Array comparative genomic hybridization based identification of key genetic alterations at 2p21-p16.3 (MSH2, MSH6, EPCAM), 3p23-p14.2 (MLH1), 7p22.1 (PMS2) and 1p34.1-p33 (MUTYH) regions in hereditary non polyposis colorectal cancer (Lynch syndrome) in the Kingdom of Saudi Arabia

Mahmood Rasool, Peter Natesan Pushparaja, Zeenat Mirza, Muhammad Imran Naseer, Heba Abusamra, Maha Alquaitia, Manal Shaabada, Abdulrahman Mohamed Saeed Sibianyd, Kalamegam Gauthaman, Mohammed Hussein Al-Qahtania, Sajjad Karima

1. Introduction

Lynch syndrome (LS) is considered to be responsible for high risk of early onset of colorectal cancer (CRC) and enhanced risk of many extra colonic malignancies including ovarian, endometrial, urinary tract, sebaceous gland, stomach, hepatobiliary, small bowel and central nervous system cancers (Barrow et al., 2013; Ligtenberg et al., 2013). LS account for approximately 1–3% of all CRC and 2–5% of endometrial cancer cases (Hampel et al., 2005). LS is mainly caused by germline pathogenic mutations in DNA mismatch repair (MMR) genes, mostly in four of the genes, MutL Homolog 1 (MLH1) (OMIM # 120436), MutS Homolog 2 (MSH2) (OMIM # 609309), MutS Homolog 6 (MSH6) (OMIM # 600678) and PMS1 Homolog 2 (PMS2) (OMIM # 600259) (Duruatto et al., 2019; Gupta and Heinen, 2019). These conditions represent characteristic features of tumors such as microsatellite instability (MSI) and expression loss of MMR proteins (one or more) in immunohistochemistry (Lynch et al., 2009). MLH1 and MSH2 accounts for majority of the pathogenic mutations (90%), while MSH6 and PMS2 accounts for less than 10% accumulative mutations underlying LS (Tutlewskia et al., 2013). Furthermore, recent studies have included EPCAM gene (OMIM: 185535) as also main
cause of LS and accounts for 1–3% of all LS patients (Kuiper et al., 2011; Cini et al., 2019).

The UK National Institute for health and care excellence (NICE) in their strategy has emphasized to screen routinely all of the patients with newly diagnosed CRC, irrespective of age for these four MMR genes (NICE, 2017). Early screening of MMR gene mutations in the relatives of the patients helps in reducing morbidity and mortality (de Jong et al., 2006). Therefore, it is necessity for health care providers to identify the LS families, screen and recommend preventive measure to decrease the cancer risk as well as the financial burden for lifelong testing and treatments (Vasen et al., 1998).

In general, germ line mutational analysis of MMR genes is done to identify LS patients, however, we attempted going to the next level for identification of copy number variants that likely change the function of the encoded MMR proteins and pathways (Xicola et al., 2019). In the present research, we have performed genome-wide high density array CGH analysis in the LS patients to characterize genomic alterations accountable for these disorders. Our current study shall pave the way further for the accumulation of LS prevalence in Saudi Arabia and its underlying genotypic rearrangements. Once the full mutational spectrum is known, then it will be easy to establish a rapid genetic testing platform and screen the patients and carriers with this syndrome.

2. Materials and methods

2.1. Sample collection and DNA extraction

We ascertained 12 samples of LS (Table 1). Clinical investigation was obtained from patients and clinicians and confirmed the LS as per revised guidelines of Bethesda and Amsterdam criteria (Vasen et al., 2013). Authorization to conduct study was taken from all patients prior to the start of study according to the Helsinki’s declaration. The study was approved from local ethical committee of Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia and conducted here during the period of December 2017 to March 2019. The DNA was extracted from tumor tissue of the samples according to the standard procedure (Qiagen, USA) and quantified (ND-1000).

2.2. Array comparative genomic hybridization profiling

Array CGH was performed on the DNA samples from colorectal tumor tissue of the patients and keeping healthy adjacent tissue as control (sureprint G3 Human CGH 2x400 K arrays kit, Agilent, USA).

2.3. Chip processing and bioinformatics data analysis

Chip scanning was done on Agilent scanner (G2505C), image analysis and data extraction was done by Feature Extraction Software (V.1.5.1.0, Agilent). CGH array profiling was performed by Cytogenomics v2.7 software (Agilent) for visualization, detection and analyzes alterations.

2.4. Quantitative PCR

For the validation of copy number variations/deletions and duplications detected by cGH array experiments, we used Real-time quantitative PCR. The primers were designed for the MLH1, MSH2, MSH6, EPCAM, PMS2 and MUTYH genes and an endogenous gene Actin Beta (ACTB) was taken as an internal control.

