PIBB 生物化学与生物物理进展
Progress in Biochemistry and Biophysics
2020,47(11):1162~1173
www.pibb.ac.cn

The Replication Transcription Collision-based Mutations and Evolutionary Implications*

DONG Qigeqi[1,2)**], QIAO Jia-Xin[1)**], SUN Hong-Wei[1)], FAN Li-Fei[1)], Morigen[1)***]

([1)]*State Key Laboratory of Reproductive Regulation & Breeding of Grassland Livestock, School of Life Sciences, Inner Mongolia University, Hohhot* 010070, China;

[2)]*College of Life Sciences, Inner Mongolia Agricultural University, Hohhot* 010011, China)

Abstract The replication and transcription machinery concurrently use the same DNA region as template so that the machineries inevitably collide with each other in the manner of either head-on or co-directional. Both head-on and co-directional collisions lead to a pause of replication fork, thereby DNA damage and genome instability. The head-on collision is more detrimental than the co-directional in respect of genome integrity. Here we review the resolving mechanisms and evolutionary impact of the replication-transcription collisions. The rate of nonsynonymous (amino-acid-changing) mutations on the lagging is higher relative to that on the leading strand and the high frequency mutagenesis in genes on the lagging strand is dependent on transcriptions and gene sizes, thus faster adaptive mutations occur on the lagging strand. Highly transcribing of head-on oriented genes increases the mutation rates responding to stress during active replication. It is likely that the replication-transcription collision no matter in the head-on or co-directional mode is a driving force for adaptive evolution.

Key words replication-transcription collisions, resolving mechanisms, collision-based mutations, evolutionary implication
DOI: 10.16476/j.pibb.2020.0107

Chromosome replication occurs once and only once per cell cycle to make sure that each daughter cell receives one set of complete genetic information[1], including a set of complex processes of initiation, elongation and termination. Replication initiates at the right time of cell cycle, through interaction between the initiator protein and the origin for replication, subsequently followed by the recruitment of DNA helicase, primase and DNA polymerase[2]. Two replisomes formed at the same origin move from the origin to the opposite directions so that replication is bi-directional[3]. Elongation of replication is highly sensitive to any disruptions, which may lead to genome instability[4]. During the elongation of replication, parental DNA duplex is separated by DNA helicase in front of moving replisome, and each single-stranded DNA serves as a template for synthesis of a new DNA chain. Transcription is also a highly regulated process, which is facilitated by transcription machinery (RNA polymerase, RNAP). Unrepaired errors in DNA can be fixed by transcription-coupled repair system during transcription[5-9]. DNA replication and transcription have to be organically combined with each other[10] to fulfill their mission to maintain the heredity and life[11].

Replication and transcription concurrently use the same DNA region as a template, then replisome and RNAP collide with each other inevitably, in

* This work was supported by grants from The National Natural Science Foundation of China (31700034) and Natural Science Foundation of Inner Mongolia Autonomous Region (2016BS0309).

** These authors contributed equally to this work.
*** Corresponding author.
Tel: 86-471-4992242, E-mail: morigenm@hotmail.com

Received: April 24, 2020 Accepted: July 10, 2020

manner of either head-on when they move in opposite directions along the same DNA template or co-directional when they move in the same direction but at different speeds. Both head-on and co-directional collisions lead to a pause of replication fork and subsequent DNA damage and genome instability[12]. The head-on collision causes more serious DNA damage relative to the co-directional collision does. These DNA damages leave the genome susceptible to mutations[13-15]. Replication-transcription collisions cause replication fork arrests, DNA damages, mutagenesis and chromosome rearrangements[16-18]. The regulatory mechanisms which deal with the DNA damages caused by replication-transcription collisions have been revealed[19-20]. It has been found that auxiliary helicases Rep, UvrD and DinG in *E. coli* and Rrm3 in yeast, transcription-repair coupling factor Mfd, transcription factor DksA and GreA/B are involved in preventing or resolving the collisions between replication and transcription[21-22].

1 Patterns of the replication−transcription collisions

Either in prokaryotes or eukaryotes, DNA replication and transcription are crucial cellular processes. Although there is a separation between replication and transcription of many genes temporally and spatially, in some cases these two processes inevitably occur simultaneously in the same DNA region, leading to transcription-replication collisions[23-24], there are two kinds of collisions between replication and transcription[17]. One is the head-on collision which occurs when replisome and RNAP move toward each other in the opposite directions along the same template. The other is the co-directional collision that happens when the fast-moving replisome tracks after the slowly-proceeding RNAP[25] (Figure 1a). In *E. coli*, a replisome moves about 1 000 bp long per second on average[26] while a RNAP processes about only 50 nucleotides per second on average[27], suggesting that the rate of replication is approximately 20-fold faster than that of transcription[28].

In the head-on collision, the RNAP is dislodged from the template to allow the replisome continue[22]. After passage of the replisome, the dislodged RNAP can be reloaded on the template to restart transcription. Such dislodgement and recruitment of the RNAP are time-consuming. By contrast, in both fast growing *E. coli* and bacteriophage T4, resolving of the co-directional collision needs only half time of that of head-on collision (1.7 s). In slow growth condition, the effect is not necessarily notable[29]. It is also found that the progression of head-on oriented replisome is severely inhibited by RNAP, but the co-directional oriented replisome is not significantly affected by transcription collision[16]. It is likely that the co-directional replisome directly contacts with RNAP in a limited range, for head-on collision, the area of direct physical interaction is larger[28]. And in human cells, DNA damage and checkpoint activation increase in head-on collisions[24]. Therefore, head-on collision is more harmful for genome stability than co-directional events both in prokaryotes and eukaryotes[19, 28]. It has been also suggested that replisome might be stalled for the collisions in both head-on and co-directional manner[30-31]. Recently, it is found that Yra1-bound RNA-DNA hybrids is a cause of transcription-replication collisions in a mode of orientation-independent and subsequent instability of telomere[32].

Chromosome replication in eukaryotes initiates simultaneously at multiple replication origins (Figure 1b) while it occurs at a single site in prokaryotes, increasing the complexity of replication-transcription collisions[33]. Two replisomes originated neighbouring origins move toward each other and collide inevitably. In turn, multiple moving replisomes in opposite direction promote the probability of replication-transcription collisions since the main obstacle to the replisome progress is transcription (Figure 1b) [34-35]. Also inefficient transcription termination, the R loop made in transcription and chromatin organized are obstacles to the replisome progress[36]. Similar to proteins found in prokaryotes, a pile of proteins in eukaryotes are involved in removing transcription barriers and nascent RNA, which are essential for replication process[33].

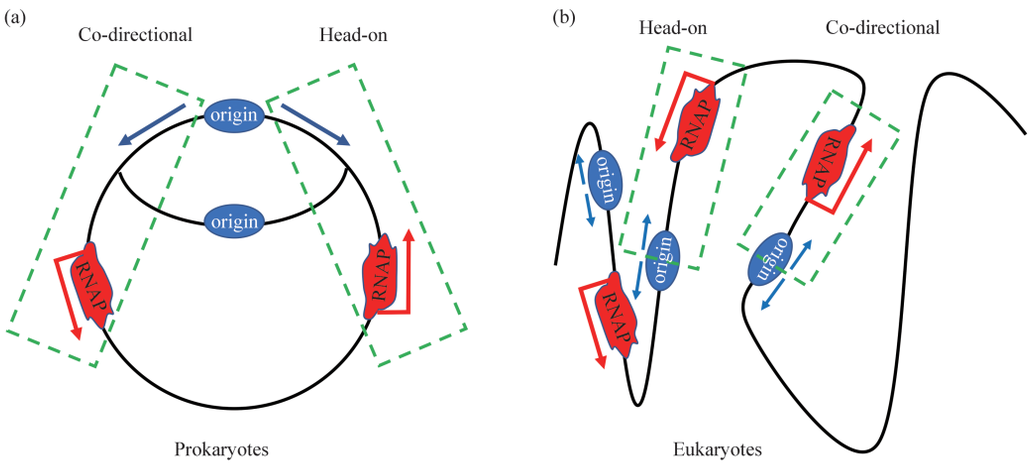

Fig. 1 Patterns of the replication−transcription collisions[37]

Chromosome replication initiates at single origin in prokaryotes (a) while these occur at multiple origins in eukaryotes (b). In both prokaryotes and eukaryotes, head-on collision occurs when replisome and RNAP move toward each other in opposite directions along the same template. Co-directional collision happens when the fast-moving replisome tracks after the slowly-proceeding RNAP. Arrows indicate direction of moving replisomes and RNAP, origin for replication (blue) and RNAP (red) are as shown.

2 Resolving mechanisms for the replication−transcription collisions

Three mechanisms have been proposed to resolve the collisions between replisomes and RNAPs. First, a RNAP and its nascent transcript are dislodged from the DNA template to allow passage of the replisome encountered[38-39]; second, a replisome passes the transcribing RNAP without displacing it from the template[29]; third, if all else fails, replisome can be rebuilt[39]. The blocked fork processing simply facilitate replication reinitiation upstream of the block, providing a second chance for a replisome to proceed successfully through the block[40].

In the situations of nutrient deprivation, proteotoxic stress and the presence of natural antibiotics, the translation can be inhibited, leading to an increased possibility of the transcription elongation complex (ECs) backtracking and subsequent DNA double-strand breaks (DSBs) [17]. During RNAP backtracking the 3'-OH terminus of RNA detaches from the catalytic site and is extruded into the secondary channel of the RNAP which is the substrate-binding pore, resulting in transient or permanent EC inactivation. Bacteria employ various strategies to avoid replisome collisions with backtracked RNAP. One is continuous translation that prevents RNAP backtracking; or if translation is disrupted, the transcription elongation factors would either prevent backtracking or reactivate backtracked ECs to suppress DSBs[17]; alternatively, termination factors could remove arrested ECs to maintain genomic stability[17]. The Rho factor suppresses backtracking by terminating transcription which uncouples from translation[41-42]. When Rho is not sufficient, the backup termination factor Mfd disrupts ECs arrested by several mechanisms[42-43]. The transcript cleavage factors (GreA and GreB)-mediated system as an additional backup pathway suppresses DSBs by restarting backtracked ECs[44]. Thus, the elongation factors GreA can prevent DNA damage by suppressing RNAP backtracking[17]. Furthermore, the Gre factors are more helpful for transcriptional fidelity than recombinant repair[45]. Under the condition of amino acid starvation, over-expression of GreA compensates for the lack of DksA in promoting the process of replicating[46]. The transcription factor DksA functions differently in resolution of the replication-transcription collisions[17, 46], being independent on its transcription initiation activity but dependent on its transcription elongation activity. The structure of DksA is well conserved, containing a coiled-coil motif and two invariant Asp residues in the C terminus. By protruding the coiled-coil motif into the RNAP secondary channel to coordinate a ppGpp bound Mg^{2+} ion with the Asp residues, DksA stabilizes the ppGpp-RNAP complex. One or two of these two conserved Asp residues are indispensable

for the effect of DksA on transcription initiation since mutation of aspartic acid residues to asparagines does not affect the role of DksA on preventing replication arrest upon starvation. This may suggest that DksA might facilitate replication by preventing transcriptional pausing but not affecting transcription initiation[46]. Finally, it is proposed that DksA ensures replication completion by dislodging RNAPs during the head-on collisions[46] by interacting with RNAP in the secondary channel. *In vitro* experiments show that Mfd dislocates RNAPs from the template to allow replisome continue when collisions occur in the head-on mode[38]. GreA and GreB prevent transcriptional pausing *via* reactivating backtracking ECs through its transcript cleavage reaction in the RNAP catalytic site to generate a new 3'-OH terminus[47-48].

In vitro, replisomes generated by bacteriophage T4 replication proteins pass a molecule of the *E. coli* RNAP moving in the same direction[29]. Thus, a model is proposed to explain how a replisome passes the replication-transcription collision without dislodging the transcribing RNAP and the nascent transcript from the template[29]. The model suggests that the RNAP ternary transcription complex has at least two DNA-interacting domains, each of which detaches from the DNA template individually without causing collapse of the complex[29]. When a replisome encounters the RNAP ternary complex from behind, the proximal DNA binding domain of the RNAP complex detaches from the template, allowing the replisome partially invade into the complex. As the replisome continues into the RNAP complex, the proximal DNA binding site re-attaches on the template while the distal binding site detaches from the template, retaining the nascent RNA in the RNAP complex. With the passage of replisome, the nascent-RNA aided reassembly of the RNAP ternary complex occurs[29]. It is obvious that the mechanism is an important way for resolving the replication-transcription collisions in the co-directional mode.

