Impediments to replication fork movement: stabilisation, reactivation and genome instability

Sarah Lambert, Antony M Carr

Impediments to replication fork movement: stabilisation, reactivation and genome instability.

Sarah Lambert[1,2] and Antony M. Carr[3]

1. Institut Curie, Bat 100, 91405, Orsay, France

2. Centre national de la recherche scientifique, UMR3348, centre universitaire, Paris Sud XI, 91405, Orsay, France

3. Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Brighton, BN1 9RQ, UK.

Correspondence:

e-mail: a.m.carr@sussex.ac.uk. Tel: 01273 678122, Fax: 01273 678121

Keywords:

Replication fork barrier; Fork arrest; Fork collapse; Chromosome rearrangement

Summary

Maintaining genome stability is essential for the accurate transmission of genetic material. Genetic instability is associated with human genome disorders and is a near-universal hallmark of cancer cells. Genetic variation is also the driving force of evolution and a genome must therefore display adequate plasticity to evolve while remaining sufficiently stable to prevent mutations and chromosome rearrangements leading to a fitness disadvantage. A primary source of genome instability are errors that occur during chromosome replication. More specifically, obstacles to the movement of replication forks are known to underlie many of the gross chromosomal rearrangements seen both in human cells and in model organisms. Obstacles to replication fork progression destabilize the replisome (replication protein complex) and impact on the integrity of forked DNA structures. Therefore, to ensure the successful progression of a replication fork along with its associated replisome, several distinct strategies have evolved. First, there are well-orchestrated mechanisms that promote continued movement of forks through potential obstacles. Second, dedicated replisome and fork DNA stabilization pathways prevent the dysfunction of the replisome if its progress is halted. Third, should stabilisation fail, there are mechanisms to ensure damaged forks are accurately fused with a converging fork or, when necessary, re-associated with the replication proteins to continue replication. Here, we review what is known about potential barriers to replication fork progression, how these are tolerated and their impact on genome instability.

Introduction:

DNA replication occurs within S-phase and is initiated at multiple replication origins. Origin activation is stochastic in most eukaryotes, but each origin has a distinct likelihood of activation during each cell cycle and also displays a preferred time of activation relative to the initiation of S phase. This gives rise to a spatio-temporal programme of replication at the population level [1-3]. Once activated, a replication origin gives birth to two replication forks which progress independently and bi-directionally. The replication fork is the point on the DNA where the two strands of the parental molecule are separated, an event that requires DNA helicase activity provided by the replicative helicase, sometimes with help from ancillary helicases.

During fork progression, nucleotide incorporation - and thus the movement of the fork-shaped DNA structure - can be impeded by a wide variety of obstacles [4]. We refer to these specific replication fork arrest sites as Replication Fork Barriers (RFBs). RFBs include damaged DNA bases (which can occur essentially randomly within the genome) and "intrinsic" RFBs that are associated with specific chromosomal features. Intrinsic RFBs embrace DNA-protein complexes (for example: centromeres and dormant origins), transcriptional units (for example tRNA genes and ribosomal DNA), DNA sequences that are prone to form non-canonical structures and a range of poorly characterised regions (for example common fragile sites and replication slow zones) that are prone to rearrangement when replication is mildly perturbed [5,6].

Our understanding of the consequences of impediments to fork movements has relied significantly on the characterization of cellular responses to global, as opposed to local RFB-dependent, replication inhibition. If replisomes are arrested using global replication inhibitors such as hydroxyurea (HU, which depletes dNTP pools) or Aphidicolin (an inhibitor of DNA polymerase α) the intra-S phase checkpoint acts to phosphorylate a variety of replisome components to maintain

replisome functionality [7,8]. We refer to such forks as "stalled forks", a definition that based largely on their ability to resume replication without further intervention once the inhibitor is removed. Conversely, if the replisome is not stabilised when it is arrested - for example when the intra-S phase checkpoint fails - then the fork cannot resume without further assistance from restart mechanisms. We refer to such non-functional forks as "collapsed forks". It is not precisely clear what occurs when a fork collapses. Early work implied that some components of the replisome dissociate from the fork when it collapses [9-12]. However, this does not now appear to be the case: the replisome remains intact and chromatin-associated [8]. Notwithstanding this, at a collapsed fork the helicase can continue to separate the parental strands, generating unscheduled single stranded regions. These become coated by the ssDNA binding protein, RPA, and subsequently are vulnerable to inappropriate recombination reactions. Similarly, while it is unclear precisely how, the nascent strand ends are no longer protected by association with the replisome and become substrates for inappropriate DNA processing activities (Figure 1) [7,13,14].

In human cells, and in yeast model organisms, low levels of replication stress (which can be tolerated by the cell) result in specific chromosomal regions becoming prone to breakage and genetic rearrangement (reviewed in [15-17]). The implication is that such loci are naturally difficult to replicate and the increase in general replication stress imposed experimentally results in additive problems that are expressed as a loci-specific increase in DNA breakage and potential genetic rearrangement. Such "fragile sites" defined some of the first loci prone to replication-associated problems. More recently, Szilard et al (2010) and others exploited the fact that yeast histone H2A is phosphorylated specifically at sites of replication arrest or DNA damage to identify sites where natural RFBs cause problems in unstressed cells. Genome-wide analysis of the chromosomal distribution of phosphorylated H2A (γ-H2A, the yeast equivalent of mammalian γ-H2AX) in unpurturbed wild-type yeast cells identified many loci where γ-H2A accumulated in a replication-dependent

manner [18]. γ-H2A-enriched loci are presumed to arise from the population average of the relatively small number of non-random replication fork arrest events that occur in each cell.

The loci identified broadly fell into three categories, each of which is covered in more detail below. One category included the many non-histone protein:DNA binding sites that had previously been identified by examining fork progression in *rrm3* mutants [19]. A second category of loci identified was transcription units, consistent with the known involvement of clashes between replication and transcription in perturbing fork progression [20,21]. Finally, a significant proportion of loci identified were novel and many of these overlapped with repressed genes. This suggests that chromatin structure might influence fork progression. A similar approach was conducted in the yeast *Schizosaccharomyces pombe* and γ-H2A was found to be enriched at known natural RFBs and at heterochromatin regions in the centromeres and telomeres during DNA replication, supporting a link to chromatin status [22].

The nature of replication fork barriers

Different forms of DNA damage and different intrinsic RFBs will arrest replication forks in different ways (Figure 2). These are likely to be distinct from how forks arrest in response to global inhibitors of replication. Different types of arrest profile will likely result in different responses, both in terms of the efficiency of checkpoint-dependent fork stabilisation (stalling) and in terms of the consequences when a fork collapses [23]. Below we discuss different RFBs and what is known about how they arrest replication.

RFBs: DNA damage

Damaged DNA bases, including abasic sites, pyrimidine dimers, alkylated bases and bulky adducts are obstacles to the replicative DNA-polymerases: replicative polymerases cannot efficiently incorporate nucleotides opposite non-canonical bases. DNA damage thus compromises DNA synthesis, but it is not necessarily the case that DNA damage *in vivo* impedes the actual movement of the replisome and DNA fork. Of particular relevance to DNA replication in the context of base damage and bulky lesions are a group of mechanisms collectively referred to as Post Replication Repair (PRR). PRR acts specifically to facilitate replication fork movement past difficult-to-replicate damaged bases [24]. The same is potentially true for some DNA secondary structures.

One of the two subsets of PRR mechanisms is defined by Trans-Lesion Synthesis (TLS) polymerases. TLS is mediated by a collection of "error prone" DNA polymerases with the ability to template lesions opposite certain damaged bases [25]. This is possible, in part, because these polymerases have more spacious active site to accommodate the distortions to the canonical base pair resulting from the damaged base and also lack proof reading exonuclease activity [26]. The spacious active site, while accommodating the incoming nucleotide pairing with a damaged base, comes at the expense of TLS polymerases being error prone when replicating non-damaged bases.

The second PRR mechanism is an Homologous Recombination (HR) protein-dependent error-free gap filling mechanism [27]. This uses the sister chromatid as a template to fill small single-stranded gaps that are left when the replisome skips over a difficult to replicate base. Until recently, the dogma assumed that such gaps were only left on the lagging strand template, and that the leading strand was always continuous. However, data from Lopes et al, suggest that single strand gaps are left in both nascent strands [28]. Analysis by 2D gel electrophoresis showed that extensive UV damage only modestly delayed replication fork progression, while EM analysis showed single strand gaps on both newly replicated sister chromatids. This

strongly implies that leading strand replication can re-prime "on the fly". The frequency of gaps also increased significantly when TLS polymerases or HR were absent, consistent with redundancy between the two mechanisms of PRR. While it has not been formally demonstrated that replication can re-prime across DNA damage on the leading strand in eukaryotes, evidence exists that the leading strand does re-prime during prokaryotic replication *in vitro* [29,30].

Due to PRR and potential repriming on leading strand, many forms of DNA damage do not necessarily arrest the progress of the replisome and fork. However, robust replication arrest in budding yeast has been reported in response to DNA damage induced by the alkylating drug, adozelin [31]. Adozelin creates bulky adducts at AT-rich sequences, within the minor groove of the DNA helix. Consistent with this, replication fork progression on a methyl methane sulfonate (MMS) treated DNA template is also perturbed [32] and exacerbated by mutations in base excision repair [33], suggesting alkylated bases directly affect fork movement. It is, however, difficult to experimentally distinguish slow replisome/fork progression from replisome/fork arrest and resumption or restart. The available data suggest that the nature of the DNA lesion, its location within the DNA template and the modes of repair and tolerance by PRR pathways significantly influence the effect that DNA base damage has on fork and replisome progression.

Discontinuities such as Double Strand Breaks (DSBs) or single strand gaps in the DNA template also affect replication fork movement. When a replication fork encounters, for example, a single strand gap, the result is predicted to be a broken fork with one arm converted into double strand break [34,35]. A fork encountering a DSB might similarly generate two equivalent double strand ends. There is, however, evidence to suggest that replisomes and forks pause before "running-off" the template [36,37], possibly to allow time for the discontinuity to be repaired. Another defined form of DNA damage, the Inter-strand Cross-Link (ICL), constitutes the most obvious direct obstacle to fork progression: ICLs prevent the unwinding of the DNA

duplex ahead of the fork [38]. When forks are arrested by an ICL, stabilisation of the replisome and fork by the intra-S phase checkpoint is not beneficial [39]. Most likely it does not efficiently occur, since the helicase can no longer unwind the template DNA and generate the amount of ssDNA required for inter-S phase checkpoint activation. In fact, extensive ssDNA is only detected once repair is initiated [40]. Following arrest at an ICL, forks thus collapse and are subsequently repaired through enzymatic processing, which includes controlled DNA cleavage by endonucleases and HR [41,42].

RFBs linked to DNA metabolism

DNA metabolism, including transcription, the establishment/maintenance of epigenetic marks and non-histone proteins that bind tightly to DNA also interfere with replication fork progress [4,24]. In addition, programmed replication fork barriers have evolved that block movement of replication forks that are approaching from one specific direction (Figure 2).

RFBs: DNA-bound proteins. Approximately a decade ago, it was estimated that budding yeast cells have to face more than 1400 natural RFBs caused by DNA-bound proteins per replication cycle [19]. This estimation was based on analysis of fork arrest by Bi-dimensional gel electrophoresis (2-DGE) in the absence of the Rrm3 helicase. Rrm3 is a 3′-5′ helicase that interacts with the catalytic subunit of DNA-polymerase ε and thus progresses with the replication fork [43]. Rrm3 is thought to dislodge non nucleosomal proteins bound to DNA in front of the fork and thus acts as a "fork clearing" motor to facilitate fork progression across natural RFBs. For example, in the absence of the Rrm3 helicase, fork arrest can be visualised by 2-DGE at non-histone protein bound sequences such as cryptic origins and centromeres [43]. In G1-phase, the Origin Recognition Complex (ORC) and the MCM2-7 complex are loaded onto replication origins [44] to form the pre-Replicative Complex (pre-RC). During S-phase, many such "licenced" replication origins do not fire. It is proposed

that the presence of unfired origins during S phase allows replication to be completed by new origin firing when the active converging forks are arrested and cannot resume [36]. The presence of the pre-RC complexes, which are by definition closely associated with the DNA, constitute potential obstacles to fork progression that are usually removed by Rrm3. A similar situation is thought to be the case at centromeres, where specialised CENP-A containing nucleosomes associate closely with kinetochore proteins to allow tethering to microtubules.

In the absence of Rrm3 or its fission homologue Pfh1, RFBs appears as hot spots of recombination, chromosomes breakages and rearrangements [45-48]. Alternative pathways such as HR then become essential to stabilize or reactivate arrested forks. Analysis of occupancy sites of DNA-polymerase ε revealed 192 fork-pausing sites in the absence of Rrm3 [49]. Together with the identification of Rrm3-dependent DNA replication pause sites by 2-DGE [43] , these data support the view that expression of multiple RFBs is supressed by Rrm3, which provides additional helicase power to the replisome, helping to maintain genome stability by providing a first line of defence against natural RFBs.

In addition to natural DNA protein complexes, the introduction of exogenous protein binding sequences into the genome, combined with the expression of the cognate binding protein, has been demonstrated to arrest replication forks in both yeast and bacterial systems [50,51]. As with natural DNA:protein barriers, the frequency of arrest and its consequences are increased when the ancillary replicative helicase is lost [50] and arrest is dependent on the strength of protein:DNA association [52].

Programmed RFBs. In a number of instances organisms have evolved specialised replication fork barriers for specific purposes. These are generally polar in nature, meaning that they are directionally determined, arresting forks coming from one direction but allowing uninterrupted passage of forks coming from the opposite direction. One common polar RFB in eukaryotes lies between ribosomal DNA

repeats (rDNA) [53]. The purpose of these ribosomal RFBs (rRFBs) is to promote unidirectional replication of the rDNA in order to prevent head-on collisions between transcription and replication. Interestingly, in prokaryotes where rDNA genes are dispersed and replication is unidirectional on the circular chromosome, the highly transcribed rDNA genes are orientated a manner that ensures they are also replicated co-directionally with transcription [54,55]. This attests to the importance of preventing head-on transcriptional clashes with the replication machinery. rRFBs contain specific DNA sequence motifs that bind proteins which promote the barrier activity. While it is unclear precisely how the barrier is activated, activity is not necessarily an intrinsic property of the DNA:protein interactions, suggesting that direct interactions between the RFB-binding proteins and regulators of the replicative and ancillary helicases (or the helicases themselves) are involved [56,57].

In addition to the rRFB, which is common to many eukaryotes, species-specific RFBs have also been identified and characterised. The most widely studied is *S. pombe* replication termination sequence 1 (*RTS1*) [58]. *RTS1*, along with a second less well characterised replication arrest sequence known as the switch activating site (SAS1) [59], was first identified due to its involvement in mating type switching in fission yeast [60]. *RTS1* is an ~850bp DNA sequence that associates with several proteins (including Rtf1, a Myb-like DNA binding protein and Rtf2, a PCNA-interacting protein) to ensure unidirectional replication of the fission yeast mating type locus [61,62]. Ectopic *RTS1*-RFB have been used in several studies to demonstrate that arrested replication forks are prone to collapse and subsequent repair by HR-dependent mechanisms that are predisposed to generate errors during the restart event due to inappropriate recombination with ectopic homologous sequences [63-69].

Mating type switching in fission yeast can be viewed as a kind of cell differentiation: two distinct cell-types express distinct sexual markers (h^+ or h^-). The switch from one sexual cell-type to another one requires two polar forks arrest events

to occur near the mating type locus, *mat1*. These are orchestrated by two polar RFBs that allow an imprint to be introduced in the lagging strand and the repair of the subsequent broken fork by a recombination-coupled replication mechanism using one of the two silent *mat2* or *mat3* cassettes [60]. Thus, the mating type switching in fission yeast provides an example of a cell differentiation process orchestrated with the support of RFBs.

Prokaryotic directional replication termination sites (known as *ter* sequences) have been characterised in some detail. In *E. coli*, these ensure that replication - which is initiated from a single origin at the "top" of a circular chromosome - terminates at the "bottom" of the chromosome. Thus, each half of the chromosome is replicated by a forks travelling in a predictable direction. This helps minimise the genetic instability associated with replication forks clashing head-on with transcription: as with the rDNA genes, the majority of transcription units in *E. coli*, and particularly those that are highly transcribed, are orientated away from the origin to ensure co-directionality of transcription and replication [70]. *E. coli* ter sequences are bound by the Tus protein and interactions between the bacterial replicative helicase and Tus appears to be the primary mechanism promoting polar fork arrest [71].

With the exception of rRFB's and a few specific replication barriers such as *RTS1*, replication termination in eukaryotes has largely been assumed to occur randomly between origins, i.e. at zones where converging forks meet stochastically. However, recent data from *S. cerevisiae* suggest that some termination zones involve a further level of organisation: overlap between fork-pausing elements or RFBs and occupancy sites for the topoisomerase 2 (Top2) has been revealed at some termination zones [72]. It is proposed that obstacles to fork progression may allow positioning of Top2, to facilitate replication completion and fork merging. A further dataset from *S. pombe* [73] also suggested that programed RFB activity is more widely spread over the genome: in addition to the Reb1-dependent rRFBs, several non-ribosomal DNA Reb1-binding sites were shown to mediate replication fork arrest.

Intriguingly, while rRFB activity does not require Reb1 protein dimerization, fork arrest at specific non-rDNA sites did. This correlated with Reb1-dependent physical contact between RFBs on chromosome II and chromosome I, leading the authors to conclude that Reb1 promotes "chromosome kissing" that, in turn, enhanced fork arrest. Distant chromosome contacts can regulate transcription [74]. Given the links between transcription and replication-induced genome instability it is thus tempting to speculate that coordination between these two aspects of chromosome metabolism may have evolved [75].

It remains unclear how polar fork arrest at DNA-bound protein occurs and why these programmed RFBs are insensitive to ancillary helicases such as Rrm3. It is currently proposed that terminator proteins bound to programmed RFB sequences are able to affect the activity of the replisome (the replicative helicase activity for example) via species-specific protein-protein interaction between the terminator and replisome components [57,71]. According to the direction of the approaching fork, it is proposed that one surface of the terminator protein is unable to interact with the converging replisome and is thus permissive for replisome passage. However, if the replisome approaches from the alternative direction, the juxtaposed surface of the terminator protein is able to interact with the replisome and thus is restrictive to fork passage. Replication-accessory factors might also regulate the ability of ancillary helicases to dislodge DNA-bound proteins. For example, it is proposed that the fork-protection complex Tof1/Csm3 in budding yeast inhibits Rrm3 at Fob1-dependent rRFBs [56].

RFBs: collision between the transcription and replication machinery. Active transcription is a well-established obstacle to the replication machinery in both prokaryotes and eukaryotes [21]. For example, ablating RFB activity at the rDNA locus causes RNA polymerase I (RNAPI) to clash head-on with replication. This results in replication fork arrest that can be visualised by 2D-Gels and correlates directly to increased genetic instability. Similarly, genes such as tRNAs that are highly transcribed by

RNAPIII represent a frequent class of RFBs. Again, these are active when transcription and replication progress in opposite directions, leading to head-on collisions [20]. Loci highly transcribed by RNAPII also cause fork arrest and increased recombination [46,49,76,77]. Interestingly, this arrest can be independent of the direction of transcription relative to replication.

Converging replication and transcription machines, both of which unwind the duplex DNA, will result in topological problems as the associated helicase activities converge and the intervening distance decreases to a point where access for topoisomerases is restricted [24]. Such problems will not occur when transcription and replication are co-directional, potentially explaining the directionality of some transcription-associated RFBs: the removal of the RNA polymerase by an ancillary helicase activity could potentially alleviate the topological issues. However, RNAPII -dependent RFBs are not entirely directional, particularly at highly transcribed loci where more than one RNAPII occupies the open reading frame [49]. The ability to perturb replication independently of transcriptional orientation could be explained by a number of issues. For example, transcription occurs within the context of topologically isolated chromosome loops [78], the nascent RNA is sequestered by hnRNPs and protein complexes such as THO/TREX facilitate correct packaging and export through the nuclear pore via a process known as "gene gating" [79-81].

Loss of THO/TREX results in an increase in the formation of RNA:DNA hybrids (R-Loops) associated with transcription. R-loop formation due to THO/TREX dysfunction results in recombinogenic structures forming at multiple loci during replication, likely due to replication fork arrest [82]. The link between transcription and the nuclear pore generates potential topological problems for replication fork progression that are relieved, at least in response to global replication stress, by direct phosphorylation of inner-basket nucleoporins by the intra-S phase checkpoint [81]. Thus, observations showing that that co-directional transcription can arrest

replication likely have many causes, ranging from arrested RNAPs, R-loops and chromatin architecture-dependent topology [24].

As mentioned above, the replicative helicase is augmented by an ancillary helicase, Rrm3, which travels with the replication fork [43] acting as an accessory helicase to the replicative helicase to help remove non-histone proteins associated with the DNA [83]. A similar accessory activity for clearing tightly associated proteins from DNA has been described for the Rep and UvrD helicases of *E. coli* [84,85]. Ablation of the Rrm3 helicase revealed a large number of fork arrest sites that are usually supressed by Rrm3 activity, including RNAPI and RNAPIII transcribed genes [49]. Interestingly, fork arrest at highly transcribed RNAPII genes was not increased when Rrm3 was ablated and, as discussed above, arrest at highly RNAPII transcribed genes can be independent of transcript orientation relative to replication. This would be consistent with a role for Rrm3 specifically in clearing protein barriers: R-loops (RNA:DNA hybrids) and topological problems promote the fork arrest independently of protein:DNA association.

RFBs: repressed genes. In analysing the genome-wide distribution of γ-H2A in unchallenged budding yeast cells, Szilard et al, identified multiple repressed genes that correlated with a peak of phosphorylated H2A [18]. Unlike many other RFBs, where DNA polymerase also accumulated, these repressed gene loci did not accumulate polymerase ε suggesting the accumulation of γ-H2A results from rare fork arrest, or possibly a decrease in fork velocity. Unlike the highly expressed genes identified as RFBs [49], the γ-H2A enriched repressed ORFs corresponded to genes with the lowest levels of RNAPII-binding, confirming that they were not being transcribed. Indeed, induction of transcription correlated with the loss of γ-H2A accumulation at these loci and γ-H2A accumulation was directly dependent on histone deacetylases (HDAC) recruitment. γ-H2A accumulation at repressed genes is increased in the absence of Rrm3 and is dependent on the two main checkpoint kinases, $Mec1^{ATR}$ and $Tel1^{ATM}$. These data distinguish the effects of replication

through repressed genes from those reported for highly transcribed RNAPII genes and establishes a connection between hypo-acetylated chromatin and replication fork progression which is intriguing in the context of human fragile sites (see below).

RFBs linked to Sequence organisation.

Studies in yeast models have defined a range of structural sequence organisations, or at risk motifs (ARMs), that can underlie sites of chromosomal rearrangement. Examples include regions that contain a variety of dispersed repeated sequences and DNA motifs that are able to form secondary DNA structures. ARMs represent a challenge for replication fork progression and thus for the maintenance of genome stability [4,86,87].

Palindromes and closely-spaced inverted repeats. A pure DNA palindrome is defined as a sequence that is the same whether read 5' to 3' on one strand or 5' to 3' on the complementary strand. A pure palindrome is thus essentially two inverted repeats joined directly at the palindrome centre. An interrupted palindrome has a number of bases separating the two inverted repeat arms. There is no formal definition of when a sequence of this sort is named as an interrupted palindrome or is known as an inverted repeat. Palindromes in dsDNA can extrude into a cruciform structure, reminiscent of a Holliday junction. A short interruption in a palindrome makes it energetically less favourable to extrude a cruciform structure due to the energy required to break the base pairing between bases within the interruption.

Pure palindromes of more than a few bases are genetically very unstable and are generally rapidly lost. Interrupted palindromes / inverted repeats are significantly more stable than pure palindromes [88]. In the human genome pure palindromes are rare while interrupted palindromes and inverted repeats occur relatively frequently. However, when analysing inverted repeats formed by common interspersed repetitive sequence elements, such as *Alu*, the number of closely spaced

inverted pairs is significantly reduced compared to that expected, suggesting negative selection for this orientation [89]. Consistent with this negative evolutionary pressure on inverted repeats, inverted closely spaced human Alu repeats have been demonstrated to be sites of replication forks arrest *in vivo* in several organisms [87]. Interestingly, the ability of the palindrome to form hairpins in ssDNA on the lagging strand during replication, as opposed to the propensity to extrude dsDNA into a cruciform, correlated with fork arrest.

If replication is arrested within an inverted repeat and the fork collapses, the homologous recombination-mediated restart of the fork is prone to generate replication errors by initiating replication at the incorrect sequence [63-65,90]. Due to the inverted orientation of the repeat sequences, this often results in the formation of dicentric and acentric palindromic chromosomes and consequent detrimental chromosomal rearrangements via breakage fusion bridge cycles. It is likely that this explains, at least in part, why closely spaced inverted repeats are underrepresented in the human genome.

Tri-nucleotide repeats. The genetic instability of tri-nucleotide repeats is responsible for a variety of human disorders, including fragile X syndrome, Huntington's disease and Friedreich's ataxia [91]. Tri-nucleotide repeats are prone to form hairpins and triplex DNA, both of which can contribute to their expansion and contractions. As with the hairpins formed by palindromes and closely-spaced inverted repeats, these appear to occur during DNA replication and the instability of tri-nucleotide repeats is linked to lagging strand DNA synthesis [92-94]. Bacterial and yeast models have been used extensively to investigate the molecular mechanisms responsible for tri-nucleotides repeat instability. We do not present a comprehensive review of tri-nucleotide repeat biology in model organisms, but focus briefly on GAA repeats in yeast as an exemplar. Large-scale expansions of GAA repeats are responsible for the human disease Friedreich Ataxia.

In budding yeast, tracts of >40 GAA/TTC repeats arrest replications forks when the GAA sequences are located on the lagging strand template, but not when the TTC sequences are present on the lagging strand template [94]. GAA repeats are capable of forming triplet (H-DNA) structures [95]. Fork arrest correlates with chromosome breakage, recombination and gross chromosomal rearrangements [94,96]. However, the instability of tri-nucleotide repeat length in budding yeast is only partially dependent on their orientation with regard to replication [97], suggesting that direct replication fork arrest is only one of several mechanism responsible for tri-nucleotide repeat instability. Nonetheless, expansion of GAA repeats in yeast was dependent on fork stabilisation/reactivation mechanisms: the fork stabilization complex proteins Tof1 and Csm3 prevented the expansions of GAA tracts whereas Sgs1, the budding yeast BLM homolog, and the HR-dependent PRR pathway promoted expansion. Thus, GAA repeat instability in yeast likely reflects several replication-dependent processes, including replication arrest and the ability of both leading and lagging strand polymerases to switch template either directly at the fork or during post-replication gap filling [98]. In human cells, while replication arrest most likely plays a role in repeat expansions, some of the more dramatic expansion of repeats number occur in non-replicating cells and are likely dependent on specific aberrant repair processes rather than RFB-promoted events [91].

Additional difficulties to replicate sequences. DNA can form a number of non-canonical (non B-form) structures [99] which can block DNA replication in vitro and, in many cases, apparently also in vivo [100]. Thus, in addition to palindromes, inverted repeats and tri-nucleotide repeats, a range of other naturally occurring DNA sequences underpin RFBs. Examples include, but are certainly not limited to, the AT-rich structure found at the human common fragile site FRA16D [101] and sequences containing 4 runs of 3G's that can form G4-DNA [102-104]. Both have been shown to be sites of replication arrest and genome rearrangements in model systems.

The lines of defence against RFBs

Given the extensive range of DNA metabolic events that can impede replication, how do cells maintain sufficient replication fidelity and efficiency? We see three general strategies: the cells first line of defence against potential RFBs is to prevent replication forks arresting at all (Figure 3). Thus, cells minimise transcriptional clashes with replication and attempt to repair DNA damage to counteract the RFB activity before replication occurs. They also have a variety of DNA damage bypass mechanisms and enzymatic activities, such as dedicated ancillary helicases, that aim to minimise fork arrest at potential RFBs. However, such mechanisms are not always sufficient to prevent potential RFBs from arresting the fork and its associated replisome. Thus, stochastically, some forks will arrest during normal DNA replication both at sites of unrepaired DNA damage and at intrinsic RFBs. The second line of defence is that the cell attempts to stabilise the fork and the replisome complex via the activity of the intra-S phase checkpoint and possibly using other as yet unidentified pathways (Figure 3). This has the advantage that the stalled fork can then either rapidly resume replication once the initial problem is resolved, or at least remains intact - protecting the nascent DNA ends - until the stalled fork merges with a converging fork.

Nevertheless, not all arrested forks can be successfully stabilised to become a stalled fork. Activation of the intra-S phase checkpoint requires the production of >100bp of single stranded DNA [105]), likely following the transient uncoupling of the replicative helicase from DNA polymerisation [106]. Thus, depending on the nature of the RFB, some arrested forks may not be able to activate the intra-S phase checkpoint at all (for example, ICLs block the helicase and do not generate ssDNA), while other fork arrest conformations may be partially refractive to rapid ssDNA production, thus being unable to activate the intra-S phase checkpoint sufficiently quickly. In such circumstances, stabilisation of the replisome and the fork may not be

successful. This will result in fork collapse, unscheduled ssDNA production and unprotected nascent ends.

Because collapsed forks and their consequent recombinogenic nature are potential instigators of chromosomal rearrangements [107], the cells final line of defence is to quickly protect them. There are two obvious options: either the collapsed fork can be held in a conformation that is capable of merging with a converging fork, or it must be restarted by reactivation or the rebuilding of the replisome and allowed to resume replication. The mechanisms of collapsed fork protection, how collapsed forks are merged with a canonical converging fork and how and when a collapsed fork is rebuilt and/or restarted are currently unclear. However, the homologous recombination machinery appears to play a key role in these events [24,108,109].

RFBs, Common Fragile Sites and genetic rearrangement

Replication errors are known to contribute significantly to the disease burden in humans, perhaps most obviously in the case of cancer: the majority of tumours are characterised by high levels of genetic instability, exhibiting gross chromosomal rearrangements (GCRs), high frequencies of inter-chromosomal copy number variations and localised rearrangements such as inversions [110-112]. Intriguingly, many of these genetic changes are associated with specific chromosomal loci that were previously identified as common fragile sites (CFS) [113-115]. CFS's were identified as loci exhibiting frequent gaps, constrictions or breaks in metaphase (condensed) chromosome spreads prepared from human cells grown under mild replicative stress [116].

CFS's have recently been associated with loci containing a paucity of replication origins, an association which appears to underpin their cell type-specific expression [117-120]. The link between a low density of replication initiation and sites of frequent genetic rearrangement is intriguing in the context of the recent

explanation for why there is a large excess of licenced - yet unused - origins of replication. This posits that these "cryptic" origins provide a reservoir of potential new initiation sites that provide a mechanism to rescue replication when two converging forks are impeded by RFBs [121-124]. In terms of CFS's, it can be proposed that, if replication fork movement is impeded within a region harbouring low origin density, the activation of cryptic origins will not be an effective route to support the completion of DNA synthesis. Thus, the following question arises : is a simple paucity of origins sufficient to define a CFS, or are additional intrinsic impediments to replication also necessary?

Some CFS's, including FRA16D, appear to contain structure-forming sequences that promote fork arrest [101,125] and it has been proposed that the combination of RFB activity and a paucity of cryptic replication origins combine to define a CFS [113]. However, most CFS's do not contain DNA structure-forming sequences and there is some evidence that not all CFS sequences present a challenge for replication fork movement. Molecular combining and fluorescence *in-situ* hybridisation approaches [118] have indicated that there is no impairment to the movement of replication forks that reach FRA3B, which suggests that this CFS may expresses fragility due simply to inefficient origin firing [126].

Notwithstanding this, many CFS's overlap with very large genes [127] and there is a positive correlation between transcription and CFS expression [128]. There is likely also a link between transcription of the genes within CFS's and the replication profile associated with cell-type specific fragility [126]. Recently, expression of long human genes lying within CFS was shown to cause interference between the late replication of the CFS and transcription, thus favouring R-loop formation and expression of CFS instability [129]. While there thus appears to be a link between increased transcription and the expression of CFS's, a recent study has also identified a relationship between histone hypo-acetylation and instability of six CFS's (FRA3B, FRA16D, FRA7G, FRAXB, FRA2G, FRA7H): relative to the

surrounding regions, hypo-acetylated chromatin was a common feature of these loci [130] and treatment with deacetylase inhibitors lead to an increase of histone H3-K9 and -K14 acetylation within the CFS's, concomitantly reducing their instability. Several of these hypo-acetylated CFS's are associated with large genes, but the presence of a large gene did not correlate with the hypo-acetylation status. It remains to be established why CFSs are hypo-acetylated and whether the hypo-acetylated chromatin structure of human fragile sites directly impacts on replication fork progression, or promotes a defective response to pre-existing replication stress. It is nevertheless intriguing that, in both yeasts and in human cells, epigenetic marks - and in particular the histone acetylation levels - are linked to the fragility of specific loci.

Conclusion:

Replication can be arrested by a wide variety of distinct RFBs. Each type of RFB potentially presents the cell with a distinct set of problems. The cells response to RFBs can be broadly categorised into three lines of defence. First, the cell first tries to prevent the expression of the potential RFB, for example by maximising DNA repair, promoting damage bypass by PRR, preventing clashes between transcription and replication and engaging ancillary helicases to remove non-histone proteins. However, some RFBs cannot be avoided. For a second line of defence, the cell next tries to maintain the correct association of the DNA and the replisome. This, if successful, prevents inappropriate processing of the nascent DNA ends and the fork structure and allows for recovery by either removing the barrier or awaiting a converging fork. However, some barriers are not amenable to stabilising the replisome (for example an ICL, see above) and others will, due to their nature be prone to replisome collapse. Thus, cells have a third string to their bow, which is to attempt to stabilise the collapsed fork and reconstruct a replisome.

Distinct types of RFB will have different consequences at each of these three potential levels of protection. These consequences will sometimes reflect the nature of the barrier itself, and at other times are likely to reflect the specific effect the RFB has on the replisome: how easy it is to stabilise? what is the consequence to the DNA structure itself if the replisome is not stabilised? Understanding how each specific RFB actually interferes with replication, and what the consequences are to the fork DNA, are thus important aspects of understanding the causes of the genetic instability which underlie cancer and human genomic diseases.

References:

1. Kelly TJ, Brown GW: **Regulation of chromosome replication**. *Annual review of biochemistry* 2000, **69**:829-880.
2. Mechali M: **Eukaryotic DNA replication origins: many choices for appropriate answers**. *Nature reviews. Molecular cell biology* 2010, **11**:728-738.
3. Lygeros J, Koutroumpas K, Dimopoulos S, Legouras I, Kouretas P, Heichinger C, Nurse P, Lygerou Z: **Stochastic hybrid modeling of DNA replication across a complete genome**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**:12295-12300.
4. Mirkin EV, Mirkin SM: **Replication fork stalling at natural impediments**. *Microbiology and molecular biology reviews : MMBR* 2007, **71**:13-35.
5. Weinert T, Kaochar S, Jones H, Paek A, Clark AJ: **The replication fork's five degrees of freedom, their failure and genome rearrangements**. *Current opinion in cell biology* 2009, **21**:778-784.
6. Lambert S, Carr AM: **Checkpoint responses to replication fork barriers**. *Biochimie* 2005, **87**:591-602.
7. Hu J, Sun L, Shen F, Chen Y, Hua Y, Liu Y, Zhang M, Hu Y, Wang Q, Xu W, et al.: **The intra-S phase checkpoint targets Dna2 to prevent stalled replication forks from reversing**. *Cell* 2012, **149**:1221-1232.
8. De Piccoli G, Katou Y, Itoh T, Nakato R, Shirahige K, Labib K: **Replisome stability at defective DNA replication forks is independent of S phase checkpoint kinases**. *Molecular cell* 2012, **45**:696-704.
9. Cobb JA, Bjergbaek L, Shimada K, Frei C, Gasser SM: **DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1**. *The EMBO journal* 2003, **22**:4325-4336.
10. Cobb JA, Schleker T, Rojas V, Bjergbaek L, Tercero JA, Gasser SM: **Replisome instability, fork collapse, and gross chromosomal rearrangements arise synergistically from Mec1 kinase and RecQ helicase mutations**. *Genes & development* 2005, **19**:3055-3069.
11. Katou Y, Kanoh Y, Bando M, Noguchi H, Tanaka H, Ashikari T, Sugimoto K, Shirahige K: **S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex**. *Nature* 2003, **424**:1078-1083.
12. Lucca C, Vanoli F, Cotta-Ramusino C, