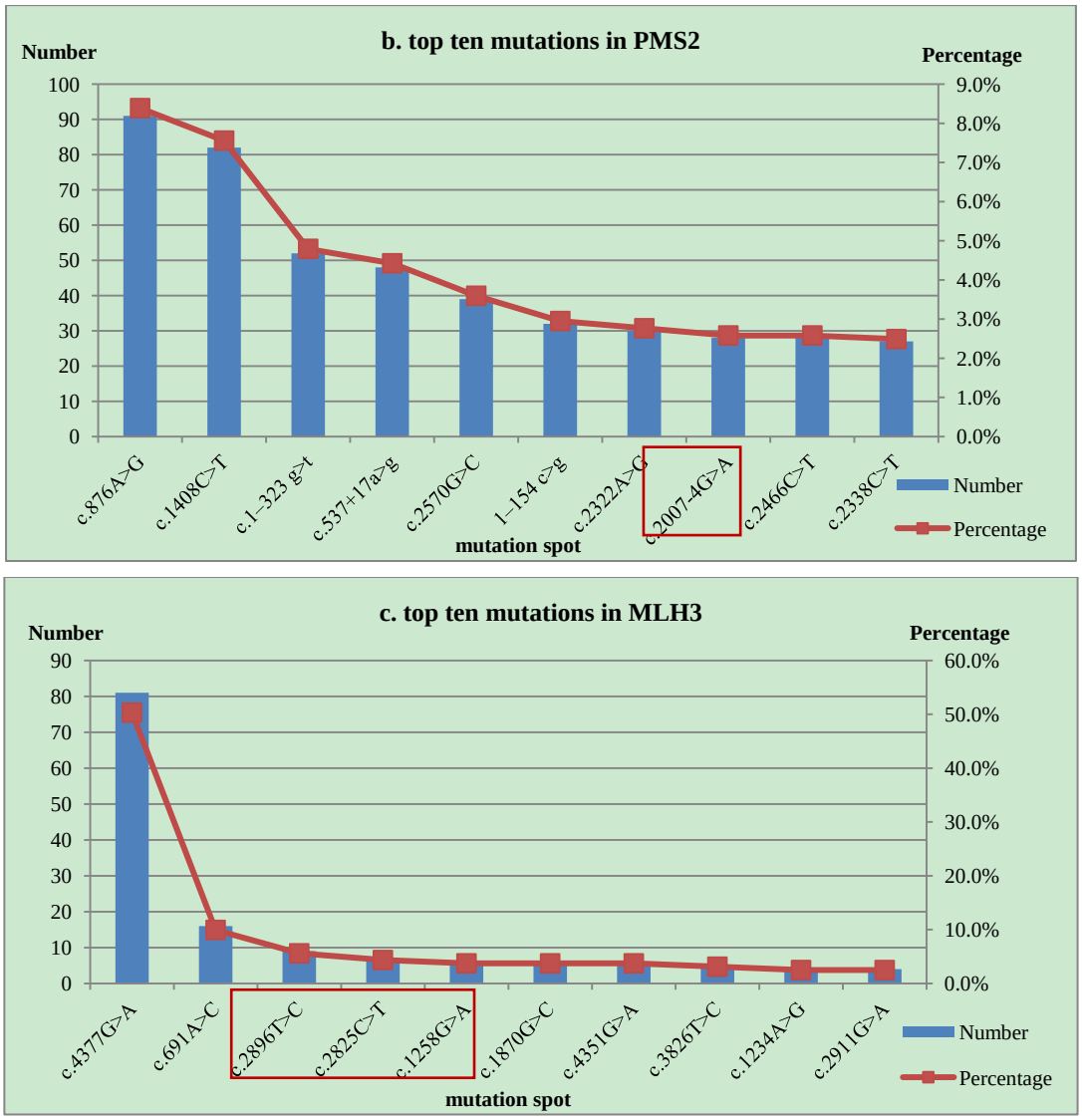
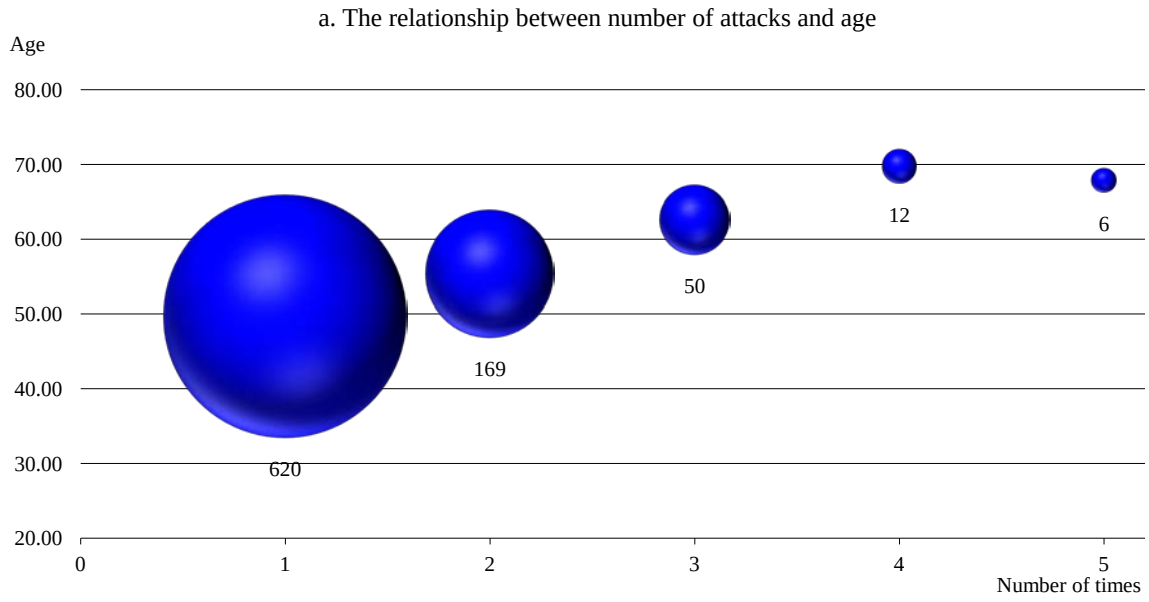


