Fork Stalling and Recombination of the Nascent DNA: Addressing the Limitations

Saeed Taheri*

The New Lahijan Scientific Foundation, Iran

*Corresponding author: Taheri S, The New Lahijan Scientific Foundation, Kargar Str, Lahijan, Iran, Email: taherimd@gmail.com

Abstract

Genomic rearrangements are hot topics for research and debates due to their complex natures as well as significant health impact in either congenital or malignant disorders. In order to explain the mechanisms responsible for their generation, models have been proposed to the literature with the most well recognized of them the non-allelic homologous recombination (NAHR), micro homology-mediated break induced recombination and fork stalling and template switching, each associated with own limitations. In an attempt to address the limitations associated with the previous models, in the current study a new model is represented and substantiated by describing some of the most frequently observed recurrent rearrangements that have been reported to the literature.

Keywords: Micro homology; Chromosomes; Rearrangement factory; Non-Allelic Homologous Recombination

Abbreviations: NAHR: Non-Allelic Homologous Recombination; SDs: Segmental Duplications; R-factory: Rearrangement factory; DSBs: Double Stranded Breakages DSBs; LCRs: Low Copy Repeats LCRs; MMBIR: Microhomology Mediated Break Induced Replication; MMR: Microhomology Mediated Recombination; FoSTeS: Fork Stalling and Template Switching; SLR: Stalled Loop Replicon; ERCEs: Early Replication Control Elements; TADs: Topologically Associated Domains
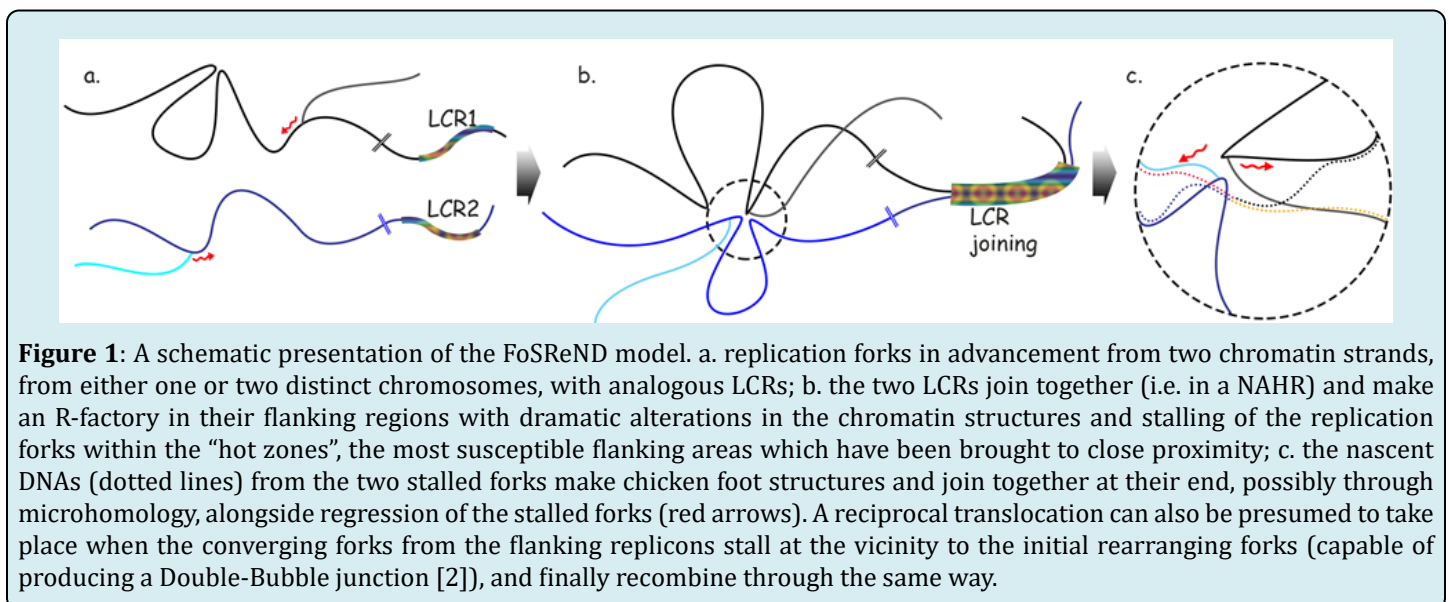
Introduction

