Journal of Proteomics & Bioinformatics - Open Access

Research Article JPB/Vol. S1/Special Issue 2008

Computational Analysis of Mutations in Colon Cancer Genes Reveals a Possible Role of Micro Satellite in Mutagenesis

Allam Appa Rao[1], G R Sridhar[2], Suresh B Mudunuri[1], E Vamsidhar[1], Gunna Kishore[1]*

[1]Department of Computer Sciences and System Engineering, Andhra University, Visakhapatnam, India

[2]Endocrine and Diabetes Centre, Visakhapatnam

*Corresponding author: Gunna Kishore

E-mail: kishore_brbm@yahoo.co.in,

Received April 20, 2008; Accepted May 15, 2008; Published May 25, 2008

Citation: Allam AR, Sridhar GR, Suresh BM, Vamsidhar E, Gunna K (2008) Computational Analysis of Mutations in Colon Cancer Genes Reveals a Possible Role of Micro Satellite in Mutagenesis. J Proteomics Bioinform S1: S041-S045. doi:10.4172/jpb.s1000007

Abstract

Computer science is a subject, which deals with the manipulation of data so that new data, implicit in the original, appear in a useful form. We have used the analogy of genome analysis and VIRUS (vital information recourse under siege) and analyzed MLH1, MSH2 and MSH6 gene which play an important role in repairing mistakes made in DNA replication in colon cancer. If the MLH1, MSH2, MSH6 proteins are mutated and therefore don't work properly, the replication mistakes are not repaired, leading to damaged DNA. The information of all the experimentally proven mutations were collected and analyzed using bioinformatics tools and software programs. We tried to find out whether the presence of or simple sequence repeats in the MLH1, MSH2, MSH6 gene has any significance in the generation of these mutations and checked whether these mutations are fallen in the regions of those microsatellites and if so is there any significance of these microsatellites in the functional domains of the each gene. Our analysis reveled that 3 of the 10 mutations of the MLH1 gene and all the 10 mutations of the MSH2 gene and the one mutation in the MSH6 gene that are existing in the microsatellite regions are fall.en in the domain regions of the respective genes and thus indicating a positive role of microsatellites in mutagenesis.

Keywords: microsatellites; mutations; functional domains

Introduction

Colon cancer is one of the most common inherited cancer syndromes known. Among the genes found to be involved in colorectal cancer are: *MSH2* and *MSH6* both on chromosome 2 and *MLH1*, on chromosome 3 (Lawes et al 2005). Normally, the protein products of these genes help to repair mistakes made in DNA replication. If the MSH2, MSH6, and MLH1 proteins are mutated and therefore don't work properly, the replication mistakes are not repaired, leading to damaged DNA (Päivi Peltomäki 2001).

Cancer occurs when cells become abnormal and divide without control or order. Like all other organs of the body, the colon and rectum are made up of many types of cells. Normally, cells divide to produce more cells only when the body needs them. This orderly process helps keep us healthy.

Apart from genes, the human genome also consists of a large number of nucleotide repeat units of size 1-6 bp repeated tandemly called Microsatellites or Simple Sequence Repeats (SSRs) or Short Tandem Repeats (STRs) (Schlotterer, C. 2000) Microsatellites are found in all the known genomes, spanning from prokaryotes, eukaryotes and viruses and are widely distributed both in coding and non-coding regions (Toth, G et al 2000 ; Sreenu.V.B.et al 2007).

Mutations in these micro satellite regions occur at much higher rate when compared with those in the rest of the genome (Ellegren, H. 2000).

Microsatellites are known to be highly polymorphic due to the high rate of mutations in their tracts (Jarne P. and Lagoda P.J.L. 1996). These mutations can be either in the form of increase / decrease of repeat units or in the form of single nucleotide substitutions/deletions/insertions and other events (Fan, H. and Chu, J.Y. 2007) . Increase or decrease of repeat units of micro satellites in coding regions might lead to shift in reading frames there by causing changes in protein product (Li Y.C. et al 2004) and in non-coding regions are known to effect the gene regulation (Martin. P. et al 2005). Point mutations (Substitutions and Indels) are also found to occur at a higher rate in micro satellites than elsewhere (Sibly.R.M. et al 2003). Micro satellite mutations with in or near certain genes are known to be responsible for some human neurodegenerative diseases. So, we made a brief study to check whether the mutations in this MLH1, MSH2 and MSH6 gene have any relation with these micro satellites repeats and the study revealed interesting results

