SciMedCentral

Annals of Clinical Pathology

Short Communication

In silico Evaluation of Nonsynonymous Single Nucleotide Polymorphisms in the *TDG* Gene, which is Involved in Base Excision Repair

Kazuya Shinmura[1]*, Hisami Kato[1], Masanori Goto[2], Chunping Du[1], Yusuke Inoue[1], Satoki Nakamura[1] and Haruhiko Sugimura[1]

[1]*Department of Tumor Pathology, Hamamatsu University School of Medicine, Japan*
[2]*Division of Cancer Development System, National Cancer Center Research Institute, Japan*

***Corresponding author**
Kazuya Shinmura, Department of Tumor Pathology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi Ward, Hamamatsu, Shizuoka 431-3192, Japan, Tel: 81534352220; Fax: 81534352225; Email: kzshinmu@hama-med.ac.jp

Submitted: 11 April 2014

Accepted: 18 May 2014

Published: 19 May 2014

Keywords
- *TDG*
- DNA glycosylase
- Nonsynonymous SNP
- *in silico*
- Base excision repair

Abstract

The human *TDG* gene encodes a DNA glycosylase protein, which is involved in base excision repair and the regulation of gene expression. Since nonsynonymous variations in two other DNA glycosylase genes, *OGG1* and *MUTYH*, are associated with an increased cancer risk, deleterious nonsynonymous variations in the *TDG* gene might also be associated with diseases, including cancer. In the present study, to identify deleterious variations in *TDG*, nucleotide variations in the coding region of the *TDG* gene were investigated using single nucleotide polymorphism (SNP) databases, and detected nonsynonymous variants were analyzed *in silico* from the standpoint of relevant protein function and stability. A total of 43 nonsynonymous SNPs consisting of 37 missense variations, 3 nonsense variations, and 3 frameshift variations were found in the *TDG* gene. Six of the 37 missense variants were predicted to be damaging or deleterious by three different software programs (PolyPhen-2, SIFT, and PROVEAN), and 28 of them were predicted to be less stable using both the I-Mutant 2.0 and MUpro software. Additionally, 6 nonsense or frameshift variants were predicted to produce a truncated *TDG* protein with a completely or partially lost DNA glycosylase domain. These results suggested that a subset of nonsynonymous SNPs in the *TDG* gene is associated with a reduced level of protein functional activity or stability.

ABBREVIATIONS

SNP: Single Nucleotide Polymorphism; MAP: *MUTYH*-Associated Polyposis; **εC:** 3,N^4-ethenocytosine; 5mC: 5-methylcytosine; 5hmC: 5-hydroxymethylcytosine; 5fC: 5-formylcytosine; 5caC: 5-carboxylcytosine; PolyPhen-2: Polymorphism Phenotyping v2; SIFT: Sorting Intolerant From Tolerant; PROVEAN: Protein Variation Effect Analyzer; HGVD: Human Genetic Variation Database

INTRODUCTION

The human thymine-DNA glycosylase (*TDG*) gene (MIM #601423) is located on chromosome 12q24.1 and encodes a 410 amino acid protein that functions as a DNA glycosylase and is a base excision repair protein [1,2]. The *TDG* protein repairs unmodified or modified bases in various mispairs in double-stranded DNA: i.e., thymine (T) and uracil (U) mispaired with guanine (G), T mispaired with O^6-methylguanine, and thymine glycol mispaired with G [3-5]. The protein is also involved in the repair of 5-halogenated derivatives of U and C, such as 5-fluorouracil and 5-bromouracil, and the exocyclic etheno-base lesion 3,N^4-ethenocytosine (εC) [6,7]. The broad range of substrates shown above enables *TDG* to efficiently stabilize genomic DNA. Recently, *TDG* protein, together with TET family proteins, has been shown to be involved in the demethylation of 5-methylcytosine (5mC) in DNA [8,9]. The 5mC bases can be oxidized to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) by TET proteins, and the resultant 5fC and 5caC base lesions are removed by *TDG*-mediated base excision repair, indicating that *TDG* is profoundly involved in DNA demethylation [8,9]. In another model of DNA demethylation, *TDG* activity is coupled with the deamination of 5mC and 5hmC by AID enzyme [10]. In addition to its role in

DNA demethylation, *TDG* protein interacts with transcription factors and transcriptional coregulators [2]. Thus, *TDG* has very important roles in not only DNA repair, but also the regulation of gene expression.

As genomic variations among people, single nucleotide polymorphisms (SNPs) exist throughout the genome and can be divided into several groups. Among the different kinds of SNPs, a nonsynonymous SNP in the coding region of a gene is important because it alters the amino acid composition; consequently, such alterations can have an impact on protein structure, function, and subcellular localization. Although pinpointing the effects of the many nonsynonymous SNPs using biochemical analyses is challenging, computational analysis tools predicting their effect on protein activity and stability have been recently developed, such as Polymorphism phenotyping v2 (PolyPhen-2) [11], Sorting Intolerant From Tolerant (SIFT) [12], Protein Variation Effect Analyzer (PROVEAN) [13], I-Mutant 2.0 [14], and MUpro [15,16] software. Since the *TDG* protein plays an important role in genome maintenance [2], a reduced functional ability of *TDG* as a result of nonsynonymous SNPs might be associated with susceptibility to diseases, including cancer. Actually, a nonsynonymous SNP in another DNA glycosylase, *OGG1* (MIM #601982), is associated with an increased risk of lung cancer [17], and biallelic nonsynonymous variations in another DNA glycosylase, *MUTYH* (MIM #604933), causes the onset of *MUTYH*-associated polyposis (MAP: MIM #608456), a hereditary disease characterized by colorectal multiple polyps and carcinoma(s) [18,19]. Thus, in the present study, we searched for nonsynonymous SNPs in the *TDG* gene using genome databases and investigated the impacts of nonsynonymous SNPs on *TDG* protein function and stability using a computational approach.

MATERIALS AND METHODS

Collection of nonsynonymous SNPs

Data on nonsynonymous variations of the *TDG* gene were collected from the database of SNPs (dbSNP) located on the homepage of the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/SNP/) and from the human genetic variation database (HGVD) in the Japanese population located on the homepage of the Kyoto University website (http://www.genome.med.kyoto-u.ac.jp/SnpDB/). The reference Transcript ID and the reference Protein ID of *TDG* are NM_003211 and NP_003202, respectively.

PolyPhen-2 prediction

PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) is a tool that predicts the possible impact of an amino acid substitution on the structure and function of a human protein [11]. This prediction is based on a number of features comprising the phylogenetic, sequence, and structural information characterizing the substitution. The PolyPhen-2 server discriminates nonsynonymous SNPs into three main categories: benign, possibly damaging (less confident prediction), or probably damaging (more confident prediction).

SIFT and PROVEAN prediction

SIFT predicts whether an amino acid substitution affects protein function based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences [12]. The SIFT scores range from 0 to 1, and scores ≤0.05 are predicted by the algorithm to be damaging amino acid substitutions, whereas scores >0.05 are considered to be tolerated. PROVEAN is a software tool that predicts whether an amino acid substitution has an impact on the biological function of a protein grounded on the alignment-based score [13]. The score measures the change in sequence similarity of a query sequence to a protein sequence homolog between without and with an amino acid variation of the query sequence. If the PROVEAN score ≤-2.5, the protein variant is predicted to have a "deleterious" effect, while if the PROVEAN score is >-2.5, the variant is predicted to have a "neutral" effect. Both types of software are available on the homepage of the J. Craig Venter Institute: the SIFT tool is at http://sift.jcvi.org, and the PROVEAN tool is at http://provean.jcvi.org.

I-Mutant 2.0 prediction

I-Mutant 2.0 (http://folding.biofold.org/i-mutant/i-mutant2.0.html) is a support vector machine-based tool for the prediction of protein stability changes upon nonsynonymous variations [14]. The tool evaluates the stability change upon nonsynonymous SNP starting from the protein structure or from the protein sequence. The DDG value (difference in free energy of mutation) is calculated from the unfolding Gibbs free energy value of the variant protein minus the unfolding Gibbs free energy value of the wild type (Kcal/mol), and scores <0 are predicted by the algorithm to indicate decreased stability, whereas scores >0 are considered to indicate increased stability.

MUpro prediction

MUpro (http://www.ics.uci.edu/~baldig/mutation.html) is also a support vector machine-based tool for the prediction of protein stability changes upon nonsynonymous SNPs [15,16]. The value of the energy change is predicted, and a confidence score between -1 and 1 for measuring the confidence of the prediction is calculated. A score <0 means the variant decreases the protein stability; conversely, a score >0 means the variant increases the protein stability.

RESULTS AND DISCUSSION

By examining SNPs in the *TDG* gene using the dbSNP and HGVD databases, a total of 43 nonsynonymous SNPs were found. These SNPs consisted of 37 missense variations, 3 nonsense variations, and 3 frameshift variations.

To determine which missense variants are damaging or deleterious, PolyPhen-2, SIFT, and PROVEAN software were applied for the 37 missense variants of the *TDG* gene (Table 1). In the PolyPhen-2 analysis, 8 (21.6%) of the 37 variants were predicted to be probably damaging, and the others were predicted to be benign or possibly damaging. When the SIFT software was used, 18 variants (48.6%) were predicted to be damaging, and the others were predicted to be tolerated. In the PROVEAN analysis, 9 variants (24.3%) were predicted to be deleterious, but the others were neutral. When variants that were common to the 8 variants in the PolyPhen-2 prediction, the 18 variants in the SIFT prediction, and the 9 variants in the PROVEAN prediction were searched, 6 *TDG* variants, namely, c.329G>A (p.Arg110His),

Table 1: PolyPhen-2, SIFT, and PROVEAN results for the 37 missense variants of the *TDG* gene.

Nucleotide[a]	Position[b]	Protein[c]	dbSNP ID	PolyPhen-2 prediction (score)	SIFT prediction (score)	PROVEAN prediction (score)
c.56C>T	g.104370728	p.Thr19Met	rs201193630	possibly damaging (0.606)	damaging(0.045)	neutral (-1.029)
c.121C>T	g.104370793	p.Pro41Ser	rs367858051	benign (0.028)	tolerated (0.101)	neutral (-0.507)
c.143C>A	g.104370815	p.Ala48Asp	rs376956993	possibly damaging (0.790)	damaging (0.011)	neutral (-0.245)
c.196A>G	g.104373638	p.Arg66Gly	rs369649741	possibly damaging (0.546)	damaging (0.009)	neutral (-0.205)
c.268A>G	g.104373710	p.Lys90Glu	rs150152878	probably damaging (0.997)	tolerated (0.054)	neutral (-0.364)
c.329G>A	g.104373771	p.Arg110His	NR[d]	probably damaging (1.000)	damaging(0.001)	deleterious (-4.407)
c.376G>A	g.104373818	p.Asp126Asn	rs149084574	probably damaging (1.000)	damaging (0.014)	deleterious (-4.485)
c.402T>G	g.104373844	p.Ile134Met	rs71466288	possibly damaging (0.673)	damaging (0.040)	neutral (-2.145)
c.431T>C	g.104374693	p.Met144Thr	rs371052913	benign (0.114)	tolerated (0.148)	deleterious (-2.691)
c.526A>G	g.104376624	p.Met176Val	rs140436257	benign (0.005)	tolerated (0.665)	neutral (-1.326)
c.527T>C	g.104376625	p.Met176Thr	rs367961832	benign (0.001)	tolerated (0.777)	neutral (-0.870)
c.595G>A	g.104376693	p.Gly199Ser	rs4135113	benign (0.432)	tolerated (0.209)	deleterious (-5.501)
c.602A>C	g.104376700	p.Lys201Thr	rs61937630	possibly damaging (0.787)	tolerated (0.121)	neutral (-1.727)
c.625C>T	g.104376924	p.Arg209Cys	NR	probably damaging (1.000)	damaging(0.001)	deleterious (-5.995)
c.674G>A	g.104376973	p.Arg225Gln	rs375015053	possibly damaging (0.762)	tolerated (0.067)	neutral (-1.157)
c.697T>C	g.104376996	p.Cys233Arg	rs368866450	possibly damaging (0.741)	tolerated (0.122)	deleterious (-3.587)
c.803T>G	g.104378537	p.Val268Gly	rs17853764	probably damaging (1.000)	damaging (0.000)	deleterious (-6.092)
c.835T>C	g.104378569	p.Phe279Leu	rs138856428	benign (0.143)	tolerated (0.365)	neutral (-0.549)
c.875T>C	g.104378609	p.Leu292Pro	rs140103994	probably damaging (1.000)	damaging (0.000)	deleterious (-6.646)
c.922G>A	g.104378656	p.Val308Ile	rs144056251	benign (0.003)	tolerated (0.453)	neutral (-0.478)
c.980T>A	g.104379396	p.Met327Lys	NR	benign (0.001)	damaging (0.006)	neutral (-1.666)
c.997A>G	g.104379413	p.Lys333Glu	rs376531574	benign (0.002)	damaging (0.023)	neutral (-0.648)
c.1006C>T	g.104379422	p.Pro336Ser	rs139405470	probably damaging (0.972)	damaging (0.004)	deleterious (-2.813)
c.1025A>G	g.104379441	p.Tyr342Cys	rs142534613	benign (0.016)	tolerated (0.054)	neutral (-1.505)
c.1036T>G	g.104379452	p.Tyr346Asp	rs61756223	possibly damaging (0.611)	damaging (0.000)	neutral (-1.937)
c.1039G>A	g.104379455	p.Gly347Arg	rs79676424	possibly damaging (0.844)	tolerated (0.117)	neutral (-0.738)
c.1048C>A	g.104379464	p.Pro350Thr	rs139535385	benign (0.004)	tolerated (0.170)	neutral (-0.582)
c.1066T>C	g.104379482	p.Cys356Arg	NR	possibly damaging (0.901)	damaging (0.003)	neutral (-1.420)
c.1081A>G	g.104379497	p.Asn361Asp	rs186233269	benign (0.000)	tolerated (0.258)	neutral (-1.631)
c.1099G>C	g.104380734	p.Val367Met	rs2888805	benign (0.074)	tolerated (0.085)	neutral (-0.593)
c.1099G>A	g.104380734	p.Val367Leu	rs2888805	benign (0.000)	tolerated (0.266)	neutral (-0.549)
c.1120G>A	g.104380755	p.Ala374Thr	rs3953598	benign (0.000)	tolerated (0.699)	neutral (0.593)
c.1136C>A	g.104380771	p.Pro379His	rs12367528	probably damaging (0.996)	damaging (0.001)	neutral (-1.513)
c.1142G>A	g.104380777	p.Gly381Glu	rs3953597	possibly damaging (0.936)	damaging (0.003)	neutral (-1.282)
c.1181C>T	g.104380816	p.Ser394Phe	rs377754877	possibly damaging (0.832)	damaging (0.003)	neutral (-1.726)
c.1187G>A	g.104380822	p.Ser396Asn	rs3953596	benign (0.000)	tolerated (1.000)	neutral (0.804)
c.1189A>C	g.104380824	p.Asn397His	rs144289190	possibly damaging (0.938)	damaging (0.005)	neutral (-1.195)

[a]Reference transcript ID, NM_003211.
[b]Reference genome, hg19/NCBI37.
[c]Reference protein ID, NP_003202.
[d]Not Registered.

c.376G>A (p.Asp126Asn), c.625C>T (p.Arg209Cys), c.803T>G (p.Val268Gly), c.875T>C (p.Leu292Pro), and c.1006C>T (p.Pro336Ser) were found. Therefore, these variants are considered to be most likely damaging or deleterious.

Next, the changes in the protein stability of the missense variants were examined using I-Mutant 2.0 and MUpro software (Table 2). A total of 28 variants (75.7%) out of the 37 missense variants, including 6 damaging or deleterious variants as determined using the PolyPhen-2, SIFT, and PROVEAN software, were predicted to be less stable using both the I-Mutant 2.0 and the MUpro software.

Table 2: I-Mutant 2.0 and MUpro results for the 37 missense variants of the *TDG* gene.

Protein[a]	I-Mutant 2.0 prediction (DDG[b])	MUpro prediction (score)
p.Thr19Met	increase (1.20)	decrease (-0.30386261)
p.Pro41Ser	decrease (-1.07)	decrease (-0.3180559)
p.Ala48Asp	decrease (-0.5)	increase (0.098690132)
p.Arg66Gly	decrease (-1.09)	decrease (-1)
p.Lys90Glu	decrease (-0.01)	decrease (-0.64510448)
p.Arg110His	decrease (-2.06)	decrease (-1)
p.Asp126Asn	decrease (-0.55)	decrease (-0.75620006)
p.Ile134Met	decrease (-1.48)	decrease (-0.50535186)
p.Met144Thr	decrease (-1.09)	decrease (-0.71375078)
p.Met176Val	decrease (-0.48)	decrease (-0.75477173)
p.Met176Thr	decrease (-0.64)	decrease (-1)
p.Gly199Ser	decrease (-0.99)	decrease (-0.29187319)
p.Lys201Thr	decrease (-0.06)	decrease (-0.11595621)
p.Arg209Cys	decrease (-1.16)	decrease (-0.82707769)
p.Arg225Gln	decrease (-0.39)	decrease (-0.38281526)
p.Cys233Arg	decrease (-1.08)	increase (0.66981316)
p.Val268Gly	decrease (-3.88)	decrease (-1)
p.Phe279Leu	decrease (-0.64)	decrease (-0.48272363)
p.Leu292Pro	decrease (-1.74)	decrease (-1)
p.Val308Ile	decrease(-0.60)	decrease (-0.66160668)
p.Met327Lys	decrease (-0.78)	decrease (-1)
p.Lys333Glu	decrease (-0.87)	decrease (-0.91871881)
p.Pro336Ser	decrease(-1.93)	decrease (-0.71066363)
p.Tyr342Cys	decrease (-0.05)	decrease (-0.19261953)
p.Tyr346Asp	decrease (-1.03)	increase (0.89760457)
p.Gly347Arg	increase (0.42)	increase (0.36647486)
p.Pro350Thr	decrease (-2.14)	decrease (-1)
p.Cys356Arg	decrease (-1.15)	increase (0.019932009)
p.Asn361Asp	decrease(-0.21)	increase (1)
p.Val367Met	decrease (-1.02)	decrease (-0.320804)
p.Val367Leu	decrease (-0.25)	decrease (-0.29335208)
p.Ala374Thr	decrease (-0.46)	decrease (-1)
p.Pro379His	decrease (-0.02)	decrease (-0.34910132)
p.Gly381Glu	decrease(-0.38)	decrease (-0.29766532)
p.Ser394Phe	increase (0.43)	decrease (-0.097516114)
p.Ser396Asn	increase (0.23)	decrease (-0.32936644)
p.Asn397His	decrease (-1.01)	decrease (-0.87856734)

[a]Reference protein ID, NP_003202.
[b]DDG, differences in the free energy.

Regarding the 3 nonsense variations and 3 frameshift variations in the *TDG* gene, all 6 variations were predicted to produce a truncated *TDG* protein (Table 3). The c.112C>T (p.Gln38*), c.272C>G (p.Ser91*), c.286_287insA (p.Ile98Asnfs*6), and c.293_294insA (p.Thr99Tyrfs*5) variants were predicted to lose the DNA glycosylase domain completely, while the c.841C>T (p.Arg281*) and c.685delT (p.Phe229Leufs*17) variants were predicted to lose it partially. These results suggested that all 6 truncated proteins arising from nonsense or frameshift variations exhibited reduced functional activity.

So far, no previous reports have investigated the difference in the repair activity and stability of *TDG* protein between wild-type protein and variant proteins based on SNPs using a biochemical analysis; thus, at present, it is unclear whether the computational prediction in this study can adequately distinguish the various *TDG* proteins based on SNPs from the

Table 3: Summary of nonsense and frameshift variations of the *TDG* gene.

Type	Nucleotide[a]	Position[b]	Protein[c]	dbSNP ID	Glycosylase domain[d]
nonsense	c.112C>T	g.104370784	p.Gln38*	rs372027681	loss
nonsense	c.272C>G	g.104373714	p.Ser91*	rs145088797	loss
nonsense	c.841C>T	g.104378575	p.Arg281*	rs149399146	partial loss
frameshift	c.286_287insA	g.104373728_104373729	p.Ile98Asnfs*6	rs151041931	loss
frameshift	c.293_294insA	g.104373735_104373736	p.Thr99Tyrfs*5	rs67803667	loss
frameshift	c.685delT	g.104376984	p.Phe229Leufs*17	rs140702710	partial loss

[a]Reference transcript ID, NM_003211.
[b]Reference genome, hg19/NCBI37.
[c]Reference protein ID, NP_003202.
[d]Catalytic domain for DNA glycosylase reaction (123-300 a.a.) [2].

standpoint of functional level and stability. However since all the computational programs used in this study are widely utilized [20-22], a concordance in the repair activities of nonsynonymous variants of the DNA glycosylase MUTYH between biochemical analyses and computational predictions has been reported [23], and more than 2 software programs were used in this study, the selection of the deleterious variants was thought to have been properly performed. However, needless to say, adding the results of future biochemical analyses of *TDG* variant proteins to the present findings would enable more solid knowledge regarding *TDG* variants.

In MAP disease, the possession of biallelic pathogenic variants of the DNA glycosylase *MUTYH* gene causes the predisposition of colorectal multiple polyps and carcinoma(s). Thus, diseases arising from biallelic deleterious variants of *TDG* may exist. Additionally, since a heterozygous *TDG* variant could be associated with an increased risk of disease, a careful investigation of the relationship between *TDG* variants and diseases will be important in the future.

CONCLUSION

A total of 43 nonsynonymous SNPs consisting of 37 missense variations, 3 nonsense variations, and 3 frameshift variations were found in the *TDG* gene by searching dbSNP and HGVD databases in this study. Six of the 37 missense variants were predicted to be damaging or deleterious by the PolyPhen-2, SIFT, and PROVEAN software programs, and 28 of the variants were predicted to be less stable by both the I-Mutant 2.0 and MUpro software programs. In addition, 6 nonsense or frameshift variants were predicted to lead to the production of a truncated *TDG* protein that had lost the DNA glycosylase domain either completely or partially. These results suggested that alleles that encode functionally reduced or less stable *TDG* proteins may exist in humans. These *TDG* alleles might be associated with an increased risk of diseases, including cancer.

ACKNOWLEDGEMENT

This work was supported in part by a Grant-in-Aid from the Ministry of Health, Labour and Welfare (21-1), the Japan Society for the Promotion of Science (25460476), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) (221S0001), the Takeda Science Foundation, the National Cancer Center Research and Development Fund, and Center of Innovation Program of Japan Science and Technology Agency of the MEXT.

REFERENCES

1. Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, et al. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. J Biol Chem. 1996; 271: 12767-12774.
2. Sjolund AB, Senejani AG, Sweasy JB. MBD4 and *TDG*: multifaceted DNA glycosylases with ever expanding biological roles. Mutat Res. 2013; 743-744: 12-25.
3. Saparbaev M, Laval J. 3,N4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase. Proc Natl Acad Sci U S A. 1998; 95: 8508-8513.
4. Yoon JH, Iwai S, O'Connor TR, Pfeifer GP. Human thymine DNA glycosylase (*TDG*) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. Nucleic Acids Res. 2003; 31: 5399-5404.
5. Cortázar D, Kunz C, Saito Y, Steinacher R, Schär P. The enigmatic thymine DNA glycosylase. DNA Repair (Amst). 2007; 6: 489-504.
6. Hardeland U, Bentele M, Jiricny J, Schär P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, Drosophila and fission yeast orthologs. Nucleic Acids Res. 2003; 31: 2261-2271.
7. Morgan MT, Bennett MT, Drohat AC. Excision of 5-halogenated uracils by human thymine DNA glycosylase. Robust activity for DNA contexts other than CpG. J Biol Chem. 2007; 282: 27578-27586.
8. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011; 333: 1300-1303.
9. Kohli RM, Zhang Y. TET enzymes, *TDG* and the dynamics of DNA demethylation. Nature. 2013; 502: 472-479.
10. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell. 2011; 146: 67-79.
11. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7: 248-249.
12. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4: 1073-1081.
13. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012; 7: e46688.

14. Bava KA, Gromiha MM, Uedaira H, Kitajima K, Sarai A. ProTherm, version 4.0: thermodynamic database for proteins and mutants. Nucleic Acids Res. 2004; 32: D120-121.

15. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. Nucleic Acids Res. 2005; 33: W72-76.

16. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins. 2006; 62: 1125-1132.

17. Xu Z, Yu L, Zhang X. Association between the hOGG1 Ser326Cys polymorphism and lung cancer susceptibility: a meta-analysis based on 22,475 subjects. Diagn Pathol. 2013; 8: 144.

18. Shinmura K, Goto M, Tao H, Sugimura H. Role of base excision repair enzyme MUTYH in the repair of 8-hydroxyguanine and MUTYH-associated polyposis (MAP). Hereditary Genet. 2012; 1: 111.

19. Yamaguchi S, Ogata H, Katsumata D, Nakajima M, Fujii T, Tsutsumi S, et al. MUTYH-associated colorectal cancer and adenomatous polyposis. Surg Today. 2014; 44: 593-600.

20. George Priya Doss C, Nagasundaram N, Tanwar H. Predicting the impact of deleterious single point mutations in SMAD gene family using structural bioinformatics approach. Interdiscip Sci. 2012; 4: 103-115.

21. Rahim F, Galehdari H, Mohammadi-Asl J, Saki N. Regression Modeling and Meta-Analysis of Diagnostic Accuracy of SNP-Based Pathogenicity Detection Tools for UGT1A1 Gene Mutation. Genet Res Int. 2013; 2013: 546909.

22. Phani NM, Acharya S, Xavy S, Bhaskaranand N, Bhat MK, Jain A, et al. Genetic association of KCNJ10 rs1130183 with seizure susceptibility and computational analysis of deleterious non-synonymous SNPs of KCNJ10 gene. Gene. 2014; 536: 247-253.

23. Shinmura K, Goto M, Tao H, Kato H, Suzuki R, Nakamura S, et al. Impaired 8-Hydroxyguanine Repair Activity of MUTYH Variant p.Arg109Trp Found in a Japanese Patient with Early-Onset Colorectal Cancer. Oxid Med Cell Longev. 2014; 2014: 617351.

Cite this article

Shinmura K, Kato H, Goto M, Du C, Inoue Y, et al. (2014) In silico Evaluation of Nonsynonymous Single Nucleotide Polymorphisms in the TDG Gene, which is Involved in Base Excision Repair. Ann Clin Pathol 2(1): 1014.