There is large consensus that the majority of genomic rearrangements are produced through replicative mechanisms, and intensive evidence have approved existence of microhomology of sequences at the rearrangement junction points. To address these observations, mechanistic models have been proposed to the literature that could be categorized into two main categories: the DSB repair-based, and the fork-stalling based mechanisms (reviewed in **Table 1**). Nevertheless, despite the high potency of the proposed models in explaining the large majority of rearrangements, still significant unattended limitations could be raised. For example, although those models are generally based on mechanisms employing microhomology to combine with the template and start replication, the microhomology lengths that have been actually found in the junctions are mostly very small (of 1 or 2 bps) and sometimes even absent. Moreover, they don't describe that even if a linkage by one or two bps is established with the template, how could it be maintained during the accumulation of the replication machinery and even though, how they start replication from that one or two bps template-linked invasion? On the other hand, events such as stalling of the replication forks or DSB generations are not exceptional events, and they constantly happen during replication process all over the genome; but genomic replication still remains astonishingly accurate. So why the proposed rearrangement mechanisms should only be happening in specific contexts or recurrently in specific genomic regions?.

Mechanisms	Short description	Ref.
DSB based		
NAHR	A DSB is induced in a chromatin region for which there are duplicated segments (i.e. within an LCR). The DSB sites, in an attempt to make a homologous recombination (HR), join homologous segments but in non-allelic homologous regions (i.e. analogous LCRs).	[3]
MMBIR	A DSB is induced on a stalled replication fork arm, and the broken end in an attempt to make break induced replication, it starts the replication from a chromatin region with just a microhomology.	[4]
MMR	A DSB is induced in a chromatin region and in an attempt to make an HR, at least one of the two DSB ends instead of joining a homologous and/or allelic region, it joins and starts the replication from a region employing just microhomology.	[5]
Fork stalling based		
FoSTeS	After the replication fork stalls, the 3' end of an Okazaki fragment starts replication on a surrounding chromatin fiber from a region with microhomology.	[6]
a.FoSTeS	After the replication fork stalls in an R-factory region, the 3' and 5' ends of the nascent DNA make a 'chicken foot' structure and screens the surrounding region for a microhomology mediated replication. Finally the rearranged strand then recombines with the converging replication fork from the neighboring replicon that getting stalled at the vicinity to the initially stalled fork (with the ability of making a 'Double Bubble junction').	[2]
FoSTeS+	A bi-sided a.FoSTeS: The converging stalled fork will also produce a 'chicken foot' junction, and makes the same process of the initial one and the two rearranged strands finally recombine on one chromatin strand.	[7]
FoSReND	Multiple forks (at least two) are stalled at the vicinity to each other in an R-factory environment, and after the 'chicken foot' is made from the nascent DNA strands, they join together through microhomology.	[Current study]

