ELUCIDATING THE PATTERNS OF EXPRESSION OF GENES RELEVANT FOR THE REPAIR OF DAMAGED DNA IN COLON CANCER TISSUES

[1]Odufuwa, Kuburat Temitope; [1]*Faponle, Abayomi Samson; Adelegan, Ayodeji Adebayo [1]; Osonuga, Oduyemi Ifabunmi[2];Atunnise, Adeleke Kazeem[3] and Salau, Bamidele Adewale[3]

[1]Department of Biochemistry, Faculty of Basic Medical Sciences, Sagamu Campus, Olabisi Onabanjo University, Ago-Iwoye, Ogun state, Nigeria.

[2]Department of Physiology, Faculty of Basic Medical Sciences, Sagamu Campus, Olabisi Onabanjo University, Ago-Iwoye, Ogun state, Nigeria.

[3]Department of Biochemistry, Faculty of Basic Medical Sciences, Redeemer's University, Ede, Osun State, Nigeria.

*Corresponding author's e-mail: abayomi.faponle@oouago-iwoye.edu.ng

ABSTRACT

Cancer is a disease that occurs as consequences of deranged functions of cells and genes. DNA repair genes are vitals for keeping intact the genome's integrity. Thus, a flop or partial failure in regulating these important repair genes is linked to carcinogenesis, progression, and resistance to therapeutic drugs against various types of cancers. Colorectal cancer (synonym: colorectal adenocarcinoma), a malignant neoplasm, has caused more of the many deaths that are of cancer origins. Gene profiling of DNA repair genes in colorectal cancers are considered to be vital for their identification at any stage of development and also in their chemotherapy. In an effort to work towards this importance, we have used the qRT-PCR techniques and comparative threshold cycle (Ct) method to quantify and then compare the levels of gene expressions among the DNA repair genes that include three *Neil* genes*, Ercc1, Mlh1, Ogg1*, and *Nthl1* relative to the reference gene, *Gapdh* used as internal control in tumour and matched normal colon tissues. The results revealed that the expression profile of the genes in the colon tumours do not follow a clear pattern but there were relatively high expressions of *Neil3*, *Ercc1*, and *Mlh1*. Since the turnovers of these genes suggestively indicate some levels of genetic instability, therefore, they have the potential to be recognized as useful biomarkers in cancer diagnosis.

Keywords: DNA repair, Gene expression, colorectal cancer, Neil3, Ercc1, Mlh1

Accepted Date: 2*0 Nov., 2020*

Introduction

Cancer is a disease that occurs as consequences of deranged functions of cells and genes(Jenkins, 2017). Colorectal cancer (CRC) is now one the foremost cancerous diseases associated with high death rate worldwide. Cancers occur by uncontrolled cell division which is a consequence of several (or accumulated) somatic and germline alterations in genetic components. Many of these alterations are, infact, mutations in DNA. These may occur because of random DNA replication errors, exposure to cancer-causing agents or deranged DNA repair processes (Abbotts *et al.,* 2014; Wolf *et al.,* 2019). Several mechanisms are deployed by cancer cells to make themselves become completely haywire - through series of regulatory instructions relayed by activated-receptorsthat act as oncogenes (Vogelstein and Kinzler, 2004)- leading to rapidly replicating DNA and, by consequence, uncontrolled cell proliferations. In addition, they up-regulate growth factor ligands, which by certain interactions, stimulate non-cancer cells to release factors that help cancer cells to grow; and receptor-proteins that confer hypersensitivity of non-cancer cells to such growth factor ligands; and as well alter the structure of receptors that helps to bring about cell divisions without dependency on a ligand (Cheng *et al.,* 2008).

Ultimately, cancers result when the integrity of the DNA is compromised. The DNA witnessed numerous alterations on a daily basis. Although some of these changes are spontaneous which include depurination and deamination, many of the DNA damages are due to endogenoeuos and exogenous chemical agents that are genotoxic. Such DNA damages can be DNA base alteration, formation of bulky substituents on bases, single/double-strand breaks, DNA crosslinkings (both intra/inter-strand).Such alterations impact DNA which may lead to permanent changes in genomic DNA sequences, and cell deaths resulting from changes in the replication fork (Friedberg *et al.,* 2005). The ultraviolet (UV) radiation, chemical agents, and chemotherapy drugs are the most common exogenous agents known to damage DNA which may results into cancerous diseases e.g. UV radiation causes skin lesions (Seebode *et al.,* 2016)and lung cancer could most possibly result from numerous chemical compounds in cigarette smoke when inhaled (Doll and Brafdford-Hill, 1950). Also, the colon epithelial cells are particularly affected by mutagenic metabolites of exogenously ingested compounds (Greenman *et al.,* 2007)which may predispose such colon cells to abnormally proliferative growth.

On the other hand, endogenous agents that cause DNA damage also come from various endogenous sources which include pool of reactive oxygen species (ROS)(includes O_2^-, H_2O_2, and OH) (Beckman and Ames, 1997; Hazra *et al.,*2007), by-products of normal aerobic metabolic processes, and some, derivatives of inflammatory cytokines that both build up a pool of oxidants leading to oxidative stress (Federico *etal.,* 2007). Although, the replicative DNA polymerases have 3'-5' "proofreading" exonuclease activity, they are still endogenous sources of DNA damage due to inefficient corrections(Pierce, 2017). The ROS are implicated in maintaining cellular homeostasis(Hancock *et al.,* 2001)but elevated ROS levels are detrimental to the cellular lipids, protein, and nucleic acids (Friedberg *et al.,* 2005), thus, annulling the antioxidative capacity of the cells. These increased ROS levels have direct consequences, through covalent modifications of DNA, ensuring an increased DNA damage. Interestingly, increased ROS levels were observed in pancreatic, prostate, breast, and colon cancers(Vaquero *et al.,* 2004; Kumar *et al.,* 2008; Acharya *et al.,* 2010; Hecht *et al.,* 2016) For instance, 8-oxo-7,8-dihydroguanine (8-oxoG) is caused by ROS attack on DNA. 8-oxoG mismatches with adenine and tranversions of GC to TA results in the DNA. 8-oxoG is non-toxic but it is a highly potent mutagen (Suzuki and Kamiya, 2016). Such mismatch pairing in somatic mutations in lungs, colorectal, breast, gastric, and ovarian cancers have been reported(Fortini *et al.,* 2003). In addition, ROS have been implicated in releases of activating transcription factors e.g. NF-kB, activator protein-1 (AP-1), and hypoxia inducible factor-1 (HIF-1α) that play crucial roles not only in cancer cell growth and survival but also in their angiogenesis, invasion, and metastasis (Gupta *et al.,* 2012).

