Trends in Biosciences 11(44), Print : ISSN 0974-8431, 00-00, 2018

REVIEW PAPER

A Review on Transcription Associated Mutagenesis in terms of Gene Regulation

SONAM MAHAWAR AND HEERA SINGH

Jayoti Vidyapeeth women's University, Jaipur, Rajasthan
email : sonammahawar29@gmail.com

ABSTRACT

Transcription is the first stage in gene expression involves synthesis of an RNA chain which represent one strand of DNA duplex. Transcribed DNA is known to be packaged into a more open chromatin conformation than non-transcribed and silent DNA, this might render highly transcribed DNA more accessible to endogenous DNA damaging agents. The causes of mutation, during cell division where a DNA fails to copy accurately, external features like chemical or radiation may leads to breakdown of DNA and also due to imperfection of DNA repair mechanism in cell. In the cell, RNA polymerase and replisome share the same template DNA in which RNAP play important role in genome integrity and mutation of collision takes place in lagging strand and in leading strand this results in understanding of mechanism of TAM where it act as a powerful genetic tool by inducing mutations in specific ways. Genetic modification associated with transcription is a global issue. The newly identified elements are associated with one another for gene expression, providing new insights into mechanism of gene regulation.

Key words *Transcription, Mutation, Mutagenesis, DNA*

The Basics of genetics at molecular level

The central dogma of molecular biology is an explanation of the flow of genetic information within a biological system i.e. genetic information is present in DNA which is transcribed to RNA and translated into protein.Transcription involves synthesis of an RNA chain representing one strand of DNA duplex. During transcription mRNA is copy of non transcribed strand. It is a cyclic process composed of 3 step initiation, elongation and termination. Initiation of transcription starts with synthesis of short RNA transcripts.

RNA polymerase creates the transcription bubble when it binds to a promoter show that RNA polymerase moves along with it and RNA chain grows longer. As RNA polymerase moves along the DNA template it unwinds the DNA duplex at the front of the bubble (the unwinding point) and rewinds the DNA at the back (rewinding point). The length of transcription bubble is ~12 to 14bp but the length of RNA-DNA hybrid region within it is 8 bp.

In eukaryotes transcription and translation separated in different cellular location i.e. in nucleus and cytoplasm respectively. So many steps for synthesis of mRNA like addition of 5'cap and poly A tail and RNA splicing therefore it is combined transcription process. In prokaryotes coupled transcription translation i.e. no gap between transcription–translation processes. The translation begins while the mRNA is still being synthesized.

Is mutation is main cause

DNA is replicated every time a cell divides. Sometime it makes mistakes most of times it can correct these mistakes but sometimes it cannot and the sequences copied DNA results different from the original. To understand the consequences of DNA mutation one has to consider two major factors. What kind of mutation has occurred and where it has occurred. The mutation if occurred in junk DNA than it might have no consequences. If it occurs in promoter region of a gene it may alter the regulation of the expression of that gene. If it occurs in the coding region of a gene it may affect the promoter sequences. The resons might be DNA fails to copy accurately or External features can create mutations. As we have known about transcription and mutation than how this transcription is associated with mutation which causes transcription associated mutagenesis.

Transcription and Mutagenesis

When DNA damage is in the transcribed strand of active gene is by passed by RNA polymerase they can miscode at the damaged site and produce mutant transcript this process is transcriptional mutagenesis.As already seen during transcription process ,transcription copies only 1 DNA strand other remain single stranded, that results in chemical reactive and vulnerable to endogenous DNA damage. In bacteria, mutagenesis is a relatively straight forward process as they were haploid organisms where any single mutation results in altered expression or impaired function. Localized changes in DNA that is associated with transcription of target sequences also known as transcription associated mutagenesis.

Effect of transcription on DNA template

In transcription process when RNAP bind to the DNA the positive supercoiling are generated ahead of transcription machinery and reflect over winding of helix behind the machinery, the corresponding under wound state of DNA leads to accumulation of negative super coiling. The SSDNA which was damaged by mutagen during negative super coiling causes endogenous damage in both strands during replication.

Transcrption – Replication interaction

Transcription and replication relation depend on collision between RNAP and replisome. There are two types of Collision between RNAP and replisome. i.e. Co-directional collision occur in leading strand and Head on collision takes place in lagging strand

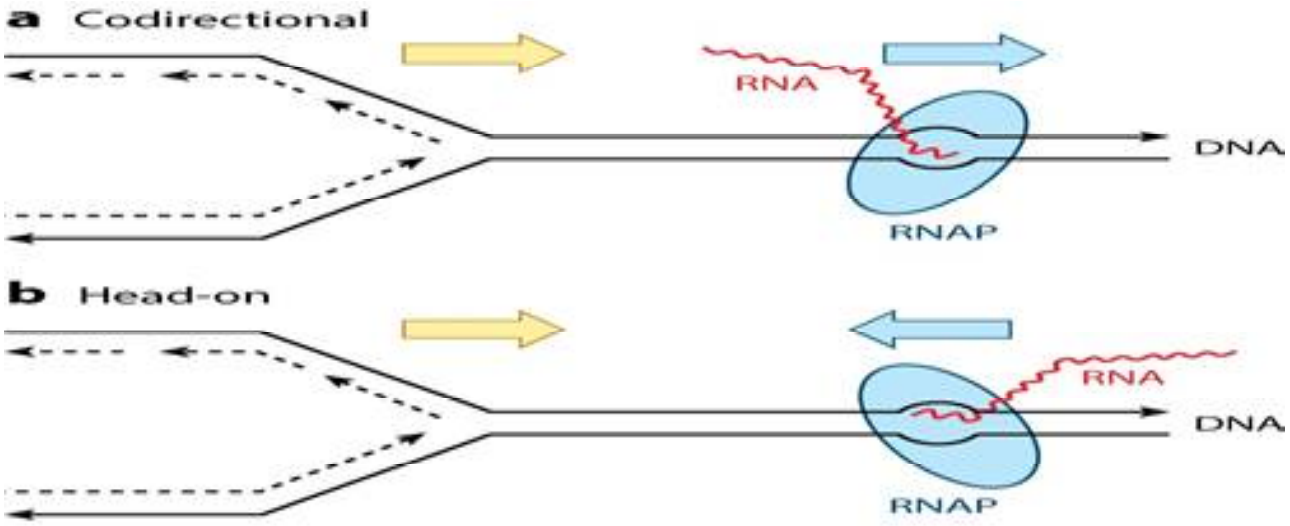

Fig. 1. Conflicts between the replication and transcription machineries. Movement of the replisome and RNAP (RNA polymerase) in the same or opposite direction can cause (*a*) codirectional or (*b*) head-on conflicts, respectively. Red and black lines represent RNA and DNA, respectively; dashed lines depict newly synthesized DNA; and blue ovals represent RNAP. Yellow and blue arrows indicate the direction of replication-fork and RNAP movement, respectively. (Robertson and bhagwat.,2014)

In co-directional orientation, replication fork progression was the same in both non-transcribed and transcribed regions where no accumulation of the replisome was observed. Interestingly, during replication of the transcribed regions RNAP molecules were absent behind the replication fork but were still present in front of it. Additionally, after the replisome had completely passed through the transcribed region repopulation by RNAP took place. These observations suggested that the replisome was not neither slowed down nor displaced from the DNA by transcribing RNAP.