Methods

All the experimental proved mutations of the genes MLH1,MSH2 and MSH6 that are falling inside the coding region and are eventually leading to phenotypic differences were collected from the Human Gene Mutation Database (HGMD) (Stenson, P.D. et al 2003). Micro satellites are obtained from the Imperfect Microsatellite Extractor (IMEx) (Mudunuri and Nagarajaram 2007) using intermediate mode with default values 10 for single 5 for di 3 tri 3 for tetra 2 for penta and 2 for hexa and obtained 14,17,24 micro satellites in MLH1,MSH2,MSH6 respectively. Since microsatellites are drawn from the nucleotide sequence and HGMD mutations are given for protein sequence we have used DNA to Amino Acid translator. We compared the regions with the mutations whether they have mutations in those regions and found some of the s have occurred in those regions. Now we analyzed whether these mutations and microsatellites have fallen in the functional domains of those genes by using Simple Modular Architecture Research Tool (SMART) (Letunic I et al 2004)and the results are as follows.

Confidently predicted domains, repeats, motifs and features: from the smart results we obtained the following domains for the MLH1 gene

NAME	BEGIN			END
Hatpase_c	23	1	5	8
Pfam:mis__dna_repair	221	3	3	5
Low complexity	362	3	7	5
Low complexity	475	486		

The codon changes (TCC-TTC) and (AGT-ATT) are fallen in the HATPase_c domain and the codon change (GAG-GGG) Is fallen in the Pfam rgion which is the region where the DNA repair mechanism takes place.

MLH1 GENE (change and phenotype)

Codon number	Codon change	Amino acid Change	Disease phenotype	References
44	TCC-TTC	Ser-Phe	Colorectal cancer,non-polyposis	Bronner CE et al 1994
46	AGT-ATT	Ser-Ile	Colorectal cancer,non-polyposis	Cai Q et al 2003
234	GAG-GGG	Glu-Gly	Colorectal cancer,non-polyposis	Kim JC et al 2001
379	TAT-TGT	Tyr-Cys	Colorectal cancer,non-polyposis	Taylor CF et al 2003
426	gCAG-TAG	Gln-Term	Colorectal cancer,non-polyposis	Bisgaard ML et al 2002
607	CTT-CAT	Leu-His	Colorectal cancer,non-polyposis	Fidalgo P et al 2000
618	AAG-ACG	Lys-Thr	Colorectal cancer,non-polyposis	Han HJ et al 1995
618	gAAG-TAG	Lys-Term	Colorectal cancer,non-polyposis	Hutter P et al 1998
622	CTT-CAT	Leu-His	Colorectal cancer,non-polyposis	Godino J, et al 2001
631	GAT-GCT	Asp-Ala	Colorectal cancer,non-polyposis	Kim et al 2001

Confidently predicted domains, repeats, motifs and features: from the smart results we obtained the following domains for the MSH2 gene.

All the codon changes of the MSH2 gene are fallen in one of the domain as indicated above the first two Codons 44 and 45 are fallen in the Pfam: MutS-I and next six codons have fallen in the MUTsd domain and the last two Codons are fallen in the MUTSac domain.

Confidently predicted domains, repeats, motifs and features: from the smart results we obtained the following do-

Name	Begin	End
Pfam: MutS_I	17	132
Pfam: MutS_II	143	290
MUTsd	321	645
MUTsac	662	849

mains for the MSH2 gene

MSH6 GENE (change and phenotype)

Codon number	Codon change	Amino acid Change	Disease phenotype	References
619	GAAg-Gac	Glu-Asp	Colorectal cancer	Plaschke J et al 2004

Name	Begin	End
Pfam: MutS_I	407	526
Pfam: MutS_II	537	704
MUTsd	753	1102
MUTsac	1127	1321

The only one change in the Codon of the MSH6 is fallen in the domain Pfam: MUTS-II.

Results and Discussion

The form of genomic instability associated with defective DNA mismatch repair in tumors is to be called instability (MSI)(Richard Boland et al 1998) and mutations in the mismatch repair (MMR) genes hMLH1 and hMSH2 can cause hereditary non-polyposis colorectal cancer(Brieger A et al 2002). s are DNA elements composed of short tandem repeats of 1–5 bp. These sequences are particularly prone to frameshift and mis sense mutations by insertion–deletion loop formation during replication. The mismatch repair system is responsible for correcting these replication errors, and mutation rates are significantly elevated in the absence of mismatch repair. (Hans Ellegren 2002) and Due to these mutations during PCR, stutter patterns may appear in the final PCR product, which hinder us from accurate genotyping (genitical information)(Yinglei Lai a and Fengzhu Sun 2004) so keeping the above things in mind we analyzed and found that Out of the ten mutations which are fallen in the regions of the Microsatillites three of them having codon numbers 44,46 and 234 have fallen in the regions of the functional domains of the MLH1 gene and for the MSH2 gene the 12 mutations which have fallen in the regions of the microsatillites are all have fallen in the functional domains of the MLH2 gene and similarly for the MSH6 the single mutation which is fallen in the region of the microsatillites is also fallen in the functional domain of the MSH6 gene. since the functional domains are the main regions responsible fot the function of that gene and any mutations in these regions may cause change in the functionality of the gene.