Interestingly, nature has devised mechanisms to repair any damage of any sorts that affects the integrity of DNA which ensure that the information carried by them is not lost but preserved. There are numerous ways the cells repair insulted DNA but five major DNA repair pathways have been identified (Friedberg, *et al.,* 2005). These are the mismatch repair (MMR), the base-excision repair (BER), the nucleotide-excision repair (NER), the non-homologous end-joining, and the homologous recombination repair (HRR). The efficiency of these repair pathways must remain high as any decrease or loss in them leads to accumulation of thousands of mutations (Abbotts *et al.,* 2014)that portends high risk of carcinogenesis. For instance, hypersensitivity to sunlight, due to absence of NER, makes individuals have high chances of either ageing rapidly or developing cancer(Friedberg, *et al.,* 2005); genetic defects in MMR and BER are underlying causes in several inherited colorectal cancer (Weren *et al.,* 2015).

Cancer cells highly express DNA repair proteins which can render the cells resistant to some chemotherapy agents. For example, colon cancer cells treated with oxaliplatin exhibit upregulation of the excision repair cross-complementation group 1 (*Ercc1*) gene but small interfering RNA (siRNA) knockdown of *Ercc1* makes the cells sensitive to the chemotherapeutic effects of the drug(Seetharam *et al.,* 2010). Also, most cancer cell lines and metastatic melanoma highly express NIEL3 (a BER DNA glycosylase) (Kauffmann *et al.,* 2008; Hildrestrand *et al.,* 2009).

In human cells, the genes *Msh2* and *Msh6* (MutS homolog 2 and 6), *Mlh1* and *Mlh3* (MutL homolog 1 and 3), and *Pms1* and *Pms2*(post-meiotic segregation increased 1 and 2) play vital parts in MMR (Harfe and Jinks-Robertson, 2000). Hence, mutations in any of them could cause microsatellite instability (MSI) and such individual could develop cancers such as CRC and cancers of the ovary(Abbotts *et al.,* 2014). Certain mutations have been implicated in the development of Lynch syndrome (hereditary nonpolyposis colorectal cancers, HNPCC), which is associated with the early stage of CRC. Such alterations occur in *Mlh1*, *Msh2* or *Pms2*. More so, alterations in Epcam (epithelial cell adhesion molecule) gene cause allele-specific *Msh2* inactivation. Notably, a high percentage of HNPCC would be associated with a number of inactivating alterations occurring in any of the MMR genes. More than 90% of these mutations occur in the human *Msh2* or *Mlh1* genes which show a significant high occurrence of phenotypic MSI(Wheeler *et al.,* 2000).Also, somatic mutations occur in MMR genes due to methylation of the promoter region of *Mlh1*. The methylations often result in CpG island methylator phenotype (CIMP) which is a subset of CRC. The hyper-methylation of promoter CpG island sites causes *Mlh1* inactivation especially in CRC with MSI-H (Haydon and Jass, 2002). In an effort to reverse *Mlh1* methylation in colon cancer, de-methylating agent 5-aza-2' -deoxycytidine (5-aza-dC) was used to induce *Mlh1* expression and also render cancer cells sensitive to 5-fluorouracil treatment (Fujita *et al.,* 2007).

The endonuclease VIII – like DNA glycosylases (NEIL) is one the important BER enzymes superfamilies. The 8-oxoguanine DNA glycosylase (Ogg1), endonuclease III homolog (Nthl1), and the three endonuclease VIII paralogs (NEIL1, NEIL2 and NEIL3) are second of their two classes and are bifunctional having apurinic/apyrimidinic (AP) lyase activity in addition to cleaving the glycosylic bond that links damaged base (Jacobs and Schär, 2012). These five bifunctional enzymes identifies and remove oxidised bases from any double/single-strand DNA (Jacobs and Schär, 2012). The two recessive oncogenes *MUTYH* and *Nthl1* (BERgenes), are, indeed, linked to high levels of polyposis and adenomatous polyposis, respectively, and have high potentials for developing CRC(Weren *et al.,* 2018).

As described *vide supra*, the interplay of the highly expressed gene products of DNA repair genes in cancers and resistance to chemotherapy agents cannot be overemphasized. As such, in this work, we determined the levels of *Neil*genes, *Ercc1*,*Ogg1*, and*Nthl1*that expressed in human CRC tumour samples and also in matched normal colon tissue. In addition, we analysed the expression profile of *Mlh1* in solid colon tumor tissues and ascertained its expression at the transcriptional level.

Materials and Methods

Total RNA was extracted from frozen human colon cancer tissue and matched normal colon tissue using the RNeasy Fibrous Tissue Mini kit (QIAGEN). The integrity of the extracted RNA was assessed by using 1% agarose gel electrophoresis. Complementary DNA (cDNA) was synthesized from the extracted RNA by reverse-transcription using QuantiNova Reverse Transcription kit (QIAGEN). The cDNA was amplified by the target gene primers (Table S1 contains the forward and reverse primers of the target genes) using the polymerase chain reaction (PCR).

The resulting cDNA sequences of the target genes (*Neil1, Neil2, Neil3, Ercc1, Mlh1, Ogg1, Nthl1 and Gapdh*) were then amplified by quantitative reverse-transcription (RT) PCR. Quantitative real-time PCR was employed toquantitatively expressthe target genes on the cDNAs using the SensivFast SYBR Lo-ROX Kit (Bioline, U.K.) with glyceraldehyde 3-phosphate dehydrogenase (*Gapdh)* as the reference gene for normalization. The results were subsequently analysed using the MJ Opticon Monitor v3.1 software and 2(-Delta Delta C(T)) method (Livak and Schmittgen, 2001), for calculating the relative expressions of target genes in cell lines through normalization of values of mRNA quantitatively expressed to those of the *Gapdh*, was used in the analysis of results.

Results

Preliminary Analysis of Test Cell Lines

Two cell lines used for the method testing include the HCT116 (human colorectal carcinoma) and the Mero25 (mesothelioma-derived). The cells gave 28S and 18S rRNA bands on electrophoresis agarose gel by the RNA kit (Bioline Isolate II) employed. Subsequently, the PCR primer pairs for each target gene gave distinct bands (Supplementary Information Figure S1) when tested on cDNA obtained from the two rRNA. Except the primer pairs 994S for *Ercc1*, 349S for *Nthl1*, and 1814S for *Mlh1* which gave no PCR products from their cDNA, all the others amplified the target segment of their unique cDNA obtained from either cell line. Similarly, the three *Neil* genes were expressed successfully with the quantitative method, although, a marked difference was seen in the expression levels in the two cell lines. However, *Gapdh* and *Neil3* showed similar expression patterns in both cell lines (Supplementary Information Figure S2).

Lanes 1 – 6, HCT116 and lanes 7-13, Mero-25. Lane 1, *Gapdh*; lane 2, *Neil1* (1079S); lane 3, *Neil1* (1071S); lane 4, *Neil2* (631S); lane 5, *Neil2* (1291S); lane 6, *Neil3* (1651S); lane7, *Gapdh*; lane 8, *Neil1* (1079S); lane 9, *Neil1* (1071S); lane 10, *Neil2* (631S); lane 11, *Neil2* (1291S); lane 12, *Neil3* (1651S); lane 13, *Neil2* (631S). H: Hyperladder 100 bp.

Colon Tissue Samples Analysis

The method employed was applied to the experimental CRC tumour and matched normal colon samples following its confirmations of specificity and suitability of the PCR primers in the preliminary tests. The colon tissue and matched normal colon tissue pairs have been designated (as shown in Table 2) in this work. Their RNA was degraded, thus, the two 18S AND 28S rRNA bands, which is a common feature of total RNA from cells (Sambrook *et al.,*1989; Sandrine *et al.,* 2005), were rarely observed in the colon tissues.

Table 1: *Designation of colon tissues and matched normal colon tissues reported in this paper*

Samples	Colon Tissues	Matched Normal Colon Tissue	Label
1	2T	1N	Pair1
2	4T	3N	Pair2
3	6T	5N	Pair3
4	8T	7N	Pair4
5	10T	10N	Pair5
6	11T	11N	Pair6
7	12T	12N	Pair7
8	13T	13N	Pair8
9	33T	33N	Pair9
10	34T	34N	Pair10
11	35T	35N	Pair11
12	36T	36N	Pair12

It is noteworthy to state that the sample pools were extended to as much as sixteen colon tissues, in part, to further test the stability of the rRNA. Attempts to reduce RNA degradation were unsuccessful. As such, we will focus on the initial twelve pairs in this paper Figure 3

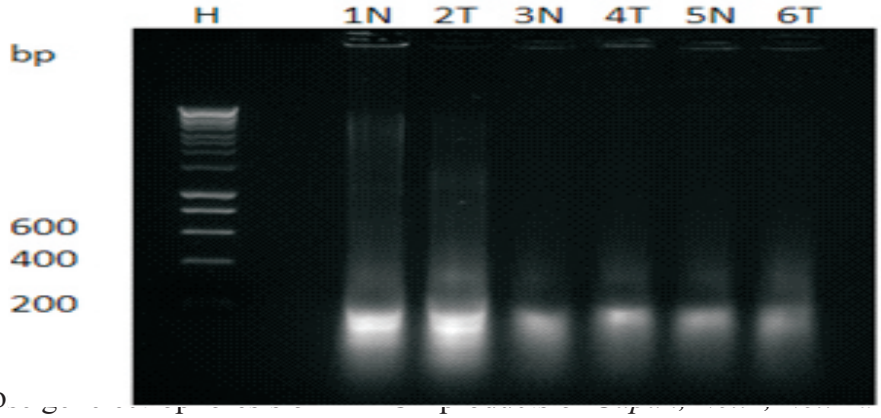

Figure 2: Agaro d *Neil3* for HCT116 and Mero-25 cells.

Figure 3: Agarose gel profile of RNA of tumour vs matched normal of colon tissues.

The Specificity of Primers

From RT-PCR on the cDNA from tissue samples, the specificity of each primer was confirmed and only the target genes were amplified. The gel profiles of each gene were determined. Figure 4 shows the gel pattern of a representative pair i.e. Pair 5

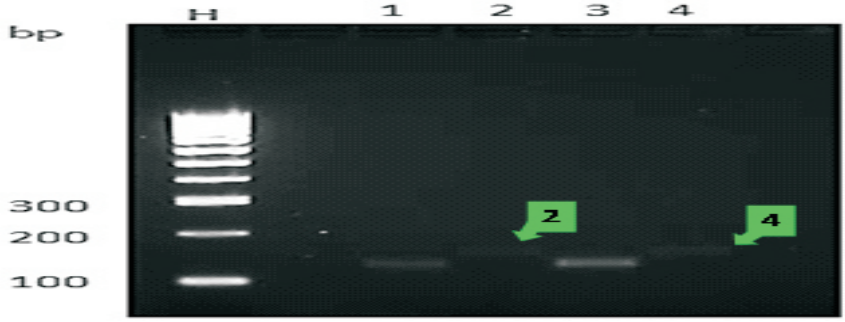

Figure 4: Agarose gel profile ofRT-PCR products of*Gapdh* and *Neil3* from 10N and 10T.

H- Hyperladder 100bp, (10N: 1 - *Gapdh*, 2 - *Neil3*), (10T: 3 - *Gapdh*, 4 - *Neil3*).

Expressions of DNA Repair Genes by Quantitative PCR

The results of the quantitative PCR (qPCR) followed the normalization procedures which took care of the experimental errors that may have arisen from the procedures of RNA extraction and the synthesis of complementary DNA, cDNA. The results were well calibrated with control experiments which involved the reference gene *Gapdh* and the genes *Neil3*, *Ercc1*, and *Mlh1* from samples 10T, 13N, and 34T, respectively. Then qPCR was carried out for real-time quantitative expressions of genes in CRC tumour, matched normal colon tissues. The expression profiles of 7 DNA repair genes that include *Neil* genes (*Neil1*, *Neil2*, *Neil3)*, *Ogg1*, *Nthl1*, *Ercc1* and *Mlh1*, in the sample pairs (Table 1) in comparison to the housekeeping gene, *Gapdh*, are reported in Figure 5. Among them, 3 target genes (*Neil3*, *Ercc1* and *Mlh1*) (Wheeler *et al.,* 2000; Haydon and Jass, 2002; Kauffmann *et al.,* 2008; Hildrestrand *et al.,* 2009;Seetharam *et al.,* 2010; Abbotts *et al.,* 2014), that are directly related or relevant to CRC via 3 repair pathways (i.e. BER, NER and MMR), are reported in different charts i.e. Figure 6 and can be easily compared.

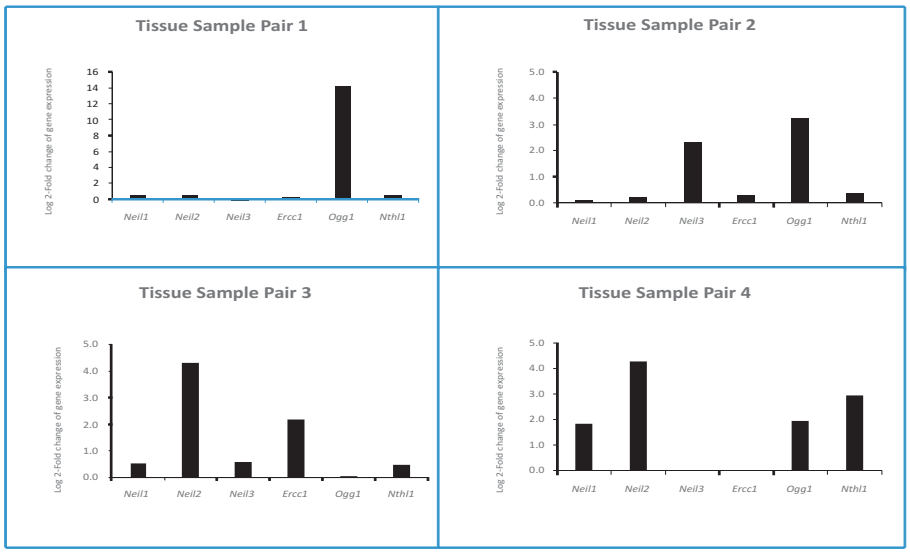

Figure 5: Levels of expressions of *Neil* genes, *Ercc1, Ogg1*, *Nthl1*in tumour/matched normal of colon tissues pairs (Pairs 1 to 4) from qPCR.

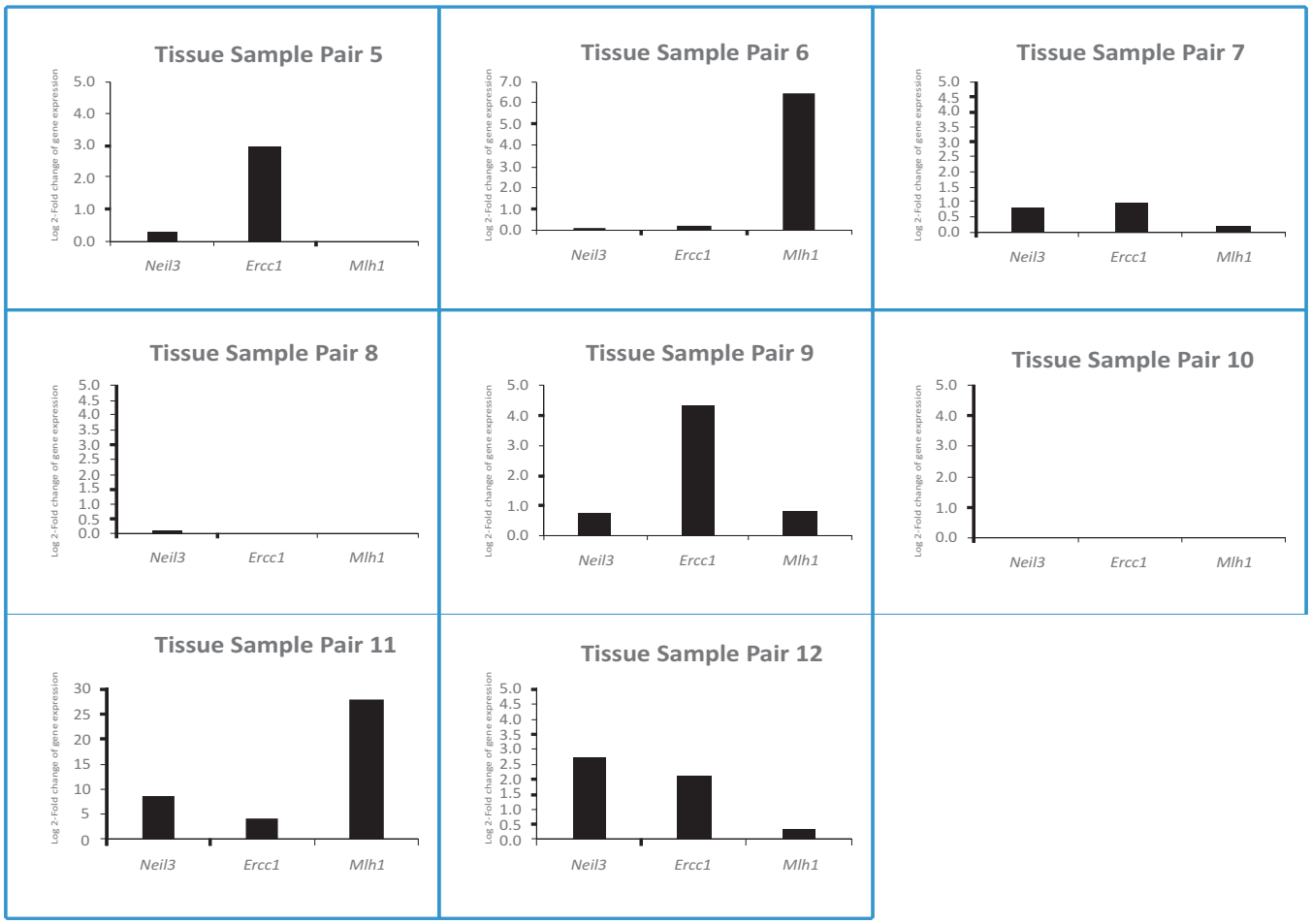

Figure 6: Gene expression levels of *Neil3*, *Ercc1* and *Mlh1*in tumour/matched normal of colon tissues pairs (Pairs 5 to 12) from qPCR