In the case of head-on collisions, RNAP was also

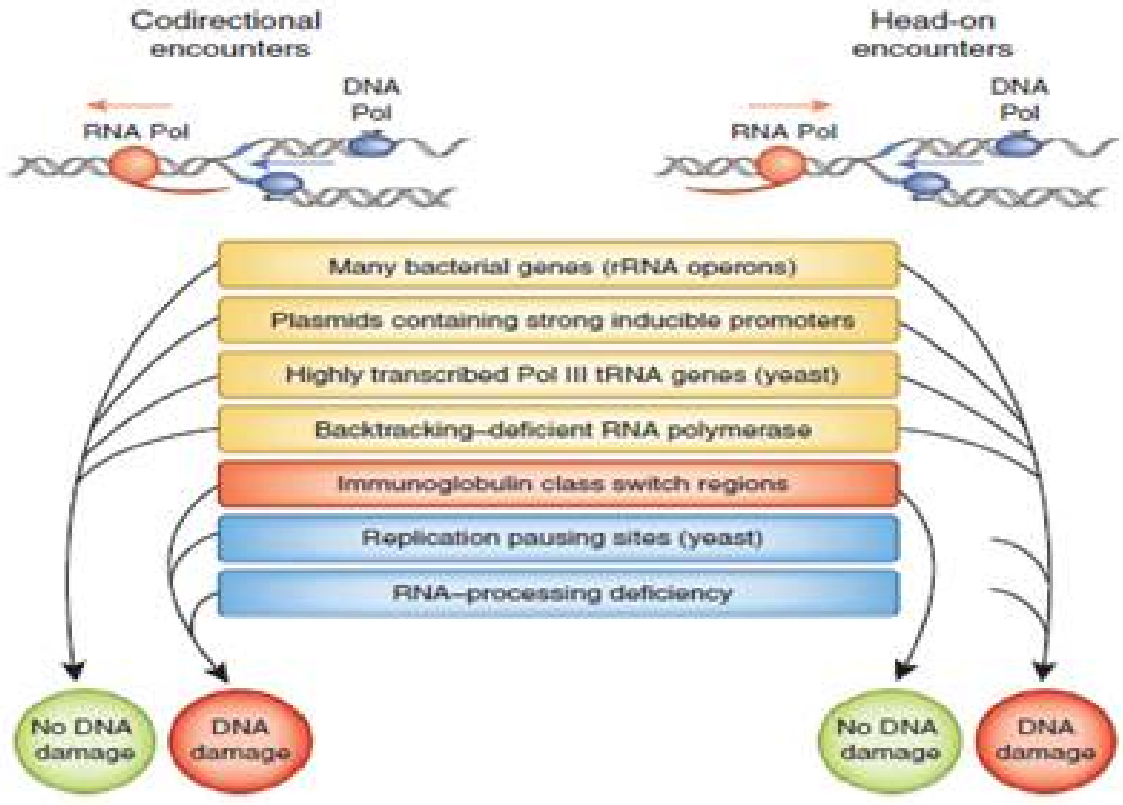

Fig. 1 (a): DNA Damage

dislodged but the replication fork progressed through the operon much more slowly than in the co-directional counterpart. The slowing down of the replication fork caused by head-on collisions could result in genomic instability due to inaccurate restart of the replisome by new collisions with transcribing RNAP molecules. Another possibility that could lead to instability is that, after collision, stalled replication fork could unwound generating a four stranded DNA structure that resembles a Holliday junction (a process known as replication fork regression) (Atkinson et al. 2009). Transcribing RNAP cannot surpass the lesions on the DNA and therefore it stalls. Head on collision are more detrimental than co-directional collision as the slowing down of replication fork caused by head on collision could result in genomic instability due to inaccurate result of the replisome by collision with transcribing RNAP molecules. Here contribution of these conflicts is considered to localize mutagenesis. A source which causes the Transcription Associated Mutagenesis (TAM) are DNA damage, Cytosine deamination on transcribed and non transcribed strand and R-loop.

DNA damage: In transcription DNA damage is of two types:

Direct damage to transcript itself:- damaged ribonucleotides can lead to altered pairing specificity of codon and anticodon recognition through that incorrect amino acid can be incorporated. Some damaged ribonucleotides with altered pairing specificity can be incorporated into the nascent mRNA by RNA polymerase. Thus leading to mutant transcript that is translated into erroneous protein.

Directionality makes the difference. Depending on the genomic location and on the state of the transcription complex, the outcomes of transcription-replication encounters can be classified into three main groups:

Those linked to DNA damage when the transcription machinery encounters the replication fork in the head-on direction (yellow boxes). Those associated with DNA damage when transcription and replication are co-directional (orange box).

Those that may cause genomic instability in either orientation (blue boxes). Orange and blue arrows represent the direction of transcription and replication, fork progression, respectively; small black arrows represent the direction of the DNA polymerases replicating each strand. Pol, polymerase.

Cytosine Deamination on Transcribed and Non Transcribed Strand

Cytosine methylation is a common form of post-replicative DNA modification seen in both bacteria and eukaryotes. Modified cytosines have long been known to act as hotspots formulations due to the high rate of spontaneous deamination of this base to thymine, resulting in a G/T mismatch.

In this process amino acid are broken down if there is excess intake of protein.

Spontaneous deamination results in the hydrolysis reaction of cytosine to uracil releasing ammonia in a process. This spontaneous deamination result from C to U generates double strand breaks and recombination.the generates U:G mismatches may be replicated creating 2 daughter strand one of which undergo C:T transition mutation.In replication process when it past by abiasic site it will result in random incorporation of nay of 4 nucleotides, possibly leading to further mutations. DC to DU mutation occur at lower level.

Deamination of cytosine on the NTS results in C > T mutations, whereas deamination of cytosine on the TS generates G > A sequence changes. Thus, by comparing the accumulation of C > T versus G > A changes, one can infer relative deamination of the NTS versus TS strand. Comparative analysis of genes in *E. coli* and *S. enterica* indicates that the cytosine deamination bias primarily reflects an asymmetry associated with transcription rather than replication.