Table 1: The list of replication-based mechanisms proposed as responsible for genomic rearrangements
DSB; double stranded breakage, LCR; low copy repeat; R-factory; rearrangement factory, NAHR; non-allelic homologous recombination, MMBIR; microhomology mediated break induced replication, MMR; microhomology mediated recombination, FoSTeS; fork stalling and template switching, a.FoSTeS; alternative FoSTeS, FoSTeS+; alternative FoSTeS plus MMR, FoSReND; fork stalling and recombination of the nascent DNA.
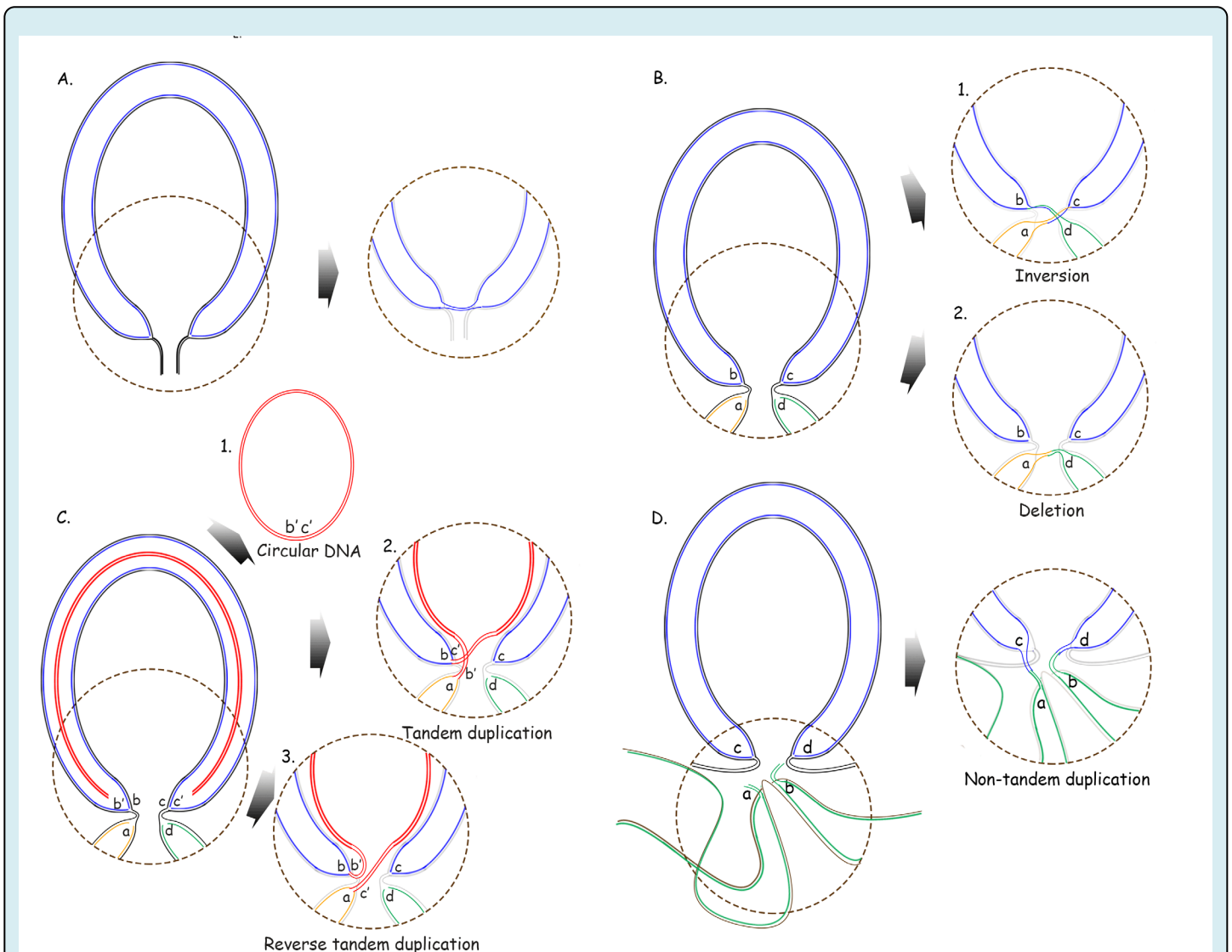

Figure 1: A schematic presentation of the FoSReND model. a. replication forks in advancement from two chromatin strands, from either one or two distinct chromosomes, with analogous LCRs; b. the two LCRs join together (i.e. in a NAHR) and make an R-factory in their flanking regions with dramatic alterations in the chromatin structures and stalling of the replication forks within the "hot zones", the most susceptible flanking areas which have been brought to close proximity; c. the nascent DNAs (dotted lines) from the two stalled forks make chicken foot structures and join together at their end, possibly through microhomology, alongside regression of the stalled forks (red arrows). A reciprocal translocation can also be presumed to take place when the converging forks from the flanking replicons stall at the vicinity to the initial rearranging forks (capable of producing a Double-Bubble junction [2]), and finally recombine through the same way.

