

Differential activity of polymerase ζ associated with replication timing and gene bodies in humans: evidence from mutational signatures

Vladimir B. Seplyarskiy^{2,3}, Georgii A. Bazykin^{1,2,3} and Ruslan A. Soldatov^{1,2}

¹*Department of Bioengineering and Bioinformatics, Moscow State University, Moscow, Russia*

²*Institute of Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia*

³*Pirogov Russian National Research Medical University, Ostrovitianov str. 1, Moscow, Russia,*

ruslansoldatov@gmail.com

Human mutation rate varies along the genome; the mechanisms underlying this variation are only partially understood. Factors correlated with the local mutation rate include replication timing, meiotic recombination rate, GC content, DNA accessibility, transcription and distance from the telomeres. Human DNA replication is a complex process which involves synthesis of DNA from multiple origins, so that replication of different large-scale regions is activated in a robust temporal order which is conserved between tissues and species [1]. Replication timing is among the strongest determinants of the mutation rate, as reflected by the evolutionary divergence of primates [2] and the somatic mutational patterns in humans [3]. Replication timing affects multiple classes of mutations, including all kinds of point mutations and copy number variation induced by nonallelic homologous recombination or nonhomologous end-joining [3].

Here, we focus on dinucleotide mutations (DNMs). DNM is a mutation event that simultaneously changes two nucleotides in a row. Mechanisms and factors governing the distribution of these events along the genome are unknown. The rates of DNMs were estimated from de-novo mutations in parent-child trios, interspecies divergence, intraspecies polymorphism and disease causing mutations [4]. The obtained estimates were rather similar, placing the rate of DNMs at about 0.4% of the single nucleotide mutation (SNM) rate. The GC->AA/TT mutation is the most frequent DNM in human polymorphism and disease data. The same DNM is strongly associated with the activity of polymerase zeta (pol ζ) [5].

Pol ζ is able to perform translesion synthesis (TLS), to bypass abasic sites, and to extend non-perfectly matched primers. It is recruited to restart stalled replication forks associated

with double strand breaks, mismatches or DNA non-B structures. Pol ζ is the only TLS polymerase that is essential for embryonic development in mammals [6]. The mutation rate of pol ζ is about 10^{-3} point mutations per site, making it substantially more error-prone than other homologous B family polymerases α , δ and ϵ which play the central role in DNA replication.

Experimental studies in yeast and mammals suggest that pol ζ tends to produce DNMs. Nucleotide excision repair (NER) deficient *S. cerevisiae* line with error-prone pol ζ demonstrates a strong overrepresentation of complementary GC->AA and GC->TT DNMs among the detected DNMs [5].

In this study, we show that the GC->AA/TT is the most common DNM in human-chimp divergence. The pol ζ signature is especially pronounced in introns, where a strong excess of GC->TT over GC->AA on the non-template strand is observed, suggesting association with gene bodies. While the overall rate of DNMs is only very weakly associated with replication timing, the pol ζ signature frequency is radically higher in late replicating regions. Together, these findings indicate the heterogeneity of the genomic effect of pol ζ , shedding light on a novel cause of mutational heterogeneity along the genome.

1. Ryba T et al. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* 2010;20: 761–770.
2. Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov GV, Mirkin SM, Sunyaev SR. Human mutation rate associated with DNA replication timing. *Nat Genet.* 2009;41: 393–395.
3. Koren A, Polak P, Nemesh J, Michaelson JJ, Sebat J, Sunyaev SR, et al. Differential relationship of DNA replication timing to different forms of human mutation and variation. *Am J Hum Genet.* 2012;91: 1033–1040.
4. Chen J-M, Cooper DN, Férec C. A new and more accurate estimate of the rate of concurrent tandem-base substitution mutations in the human germline: ~0.4% of the single-nucleotide substitution mutation rate. *Hum Mutat.* 2014;35: 392–394.

5. Stone JE, Lujan SA, Kunkel TA, Kunkel TA. DNA polymerase zeta generates clustered mutations during bypass of endogenous DNA lesions in *Saccharomyces cerevisiae*. *Environ Mol Mutagen.* 2012;53: 777–786.

6. Lange SS, Wittschieben JP, Wood RD. DNA polymerase zeta is required for proliferation of normal mammalian cells. *Nucleic Acids Res.* 2012;40: 4473–4482.