Transcription mediated R-LOOP:

R loop structure was first characterized by Thomas *et al*., 1976 and R loop exists in vivo is demonstrated by

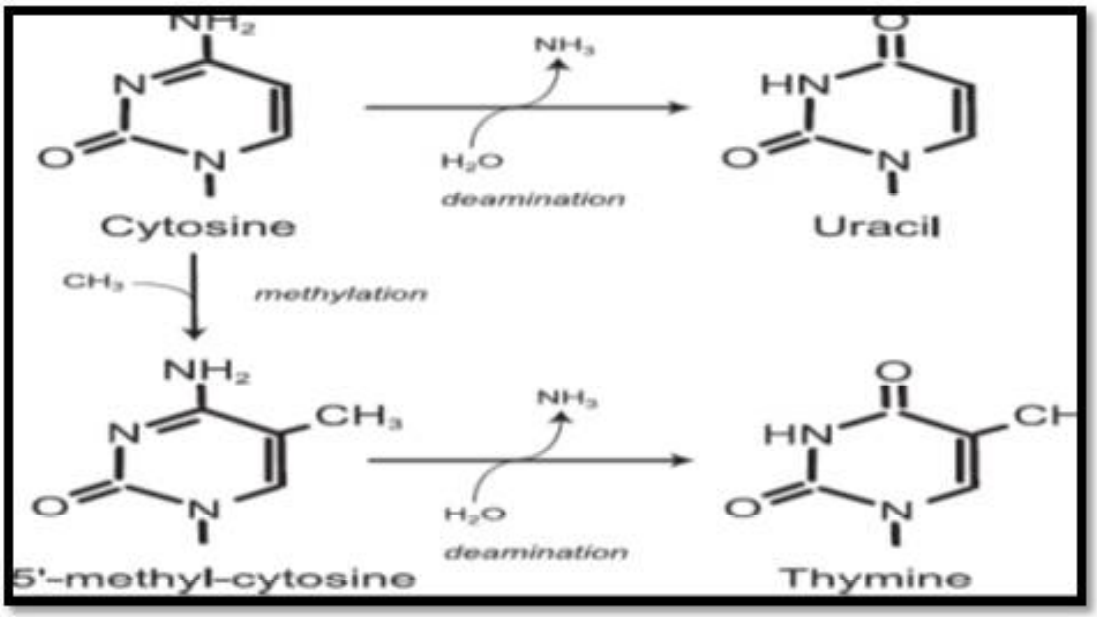

Fig.1. (b): Cytosine methylation

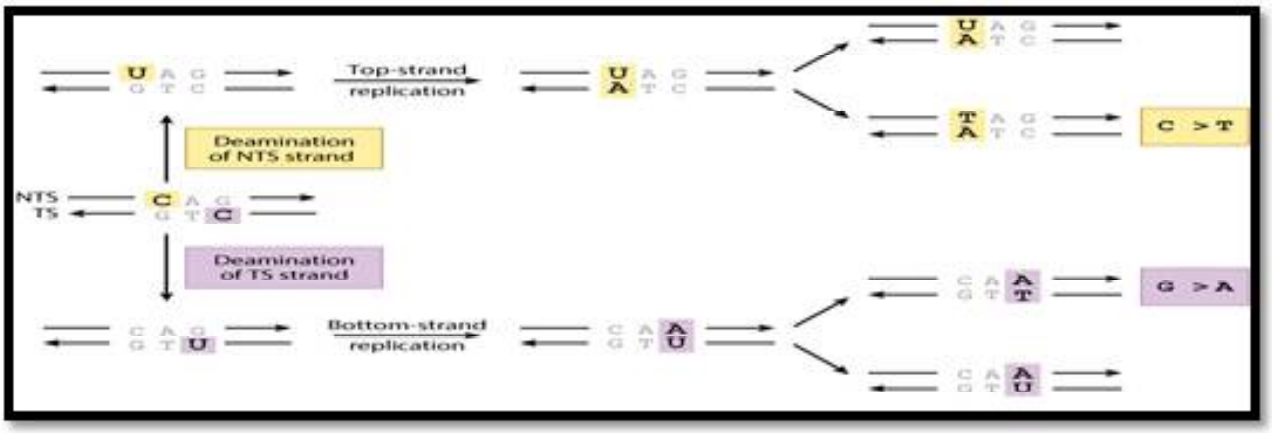

Fig.1 (c): Yellow and purple boxes indicate consequences of cytosine deamintion on the non transcribed and transcribed strand respectively.

Crouch and colleagues in 1995. They show R loop formation occur in abacterial cell and is a consequences of transcription process. R loop is a 3 strand nucleic acid structure formed by RNA :DNA hybrid plus a displaced DNA strand (ssDNA) identical to the RNA molecule.

There are two model to demonstrate formation of R loop:

1. RNA-DNA hybrid model: According to RNA DNA hybrid model the RNA / DNA hybrid duplex could be result of an exetension of the usual 8bp RNA/DNA hybrid within transcription bubbles as Pol II elongates but it is not accepted.
2. Threading back model: RNA/DNA hybrid could arise by threading back the RNA transcripts before the 2 strands of the DNA duplex anneal as soon as it exits RNAP R loop depends on 3 features: high G density, negative supercoiling and DNA nicks (Roy *et al.*, 2010)

Impact of repair on TAM

As transcription-associated mutations occurs asymmetrically on the transcribed and non-transcribed strands, when all the transcription-associated mutations that occur in a cell are not repaired, its two daughter cells have different genomic DNA. Transcription-coupled DNA repair act selectively on the transcribed strand to understanding the mechanistic basis of TAM the impact of individual DNA repair pathways on this process has been shown in yeast