Discussion

The outset of carcinogenesis is somewhat characterized by high level of genetic instability. Such instability could be a result of accumulation of many unresolved DNA damages. Invariably, the cells try to bring most pre-oncogenic molecular events under control through various mechanisms. Certainly, the rate of cellular expressions of DNA repair genes is one the consequences of the cells failure to forestall the damages to DNA. Moreover, despite the reparative role of DNA repair proteins, it is surprising that some of the genes coding for them were associated with tumor development and drug resistance in certain cancer types. Their involvements are characterized by high level of cellular expressions e.g. *Ercc1* was highly expressed in colorectal cancer patients who were being administered combined chemotherapy of oxaliplatin and 5-fluorouracil (Shirota *et al.,* 2001), and *Neil3* highly expresses in melanoma tumor and many cell lines (Taylor *et al.,* 2015; Shinmura *et al.,* 2016).

The genetic expression profiling of a wide range of DNA repair genes that are closely associated with the major DNA repair pathways have been carried out on colon cancer tissues in particular in this study, most importantly, to elucidate or determine if these genes could be targeted as biomarkers in CRC patients based on the patterns of expected outcomes. Gene analyses were done on the various colon tumour tissues, including matched normal colon tissue samples for comparison; twelve different samples in the first instance, whose results are reported here. Although, colon tumours tissues from other different sources were analyzed the same way but the expression patterns were not dramatically different from the initial results. It is

pertinent to note that the total RNA degraded in samples of both sources as the two 18S AND 28S rRNA bands were rarely observed in them, in spite of application of different RNA extraction methods and careful addition of RNase inhibitors to the buffers. By contrast, ribosomal RNA bands are commonly observed following RNA extraction from cells (Sambrook *et al.,* 1989; Sandrine *et al.,* 2005). Importantly, this lacking in two rRNA bands eliminated the fear that the initial findings were due to artifacts. Perhaps the absence of rRNA bands in these colon tissues could be a common outcome in spite of any RNA extraction protocols deployed.

Broadly, the BER DNA glycosylase genes which consist of the *Neil* genes (*Neil1*, *Neil2*, *Neil3)*, *Ogg1*, *Nthl1*, *Ercc1* (that is also a component of the XPF/ERCC1 lesion-specific endonuclease of the NER) and the MMR genes *Mlh1* were initially expressed only in four sample pairs (Pairs 1 to 4 in Table 1). Nonetheless, due to practical constraints, only three DNA repair genes; *Neil3*, *Ercc1* and *Mlh1*, representing each of the DNA repair pathways (although, in particular, these genes are CRC relevant), were subsequently quantitatively expressed in rest of the pairs. Our choice of these genes was informed by their links to cancer or metastatic tumours e.g. increased expressions of *Neil3*(Shinmura *et al.,* 2016); and *Ercc1* following CRC treatment with oxaliplatin (Shirota *et al.,* 2001; Seetharam *et al.,* 2010). More so, platinum-based drug – induced DNA damage was repaired via *Ercc1* and NER pathway (Reed, 1998; Vilmar and Sørensen, 2009; Rao *et al.,* 2019); mouse cells were sensitized by cisplatin in the absence of *Neil3* (Rolseth *et al.,* 2013); and *Neil3* and *Neil1*are capable of repairing interstrand crosslinks in DNA (Semlow *et al.,* 2016; Martin *et al.,* 2017).The expression protocols used the *Gapdh* as the internal control. Invariably, the pattern of *Gapdh* bands conforms to that of a housekeeping gene in both normal and tumour tissues in all cases.

The double-delta threshold cycle ($\Delta\Delta C_t$) method (Livak and Schmittgen, 2001) was used. The C*t* were recorded in triplicates from which the mean values and standard errors were calculated. Thus, the normalized mean C*t* were obtained by subtracting the *Gapdh* mean C*t* from that of the target genes. The base 2 logarithm fold change gene expression derived from the ratio of normalized C*t* values of tumour to that of normal tissues. Overall, the results reveal no clear expression pattern of the target genes (the DNA repair genes) in the qPCR experiments but few of the genes that include *Ogg1*, *Neil* genes, *Ercc1*, and Nthl1 were commonly quantifiable with amount more than quadrupled that of the *Gapdh* housekeeping gene ($\geq$1-fold) in about 67% of the sample pairs.

In the initial analyses that considered the six BER DNA gycosylase genes in four sample pairs (Figure 3), *Ogg1* had 14.07-, 3.27-, and 1.94-fold changes in sample pair 1, pair 2, and pair 4, respectively. Sample pair 1 had the overall highest expression of all the sample pairs analyzed in the qPCR. The second highest expressions were quantified for *Neil* genes, particularly, for *Neil2* which had 4.30-fold change in sample pairs 3 and 4. Sample pair 2 had 2.32-fold change of *Neil3* expression while 1.83-fold change was calculated for *Neil1* in sample pair 4. The subsequent qPCR quantification involved three of the target genes, that is *Neil3*, *Ercc1*, and *Mlh1*, which are considered highly implicated colorectal cancer(Wheeler *et al.,* 2000; Haydon and Jass, 2002; Kauffmann *et al.,* 2008; Hildrestrand *et al.,* 2009;Seetharam *et al.,* 2010; Abbotts *et al.,* 2014).*Niel3* does seem ubiquitous to the target genes, and by implications to colon cancer tumours, even as < 1-fold change were observed for sample pair 6 (0.05-fold), pair 8 (0.12-fold) and pair 10 (0.04-fold). *Ercc1* and the MMR gene *Mlh1* were highly expressed to varying extents (Figure 4). The fold changes were high for only *Ercc1* (2.93-fold) in sample pair 5 - *Mhl1* lacked detectable gene product in this sample pair. *Ercc1* had low expression (0.24-fold) but *Mhl1* was highly expressed (6.44-fold) in sample pair 6. The three genes had < 1-fold in pair 7. They were nearly non-detectable in pairs 8 and 10. Only *Ercc1* was prominently expressed in pair 9 with 4.33-fold change that almost doubles its expression in sample pair 5. All the three genes expressed very high in sample pair 11 with *Neil3* - 8.86-fold, *Ercc1* - 4.05-fold, and exceedingly high 27.86-fold change of *Mlh1* gene expression. However, in sample pair 12, the fold change of *Mhl1* gene expression is low (< 1-fold) compared to > 2-fold change calculated for *Neil3* and *Ercc1* expressions. Clearly, the outcome of these investigations revealed no constant levels of expressions of any of the DNA repair genes in colon tissues. It may have indicated how heterogeneous the tumour tissues samples are and