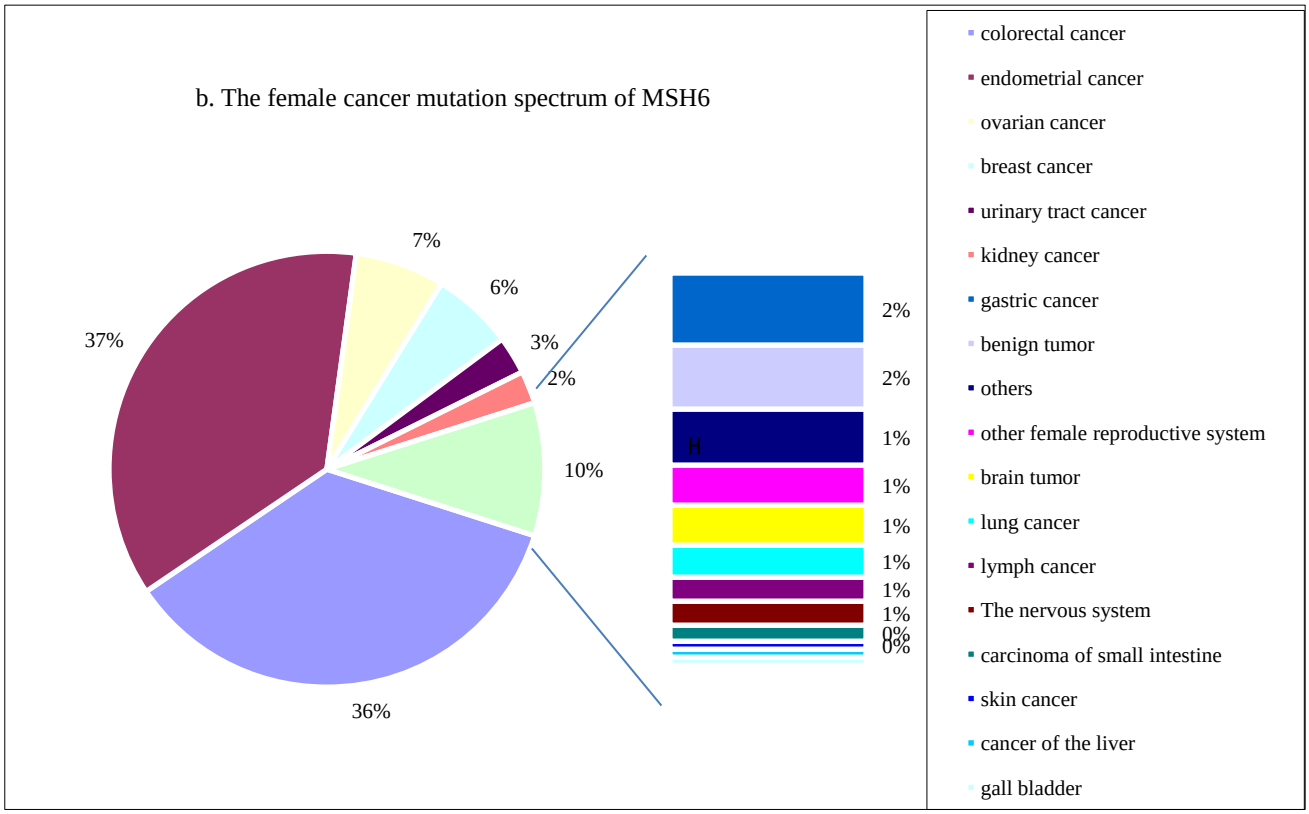
Fig. 2 Distribution of top ten mutations in gene MSH6 (a), PMS2 (b) and MLH3(c) in all samples. The mutation spot marked with "□" in Fig. 2a is also a top mutation in Asians, and in Fig. 2c the three mutations are not only top mutations but also the specific mutations in Asians (see Table 3)