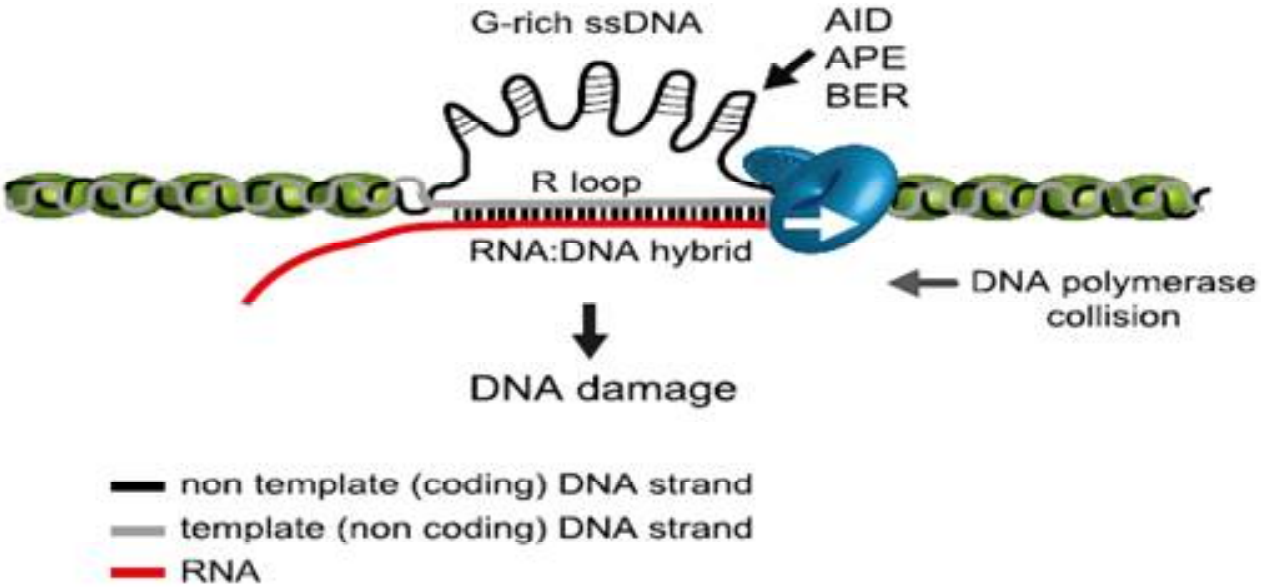

Fig. 2. Initial R Loop formation is favourable by G cluster and DNA nicks downstream from the promoter on the non template DNA strand nicks whereas subsequent RNA/DNA hybrid extension and stabilization are enhanced by high G density and negative supercoiling. The displaced ssDNA in a fig can act as a substrate to DNA damaging agent , AID (deaminase), Repair enzymes (BER) leading to DNA lesion and nicks. Transcriptional machinery impeded by R loops can cause replication- transcription collision.

Three major DNA repairing mechanism for a single strand damage

a. Base excision repair
b. Nucleotide excision repair
c. Mismatch repair

a. Base excision repair : it correct the damages that arise due to chemically induced deamination of cytosine that result in the U:G mismatch so excision of uracil is important to avoid GC: AT transition mutation in replication.Steps involved shown in Fig.3.(a)

DNA glycosylase recognizes a damaged base and cleaves between the base The primary and initiating event of BER is the hydrolysis of the N-glycosyl bond linking a nitrogenous base to the deoxyribose-phosphate chain, thereby releasing the free base.

AP endonuclease cleaves the phosphodiester backbone near the AP site man cells possess for the repair of mismatched bases the excision of bases from duplex DNA generates apurinic or apyrimidinic (AP) sites (Fig. The repair of these sites of base **loss** utilizes a specific class of endonucleases designated AP endonucleases.

Prokaryotes such as ***Escherichia coli*** have at least two such enzymes. However, studies of yeast and mammalian cells have thus far resulted in the purification and characterization of a single major endonuclease that catalyzes the incision of phosphodiester linkages exclusively *5* to AP sites, generating 3 ***I*** -OH and *5* ***I*** – deoxyribosephosphate residues. This enzyme also has 3'-phosphatase activity, and, in some organisms, a weak 3 ***t*** - 4 ***t*** exonuclease activity. In mammalian cells, the gene that encodes this AP endonuclease is variously called ***BAPl*** (bovine AP endonuclease), The completion of BER requires the removal of the 5'-terminal deoxyribose-phosphate residue generated by the AP endonuclease, followed by repair synthesis and DNA ligation. A 47-kD enzyme activity designated DNA deoxyribophosphodiesterase (dRpase) has been identified in human cells and can remove such sugar-phosphate residues from duplex DNA (Price and Lindahl, 1991). If this enzyme is indeed utilized in vivo, a single nucleotide gap would be generated that can be filled in by a DNA polymerase. In vitro, such repair synthesis is efficiently catalyzed by DNA polymerase-P (Dianov *et al.* 1992). It has been suggested that because of its limited fidelity, polymerase-P may be utilized in this particular very short patch mode of repair synthesis during BER and not in repair synthesis associated with longer repair patches. The nick remaining after DNA polymerase I is sealed by DNA ligase.

Nucleotide excision repair: Fig. 3. (b) it corrects the damage that are particularly due to environment mutagenesis and carcinogenic agent. Three exonucleases bind DNA at the