how random the accumulation of mutations could be (Loeb, 2001).

Despite the lack of unique expression patterns, certain outcomes are obvious with respect to CRC. These include evidently showing the over-expression of, at least, one of the key DNA damage repair genes. Notably, in particular, the observed over-expressions of *Neil3*, *Ercc1* - as was also observed in melano-carcinoma cell lines (Seetharam *et al.,* 2010), and *Mlh1* could be seen as indicative of underlying or prevailing colon tumourigenesis. Thus, they are suggested as potential biomarkers of colon cancer. Some of the DNA repair genes that either were low expressed or expressed gene products that were below detectable limit could be due to the stage of development of the cancer at the point of surgical removal.

Conclusion

In this study, we elucidated the expression profiles of important DNA repair genes with various gene analysis tools, which include agarose gel electrophoresis and qRT-PCR techniques on colon tumour tissues. The ribosomal RNA of colon tumour samples were rapidly degraded and lacked the signature two bands 18s and 28s rRNA. Although, there is a complex level of interactions between the various DNA repair genes in colon cancer, no clear pattern of expressions was observed for these genes in the sample pairs (tumour/matched normal). However, the results revealed that over-expression of at least one key DNA damage repair gene is a defining characteristic of colon cancer. Notably, the expressions of *Neil3*, *Ercc1*, and *Mlh1* genes were characterized with markedly high fold changes relative to the housekeeping gene, *Gapdh*. Thus, these highly expressed genes could serve as useful biomarkers for the diagnosis of colorectal cancer.

Acknowledgements

OKT thanks the Tertiary Education Trust Fund (TETFund) for granting studentship.

References

Abbotts R., Thompson N. & Madhusudan S. (2014). DNA repair in cancer: emerging targets for personalized therapy. *Cancer Manag Res,* 6: 77-92.

Acharya, A., Das, I., Chandhok, D. & Saha, T. (2010). Redox regulation in cancer: a double-edged sword with therapeutic potential. *Oxid. Med. Cell Longev*., 3:23-34.

Beckman, K.B. & Ames, B.N. (1997). Oxidative Decay of DNA. *J. Biol. Chem*., 272:19633-19636.

Cheng, N., Chytil, A., Shyr, Y., Joly, A. & Moses, H.L. (2008). Transforming growth factor-beta signaling-deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering and invasion. *Mol. Cancer Res.,*6:1521–1533.

Doll, R. & Brafdford-Hill, A. (1950). Smoking and carcinoma of the lung. *Br. Med. J.,* 2(4682):739–748.

Federico, A., Morgillo, F., Tuccillo, C., Ciardiello, F. & Loguercio, C. (2007). Chronic inflammation and oxidative stress in human carcinogenesis. *Inter. J. Cancer*, 121:2381-2386.

Fortini, P., Pascucci, B., Parlanti, E., D'Errico, M., Simonelli, V. & Dogliotti, E. (2003). 8-Oxoguanine DNA damage: at the crossroad of alternative repair pathways. *Mutat. Res*. 531:127-139.

Friedberg, E.C., Walker, G.C., Siede, W. & Wood, R.D. (2005). DNA Repair and Mutagenesis. *American Society for Microbiology Press*, Washington DC.

Fujita, H., Kato, J., Horii, J., Harada, K., Hiraoka, S., Shiraha, H., Sakaguchi, K. & Shiratori, Y. (2007). Decreased expression of hMLH1 correlates with reduced 5-florouracil-mediatedapoptosis in colon cancer cells. *Oncol. Rep*., 18:1129-1137.

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., & Edkins, S. (2007). Patterns of somatic mutation in human cancer genomes. *Nature,* 446:153-158.

Gupta, S.C., Hevia, D., Patchva, S., Park, B., Koh, W. & Aggarwal B.B. (2012). Upsides and downsides of reactive oxygen species for

cancer: the roles of reactive oxygen species in tumorigenesis, prevention, and therapy. *Antioxid. Redox Signal*. 16:1295-1322.

Hancock, J.T., Desikan, R. & Neill, S.J. (2001). Role of reactive oxygen species in cell signalling pathways. *Biochem. Soc. Trans*., 29:345-350.

Harfe, B.D. & Jinks-Robertson, S. (2000). DNA mismatch repair and genetic instability. *Annu. Rev. Genet*. 34:359-399.

Haydon, A.M. & Jass, J.R. (2002). Emerging pathways in colorectal-cancer development. *Lancet Oncol*. 3:83-88.

Hazra, T.K., Das, A., Das, S., Choudhury, S., Kow, Y.W. & Roy, R. (2007). Oxidative DNA damage repair in mammalian cells: a new perspective. *DNA Repair*, 6:470-480.

Hecht, F., Pessoa, C., Gentile, L.B., Rosenthal, D., Carvalho, D.P. & Fortunato, R.S. (2016). The role of oxidative stress on breast cancer development and therapy. *Tumor Biol*., 37:4281.

Hildrestrand, G. A., Neurauter, C. G., Diep, D. B., Castellanos, C. G., Krauss, S., Bjørås, M., & Luna, L. (2009). Expression patterns of NEIL3 during embryonic brain development and neoplasia. *BMC Neurosci*., 10:45.