Figure 2: Illustration of generation of different types of genomic rearrangements through FoSReND. A. producing a circular DNA through FoSReND. Both diverging forks of a replicon (or flanking replicons) stall at the spatial vicinity to each other (i.e. a loop anchor sites), producing a structure termed as the bilateral stalled loop replicon (SLR) [7]. The nascent DNAs from the stalled forks produce 'chicken foot' structures, and recombine at their ends; B. A bilateral SLR as described above, with the converging forks from the flanking replicons stalling at the vicinity to the SLR forks ('*Factory Cross*' structure-encircled), making a *Double Bubble* junction (in case the converging fork stalls at the immediate vicinity to the first stalled fork – illustrated for both a-b and c-d forks) [2]; 1. Non-counterpart forks, one from the SLR and the other from the distal (converging) replicons recombine through FoSReND, making an inversion of the looped region in the final rearrangement; 2. The two forks from the distal replicons recombine in FoSReND, and the intervening looped region is deleted; C. in specific circumstances, it has been predicted that a replication firing may take place within an already stalled replicon, making multiple copies of the region (red lines); now these extra copies can make 1. a circular DNA, through joining of their two ends; 2. a tandem duplication, and 3. a reverse tandem duplication; D. producing non-tandem duplication through FoSReND: when a bilateral SLR comes into spatial vicinity to a chromatin region with two converging replication forks stalling at the overall vicinity to each other (i.e. a Double Bubble junction, as shown), in case forks from the two non-counterpart regions recombine (in this case, fork 'a' with either 'c' or 'd', and fork 'b' with the other one), the sequences from the SLR could be translocated to another region, and since the genomic region respective to the SLR could be replaced with either a new replication firing origin or through recombination of the flanking replicon, the original sequence would be retained beside the translocated one.

Saeed Taheri. Fork Stalling and Recombination of the Nascent DNA: Addressing the Limitations. Genom & Gene Ther Int 2020, 4(1): 000114.

Copyright© Saeed Taheri.

Genomic segmental duplications (SDs), i.e. the LCRs, have constantly been found in the regions mapped for rearrangements and have largely been implicated as players of significant roles in this context. But except for the NAHR, their exact role has been a matter of uncertainty. The 'rearrangement factory' (R-factory is a recently introduced) model has in which joining of the SDs residing on either the same or different chromosomes in specific genomic regions, could result in profound alterations in the physiological states of the flaking chromatin regions (including within the LCR sequences themselves) [1]. These alterations are supposed to be inducing a broad array of structural and functional eccentricities that may include generation of non-B DNA conformations, changes in the physiological chromatin folding patterns and lamina-associated compartmentalization, replication timing, replication fork stalling and double stranded breakages (DSBs). In order to address the mentioned limitations associated with the proposed models and based on the predictions made by the R-factory model, a new model is developed that attempts to eliminate the mentioned concerns.

According to the R-factory model [1], joining of some of the genomic SDs of enough length (i.e. the low copy repeats (LCRs)) for any reason (most notably NAHR), could result in significant alterations in the genomic structures (i.e. chromatin folding), and functions (i.e. replication). These alterations could take place in a broad range from within the LCRs' sequences to over a million base pair far from each. But at least in case of some specific LCRs, there are some narrow and predictable ranges of the genomic sequences within which these alterations has the highest intensity (termed here as R-factory hot zones), making it possible to induce recurrent rearrangements. Now if we assume that generation of an R-factory could result in multiple replicon forks to stall in the hot zones, and the alteration of chromatin folding brings some of these stalled forks to close proximity, the subsequent rearrangement of the nascent DNAs from these stalled forks (i.e. through end joining of the chicken foot structures), could result in genomic rearrangements with copy number variations (CNVs) **(Figure 1)**. In case one of the stalled replicons is short enough, the whole replicon sequence could be released from the parental chromatin through regression of the chicken foot structure, making small insertions at the rearrangement junction points that have largely been reported in different series **(Figure 2)** illustrates generation of some of the best known types of genomic rearrangements through FoSReND model.