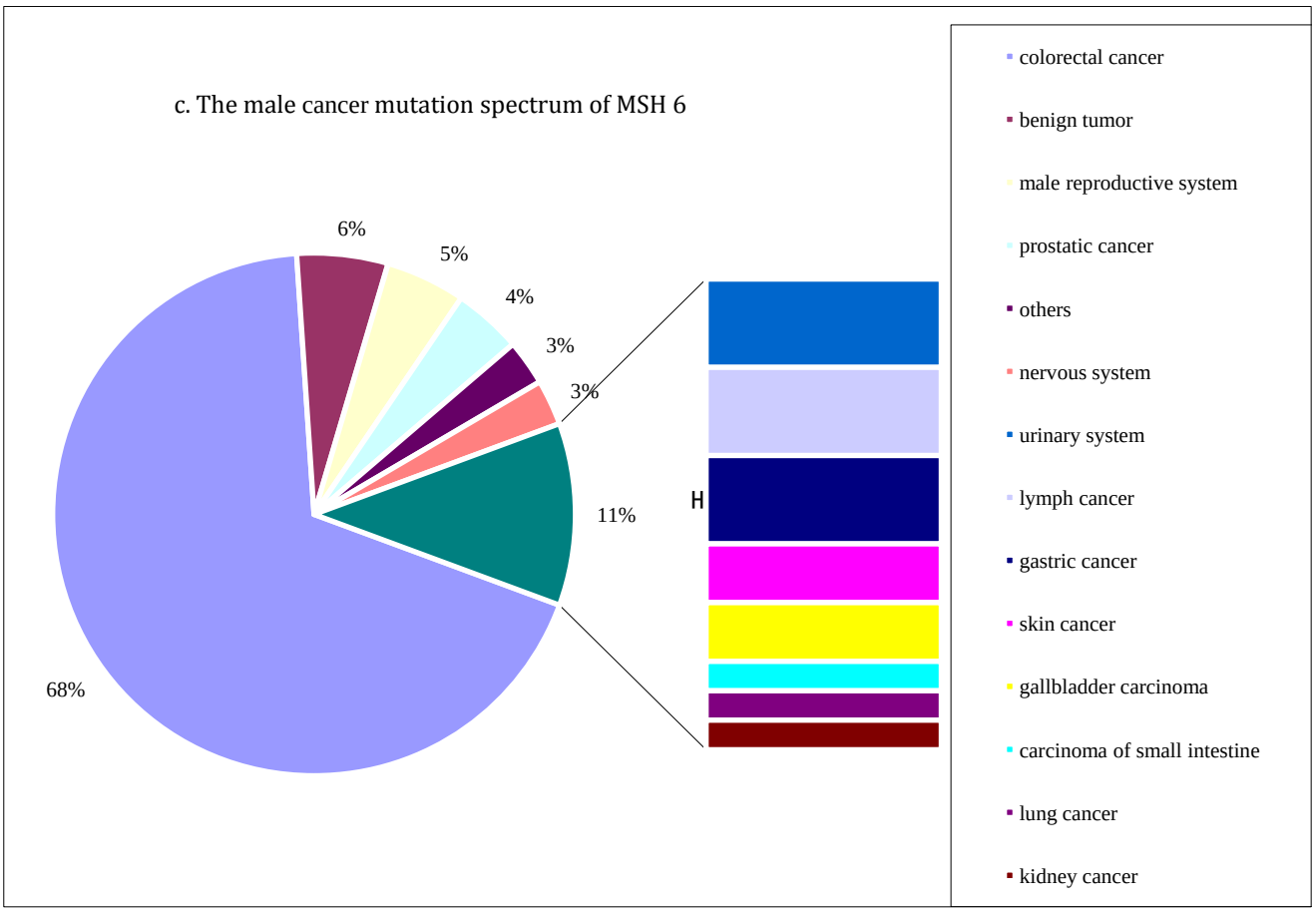
TABLE 3 MUTATION PREFERENCE IN ASIAN POPULATION

gene		Mutation	Exon	n (%) in Asians
MSH6	Specific mutation	c.440T>A	2	
		c.3200C>A	5	
		c.3246T>A	5	
	Top mutation	c.3261dupC	5	5(11.1%)
		c.116G>A	1	3(6.7%)
PMS2	Specific mutation	c.705G>A	6	
MLH3	Specific mutation	c.691A>C	2	
		c.2896T>C	2	
		c.2825C>T	2	
	Top mutation	c.691A>C	2	16(26.7%)
		c.2896T>C	2	9(15.0%)

Notes: Specific mutation means mutations only occur in Asian population. And as the sample size of Asian population is not so large, we only listed two top mutations here. For gene PMS2, there is in total only one mutation in Asians.




b. The female cancer mutation spectrum of MSH6
7%
6%
3%
2%
37%
10%
36%
2%
2%
1%
1%
1%
1%
1%
1%
0%
0%
colorectal cancer
endometrial cancer
ovarian cancer
breast cancer
urinary tract cancer
kidney cancer
gastric cancer
benign tumor
others
other female reproductive system
brain tumor
lung cancer
lymph cancer
The nervous system
carcinoma of small intestine
skin cancer
cancer of the liver
gall bladder


c. The male cancer mutation spectrum of MSH 6
6%
5%
4%
3%
3%
11%
68%
colorectal cancer
benign tumor
male reproductive system
prostatic cancer
others
nervous system
urinary system
lymph cancer
gastric cancer
skin cancer
gallbladder carcinoma
carcinoma of small intestine
lung cancer
kidney cancer