Jacobs, A.L. & Schär, P. (2012). DNA glycosylases: in DNA repair and beyond. *Chromosoma*, 121:1-20.

Jenkins, R. W. (2017). Introduction to Ex Vivo Cancer Models. *In Ex Vivo Engineering of the Tumor Microenvironment.Humana Press,* Cham.1-12.

Kauffmann, A., F. Rosselli, F., Lazar, V., Winnepenninckx, V., Mansuet-Lupo, A., Dessen, P., & Sarasin, A. (2008). High expression of DNA repair pathways is assocated with metastasis in melanoma patients. *Oncogene,*27:565–573.

Kumar, B., Koul, S., Khandrika, L., Meacham, R.B. & Koul, H.K. (2008). Oxidative stress is inherent in prostate cancer cells and is required for aggressive phenotype. *Cancer Res*., 68:1777-1785.

Livak, K.J. & Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*, 25:402-8.

Loeb, L.A. (2001). A mutator phenotype in cancer. *Cancer Res.*61:3230-9.

Martin, P.R., Couvé, S., Zutterling, C., Albelazi, M.S., Groisman, R., Matkarimov, B.T., Parsons, J.L., Elder, R.H. & Saparbaev, M.K. (2017). The Human DNA glycosylases NEIL1 and NEIL3 Excise Psoralen-Induced DNA-DNA Cross-Links in a Four-Stranded DNA Structure. *Sci. Rep*. 7(1):17438.

Pierce, B.A. (2017). Genetics: A conceptual approach. *6th Edition, W.H. Freeman*, New York.

Rao D., Mallick A. Basu, Augustine T., Daroqui C., Jiffry J., Merla A., Chaudhary I., Seetharam R., Sood A., Gajavelli S., Aparo S., Rajdev L., Kaubisch A., et al. (2019). Excision repair cross-complementing group-1 (ERCC1) induction kinetics and polymorphism are markers of inferior outcome in patients with colorectal cancer treated with oxaliplatin. *Oncotarget*., 10: 5510-5522.

Reed E. (1998). Platinum-DNA Adduct, Nucleotide Excision Repair and Platinum Based Anti-Cancer Chemotherapy. *Cancer Treat Rev*., 24(5):331-44.

Rolseth, V., Krokeide, S.Z., Kunke, D., Neurauter, C.G., Suganthan, R., Sejersted, Y., & Luna, L. (2013). Loss of Neil3, the major DNA glycosylase activity for removal of hydantoins in single stranded DNA, reduces cellular proliferation and sensitizes cells to genotoxic stress. *Biochim, Biophys, Acta* 1833:1157-1164.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual, 2nd edn*, NY Cold Spring Harbor Laboratory Press

Sandrine Imbeaud, Esther Graudens, Virginie Boulanger, Xavier Barlet, Patrick Zaborski, Eric Eveno, Odilo Mueller, Andreas Schroeder, Charles Auffray. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces, *Nucleic Acids Research*, 33(6):e56

Seebode, C., Lehmann, J. & Emmert, S. (2016). Photocarcinogenesis and Skin Cancer Prevention Strategies. *Anticancer Res*., 36:1371-1378.

Seetharam, R.N., Sood, A., Basu-Mallick, A., Augenlicht, L.H., Mariadason, J.M. & Goel, S. (2010). Oxaliplatin resistance induced by ERCC1 up-regulation is abrogated by siRNA-mediated gene silencing in human colorectal cancer cells. *Anticancer Res*. 30:2531-2538.

Semlow, D.R., Zhang, J., Budzowska, M., Drohat, A.C. & Walter, J.C. (2016). Replication-dependent unhooking of DNA interstrand cross-links by the NEIL3 glycosylase. *Cell*, 167:498–511.

Shinmura, K., Kato, H., Kawanishi, Y., Igarashi, H., Goto, M., Tao, H., & Sugimura, H. (2016). Abnormal expressions of DNA glycosylase genes NEIL1, NEIL2, and NEIL3 are associated with somatic mutation loads in human cancer. *Oxid. Med. Cell. Longev.* 2016:1546392.

Shirota, Y., Stoehlmacher, J., Brabender, J., Xiong Y.P., Uetake, H., Danenberg, K.D., & Lenz, H.J. (2001). ERCC1 and thymidylate synthase mRNA levels predict survival for colorectal cancer patients receiving combination oxaliplatin and fluorouracil chemotherapy. *J. Clin. Oncol*. 19:4298-304.

Suzuki, T. & Kamiya, H. (2016). Mutations induced by 8-hydroxyguanine (8-oxo-7,8-dihydroguanine), a representative oxidized base, in mammalian cells. *Genes Environ*. 39:2.

Taylor, J., Ferry, N. & Elder, R.H. (2015). RNAi Knockdown of NEIL3 Sensitizes HCT116 Colorectal Cancer Cells to Oxaliplatin. *Tomas Lindahl Conference on DNA Repair*, Oslo, Norway.

Vaquero, E.C., Edderkaoui, M., Pandol, S.J., Gukovsky, I. & Gukovskaya, A.S. (2004). Reactive oxygen species produced by NAD(P)H oxidase inhibit apoptosis in pancreatic cancer cells. *J. Biol. Chem*., 279:34643–34654.

Vilmar A, Sørensen JB. (2009). Excision repair cross-complementation group 1 (ERCC1) in platinum-based treatment of non-small cell lung cancer with special emphasis on carboplatin: a review of current literature. *Lung Cancer*., 64(2):131-9.

Vogelstein, B. & Kinzler, K.W. (2004). Cancer genes and the pathways they control. *Nature Med*., 10:789-799.

Weren, R.D., Ligtenberg, M.J., Geurts van Kessel, A., De Voer R.M., Hoogerbrugge, N. & Kuiper, R.P. (2018). NTHL1 and MUTYH polyposis syndromes: two sides of the same coinΔ*J. Pathol*. 244:135-142.

Weren, R.D., Ligtenberg, M.J., Kets, C.M., de Voer, R.M., Verwiel, E.T., Spruijy, L, & Hoogerbrugge, N. (2015). A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat. Genet*. 47:668-671.