Conclusion

FoSReND model of genomic rearrangements is a prediction from the R-factory model. It not only is able to describe genomic rearrangements that have been reported

to the literature but also, it eliminates the limitations that have been recounted for the previously introduced models. It also fits with the structural and functional features that have been described for the chromatin. Herein, some of the significant aspects of the FoSReND model for the genomic rearrangements would be discussed.

Addressing the Limitations

FoSReND responds to the question 'why replication fork stalling in the physiological states doesn't lead to genomic rearrangements' through generation of an R-factory in which, multiple forks (vs an isolated one in the former models) stall in spatial vicinity to each other for which physiological ways of resolution won't work (i.e. due to the R-factory environment). It also easily describes why the reciprocal translocations are not exactly reciprocal and there usually are either micro-deletions or –insertion/duplications at the junction points. Moreover, it fairly addresses maybe a most important limitation of the microhomology-mediated replicative mechanisms about the existence of very limited sequences of microhomology or none (as mentioned in the Introduction section) at the rearrangement junction-points, since the chicken foot structures from the different stalled forks can recombine in either a microhomology mediated or simply non-homologous end-joining models.

R-Factory Hot Zones

Reports indicate that in specific regions of the genome, genomic rearrangements take place recurrently for which the junction points are more or less clustered in limited ranges of the chromatin sequences, sometimes almost pinpointed (i.e. in PLP1 gene rearrangements [8]). It compels that replication forks stall almost precisely at reproducible regions (i.e. R-factory hot zones) and the subsequent events which could be best explained by FoSReND. Evidence suggests that these hot zone clusters have special markers making them not just stochastically chosen points. Here their associations with structural and functional chromatin domains would be discussed.

Fosrend, Chromatin Structures & Replication Timing

It is a well-known idea that specific types of genomic rearrangements are likely to take place in specific cell types suggesting cell-type chromatin structures as significant players in predisposition to any rearrangement type. Early replication control elements (ERCEs) are cis-regulatory elements that form CTCF-independent loops and are in charge of controling the replication timing [9]. They have also been implicated as the upstream controlling centers for a wide array of genomic functions and architecture including the A/B compartmentalization and the topologically

associated domains (TADs) boundaries, and transcription. ERCEs have also been shown to be enriched in the features of accessibility (enhancer/promoters) including DNase I hypersensitivity, H3K27ac, H3K4m1, H3K4m3 [9]. On the other hand, the same markers also have been shown to be significantly enriched in the chromatin regions most vulnerable to rearrangements [10,11]. Moreover, it has been well established that the nuclear matrix attachment regions are highly enriched in the rearrangement junction-points [12]. Putting together, it seems that R-factory hot zones fairly fit to either the ECRCe sites or specific structural features that are highly under its control. And in case upstream R-factory generation events (i.e. LCRs joining) affect ERCEs, it very well explains all the observations associated with the genomic rearrangements; though it still needs to be confirmed by scientific experiments.

Chromothripsis

Chromothripsis is defined as extensive genomic rearrangements with an oscillating pattern of DNA CNVs, all usually taking place in a range of a unique or sometimes a small number of chromosomes. The mechanism for chromothripsis is unknown, and the proposed mechanisms include fragmentation of an isolated chromosome and subsequent end joining, collapse of a replication fork and subsequent MMBIR events, and the former hypothesis in the context of micronuclei [13]. FoSReND can provide a fabulous mechanism through generation of an extensive number of R-factory hot zones due to combining of two or multiple number of LCRs around the area, inducing multiple replication forks stalling in a highly crowded and dynamic area of chromatin folding, and subsequent FoSReND processes taking place. This probably provides a much better mechanism because the extensive fragmentation of one chromosome without affecting the others is not very likely, and yet it could induce apoptosis.