In bacteria, blocked replisome losses its replicative function for a while[49-50]. Loss of replicative function of the replisome leads to recruitment of restarting proteins PriA or PriC in *E. coli*[51], the Pri protein recognizes stalling replication fork and reassembles replisome outside of *oriC*[52]. Indeed, cells lacking PriA are very sick in rich media[53-54] and a mutation in RNAP that reduces backtracking can suppress the sickness[17, 55]. Therefore, it is clear that replisome collapses due to replication-transcription collision drastically requires replisome reloading[56] even in the presence of the other mechanisms that may resolve the collisions although the detailed mechanisms are not known.

Replication-transcription collisions have been described in many organisms from bacteria to human cells[33]. Long-term stagnation of replication activates the cell cycle checkpoints as a DNA damage response and may eventually lead to the formation of double-strand breaks[57]. Such DNA damages caused by stalling replicons *in vivo* is mostly because of RNAP[58]. In principle, stalled RNAP might be removed from chromatin to avoid DNA damages[59]. The removal of RNAPII after replication stress in budding yeast involves replication checkpoints, chromatin remodeling complex INO80C and PAF transcription complex[60]. Indeed, the replisome progress requires cofactors such as chromatin remodeling agents and histone chaperones *in vitro*[61] and RNAP removal results in prevention of replication-transcription collisions[62]. In *Drosophila* and human cells, depletion of histone H1 causes replication stress and DNA damages resulted from replication-transcription collisions[63]. Senataxin may be recruited as a DNA-RNA helicase to resolve replication-transcription collisions, or it may promote the release of RNAPII at the collision sites to resolve the collisions[59]. Experimentally, T7 helicase strongly interacts with non-replicating T7 DNA polymerase (DNAP) at the replisome. This DNAP and helicase can indeed remove the stalled transcriptional elongation complex, and then use RNA transcripts as primers to initiate replication[64].

When DNA replication is disrupted, cells trigger a replication stress response[65], activate ATR checkpoint kinase and subsequently stall progress of the cell cycle[66] to solve the problems in replication. High frequency of R-loop formation during transcription also activate ATR kinases for survival, depending on MUS81 endonuclease. The ATR activation protects the genome integrity by inhibiting transcription-replication collisions, promoting replisome recovery and enforcing G2/M cell cycle arrest[67]. ATM and ATR are the two central kinases in the DNA damage response pathway and also help limit R-loop accumulation[68]. The cellular response for transcription-replication collision depends on the basal activity of ATR kinase without inducing

excessive ATR activation. And the specific abrogation of the transcription-replication collision response causes DNA damage in mitosis and subsequently promote chromosomal instability and cell death[69]. The ATR homolog in yeast, Mec1, uncouples the transcribed gene from the nuclear pore by phosphorylating the nucleoporin Mlp1 to neutralize the formation of topological tension[70]. The effector Mrc1, a downstream component of the Mec1/ATR pathway, is phosphorylated by Hog1. This phosphorylation delays onset of the origins, thereby preventing transcription-replication collision response and transcription-related recombination[71]. Mec1 also responds to replication pressure by phosphorylating Maf1[72].

3 Transcription-coupled DNA damage repair

DNA template damage resulting from exogenous (*e.g.* UV and gamma irradiation) and endogenous (*e.g.* active oxidized substances produced by respiration) stress which blocks the elongation of RNAP[73]. A more detailed pathway in which the UvrD protein pulls RNAP backwards in replication-transcription collision to facilitate DNA repair is verified[74]. UvrD is a helicase required for nucleotide excision repair[75], binding RNAP *in vitro* and *in vivo* to promote RNAP backtracking at the DNA site damaged[74]. It is proposed that UvrD and NusA cooperatively pull the RNAP stalled with the energy from ATP hydrolysis to expose the DNA lesion where UvrABC recruit, and repair the damage with assistance of Pol I and ligase[74] (Figure 2a). It is therefore suggested that RNAP is a global genome surveillance vehicle to screen the damages along chromosomes during transcription.

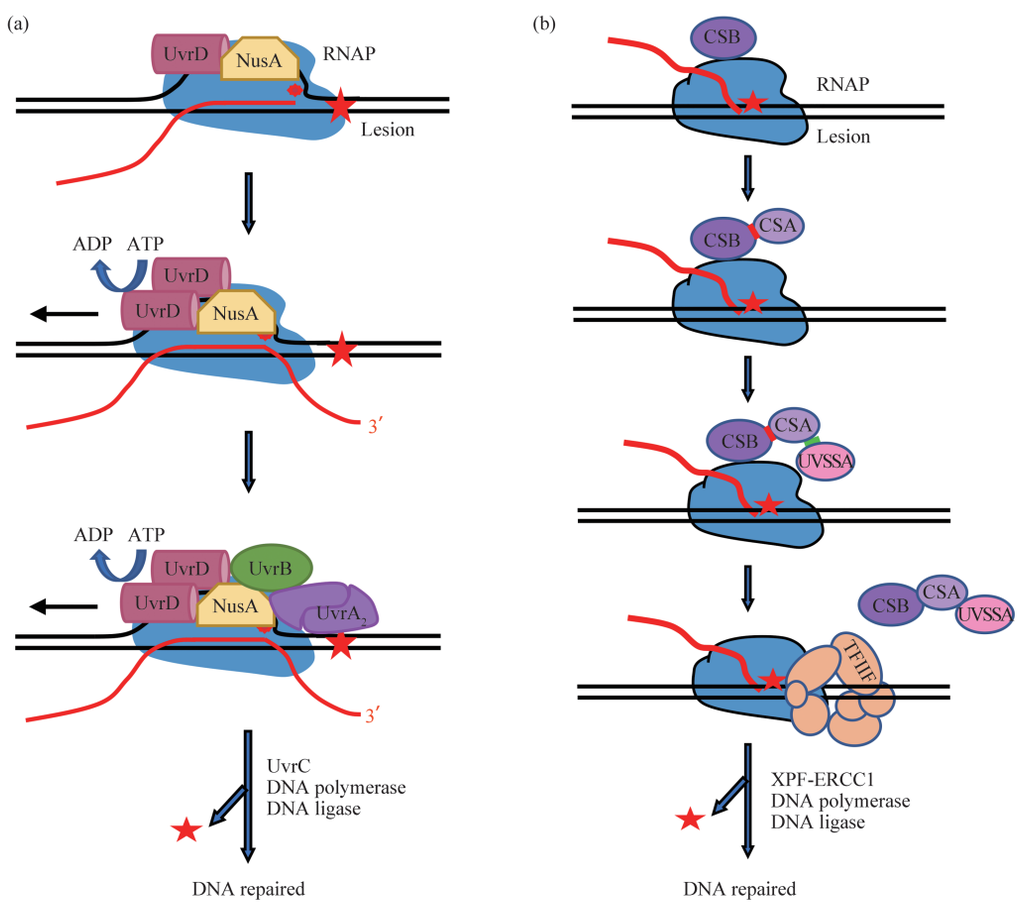

Fig. 2　Models for the transcription-coupled repair pathway[48,55,74]