Conclusion

Microsatillites are known for their higher rate of mutations and are known to be associated with various diseases. So, we analyzed the MLH1, MSH2 and MSH6 gene mutations and their possible association with the micro satellites. These mutations from HGMD database are mapped on to the micro satellite tracts and the results seem to indicate that micro satellites play an important role in mutagenesis and by mapping the same with the functional domains we can say that these can cause functionality changes of those genes. Extending this work on a large scale by analyzing large number of genes might give a better evidence of the role of micro satellites in generating mutations.

Acknowledgement

The authors are thankful for partial financial support from IIT upgradation gratns of AUCE (A).

References

1. Bisgaard ML, Jäger AC, Myrhoj T, Bernstein I, Nielsen FC (2002) Hereditary non-polyposis colorectal cancer (HNPCC): phenotype-genotype correlation between patients with and without identified mutation. Hum Mutat 20: 20-7. » CrossRef » Pubmed » Google Scholar

2. Brieger A, Trojan J, Raedle J, Plotz G, Zeuzem S (2002) Transient mismatch repair gene transfection for functional analysis of genetic hMLH1 and hMSH2 variants. Gut 51: 677-84. » CrossRef » Pubmed » Google Scholar

3. Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, et al. (1994) Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 368: 258-61. » CrossRef » Pubmed » Google Scholar

4. Cai Q, Sun MH, Fu G, Ding CW, Mo SJ, etal. (2003) [Mutation analysis of hMSH2 and hMLH1 genes in Chinese hereditary nonpolyposis colorectal cancer families] Zhonghua Bing Li Xue Za Zhi 32: 323-8. » Pubmed » Google Scholar

5. Ellegren H (2000) Heterogeneous mutation processes in human DNA sequences. Nat Genet 24: 400-402. » CrossRef » Pubmed » Google Scholar

6. Fan H, Chu JY (2007) A brief review of short tandem repeat mutation. Genomics Proteomics Bioinformatics 5: 7-14. » CrossRef » Pubmed » Google Scholar

7. Fidalgo P, Almeida MR, West S, Gaspar C, Maia L, etal. (2000) Detection of mutations in mismatch repair genes in Portuguese families with hereditary non-polyposis colorectal cancer (HNPCC) by a multi-method approach. Eur J Hum Genet 8: 49-53.» CrossRef » Pubmed » Google Scholar

8. Godino J, de La HM, Diaz RE, Benito M, Caldés T (2001) Eight novel germline MLH1 and MSH2 mutations in hereditary non-polyposis colorectal cancer families from Spain Hum Mutat 18: 549.» CrossRef » Pubmed » Google Scholar

9. Han HJ, Maruyama M, Baba S, Park JG, Nakamura Y (1995) Genomic structure of human mismatch repair gene, hMLH1, and its mutation analysis in patients with hereditary non-polyposis colorectal cancer (HNPCC) Hum Mol Genet 4: 237-42.» CrossRef » Pubmed » Google Scholar

10. Hans (2002) Ellegren Mismatch repair and mutational bias in DNA Trends in Genetics Volume 18 Issue 11 1 November Page 552. » CrossRef » Pubmed » Google Scholar

11. Hutter P, Couturier A, Membrez V, Joris F, Sappino AP, etal. (1998) Excess of hMLH1 germline mutations in Swiss families with hereditary non-polyposis colorectal cancer. Int J Cancer 78: 680-4. » CrossRef » Pubmed » Google Scholar

12. Jarne P, Lagoda PJL (1996) s, from molecules to populations and back. Trends Ecol Evol 11: 424-429. » CrossRef » Google Scholar

13. Kim JC, Kim HC, Roh SA, Koo KH, Lee DH, etal. (2001) hMLH1 and hMSH2 mutations in families with familial clustering of gastric cancer and hereditary nonpolyposis colorectal cancer. Cancer Detect Prev 25: 503-10. » Pubmed » Google Scholar

14. Lawes DA, Pearson T, Sengupta S, Boulos PB (2005) The role of MLH1, MSH2 and MSH6 in the development of multiple colorectal cancers.93: 472-7. » CrossRef » Pubmed » Google Scholar

15. Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, et al. (1993) Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 75: 1215-25. » CrossRef » Pubmed » Google Scholar

16. Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, etal. (2004) SMART 4.0: towards genomic data integration. Nucleic Acids Res 32: D142-4. » CrossRef » Pubmed » Google Scholar