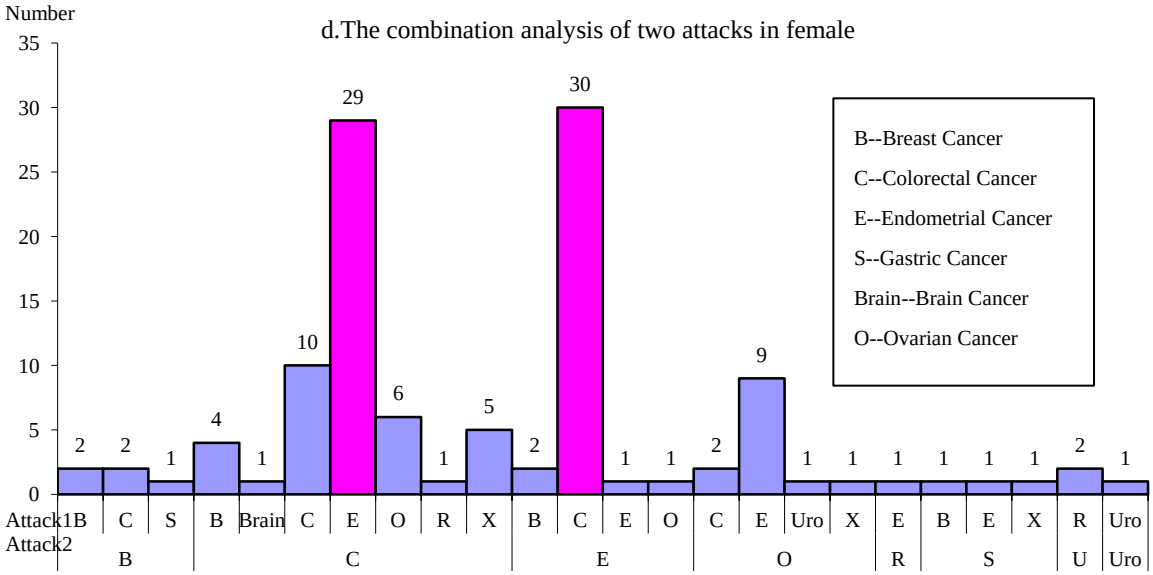


Fig. 3 The further analysis of MSH6 with complete information. In Fig. 3a, the first circular indicates the average age of first attack for all 620 patients, and the second indicates the average age for all the second attack, and so on. In Fig. 3d, the red bars represent the female patients who got both colorectal cancer and endometrial cancer, which are in total 59 patients, accounting for the majority (51.3%) of the total number

IV DISCUSSION

In this study, we downloaded all the relevant literature and extracted the key information including patients' mutations, protein, races, disease information and pathogenicity, etc. Since the information in the InSight database is incomplete, we constructed our own data based on these thousands of data items.

Only a few studies have investigated the effects of race in HNPCC or related diseases [10, 11]. Some research papers show that Caucasian and Asian populations have differences in the distributions of both individuals and mutation types in MLH1 and MSH2 [12]. In many other diseases, differences have been found between races. For example, Muhammad U et al. analysed 1631 cases from the SEER Database and found that Caucasians have a 9-fold higher incidence of Ewing sarcoma compared with African American [13]. An analysis of a multiple sclerosis (MS) centre population from the North-eastern United States showed a higher proportion of non-Caucasians in the younger cohort to have paediatric-onset MS [14]. Comparing the out-of-hospital cardiac arrest (OOHCA) in South Asian and Caucasian population in London, Shah AS et al. found that South Asians have an OOHCA at a significantly younger age [15]. Kabir AA et al. even found racial differences in Caesareans, although the differences were potentially unnecessary [16]. These studies only focus on the differences between Africa-American and Caucasian populations in USA, and are based on limited populations. In this study, we focused on the mutations, the more essential cause of HNPCC, on a worldwide scale to find the distinctions of patients with different age, gender and nationalities. Also, we focused on the genes MSH6, PMS2 and MLH3 to provide indications and suggestions to clinical process.

According to our study, the total number of mutations in MLH3 is relatively rare, which might imply that MLH3 plays a minor role in HNPCC. An analysis of 30 Finnish CRC cases has been done and the results showed that while it is a difficult task to exclude a role of MLH3 in HNPCC, the study could not confirm a role for MLH3 in CRC predisposition [17]. The top one mutation, c.4377G>A, was found in 81 individuals though, causing no changes in the structure of protein. However, there are some important spots that worth attention. Compared to PMS2 and considering the small number of samples, MLH3 gene has a relatively large amount of mutations in Asians. The three specific mutations of MLH3 in Asian people are also important, for that only these three mutations, c.691A>C, c.2896T>C and c.2825C>T, occurred in the top mutations of the total sample and many families, and caused cancers in most cases. For the gene PMS2, there are in total 2 individuals of Asian people reported with this mutation. And for the mutation type of large deletion/duplication, the number of Asian people is zero, while the number in total is 23(11.1%). Yet this cannot be the reason to ignore the detection of large deletion/ duplication in these genes during clinical process, for that the number of studies in Asian area is far less than that in Caucasians, as we can see from the statistics in Table 1.