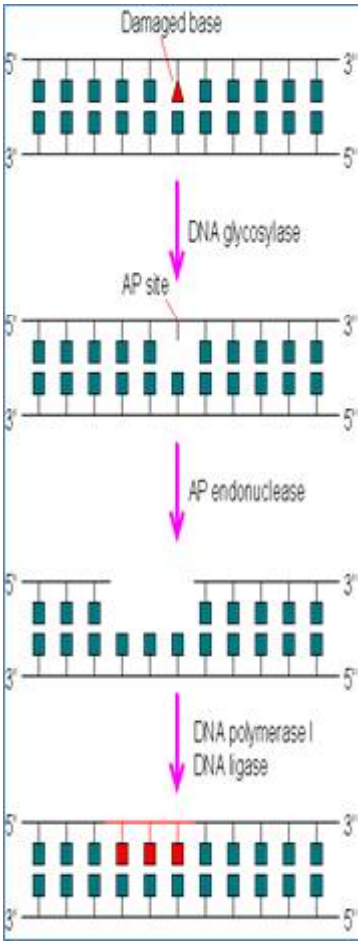

Fig. 3(a): Base excision repair

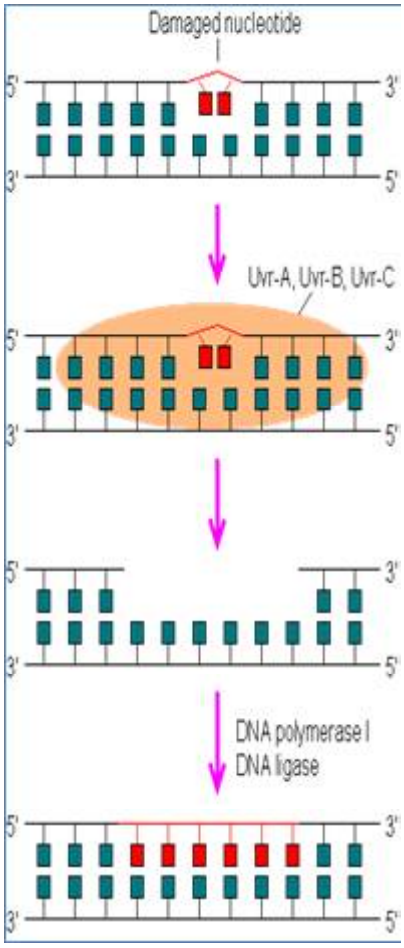

Fig.3.(b): Nucleotide excision repair

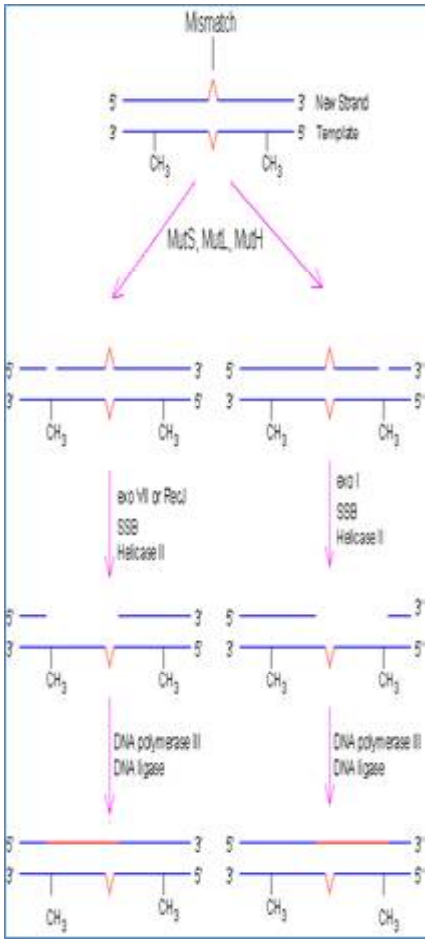

Fig.3(c): Mismatch repair

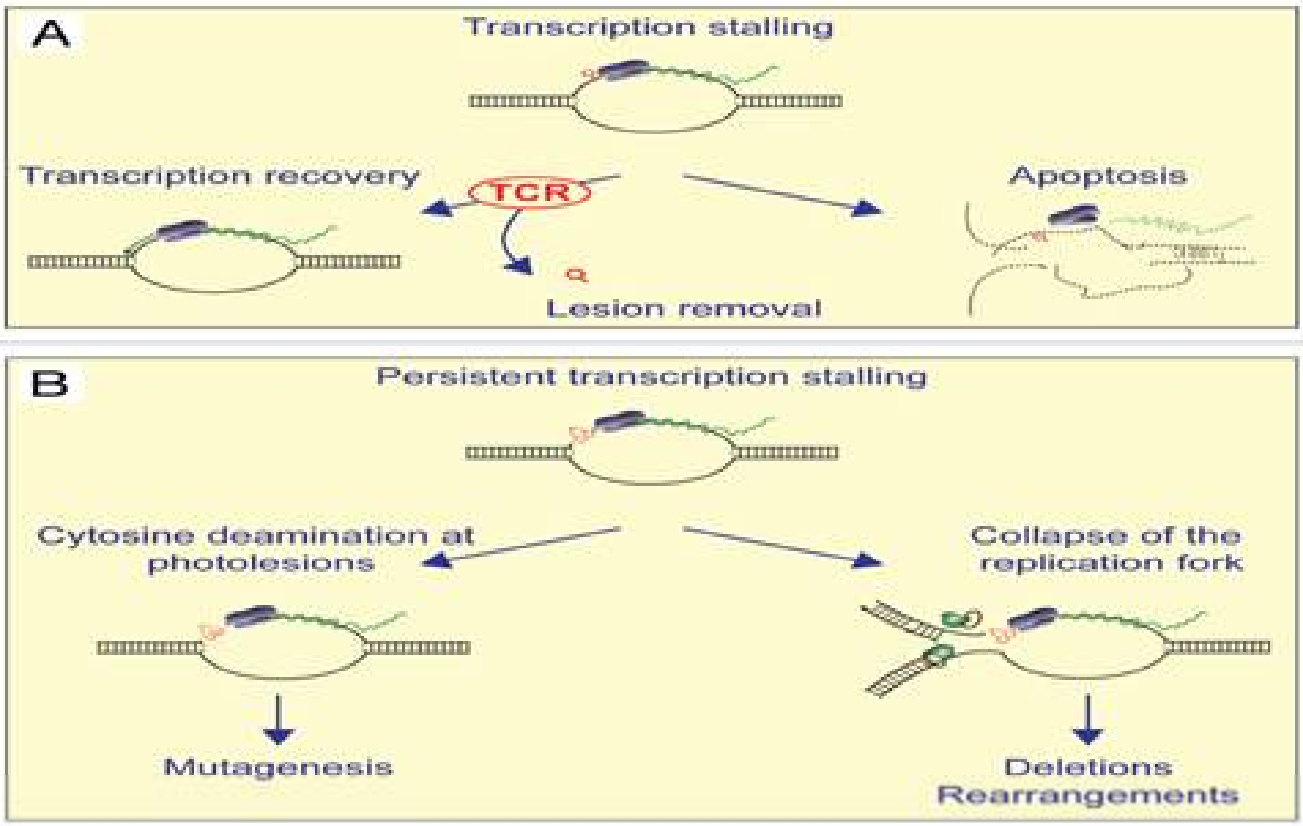

Fig. 4. Model for the roles of TCR and transcription arrest-associated apoptosis in protection against UV-TAM

site of bulky lesion in which UvrA bind first**.** UvrB binds to the UvrA-DNA complex and increases specificity of complex for irradiated DNA. UvrC remove DNA 8 bases upstream and 4 or 5 bases downstream of dimer. The gap is filled by DNA polymerase I**.**The remaining nick is sealed with DNA ligase

Mismatch repair: Fig. 3. (c) Mismatch repair correct error occur in DNA replication and recombination immediately after replication newly synthesized strand and template strand. Newly synthesized strand not methylated yet.

Other repair for TAM: Fig. 9. Model for the roles of TCR and transcription arrest-associated apoptosis in protection against UV-TAM. (A) TCR is initiated by stalling of the elongating RNA polymerase and results in removal of the transcription-blocking photolesion, thereby allowing transcription recovery.

In case the load of photolesions exceeds the capacity of TCR or when TCR is impaired, arrested transcription complexes induce apoptosis when the cell enters S phase. (B) Persistent stalling of the transcription complex at a deoxycytidine-containing-CPD at the template strand results in its increased spontaneous deamination frequency. In case the lesion is not timely removed by TCR, or when the cell escapes from apoptosis, it will result in a C to T transition mutation during replication. Alternatively, persistent stalling of the transcription complex at a photolesion may lead to collapse of an approaching replication fork, resulting in DSBs and genome instability.

Detection of mutation can be done by:

Forward mutagenesis assay which inactivate a functional gene and Reversion mutagenesis assay, in which reversion by frame-shift mutation is essentially the reciprocal event, in which a mutation restores the normal function.

An example of Mechanism of TAM in yeast

TAM in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. This study was done using Tetracycline regulated LYS 2 reporter system was developed to modulate the transcriptional level over a broad range in *S. cerevisae.*

TAM under stress in *E. coli*

Transcription associated mutations are considered to occur regardless of specific secondary structures. Transcription-associated mutagenesis becomes active under stress and occurs in both genomic DNA and plasmid DNA in *E.coli* .Transcription associated mutagenesis is considered to be an intrinsic source of mutations. The mutation rate is one base pair per genome per replication per TAM. TAM not affect on other genomic strand.TAM can be considered a safe way for dividing cells to rapidly increase the sequence diversity of the next generation. This study, evaluated the selection pressure on MI, which represents the relative potential for transcription-associated mutagenesis per transcription event under stress. results suggest that the majority of protein-coding sequences have evolved to increase protein sequence diversity by controlling transcription associated mutagenesis under stress and that transcription associated mutagenesis produces protein sequence diversity more effectively than does random mutagenesis. Therefore, transcription-associated mutagenesis will confer faster protein

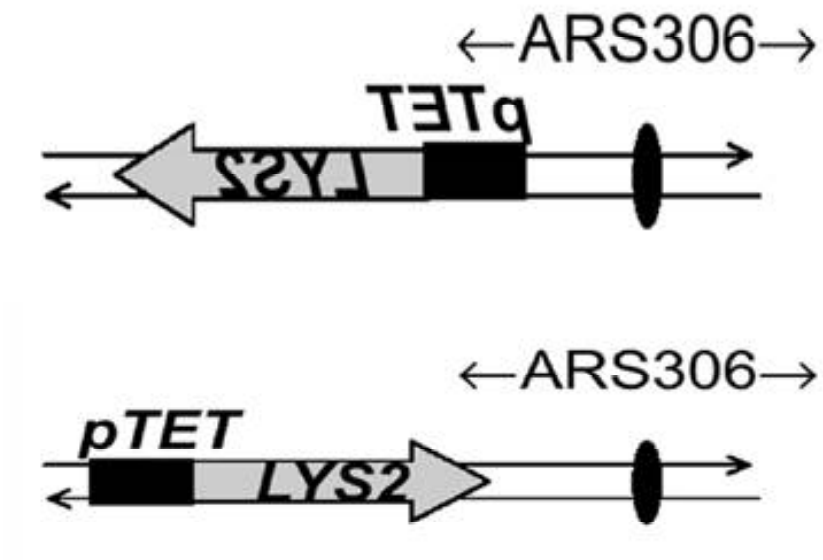

Fig. 5. Schematic drawings of a portion of the right arm of chromosome III with *LYS2* integrated in (A) SAME orientation or (B) OPPOSITE orientation. The black oval indicates the replication origin *ARS306*, the black box indicates *the pTET* promoter cassette and the shaded arrow represents the *LYS2 gene and its direction of* transcription. This suggests spontaneous mutation rate is directly proportional to the transcriptional level, suggesting that movement of RNA polymerase through the target initiates a mutagenic process.

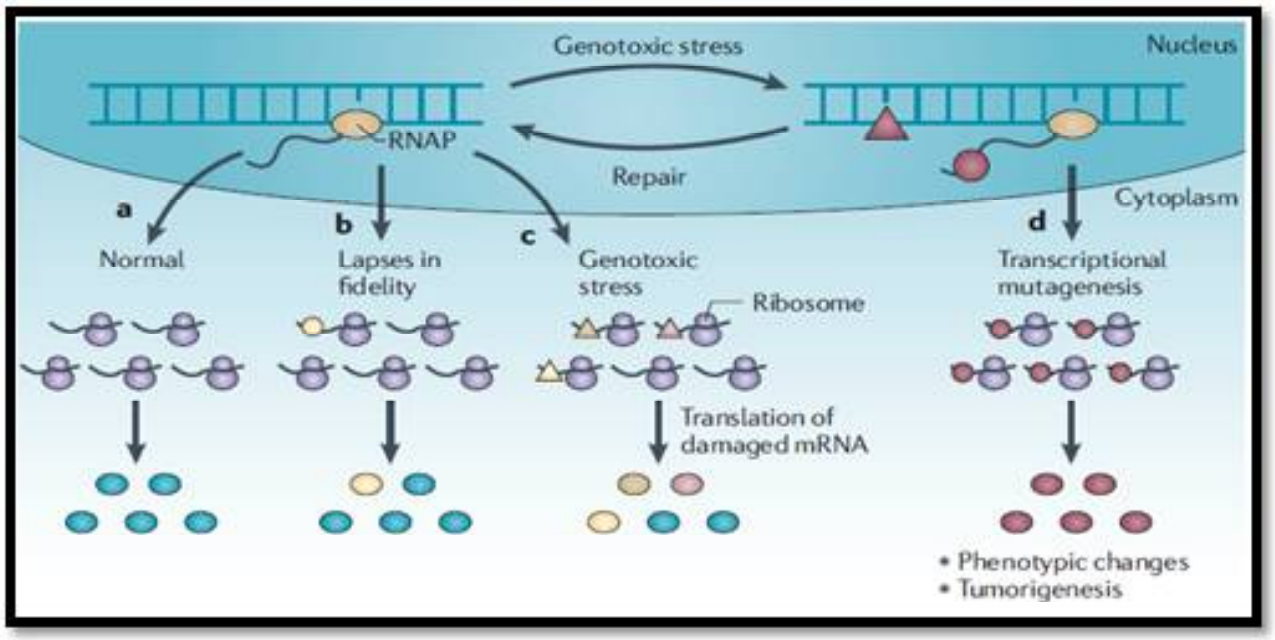

Fig 6: DNA replication-independent production of erroneous proteins. Under normal conditions (**a**), transcription in the nucleus produces error-free mRNAs that are translated by ribosomes to normal proteins (blue ovals) in the cytoplasm. In some cases (**b**), lapses in RNA polymerase (RNAP) fidelity can generate aberrant transcripts (yellow circle) that are translated into erroneous proteins (yellow oval). This random, low-frequency production of erroneous proteins can also be caused by lapses in ribosome fidelity. When exposed to a genotoxic agent (**c**), RNA molecules in a cell may contain various lesions (triangles) that could induce the production of erroneous proteins during translation because of their potentially altered codon–anticodon pairing during tRNA selection. DNA is the other target for genotoxic stress (**d**). RNAP can bypass numerous unrepaired damaged deoxyribonucleotides on the transcribed strand of a gene (red triangle) that can result in misincorporation events in the transcript sequence (red circles) as long as the DNA damage is not removed by one of the DNA repair pathways. Transcriptional mutagenesis results in the production of a mostly homogenous mutant transcript population, which in turn leads to the production of high levels of erroneous proteins that possess the same mutant sequence, and that could alter the phenotype of the cell.