17. Li YC, Korol AB, Fahima T, Nevo E (2004) s within genes: structure, function, and evolution. Mol Biol Evol 21: 991-1007. » CrossRef » Pubmed » Google Scholar

18. Lu SL, Akiyama Y, Nagasaki H, Nomizu T, Ikeda E, etal. (1996) Loss or somatic mutations of hMSH2 occur in hereditary nonpolyposis colorectal cancers with hMSH2 germline mutations. Jpn J Cancer Res 87: 279-87.» Pubmed » Google Scholar

19. Mangold E, Pagenstecher C, Friedl W, Mathiak M, Buettner R, etal. (2005) Spectrum and frequencies of mutations in MSH2 and MLH1 identified in 1,721 German families suspected of hereditary nonpolyposis colorectal cancer. Int J Cancer 116: 692-702. » CrossRef » Pubmed » Google Scholar

20. Martin P, Makepeace K, Hill SA, Hood DW, Moxon ER (2005) instability regulates transcription factor binding and gene expression. PNAS 102: 3800-3804. » CrossRef » Pubmed » Google Scholar

21. Mudunuri SB, Nagarajaram HA (2007) IMEx: Imperfect Extractor. Bioinformatics 23: 1181-1187.» CrossRef » Pubmed » Google Scholar

22. Nomura S, Sugano K, Kashiwabara H, Taniguchi T, Fukayama N, etal.(2000) Enhanced detection of deleterious and other germline mutations of hMSH2 and hMLH1 in Japanese hereditary nonpolyposis colorectal cancer kindreds. Biochem Biophys Res Commun. » CrossRef » Pubmed » Google Scholar

23. Ollila S, Sarantaus L, Kariola R, Chan P, Hampel H, etal. (2006) Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. Gastroenterology. 131: 1408-17. » CrossRef » Pubmed » Google Scholar

24. Otway R, Tetlow N, Hornby J, Doe WF, Kohonen CMR (2005) Gene symbol: MSH2. Disease: Hereditary

nonpolyposis colorectal cancer. Hum Genet 116: 538. » Pubmed » Google Scholar

25. Päivi P (2001) Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum Mol Genet 10: 735-40. » CrossRef » Pubmed » Google Scholar

26. Plaschke J, Engel C, Kriiger S, Holinski FE, Pagenstecher C, etal. (2004) Lower incidence of colorectal cancer and later age of disease onset in 27 families with pathogenic MSH6 germline mutations compared with families with MLH1 or MSH2 mutations: the German Hereditary Nonpolyposis Colorectal Cancer Consortium J Clin Oncol 22: 4486-94. » CrossRef » Pubmed » Google Scholar

27. Richard B, 1 Stephen NT, Stanley R, Hamilton, David S, etal. (1998) National Cancer Institute Workshop on Instability for Cancer Detection and Familial Predisposition: Development of International Criteria for the Determination of Instability in Colorectal Cancer CANCER RESEARCH 58: 5248-5257. » CrossRef » Pubmed » Google Scholar

28. Schlotterer C (2000) Evolutionary dynamics of DNA. Chromosoma 109: 365-371. » CrossRef » Pubmed » Google Scholar

29. Sibly RM, Meade A, Boxall N, Wilkinson MJ, Corne DW, etal. (2003) The structure of interrupted human AC s. Mol Biol Evol 20: 453-9. » CrossRef » Pubmed » Google Scholar

30. Sreenu VB, Kumar P, Nagarajaram HA (2007) Simple sequence repeats in mycobacterium genomes. J Biosci 32: 3-15. » CrossRef » Pubmed » Google Scholar

31. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, etal. (2003) The Human Gene Mutation Database (HGMD®): 2003 Update. Hum Mutat 21: 577-581. » CrossRef » Pubmed » Google Scholar

32. Taylor CF, Charlton RS, Burn J, Sheridan E, Taylor GR (2003) Genomic deletions in MSH2 or MLH1 are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA. Hum Mutat 22: 428-33. » CrossRef » Pubmed » Google Scholar

33. Toth G, Gaspari Z, Jurka J (2000) s in different eukaryotic genomes: survey and analysis. Genome Res 10: 967-981. » CrossRef » Pubmed » Google Scholar

34. Yamada K, Zhong X, Kanazawa S, Koike J, Tsujita K, etal. (2003) Oncogenic pathway of sporadic colorectal cancer with novel germline missense mutations in the hMSH2 gene.Oncol Rep 10: 859-66. » Pubmed » Google Scholar

35. Yinglei L, Fengzhu S (2004) Journal of Theoretical Biology Volume 228 Issue 2 21 May Pages 185-194.