In this study, we listed the top ten frequent mutations in MSH6, PMS2 and MLH3 (Fig. 2). In gene PMS2, four of these top spots take place all over the world (European, Australasia and the north America): c.1454C>A, c.2322A>G, c.1621G>A and c.2570G>C, in the exon 1 and 11, and other hot spots exist in exon 14 and 15. This outcome could be a suggestion for the clinical diagnosis for Asian people, considering the rare examples in PMS2 in Asian population. Also, we displayed the number of patients with each mutation type in each exon in Fig. 1. In gene MSH6, the exon 4 seems to be the most important because it has a much higher frequency of mutations than others. However, in Asian population we should emphasize on exon 5 as well as exon 4. In PMS2, most mutations take place in exon 11 with the type of missense, which is an implication during

detection. But when dealing with Asians, this gene seems to play a minor role in HNPCC. And in gene MLH3 exon 2 would be the most important during diagnosis, as well as in Asian people.

In the further analysis of some MSH6 data with complete information, we could see the different strategies are necessary not only between different races but also between different genders and ages. According to Fig. 3a, the average age of patients with MSH6 mutations is bigger, and a large number of patients may get more than one attack, which means we need follow-up visits. The majority of female patients would get both colorectal cancer and endometrial cancer, which implies that when diagnosing colorectal cancer, it is necessary to examine the female reproductive system such as the uterus and ovary. In a reported study, the observed association of PMS2 rs7797466 with ovarian cancer warrants confirmation in an independent study [18]. A study of Polish MSH6 family shows that besides colorectal and endometrial cancer, the late-onset endometriosis type of ovarian cancer can also be a feature of families with MSH6 germline mutations [19]. Thus this result has provided a prognosis and therapeutic implications for women. It is also better for male patients to have their male reproductive system examined, such as prostate glands, although the percentage is much lower than that of female.

In clinical detection and diagnosing, these statistical numbers and features suggest that when screening mutations of a particular gene, mutations or exons with high frequency in this gene are primary. Also when dealing with patients from different areas, we might take different strategies. The mechanism behind these differences is not discussed in our study and requires further research.

V CONCLUSION

Our study provided some suggestions to the clinical detection. Our results indicate that each gene of MSH6, PMS2 and MLH3 has its own hot spots and exons, which we might pay special attention to when processing the clinical detection and diagnosis. And although Asian population accounts for the minority of the total patients, there are still some important hot spots and exons that are different from the total sample and require special attention. For the patients with mutations of MSH6, many of them get more than one attack and have a relatively later age (around 50) of incidence. And it is worth mentioning that most women (51.3%) got both colorectal cancer and endometrial cancer, which implies that when diagnosing colorectal cancer, it is necessary for women to have their female reproductive system examined, such as the uterus and ovary.

ACKNOWLEDGMENT

This study was supported by the grants from the 863 Hi-Tech Program of China (2009AA02Z308 and 2012AA02A602).

REFERENCE

[1] H. T. Lynch and T. Smyrk, “Hereditary nonpolyposis colorectal cancer (Lynch syndrome): An updated review”, *Cancer*, vol. 78, no. 6, pp. 1149-1167, 1996.

[2] H. T. Lynch and A. de la Chapelle, “Genetic susceptibility to non-polyposis colorectal cancer”, *Journal of Medical Genetics*, vol. 36, no. 11, pp. 801-818, 1999.

[3] P. Peltomaki and H. Vasen, “Mutations associated with HNPCC predisposition: Update of ICG-HNPCC/INSiGHT mutation database”, *Disease Markers*, vol. 20, iss. 4-5, pp. 69-276, 2004.

[4] M. A Jenkins, “Role of MSH6 and PMS2 in the DNA Mismatch Repair Process and Carcinogenesis,” *Surgical Oncology Clinics of North America*, vol. 18, no. 4, pp. 625-36, 2009.

[5] Herkert, J. C. Niessen, R. C. Olderode-Berends, M. J. W. Veenstra-Knol, H. E. Vos, Y. J. van der Klift, H. M. et al, “Paediatric intestinal cancer and polyposis due to bi-allelic PMS2 mutations: Case series, review and follow-up guidelines,” *European Journal of Cancer*, vol. 47, iss. 7, pp. 965-982, 2011.

[6] L. Raskin, F. Schwenter, M. Freytsis, M. Tischkowitz, N. Wong, G. Chong, et al, “Characterization of two Ashkenazi Jewish founder mutations in MSH6 gene causing Lynch syndrome,” *Clinical Genetics*, vol. 79, no. 6, pp. 512-522, 2011.