Mfd-independent[74], transcription-coupled repairs in *E. coli* (a), and human (b) TCR repair[48,55] are as shown. Proteins involved and recruitment order of the proteins in the process are as indicated.

Alternatively, after head-on collision between the replisome and stalled RNAP, Mfd dissociates RNAP from the template, thus promotes replisome restart *in vitro*[38]. In an updated model for transcription coupled repair (TCR), RNAP stalls at damage sites of template, and the stalled complex recruits Mfd rapidly, then the loaded Mfd releases the nascent transcript and dissociates RNAP from the template due to its translocase action. Further, Mfd recruits $UvrA_2B_1$ to the damage sites by interacting with UvrA while RNAP is restrained in the Mfd – DNA complex. The recruitment is coupled with loading of UvrB onto the transcription-blocking damage and releasing of RNAP, Mfd and UvrA. Then, UvrC binds to the UvrB-DNA complex and makes dual incisions, which is followed by displacement of the oligomer excised and UvrB and UvrC from the site repairing by the UvrD helicase[76-77]. And transcription-coupled nucleotide excision repair (TC-NER) is proposed for the increased mutagenesis of lagging strand genes due to the interplay between replication and transcription on lagging-strand genes. This process, at least partially, increases the lagging-strand gene mutations[78].

In human cells, transcription coupled repair (TCR) removes DNA damages as that in *E. coli* at actively transcribing genes, but the number of proteins involved and their interactions are more complicated[79]. Transcription stalling at the DNA lesions not only leads to gene damage but also to overall shutdown of transcription[80]. Highly specialized TCR pathway repairs the lesion blocked transcription, thereby restoring transcription and gene expression[81]. When RNAP is stalling at the DNA lesion, CSB is recruited first through interacting with RNAPIIo for further recruitment of TCR component. Such recruitment of CSB leads to its own conformational change to stimulate the CSA recruitment through CIM (the newly discovered CSA interaction motif). Then the XPC and DDB2 damage recognition proteins are recruited to confirm the lesion DNA. Further, UVSSA loads to the RNAPIIo complex at the lesion site by targeting CSA and subsequently regulates the TFIIH complex recruitment. In the way of CSB and CSA stimulation, UVSSA is the key factor for recruiting TFIIH complex. The TFIIH loaded replaces RNAPIIo and CSB/CSA/UVSSA complex, then with the cooperation of XPF-ERCC1 heterodimer, DNA polymerase and DNA ligase, repairs the lesion[82] (Figure 2b).

4 The replication–transcription collision–based mutations

Mutations cause genetic diseases and also drive evolution by changing gene coding sequences or non-coding elements that regulate gene expression. There are many mechanisms leading to mutagenesis: DNA replication errors, error-prone repair, transcription-related mutagenesis, replication stall-mediated template switching[83]. When transcription is active, the gene orientation can be switched between the two collision types. And the gene orientation can increase or decrease the mutation rate in a gene-specific manner[84]. This mechanism is driven by the inevitable collision between DNA replication and transcription machinery. Transcription-replication collisions have different consequences for the genes encoded on the leading and lagging strands of the replisome[85]. The mutation rate on the lagging strand is higher. The differences in the mutagenesis rates of head-on (lagging strand) genes and co-directional (leading strand) genes are transcription-dependent, indicating that the collision between replication and transcription is the driving force of mutations, and these collisions will greatly increase possibility of the adaptive structural variation of the encoded protein[78, 84]. The increase in rate of positive mutagenesis depends on the R loop, which plays a key role in replication-transcription collisions[86]. The increased mutation rate of head-on genes may provide an adaptive advantage, but not all genetic changes induced by collisions are adaptive[84]. In head-on genes, different sequence contexts are suggested to be related with a higher rate of spontaneous mutations, thereby promoting their accelerated evolution[87]. However, the head-on genes accelerate evolution independently of sequence context[88]. It would be noted that mutations caused by the collisions can change gene expression patterns[89].

Occurrence of mutations is probabilistic, associating with a variety of physicochemical parameters, the latter depends on environment and the cell physiology[90-91]. The cell-dependent physiological and environment-dependent physical and chemical parameters affect the frequency, nature, and location of mutations[92], increasing the probability of genetic adaptability[90-91]. Indeed, physiological adaptation promotes evolution in an environment-dependent

manner. Also increase in the global mutation rate during stress promotes the chance of adaptive mutations, thus the cell can deal with the environmental changes[85].

5 The replication−transcription collisions is a driving force for adaptive evolution

In replication-transcription collisions, co-directed collisions occur in highly transcribed regions or RNAP backtracking sites, causing replication to stall and eventually lead to replication to restart[56] and the head-on collisions cause more serious consequences such as replisome disintegration and DNA breaks[58]. The orientation of genes determines the severity of replication-transcription collisions[23]. Unequal mutagenesis on lagging strand and leading strand[93-94] results in differences in evolution rates of genes on the lagging and the leading strands[95]. Interestingly, over a half of genes (mainly tRNA genes, protein coding genes) aligned co-directionally with replication in many bacteria[96]. In *Bacillus subtilis*, transcription of 75% genes is co-directional oriented with replication on leading strand and such a co-directional orientation for highly expressed genes might confer an advantage by reducing blockage of replisome progression[97-98] (Figure 3a). The co-directional orientation bias has also been shown to be prominent for highly transcribed genes in most bacteria and eukaryotes[8].

However, a large number of co-directional genes can be inverted to head-on genes during evolution, indicating that evolutionary pressure will prevent complete co-directional genes and will drive the maintenance of the head-on genes, increase the frequency of head-on replication-transcription collisions[86]. These results show that spontaneous gene inversion can increase bacterial evolutionary ability through head-on replication-transcription collisions[23, 99]. Interestingly, rate of synonymous (or silent) mutations on the lagging strand is 2% higher than that on the leading strand while the rate of nonsynonymous (amino-acid-changing) mutations on the lagging stand is 42% higher relative to that of the genes on the leading strand[84]. Convergent mutations are in 24% of core genes on the lagging strand compared to only 11% of the genes on the leading strand, and the increased rates of mutagenesis in the lagging strand genes (head-on oriented genes) is in manner of transcription-dependent. Clearly, both nonsynonymous and convergent mutations in the head-on oriented genes are high, suggesting that faster adaptive mutations occurs in the head-on oriented genes[84] (Figure 3b). Furthermore, genes on the lagging strand are shorter on average while those on the leading strand are 48% larger; especially, only 24% of genes coding for proteins with more than 200 amino acids are on the lagging strand whereas 48% of genes coding proteins exceeded 200 amino acids are on the leading strand[84]. The data suggest that selection for genes on the lagging strand to be shorter is likely to decrease the possibility of head-on collision. Indeed, the higher nonsynonymous mutation rates occur in longer genes on the lagging strand[84]. And an increased rate of spontaneous mutations in the head-on oriented genes is believed to drive an accelerated rate of evolution[85].

Two models have been proposed to explain the reasons behind the head-on orientation of genes. One model explains that the head-on orientation of genes is harmful for genome stability by increasing mutagenesis[28] while the other suggests that the head-on orientation is the reason of increased rate of gene transcript truncations[97]. In *B. subtilis*, *E. coli* and *Saccharomyces cerevisiae*, it is found that genomic instability is likely resulted from the head-on collisions[13, 100-101]. Further, the strong co-orientation bias of transcription is due to selective pressure for processive, efficient, and accurate replication[102]. During normal growth, head-on oriented genes are transcribed to some degree and can be highly induced responding to stress exposure[103-105]. Actively transcribing of the head-on oriented genes under stress causes severe collisions during active replication and increases the mutation rate of head-on oriented genes, providing the chances for cells to survive in a changing environment. This also creates possibility of acquiring lethal mutations[85] and subsequent cell death. It seems that transcription orientation of genes is a gene-specific manner to increase or decrease mutation rates[84]. Evidence discussed above, therefore, allow us to assume that the replication-transcription collisions no matter in the head-on or co-directional mode might be a driving force for adaptive evolution.

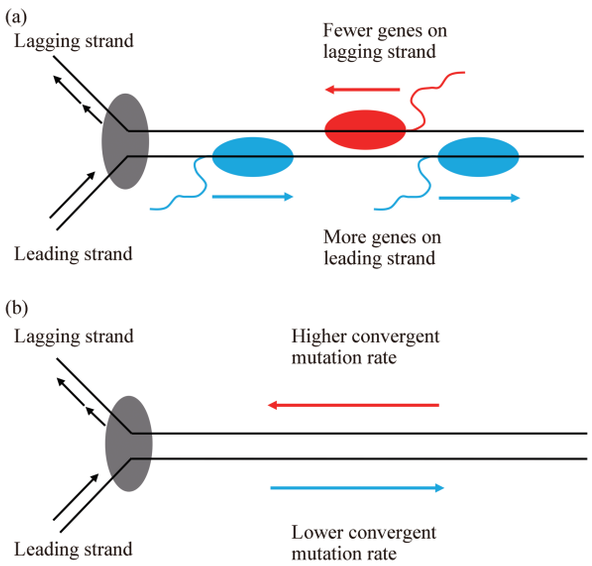

Fig. 3 A model for adaptive evolution through the replication–transcription collisions

(a) Most genes in bacteria are encoded on the leading strand of replication to avoid the potentially harmful head-on collisions that occur between the replication and transcription machineries when genes are encoded on the lagging strand. (b) Lagging strand genes are transcribed in the head-on orientation with respect to DNA replication, leading to stalled replication which shape genomes and influence evolution. The head-on encounters between replication and transcription increase convergent mutagenesis in lagging strand genes, indicating faster adaptive evolution in many genes in the head-on orientation. Ovals indicate direction of moving replisome (dark gray) and RNAPs (blue and red) with arrows.

6 Prospective

It is most likely that all organisms have a large number of genes on the lagging strand. The fact indicates that replication-transcription collisions in head-on mode is unavoidable, and this is a universal strategy to link gene expression and evolution rate under selection. It is necessary to measure the number of genes on the lagging strand in a number of organisms, as well as rates of mutagenesis and evolutionary impact. Hopefully we can develop some easier and economical methods to show the biological insight into adaptation of organisms by the replication-transcription collisions.

References

[1] Boye E, Løbner-Olesen A, Skarstad K. Limiting DNA replication to once and only once. EMBO Rep, 2000, **1**(6): 479-483

[2] Klein A, Bonhoeffer F. DNA replication. Annu Rev Biochem, 1972, **41**(10): 301-332

[3] Prescott D M, Kuempel P L. Bidirectional replication of the chromosome in *Escherichia coli*. Proc Natl Acad Sci USA, 1972, **69**(10): 2842-2845

[4] Mirkin E V, Mirkin S M. Replication fork stalling at natural impediments. Microbiol Mol Biol R, 2007, **71**(1): 13

[5] Donahue B A, Yin S, Taylor J S, *et al*. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. Proc Natl Acad Sci USA, 1994, **91**(18): 8502-8506

[6] Chambers A L, Smith A J, Savery N J. A DNA translocation motif in the bacterial transcription--repair coupling factor, Mfd. Nucleic Acids Res, 2003, **31**(22): 6409-6418

[7] Lainé J P, Egly J M. When transcription and repair meet: a complex system. Trends Genet, 2006, **22**(8): 430-436

[8] Kuo H K, Krasich R, Bhagwat A S, *et al*. Importance of the tmRNA system for cell survival when transcription is blocked by DNA-protein cross-links. Mol Microbiol, 2010, **78**(3): 686-700

[9] Daulny A, Tansey W P. Damage control: DNA repair, transcription, and the ubiquitin-proteasome system. DNA Repair, 2009, **8**(4): 444-448

[10] Vulliémoz D, Cordey S, Mottet-Osman G, *et al*. Nature of a paramyxovirus replication promoter influences a nearby transcription signal. J Gen Virol, 2005, **86**(1): 171-180

[11] Meyers R A. Encyclopedia of Melecular Cell Biology and Molecular Medicine. 2nd. Weinheim, Germany: John Wiley and Sons, 2004: 499-475

[12] Lin Y L, Pasero P. Interference between DNA replication and transcription as a cause of genomic instability. Curr Genomics, 2012, **13**(1): 65-73

[13] Srivatsan A, Tehranchi A, MacAlpine D M, *et al*. Co-orientation of replication and transcription preserves genome integrity. Plos Genet, 2010, **6**(1): e1000810

[14] Poveda A M, Le Clech M, Pasero P. Transcription and replication: breaking the rules of the road causes genomic instability. Transcription, 2010, **1**(2): 99-102

[15] Hendriks G, Jansen J G, Mullenders L H, *et al*. Transcription and replication: far relatives make uneasy bedfellows. Cell Cycle, 2010, **9**(12): 2300-2304

[16] French S. Consequences of replication fork movement through transcription units *in vivo*. Science, 1992, **258**(5086): 1362-1365

[17] Dutta D, Shatalin K, Epshtein V, *et al*. Linking RNA polymerase backtracking to genome instability in *E*. coli. Cell, 2011, **146**(4): 533-543

[18] Sollier J, Stork C T, García-Rubio M L, *et al*. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. Mol Cell, 2014, **56**(6): 777-785

[19] Prado F, Aguilera A. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. EMBO J, 2005, **24**(6): 1267-1276

[20] Rudolph C J, Dhillon P, Moore T, *et al*. Avoiding and resolving conflicts between DNA replication and transcription. DNA Repair,

2007, **6**(7): 981-993

[21] Pomerantz R T, O'Donnell M. Polymerase trafficking: a role for transcription factors in preventing replication fork arrest. Transcription, 2010, **1**(3): 136-139

[22] Merrikh H, Zhang Y, Grossman J D, *et al*. Replication-transcription conflicts in bacteria. Nat Rev Microbiol, 2012, **10**(7): 449

[23] Lang K S, Merrikh H. The clash of macromolecular titans: replication-transcription conflicts in bacteria. Annu Rev Microbiol, 2018, **72**: 71-88

[24] Hamperl S, Bocek M J, Saldivar J C, *et al*. Transcription-replication conflict orientation modulates R-Loop levels and activates distinct DNA damage responses. Cell, 2017, **170**(4): 774-786

[25] Soultanas P. The replication-transcription conflict. Transcription, 2011, **2**(3): 140

[26] Hirose S, Hiraga S, Okazaki T. Initiation site of deoxyribonucleotide polymerization at the replication origin of the *Escherichia coli* chromosome. Mol Genet Genomics, 1983, **189**(3): 422-431

[27] Gotta S L, Jr M O, French S L. rRNA transcription rate in *Escherichia coli*. J Bacteriol, 1991, **173**(20): 6647

[28] Mirkin E V, Mirkin S M. Mechanisms of transcription-replication collisions in bacteria. Mol Cell Biol, 2005, **25**(3): 888-895

[29] Liu B, Mei L W, Tinker R L, *et al*. The DNA replication fork can pass RNA polymerase without displacing the nascent transcript. Nature, 1993, **366**(6450): 33

[30] Elías-Arnanz M, Salas M. Bacteriophage φ29 DNA replication arrest caused by codirectional collisions with the transcription machinery. EMBO J, 1997, **16**(18): 5775-5783

[31] Elíasarnanz M, Salas M. Resolution of head-on collisions between the transcription machinery and bacteriophage phi29 DNA polymerase is dependent on RNA polymerase translocation. EMBO J, 1999, **18**(20): 5675-5682

[32] García-Rubio M, Aguilera P, Lafuente-Barquero J, *et al*. Yra1-bound RNA-DNA hybrids cause orientation- independent transcription-replication collisions and telomere instability. Gene Dev, 2018, **32**(13-14): 965-977

[33] Brambati A, Colosio A, Zardoni L, *et al*. Replication and transcription on a collision course: eukaryotic regulation mechanisms and implications for DNA stability. Front Genet, 2015, **6**: 166

[34] Bermejo R, Lai M S, Foiani M. Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. Mol Cell, 2012, **45**(6): 710-718

[35] Gaillard H, Aguilera A. Transcription as a threat to genome integrity. Annu Rev Biochem, 2016, **85**: 291-317

[36] Santos-Pereira J M, Aguilera A. R loops: new modulators of genome dynamics and function. Nat Rev Genet, 2015, **16**(10): 583-597

[37] Hamperl S, Cimprich K A. Conflict resolution in the genome: how transcription and replication make it work. Cell, 2016, **167**(6): 1455-1467

[38] Pomerantz R T, O'Donnell M. Direct restart of a replication fork stalled by a head-on RNA polymerase. Science, 2010, **327**(5965): 590-592

[39] Mcglynn P, Savery N J, Dillingham M S. The conflict between DNA replication and transcription. Mol Microbiol, 2012, **85**(1): 12-20

[40] Payne B T, van Knippenberg I C, Bell H, *et al*. Replication fork blockage by transcription factor-DNA complexes in *Escherichia coli*. Nucleic Acids Res, 2006, **34**(18): 5194

[41] Richardson J P. Preventing the synthesis of unused transcripts by Rho factor. Cell, 1991, **64**(6): 1047-1049

[42] Roberts J W, Shankar S, Filter J J. RNA polymerase elongation factors. Annu Rev Microbiol, 2008, **62**(1): 211

[43] Washburn R S, Wang Y, Gottesman M E. Role of *E. coli* transcription-repair coupling factor Mfd in Nun-mediated transcription termination. J Mol Biol, 2003, **329**(4): 655

[44] Borukhov S, Lee J, Laptenko O. Bacterial transcription elongation factors: new insights into molecular mechanism of action. Mol Microbiol, 2005, **55**(5): 1315

[45] Sivaramakrishnan P, Sepúlveda L A, Halliday J A, *et al*. The transcription fidelity factor GreA impedes DNA break repair. Nature, 2017, **550**(7675): 214-218

[46] Tehranchi A K, Blankschien M D, Zhang Y, *et al*. The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. Cell, 2010, **141**(4): 595-605

[47] Artsimovitch I, Landick R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc Natl Acad Sci USA, 2000, **97**(13): 7090-7095

[48] Marr M T, Roberts J W. Function of transcription cleavage factors GreA and GreB at a regulatory pause site. Mol Cell, 2000, **6**(6): 1275

[49] Marians K J, Hiasa H, Kim D R, *et al*. Role of the core DNA polymerase III subunits at the replication fork. Alpha is the only subunit required for processive replication. J Biol Chem, 1998, **273**(4): 2452-2457

[50] McGlynn P, Guy C P. Replication forks blocked by protein-DNA complexes have limited stability *in vitro*. J Mol Biol, 2008, **381**(2): 249-255

[51] Heller R C, Marians K J. Replisome assembly and the direct restart of stalled replication forks. Nat Rev Mol Cell Biol, 2006, **7**(12): 932-943

[52] Michel B, Sandler S J. Replication restart in bacteria. J Bacteriol, 2017, **199**(13)

[53] Lee E H, Kornberg A. Replication deficiencies in priA mutants of *Escherichia coli* lacking the primosomal replication n' protein. Proc Natl Acad Sci USA, 1991, **88**(8): 3029-3032

[54] Nurse P, Zavitz K H, Marians K J. Inactivation of the *Escherichia coli* priA DNA replication protein induces the SOS response. J Bacteriol, 1991, **173**(21): 6686-6693

[55] Mahdi A A, Buckman C, Harris L, *et al*. Rep and PriA helicase activities prevent RecA from provoking unnecessary

recombination during replication fork repair. Genes Dev, 2006, **20**(15): 2135-2147

[56] Merrikh H, Machón C, Grainger W H, *et al*. Co-directional replication-transcription conflicts lead to replication restart. Nature, 2011, **470**(7335): 554-557

[57] D'Alessandro G, d'Adda di Fagagna F. Transcription and DNA damage: holding hands or crossing swords?. J Mol Biol, 2017, **429**(21): 3215-3229

[58] Mangiameli S M, Merrikh C N, Wiggins P A, *et al*. Transcription leads to pervasive replisome instability in bacteria. Elife, 2017, **6**: e19848

[59] Gómez-González B, Aguilera A. Transcription-mediated replication hindrance: a major driver of genome instability. Genes Dev, 2019, **33**(15-16): 1008-1026

[60] Poli J, Gerhold C B, Tosi A, *et al*. Mec1, INO80, and the PAF1 complex cooperate to limit transcription replication conflicts through RNAPII removal during replication stress. Genes Dev, 2016, **30**(3): 337-354

[61] Devbhandari S, Jiang J, Kumar C, *et al*. Chromatin constrains the initiation and elongation of DNA replication. Mol Cell, 2017, **65**(1): 131-141

[62] Felipe-Abrio I, Lafuente-Barquero J, García-Rubio M L, *et al*. RNA polymerase II contributes to preventing transcription-mediated replication fork stalls. EMBO J, 2015, **34**(2): 236-250

[63] Bayona-Feliu A, Casas-Lamesa A, Reina O, *et al*. Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. Nat Commun, 2017, **8**(1): 283

[64] Sun B, Singh A, Sultana S, *et al*. Helicase promotes replication re-initiation from an RNA transcript. Nat Commun, 2018, **9**(1): 2306

[65] Zeman M K, Cimprich K A. Causes and consequences of replication stress. Nat Cell Biol, 2014, **16**(1): 2-9

[66] Técher H, Koundrioukoff S, Nicolas A, *et al*. The impact of replication stress on replication dynamics and DNA damage in vertebrate cells. Nat Rev Genet, 2017, **18**(9): 535-550

[67] Matos D A, Zhang J M, Ouyang J, *et al*. ATR protects the genome against R Loops through a MUS81-triggered feedback loop. Mol Cell, 2020, **77**(3): 514-527

[68] Barroso S, Herrera-Moyano E. The DNA damage response acts as a safeguard against harmful DNA-RNA hybrids of different origins. EMBO Rep, 2019, **20**(9): e47250

[69] Shao X, Joergensen A M, Howlett N G, *et al*. A distinct role for recombination repair factors in an early cellular response to transcription-replication conflicts. Nucleic Acids Res, 2020, **48**(10): 5467-5484

[70] Bermejo R, Capra T, Jossen R, *et al*. The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. Cell, 2011, **146**(2): 233-246

[71] Duch A, Felipe-Abrio I, Barroso S, *et al*. Coordinated control of replication and transcription by a SAPK protects genomic integrity. Nature, 2013, **493**(7430): 116-119

[72] Nguyen V C, Clelland B W, Hockman D J, *et al*. Replication stress checkpoint signaling controls tRNA gene transcription. Nat Struct Mol Biol, 2010, **17**(8): 976-981

[73] Tornaletti S, Hanawalt P C. Effect of DNA lesions on transcription elongation. Biochimie, 1999, **81**(2): 139-146

[74] Epshtein V, Kamarthapu V, McGary K, *et al*. UvrD facilitates DNA repair by pulling RNA polymerase backwards. Nature, 2014, **505**(7483): 372-377

[75] Kumura K, Sekiguchi M. Identification of the uvrD gene product of *Escherichia coli* as DNA helicase II and its induction by DNA-damaging agents. J Bio Chem, 1984, **259**(3): 1560-1565

[76] Fan J, Leroux-Coyau M, Savery N J, *et al*. Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. Nature, 2016, **536**(7615): 234-237

[77] Adebali O, Chiou Y Y, Hu J, *et al*. Genome-wide transcription-coupled repair in *Escherichia coli* is mediated by the Mfd translocase. Proc Natl Acad Sci USA, 2017, **114**(11): 2116-2125

[78] Million-Weaver S, Samadpour A N, Moreno-Habel D A, *et al*. An underlying mechanism for the increased mutagenesis of lagging-strand genes in *Bacillus subtilis*. Proc Natl Acad Sci USA, 2015, **112**(10): 1096-1105

[79] Hanawalt P C, Spivak G. Transcription-coupled DNA repair: two decades of progress and surprises. Nat Rev Mol Cell Biol, 2008, **9**(12): 958-970

[80] Gregersen L H, Svejstrup J Q. The cellular response to transcription-blocking DNA damage. Trends Biochem Sci, 2018, **43**(5): 327-341

[81] Pani B, Nudler E. Mechanistic insights into transcription coupled DNA repair. DNA Repair, 2017, **56**: 42-50

[82] Marteijn J A, Lans H, Vermeulen W, *et al*. Understanding nucleotide excision repair and its roles in cancer and ageing. Nat Rev Mol Cell Biol, 2014, **15**(7): 465-481

[83] Kim N, Jinks-Robertson S. Transcription as a source of genome instability. Nat Rev Genet, 2012, **13**(3): 204-214

[84] Paul S, Million-Weaver S, Chattopadhyay S, *et al*. Accelerated gene evolution through replication-transcription conflicts. Nature, 2013, **495**(7442): 512-515

[85] Merrikh H. Spatial and temporal control of evolution through replication-transcription conflicts. Trends Microbiol, 2017, **25**(7): 515-521

[86] Lang K S, Hall A N, Merrikh C N, *et al*. Replication-transcription conflicts generate R-Loops that orchestrate bacterial stress survival and pathogenesis. Cell, 2017, **170**(4): 787-799

[87] Schroeder J W, Hirst W G, Szewczyk G A, *et al*. The effect of local sequence context on mutational bias of genes encoded on the leading and lagging strands. Curr Biol, 2016, **26**(5): 692-697

[88] Merrikh C N, Weiss E, Merrikh H. The accelerated evolution of lagging strand genes is independent of sequence context. Genome Biol Evol, 2016, **8**(12): 3696-3702

[89] Sankar T S, Wastuwidyaningtyas B D, Dong Y, *et al*. The nature of mutations induced by replication – transcription collisions. Nature, 2016, **535**(7610): 178-181

[90] Noble R, Noble D. Was the watchmaker blind? Or was she one-eyed?. Biology, 2017, **6**(4): 47

[91] Yona A H, Frumkin I, Pilpel Y. A relay race on the evolutionary adaptation spectrum. Cell, 2015, **163**(3): 549-559

[92] Tomkova M, Tomek J, Kriaucionis S, *et al*. Mutational signature distribution varies with DNA replication timing and strand asymmetry. Genome Biol, 2018, **19**(1): 129

[93] Maliszewska-Tkaczyk M, Jonczyk P, Bialoskorska M, *et al*. SOS mutator activity: unequal mutagenesis on leading and lagging strands. Proc Natl Acad Sci USA, 2000, **97**(23): 12678-12683

[94] Khrustalev V V, Barkovsky E V. The probability of nonsense mutation caused by replication-associated mutational pressure is much higher for bacterial genes from lagging than from leading strands. Genomics, 2010, **96**(3): 173-180

[95] Szczepanik D, Mackiewicz P, Kowalczuk M, *et al*. Evolution rates of genes on leading and lagging DNA strands. J Mol Evol, 2001, **52**(5): 426-433

[96] McLean M J, Wolfe K H, Devine K M. Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. J Mol Evol, 1998, **47**(6): 691-696

[97] Rocha E P, Danchin A. Gene essentiality determines chromosome organisation in bacteria. Nucleic Acids Res, 2003, **31**(22): 6570-6577

[98] Rocha E P C, Danchin A. Essentiality, not expressiveness, drives gene-strand bias in bacteria. Nature Genetics, 2003, **34**(4): 377-378

[99] Merrikh C N, Merrikh H. Gene inversion potentiates bacterial evolvability and virulence. Nat Commun, 2018, **9**(1): 4662

[100] Kim N, Abdulovic A L, Gealy R, *et al*. Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. DNA Repair, 2007, **6**(9): 1285-1296

[101] De Septenville A L, Duigou S, Boubakri H, *et al*. Replication fork reversal after replication-transcription collision. Plos Genet, 2012, **8**(4): e1002622

[102] Wang J D, Berkmen M B, Grossman A D. Genome-wide coorientation of replication and transcription reduces adverse effects on replication in Bacillus subtilis. Proc Natl Acad Sci USA, 2007, **104**(13): 5608-5613

[103] Merrikh C N, Brewer B J, Merrikh H. The B. subtilis accessory helicase PcrA facilitates DNA replication through transcription units. Plos Genetics, 2015, **11**(6): e1005289

[104] Guariglia-Oropeza V, Helmann J D. Bacillus subtilis σ(V) confers lysozyme resistance by activation of two cell wall modification pathways, peptidoglycan O-acetylation and D-alanylation of teichoic acids. J Bacteriol, 2011, **193**(22): 6223-6232

[105] Nicolas P, Mäder U, Dervyn E, *et al*. Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis. Science, 2012, **335**(6072): 1103-1106

基于复制–转录碰撞的突变和进化意义*

董其格其[1,2)**]　乔佳鑫[1)**]　孙宏伟[1)]　范丽菲[1)]　莫日根[1)***]

([1)] 内蒙古大学生命科学学院，省部共建草原家畜生殖调控与繁育国家重点实验室，呼和浩特 010070；
[2)] 内蒙古农业大学生命科学学院，呼和浩特 010011）

摘要　复制和转录机器会同时使用相同的DNA区域作为模板，因此复制和转录不可避免地以头对头或追尾方式相互碰撞. 头对头碰撞和追尾碰撞均会导致复制机器停留，从而造成DNA损伤和基因组不稳定. 就基因组完整性而言，头对头碰撞比追尾碰撞的后果更严重. 本文回顾总结了复制-转录冲突的解决机制和进化影响. 相对于前导链，滞后链上非同义（氨基酸改变）突变的发生率更高，并且滞后链上基因的高频诱变取决于转录本和基因大小，因此较快的适应性突变发生在滞后链上. 头对头基因的高度转录增加了复制过程中响应压力的突变率. 无论是头对头还是追尾模式，复制-转录冲突都可能是适应性进化的驱动力.

关键词　复制-转录冲突，解决机制，冲突介导的突变，进化意义

中图分类号　Q75，Q36　　**DOI**：10.16476/j.pibb.2020.0107

* 国家自然科学基金（31700034）和内蒙古自治区自然科学基金（2016BS0309）资助项目.

** 并列第一作者.

*** 通讯联系人.

Tel：0471-4992242，E-mail：morigenm@hotmail.com

收稿日期：2020-04-24，接受日期：2020-07-10