3. Results

3.1. Clinicopathological study

We included twelve LS patients under the age of 40 years, seven of them were males and five were females. Histopathological study of LS patients revealed around 65% cases in grade 2 (moderate grade) followed by grade 3 (poor grade) of 30%. Surprisingly, 80% tumors were located on right side of the body. Tumor size were variable ranging from 2 cm to 12 cm (Table 1).

3.2. Array CGH study

We identified more than hundred alterations (gain or loss) with mean log2 ratio cut off value of ±0.25. For single copy number deletion or amplification cut off mean log ratio was 1.0 and 0.58 respectively. We filtered the result to focus on genomic alterations associated with LS and identified four alterations, all were structural losses, encompassing LS genes (Table 2). We detected 2p21-p16.3 (MSH2, MSH6, EPCAM) in eight (Fig. 1), 3p23-p14.2 (MLH1) in five (Fig. 2), 7p22.1 (PMS2) in two (Fig. 3), and 1p34.1-p33 (MUTYH) in one LS patients. (Fig. 4).

3.3. Validation by qPCR

Genomic alterations and genes located on these regions were confirmed by real time quantitative PCR technique by calculating

| Sample No. | Age | Gender | Grade | Tumor location | Tumor size (cm) | Tumor stage | LN status |
|------------|-----|--------|-------|----------------|----------------|-------------|-----------|
| P1         | 34  | M      | 2     | 1              | 7.0            | 3           | 1         |
| P2         | 47  | F      | 3     | 2              | 8.0            | 3           | 1         |
| P3         | 32  | F      | 3     | 2              | 7.5            | 4           | 2         |
| P4         | 39  | M      | 3     | 2              | 8.0            | 3           | 1         |
| P5         | 31  | F      | 2     | 2              | 2.0            | 3           | 1         |
| P6         | 44  | F      | 2     | 2              | 2.0            | 2           | 1         |
| P7         | 43  | M      | 2     | 1              | 9.0            | 3           | 2         |
| P8         | 36  | M      | 2     | 2              | 3.0            | 3           | 2         |
| P9         | 35  | M      | 2     | 2              | 3.5            | 3           | 1         |
| P10        | 43  | F      | 2     | 1              | 12.0           | 3           | 1         |
| P11        | 49  | M      | 3     | 2              | 7.5            | 3           | 2         |
| P12        | 38  | M      | 1     | 1              | 1.5            | 3           | 1         |

* Grades: 1 = well, 2 = moderate, and 3 = poor.
* Tumor location: 1 = Right and 2 = Left.
* LN status: 1 = negative and 2 = positive.
Fig. 1. Whole genome array CGH analysis (2x 400 K) showing structural loss in 2p21-p16.3 region, a locus for MSH2, MSH6 and EPCAM genes. Copy number variation is detected by cut off mean log2 ratio of ±0.25, where red and blue color indicates loss and gain respectively.

Table 2

| Cytoband   | Size (kb) | Probes | Mean Log Ratio | Copy No | Type | p-val | Gene Name                                                                 |
|------------|-----------|--------|----------------|---------|------|-------|---------------------------------------------------------------------------|
| 2p21-p16.3 | 828.608   | 154    | -0.608         | 1       | Loss | 1.37E-20 | EPCAM, MIR559, MSH2, KCNK12, HCG2040054, MSH6, FBX011...                 |
| 3p23-p14.2 | 27363.2   | 4574   | -0.259         | 1       | Loss | 2.76E-216 | STAC, DCLK3, TRANK1, TPM2AIP1, MLH1, LRRFIP2, GOLGA4...                |
| 7p22.1     | 830.844   | 128    | -0.786         | 1       | Loss | 7.65E-14 | RSPH10B2, PMS2, AIMP2, EIF2AK1, ANKRD61, CYTH3, RAC1...                   |
| 1p34.1-p33 | 1610.385  | 309    | -0.596         | 1       | Loss | 1.76E-13 | HECTD3, UROD, ZSWIM5, LINC01144, HPDL, MUTYH, TOE1, TESK2, CCDC163, MMACHC... |

Bold is made to signify the pathological importance of key genes associated with Lynch syndrome.

Fig. 2. Whole genome array-CGH analysis (2x 400 K) showing structural loss in 3p23-p14.2 region, a locus of MLH1 gene. Copy number variation is detected by cut off mean log2 ratio of ±0.25, where red and blue color indicates loss and gain respectively.
mean Rq values of genes; MLH1 (0.725), MSH2 (0.611), MSH6 (0.667), EPCAM (0.510), PMS2 (0.359) and MUTYH (0.593) (Fig. 5). This result confirms that genes causing LS were less expressed as detected in array CGH.

4. Discussion

Disturbance to genomic stability or errors in DNA replication are significant threat to regular cell division process. MMR genes, proteins and pathways correct any unwanted changes and work as tumor suppressor. However, loss of function in MMR leads to sporadic cancer and germline mutation in MMR genes leads to hereditary cancer like LS (Veigl et al., 1998; Li, 2008; Jiricny, 2013; Heinen, 2014; Lynch et al., 2015). MLH1 and MSH2 were identified as most mutated MMR genes associated with LS (Lynch, 2009; Liccardo et al., 2017). Several mutations identified in LS falls under missense or nonsense categories, however other pathogenic variants are small insertions/deletions or large genetic rearrangements (large deletions/insertions) (Duraturo et al., 2013; Liccardo et al., 2018). We conducted array-CGH based study of LS to identify the chromosomal abnormalities in genome and confirm the alteration in MMR genes regions.

Array CGH analysis detected structural loss in the genomic regions of MLH1 (3p23-p14.2), MSH2, MSH6, EPCAM (2p21-p16.3), PMS2 (7p22.1) and MUTYH (1p34.1-p33) in LS patients. MLH1 (homolog to MutL gene of E. coli) is involved in DNA mismatch repair and positioned at 3p22.2. The mutations in this gene were first described by Papadopoulos et al. (1994) to cause LS (Papadopoulos et al., 1994). MSH2 (homolog to the E. coli MutS gene) is located on chromosome 2p21-p16 and is involved in DNA mismatch repair (Fishel et al., 1994). MSH6 is located at 2p16.3 and also involved as DNA MMR gene (Gradia et al., 1997). The EPCAM gene is located on 2p21 and encodes a carcinoma-associated antigen (Munz et al., 2009). PMS2 is positioned at 7p22.1, is also a MMR gene and showed to be involved in hereditary colon cancer (Papadopoulos et al., 1994). MUTYH (Homology of E. coli mutY) is a DNA repair gene located on chromosome 1 at position 34.1 (Slupska et al., 1996).

MLH1, MSH2, PMS1, PMS2, and MSH6 proteins interact with each other, however, defects in protein-protein interactions (PPI) have shown strong link with missense mutations in specific regions of these genes (Yuan et al., 2002; Nakagawa et al., 2004). The loss of function of any one of MMR protein prevents proper function of mismatch repair’s complex and these variants disrupting the crucial PPI may increase risk for tumorigenesis in LS (Vasen et al., 1999; Yuan et al., 2002).

Current study is very useful in Saudi perspective as consanguineous marriages are common in the society and family is generally big. Therefore, the risks for transferring genetic alterations to the next generations in the region is higher as compared to Europe and America, where the consanguineous marriages are lower and family size is generally small. Once the mutation is established in an individual it also allows screening for the same alteration in extended family members. So, the identification of the carriers of the MMR gene mutation is very important because it will allow them to undergo early testing and routine surveillance of LS. Regular colonoscopy screening has helped to reduce the mortality rate from CRC in mutation carriers (Jarvinen et al., 2000).

5. Conclusion

In the present study we did genome-wide profiling of 12 Saudi Lynch syndrome patients with healthy controls. We have reported loss of copy numbers with Lynch syndrome genes at 2p21-p16.3 (MSH2, MSH6, EPCAM), 3p23-p14.2 (MLH1), 7p22.1 (PMS2), and 1p34.1-p33 (MUTYH). We found lower expression of MLH1, MSH2, MSH6, PMS2, EPCAM and MUTYH genes that confirms the loss of copy number in affected regions. Our finding in genome-wide profiling might lead to novel therapeutical intervention in the future and reducing the health care burden in the society.
Fig. 4. Whole genome array CGH (2x 400 K) analysis showing structural loss in 1p34.1-p33 region, a locus for MUTYH gene. Copy number variation is detected by cut off mean log₂ ratio of ±0.25, where red and blue color indicates loss and gain respectively.

Fig. 5. Real time quantitative PCR analysis of MSH2, MSH6, EPCAM, MLH1 and MUTYH genes showing down expression in Lynch Syndrome. Gene expression is measured in term of mean relative quantification (Rq) values cut off ±1, where 1 is normal expression and Rq value more or less than 1 are indicator of up or down expression.
Declaration of Competing Interest

None

Acknowledgement

This project was funded by NSTIP strategic technologies program in the Kingdom of Saudi Arabia – Project No. 12-MED3078-03. The authors acknowledge with thanks Science and technology Unit, King Abdulaziz University for technical support. We also acknowledge the support of “AZIZ” Supercomputing Facility, FCIT, King Abdulaziz University, Jeddah, Saudi Arabia for data storage and analysis. We also thanks to staffs of KAU Hospital and Center of Excellence in Genomic Medicine Research, King Abdulaziz University for clinical and research facilities.