[7] H. Okkels, K. Lindorff-Larsen, O. Thorlasius-Ussing, M. Vyberg, J. Lindebjerg, L. Sunde, et al, “MSH6 Mutations are Frequent in Hereditary Nonpolyposis Colorectal Cancer Families With Normal pMSH6 Expression as Detected by Immunohistochemistry,” *Applied Immunohistochemistry & Molecular Morphology*, vol. 20, no. 5, pp. 470-477, 2012.

[8] J. H. Ou, M. Rasmussen, H. Westers, S. D. Andersen, P. O.Jager, K. A. Kooi, et al, “Biochemical Characterization of MLH3 Missense Mutations Does Not Reveal an Apparent Role of MLH3 in Lynch Syndrome,” *Genes Chromosomes & Cancer*, vol. 48, no. 4, pp. 340-350, 2009.

[9] W. q. Wei, L. Liu, J. Chen, K. Jin, F. Jiang, F. q. Liu, et al, “Racial differences in MLH1 and MSH2 mutation: An analysis of yellow race and white race based on the InSiGHT database,” *Journal of Bioinformatics and Computational Biology*, vol. 8, Suppl. 1, pp. 111-125, 2010.

[10] H. -S. Hsu, I. H. Lee, W. -H. Hsu, W. -T. Kao, Y.-C.Wang, “Polymorphism in the hMSH2 gene (gISV12-6T > C) is a prognostic factor in non-small cell lung cancer”, *Lung Cancer*, vol. 58, pp. 123-130, 2007.

[11] T. Chitnis, B. Glanz, S. Jaffin, and B. Healy, “Demographics of pediatric-onset multiple sclerosis in an MS center population from the Northeastern United States”, *Multiple Sclerosis*, vol. 15, pp. 627-631, 2009.

[12] W. q. Wei, L. Liu, J. Chen, K. Jin, F. Jiang, F. q. Liu, et al, “Racial differences in MLH1 and MSH2 mutation: An analysis of yellow race and white race based on the InSiGHT database,” *Journal of Bioinformatics and Computational Biology*, vol. 8, Suppl. 1, 111-125,

2010.

[13] M. Nyström-Lahti, P. Sistonen, J. P. Meckli, L. Pylkkänen, L. A. Aaltone, H. Järvinen, et al, "Close linkage to chromosome 3p and conservation of ancestral founding haplotype in hereditary nonpolyposis colorectal cancer families", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, pp. 6054-6058, 1994.

[14] A. C. Jager, M. L. Bisgaard, T. Myrhoj, I. Bernstein, J. F. Rehfeld, and F. C. Nielsen, "Reduced frequency of extracolonic cancers in hereditary nonpolyposis colorectal cancer families with monoallelic hMLH1 expression", *American Journal of Human Genetics*, vol. 61, pp. 129-138, 1997.

[15] A. S. Shah, R. Bhopal, S. Gadd, and R. Donohoe, "Out-of-hospital cardiac arrest in South Asian and white populations in London: database evaluation of characteristics and outcome", *Heart*, vol. 96, pp. 27-29, 2010.

[16] A. A. Kabir, G. Pridjian, W. C. Steinmann, E. A. Herrera, and M. M. Khan, "Racial differences in caesareans: an analysis of U.S. 2001", *National Inpatient Sample Data*, Obstet Gynecol, vol. 105, pp. 710-718, 2005.

[17] T. Hienonen, F. C. Nielsen, J. H. Kleibeuker, "Little evidence for involvement of MLH3 in colorectal cancer predisposition," *International Journal of Cancer*, vol. 106, no. 2, pp. 292-296, 2003.

[18] H. Song, L. Klarskov, S. Holck, "Common variants in mismatch repair genes and risk of invasive ovarian cancer," *Carcinogenesis*, vol. 27, no. 11, pp. 2235-2242, 2006.

[19] J. Suchy. R. M. W. Hofstra, J. H. Kleibeuker, "Ovarian cancer of endometrioid type as part of the MSH6 gene mutation phenotype," *Journal of Human Genetics*, vol. 47, no. 10, pp. 529-531, 2002.