Wheeler, J.M., Loukola, A., Aaltonen, L.A., Mortensen, N.J. & Bodmer, W.F. (2000). The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadiccolorectal cancers. *J. Med. Genet.*, 37:588-592.

Wolf, D. C., Cohen, S. M., Boobis, A. R., Dellarco, V. L., Fenner-Crisp, P. A., Moretto, A., & Doe, J. E. (2019). Chemical carcinogenicity revisited 1: A unified theory of carcinogenicity based on contemporary knowledge. *Regulatory Toxicology and Pharmacology*, *103*: 86-92.

Table S1. Primers used for Quantitative Polymerase Chain Reaction.

Target Gene	DNA Sequence	Tm (°C)	GC-Content (%)	PCR product (bp)
Gapdh 1014	GGTGGTCTCCTCTGACTTCAACA	61.8	52.2	127
Gapdh 1140	GTTGCTGTAGCCAAATTCGTTGT	60.5	43.5	
Ercc1 884S	CAAAACGGACAGTCAGACCCT	59.8	52.4	146
Ercc1 1029AS	TCAAGAAGGGCTCGTGCAG	58.8	57.9	
Neil1 1071S**	AGAAGATAAGGACCAAGCTGC	57.9	47.6	212
Neil1 1283AS**	GATCCCCCTGGAACCAGATG	61.4	60.0	
Neil1 1079S	AGGACCAAGCTGCAGAATCC	60.0	55.0	125
Neil1 1203AS	GCTCGAAAGGCAGCAAAGTC	60.1	55.0	
Mlh1 2276S^	AGGAGTCGACCCTCTCAGG	61.0	63.2	66
Mlh1 2342AS^	GTCCACTTCCAGGAGTTTGG	59.4	55.0	
Neil2 631S	GAAGCTTCCCCGTAGAAGAGG	61.8	57.1	122
Neil2 773AS	TGTAGCTTCTTACTGCTGCCC	59.8	52.4	
Neil2 1291S^^	GCCTTAGAAGCTCTAGGCCA	59.4	55.0	
Neil2 1436AS^^	GCACTCAGGACTGAACCGAG	60.2	54.0	145
Neil3 1651S	CGCCTCTGCATTGTCCGAGT	62.3	62.3	
Neil3 1798AS	TGGAACGCTTGCCATGGTTG	61.8	61.8	147
Nthl1 679S*	GATGGCACACCTGGCTATG	58.8	57.9	
Nthl1 844AS*	CCACAGCTCCCTAGGCAG	60.5	66.7	165
Ogg1 1020S	AGCAGCTACGAGAGTCCTCA	59.4	55.0	137
Ogg1 1156AS	CATATGGACATCCACGGGCA	59.4	55.0	

**Shinmura *et al*., (2004); *Goto *et al*., (2009); ^Jensen *et al*., (2013); ^^Mandal *et al*., (2012).

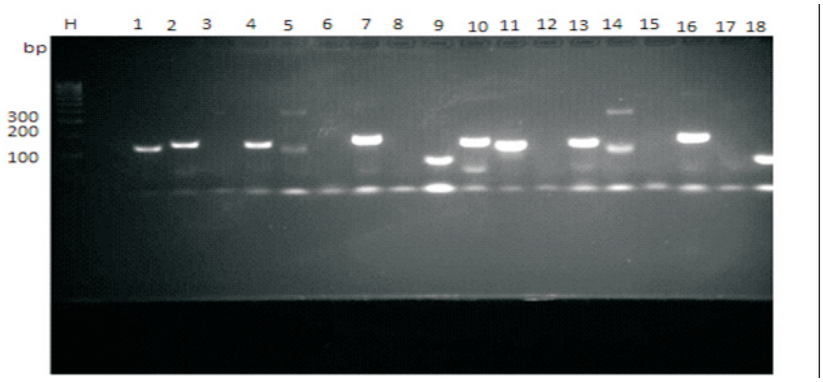

Figure S1: Agarose gel electrophoresis of RT-PCR products of *Gapdh, Ercc1, Ogg1*, *Nthl1* and *Mlh1* genes for HCT116 and Mero-25 cell lines.

Lane 1, *Gapdh*; lane 2, *Ercc1* (884S); lane 3, *Ercc1* (994S); lane 4, *Ogg1* (1020S); lane 5, *Ogg1* (1258S); lane 6, *Nthl1* (349S); lane 7, *Nthl1* (679S); lane 8, *Mlh1* (1814S); lane 9, *Mlh1* (2276S); lane 10, *Ercc1* (884S); lane 11, *Gapdh*; lane 12, *Ercc1* (994S); lane 13, *Ogg1* (1020S); lane 14, *Ogg1* (1258S); lane 15, *Nthl1* (349S); lane 16, *Nthl1* (679S); lane 17, *Mlh1* (1814S); lane 18, *Mlh1* (2276). H represents Hyperladder 100bp.

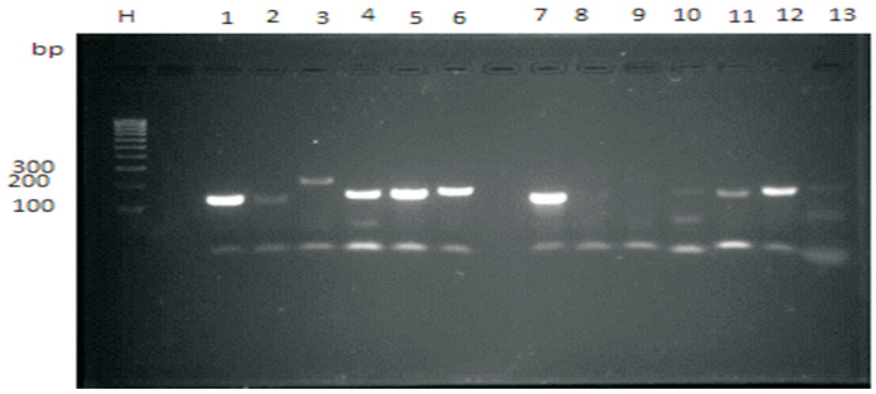

Figure S2: Agarose gel electrophoresis of RT-PCR products of *Gapdh, Neil1, Neil2* and *Neil3* for HCT116 and Mero-25 cells.