References

(NICE), N. I. F. H. a. C. E. (2017). “Molecular testing strategies for Lynch syndrome in people with colorectal cancer.” 2019, from <https://www.nice.org.uk/guidance/dg27>.

Barrow, P., Khan, M., Lalloo, F., Evans, D.G., Hill, J. 2013. Systematic review of the impact of registration and screening on colorectal cancer incidence and mortality in familial adenomatous polyposis and Lynch syndrome. British J. Surgery 100 (13), 1719–1731.

Cini, G., Quaia, M., Canzonieri, V., Fornasarig, M., Maestro, R., Morabito, A., D’Elia, A, V., Ursu, E.D., Mammi, I., Viel, A. 2019. Toward a better definition of EPcAM deletions in Lynch Syndrome: Report of new variants in Italy and the associated molecular phenotype. Mol. Genet. Genomic Med. 7 (5), e587.

de Jong, A.E., Hendriks, Y.M., Klebecker, J.H., de Boer, S.Y., Cats, A, Griffinon, G., Nagengast, F.M., Nelis, F.G., Rookus, M.A., Vasen, H.F., H. 2006. Decrease in mortality in Lynch syndrome families because of surveillance. Gastroenterology 130 (3), 665–671.

Duraturo, F., Cavallo, A, Liccardo, R., Cudia, B., De Rosa, M., Diana, G., Izzo, P. 2013. Contribution of large genomic rearrangements in Italian Lynch syndrome patients: characterization of a novel alu-mediated deletion. BioMed. Res. Int. 2013, 219897.

Duraturo, F., Liccardo, R., De Rosa, M., Izzo, P., 2019. Genetics, diagnosis and treatment of Lynch syndrome: Old lessons and current challenges. Oncol. Lett. 17 (3), 3048–3054.

Fishel, R., Ewel, A., Lee, S., Lescoe, M., Griffith, J. 1994. Binding of mismatched microsatellite DNA sequences by the human MSH2 protein. Science (New York, N.Y.) 266 (5189), 1403–1405.

Gradia, S., Acharaya, S., Fishel, R. 1997. The human mismatch recognition complex hMHS2-hMHS6 functions as a novel molecular switch. Cell 91 (7), 995–1005.

Gupta, D., Heinen, C.D., 2014. Translating mismatch repair mechanism into cancer care. Cold Spring Harbor Perspect. Biol. 5 (4), a012633.

Kuiper, R.P., Vissers, L.E., Van Koff, A., Bodmer, D., Hoenschl, E., Goossens, M., Hauf, A., Kopp, E., Nielsen, R.C., Hogervorst, F.B., Gillie, J.J., Redeker, B., Tops, C.M., van Gijn, M.E., de van Ouweland, A.M., Rahner, N., Strielt, V., Kahl, P., Holinski-Feder, E., Morak, M., Kloor, M., Stemmler, S., Betz, B., Hutter, P., Bunyan, D.J., Syngal, S., Gulver, J.O., Graham, T., Can, T.L., Nagle, T.E., van Krieken, J.H., Schackert, H.K., Hogerurghue, N., van Kessel, A.C., Ligneten, M.J., 2011. Recurrence and variability of germline EPcAM deletions in Lynch syndrome. Human mutation 32 (4), 407–414.

Li, G.M., 2008. Mechanisms and functions of DNA mismatch repair. Cell Res. 18 (1), 85–98.

Liccardo, R., De Rosa, M., Rossi, G.E., Carlomagno, N., Izzo, P., Duraturo, F. 2017. Incomplete segregation of MSH6 frameshift variants with phenotype of Lynch syndrome. Int. J. Mol. Sci. 18 (5).

Liccardo, R., De Rosa, M., Rossi, G.E., Rigler, G., Izzo, P., Duraturo, F., 2018. Characterization of novel, large duplications in the MSH2 gene of three unrelated Lynch syndrome patients. Cancer Genet. 221, 169–174.

Ligneten, M.J., Kuper, R.P., Geurts van Kessel, A., Hoogerurghue, N., 2013. EPcAM deletion carriers constitute a unique subgroup of Lynch syndrome patients. Familial cancer 12 (2), 169–174.