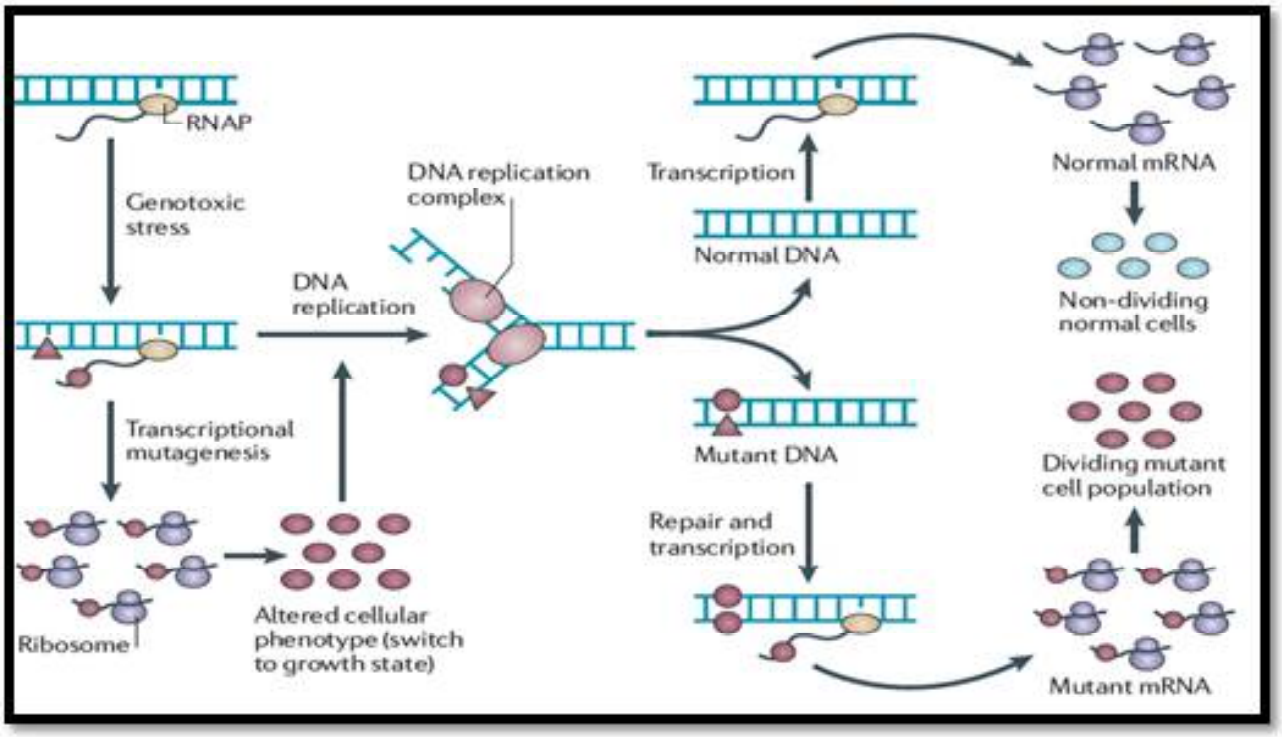

Fig. 7: The potential role of transcriptional mutagenesis in tumour development. Following genotoxic stress, a DNA lesion (red triangle) can appear on the transcribed strand of a gene, resulting in the production of high levels of erroneous protein by transcriptional mutagenesis (Fig.6). The resulting mutant proteins (red ovals) may have the ability to alter the phenotype of the cell in such a way that a growth advantage is conferred, leading to initiation of DNA replication. If left unrepaired, the DNA lesion will subsequently be encountered by the replication machinery and will probably cause similar miscoding during DNA synthesis, which will result in the fixation of the mutation into the genome of one of the replicated progeny (red circle). Subsequent rounds of replication of the mutated chromosome could lead to tumour development. RNAP, RNA polymerase

evolvability under stress, and will improve the chance of survival of E. coli. Nondividing- nonrevertant, but engineered *E.coli.* for dividing cell- stress applied- detect colonies of mutation .

CSR and SHM in transcription associated mutagenesis

Until very recently CSR and SHM were believed to be regulated by different mechanisms. However, a putative RNA editing enzyme, activation-induced cytidine deaminase (AID), has been shown to regulate both CSR and SHM in mouse and human.

The regulation of two different types of genetic alteration mechanism by AID indicates that mammals are equipped with surprisingly sophisticated and complex layers of the genetic alteration mechanisms to diversify our genomic information.Vertebrate antibody gene undergo 3 genetic alterations. As discussed about CSR and SHM they were regulated by different mechanisms

Although Somatic hypermutation (SHM) and class switch recombination (CSR) cause distinct genetic alterations at different regions of immunoglobulin genes in B lymphocytes: point mutations invariable regions and large deletions in S regions, respectively. Yet both depend on activation induced deaminase (AID), the function of which in the two reactions has been an enigma. RNA editing enzyme, activation-induced cytidine deaminase (AID), regulates both in mouse and human. Regulation of two different types of genetic alteration mechanism by AID indicates that mammals are equipped with surprisingly sophisticated and complex layers of the genetic alteration mechanisms to diversify our genomic information Mutation in Ig genes start from 150bp downstream of Ig promoter. AID –B cell specific deaminase converts cytosine to uracil in ssDNA –initiates SHM and CSR.SHM-induced GC>TA transitions and CSR induced double strand breaks lastly AID interact with DNA in stalled transcription bubble.

Case study shows Gene transcription increases DNA damage induced mutagenesis in mammalian stem cells

In mammalian stem cell ,mutation in transcribed gene underlie genetic diseases including cancer. After UV irradiation the RNA polymerase stalls at UV lesions and remains stably associated with the DNA template, forming a potential block for replication forks. *In vitro,* stalled RNA polymerases do not pose any barrier for DNA replication. Fork progression in the case of co-directional collision and only temporal pausing of replication when there is head on collision.this reveals that gene transcription during