Reciprocal Translocations

Current models describing genomic rearrangements usually have limited capacity for describing the unexpected features of the reciprocal translocations like micro deletions as well as direct duplications at their breakpoints [14]. However, FoSReND model can easily explain them with the former as a consequence of stalling of the two converging forks at some sequence distance, failing to make a Double Bubble junction, and the latter through the regression of the primary stalled fork after its chicken foot structure recombines in its first FoSReND process and as a result, the second (converging) fork could proceed to the sequences that had been previously replicated by the primary fork, leading to regional duplication.

Limitations

Despite the advantages mentioned for the FoSReND model, it should be noted that there are still limitations that have not been well resolved including the existence of non-templated sequences reported at the breakpoint mappings of genomic rearrangements, or small insertions at the respective sites which are not much likely to be from replicons with the ability to make looped structures. Although the latter might be explained through resolution of a short stalled replicon in its whole length after only one side of it is recombined to a larger replicon, before the subsequent main FoSReND event.

Future Perspectives For Research

Plenty of opportunities for research could be derived from the FoSReND model, including imaging studies to confirm the chromatin structural predictions like the Double Bubble junction and the FoSReND Cross. Moreover, a broad range of opportunities for research is provided for attempts in determining the R-factory generation as a consequence of specific SDs' joining, the characteristics and configurations of the LCRs that could make it, and their potential effects on the genomic structural and functional aspects including the compartmentalization, replication timing and ERCEs' features.

References

1. Taheri S (2019) On Complex Chromosomal Aberrations: 'Rearrangement Factory' Generation at Flanking Regions to the Interacting Genomic Duplicons. N Lahij Med J 3: 1-13.

2. Taheri S (2019) A Leading Strand-Based Alternative for the Fork Stalling & Template Switching (a.FoSTeS): A Prediction from the 'Rearrangement Factory' Model. N Lahij Med J 3: 14-17.

3. Lupski JR, Stankiewicz P (2005) Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS genetics, 1(6): e49.

4. Hastings PJ, Ira G, Lupski JR (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS genetics *5*(1): e1000327.

5. Taheri S (2019) Micro-homology Mediated Recombination: Introduction to a Theory Model. N Lahij Med J 3: 18-25.

6. Lee JA, Carvalho CM, Lupski JR (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. cell 131(7): 1235-1247.

7. Taheri S (2019) Reconciling Alternative FoSTeS with the Microhomology Mediated Recombination: Promising an M-Theory in the Genomic Rearrangement Studies. N Lahij Med J 3: 26-29.

8. Beck CR, Carvalho CM, Banser L, Gambin T, Stubbolo D, et al. (2015) Complex genomic rearrangements at the PLP1 locus include triplication and quadruplication. PLoS genetics 11(3): e1005050.

9. Sima J, Chakraborty A, Dileep V, Michalski M, Klein KN (2019) Identifying cis elements for spatiotemporal control of mammalian DNA replication. Cell 176(4): 816-830.

10. Canela A, Maman Y, Jung S, Wong N, Callen E, et al. (2017) Genome Organization Drives Chromosome Fragility. Cell 170(3): 507-521.

11. Hogenbirk MA, Heideman MR, de Rink I, Velds A, Kerkhoven RM, et al. (2016) Defining chromosomal translocation risks in cancer. Proceedings of the National Academy of Sciences 113(26): E3649-E3656.

12. Gutierrez S, Javed A, Stein J, Stein G, Nicovani S, et al. (2011) Epigenetic Changes Associated with Chromosomal Translocation in Leukemia. Myeloid Leukemia: Basic Mechanisms of Leukemogenesis: 449-464.

13. Zhang CZ, Spektor A, Cornils H, Francis JM, Jackson EK, et al. (2015) Chromothripsis from DNA damage in micronuclei. Nature 522(7555): 179-184.

14. Bakhshi A, Wright JJ, Graninger W, Seto M, Owens J, et al. (1987) Mechanism of the t (14; 18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. Proceedings of the National Academy of Sciences of the United States of America 84(8): 2396-2400.

Saeed Taheri. Fork Stalling and Recombination of the Nascent DNA: Addressing the Limitations. Genom & Gene Ther Int 2020, 4(1): 000114.

Copyright© Saeed Taheri.