Lynch, H.T., Lynch, P.M., Lanspa, S.J., Snyder, C.L., Lynch, J.F., Boland, C.R., 2009. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. Clin. Genet. 76 (1), 1–18.

Lynch, H.T., Snyder, C.L., Shaw, T.G., Heinen, C.D., Hitchins, M.P., 2015. Milestones of Lynch syndrome: 1895–2015. Nature Rev. Cancer 15 (3), 181–194.

Lynch, P.M., 2009. The hMSH2 and hMLH1 genes in hereditary nonpolyposis colorectal cancer. Surgical Oncol. Clin. North America 18 (4), 611–624.

Munz, M., Baeuerle, P.A., Gires, O. 2009. The emerging role of EPcAM in cancer and stem cell signaling. Cancer Res. 69 (14), 5627–5629.

Nakagawa, H., Lockman, J.C., Frankel, W.L., Hampel, H., Steenblock, K., Burgart, L.J., Thibodeau, S.N., de la Chapelle, A. 2004. Mismatch repair gene PM2: disease-causing germline mutations are frequent in patients whose tumors stain negative for PM2 protein, but paralogous genes obscure mutation detection and interpretation. Cancer Res. 64 (14), 4721–4727.

Papadopoulos, N., Nicolai, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., et al., 1994. Mutation of a mutL homolog in hereditary colon cancer. Science (New York, N.Y.) 263 (5153), 1625–1629.

Sipuska, M.M., Raikakov, C, Luther, W.M., Chang, J.H., Wei, Y.F., Miller, J.H., 1976. Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutL gene whose function is required for the repair of oxidative DNA damage. J. Bacterial. 178 (13), 3885–3892.

Tutlewiska, K., Lubinski, J., Kurzawski, G. 2013. Germline deletions in the EPcAM gene as a cause of Lynch syndrome - literature review. Hered. Cancer Clin. Practice 11 (1), 9.

Vasen, H.F., Blanco, L., Akan-Can, K., Kopje, J.P., Alonso, A., Aretz, S., Bernstein, L., Bertario, L., Burn, J., Capella, G., Colas, C., Engel, C., Frayling, I.M., Genuardi, M., Heinmann, K., Hes, F.J., Hodgson, S.V., Karyagisami, J.A., Laloo, F., Lindblom, A., Mecklin, J.P., Poller, M., Myrohi, T., Nagengast, F.M., Parc, Y., Ponz de Leon, M., Renkonen-Sinisalo, L., Sampson, J.R., Stormorken, A., Sijmons, R.H., Teijaz, S., Thomas, H.J., Rahner, N., Wijnen, J.T., Jarvinen, H.J., Moslein, G., 2013. Revised guidelines for the clinical management of Lynch Syndrome (HNPCC): recommendations by a group of European experts. Gut 62 (6), 812–823.

Vasen, H.F., van Ballegooijen, M., Buskens, E., Klebecker, J.K., Taal, B.G., Griffinon, G, Nagengast, F.M., Menko, F.H., Meera Khan, P., 1998. A cost-effectiveness analysis of colorectal screening of hereditary nonpolyposis colorectal carcinoma gene carriers. Cancer 82 (9), 1632–1637.

Vasen, H.F., van Ballegooijen, M., Buskens, E., Klebecker, J.K., Taal, B.G., Griffinon, G., Menko, F.H., Meera Khan, P., 1998. A cost-effectiveness analysis of colorectal screening of hereditary nonpolyposis colorectal carcinoma gene carriers. Cancer 82 (9), 1632–1637.

Veigl, M.I., Kastub, I., Olechnowicz, J., Ma, A.H., Lutterbaugh, J.D., Periyasamy, S., Li, G.M., Drummond, J., Modrich, P.L., Sedwick, W.D., Markowitz, S.D., 1998. Biallelic inactivation of IMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. In: Proceedings of the National Academy of Sciences of the United States of America, pp. 8988–8702.

Xicola, R.M., Clark, J.R., Carroll, T., Alvikas, J., Marwaha, P., Regan, M.R., Lopez-Giraldez, F., Conde, J., Emadi, R., Alagiaian-Angelova, V., Kupper, S.S., Ellis, N.A., Ltor, X., 2012. Implication of DNA repair genes in Lynch-like syndrome. Familial cancer.

Yuan, Z.Q., Gottlieb, B., Beitel, L.K., Wong, N., Gordon, P.H., Wang, Q., Puisieux, A., Foulkes, W.D., Trifiro, M., 2002. Polymorphisms and HNPCC: PM2-MLH1 protein interactions diminished by single nucleotide polymorphisms. Human mutation 19 (2), 108–113.