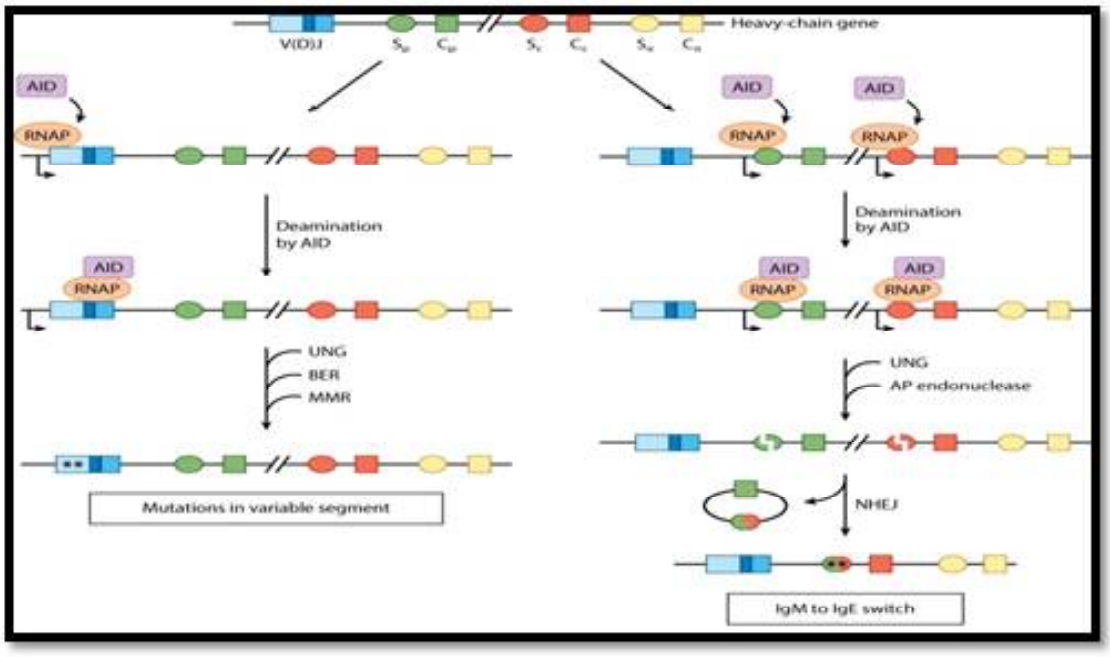

Fig. 8: A model for somatic hypermutation (SHM) and class-switch recombination (CSR) during antibody maturation. Different sequence elements are shown as rectangles or ovals of different colors. The direction of transcription is indicated by a rightward arrow. AID travels with RNAP (RNA polymerase) II as it transcribes DNA and converts cytosine to uracil on each DNA strand. Uracil is excised by uracil *N*-glycosylase (UNG) and processed by error-prone base-excision repair (BER) or by mismatch repair (MMR) to introduce point mutations, which are indicated by asterisks. To initiate CSR, double-strand breaks are mainly generated by AP endonuclease incision at UNG-generated apurinic/apyrimidinic (AP) sites. Broken ends are ligated by the nonhomologous end-joining (NHEJ) pathway, and the intervening DNA is released as a switch circle. Abbreviations: AID, activation-induced deaminase; C, constant segment (only the ì, å, áregions are shown); S, switch region preceding each C segment; V(D)J, variable segment.

exposure to UV increases both intragenic deletions and nucleotide substitution frequencies (Giel Hendriks *et al., 2008)*

The Tetracycline-controlled transactivator gene *tTA2 (Tet-off),*cloned between a CAG promoter and SV40 and BGH poly(A). The mouse *Hprt minigene, obtained from* PGK-Hprt [18], was cloned under control of the Tetracycline responsive. TRE2 promoter and -globin poly(A) signal into the pKO-LC1 vector.

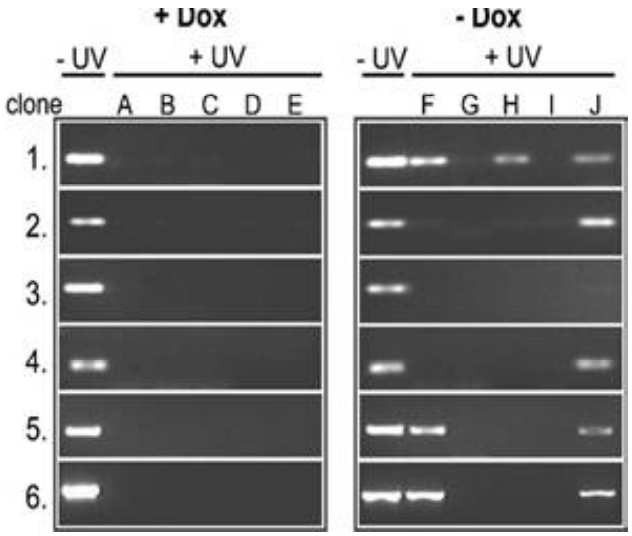

Fig. 9:PCR analysis of genomic DNA from untreated (“UV) and UV-irradiated Hprt-mg2 ES cells (+UV) to identify intragenic deletions of the *Hprt minigene*.

Cells were irradiated in the absence of transcription (6-TGR clones A–E) or presence of transcription (6-TGR clones F–J).

CONCLUSION

By inducing mutations in specific ways and then observing the phenotype of the organism the function of genes and even individual nucleotides can be determined**.** There are new factors and mechanisms involved in basal transcription that need to be identified. Identifying molecular mechanism in which R loops promote termination provide new insights to regulates gene expression at transcription level.

LITERATURE CITED

Brégeon D and Doetsch P.W.2011 Transcriptional mutagenesis: and involvement in tumour development *Nat* **11**.

Gottipati P and HelledayT.2009 Transcription-associated recombination in eukaryotes: link between transcription, replication and recombination *Mutagenesis* **24**: 203–210.

Honjo, T., Kinoshita, K. and Muramatsu, M., 2002, Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu. Rev. Immunol.* **20**:165–96.

Kim N., Abdulovic, L., A., Gealy, R., Lippert, L., M. and Robertson**,** S. J., 2007**,** Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. *DNA Repair*. 6**(9)**: 1285–1296.

Kim, H., Tomita, M. and Kanai, A., 2010, Transcription-associated mutagenesis increases protein sequence diversity more effectively than does random mutagenesis in *Escherichia coli. DNA Repair.***5**: 5p.

Morey N. J., Greene C. N. and Robertson S. J., 2000, Genetic analysis of transcription-associated mutation in *Saccharomyces cerevisiae*. *Genetics.* **154**: 109–120.

Park.C., Qian W. and Zhang G., 2012Genomic evidence for elevated mutation rates in highly expressed genes EMBO reports 13, 1123–1129.

Roa, D. and Zenkin, N., 2011, Relations between replication and transcription. *Fundamental Aspects of DNA Replication* 290-306.

Robertson S.J and Bhagwat A.S 2014. Transcription-Associated Mutagenesis *Rev. Genet.* **48**:341–59.

Williams T.M, Fabbri R.M, Reeves J.W and Crouse F.G 2005 A New Reversion Assay for Measuring All Possible Base Pair Substitutions in *Saccharomyces cerevisiae Genetics* **170:** 1423–1426.

Received on 12-09-2018 Accepted on 29-11-2018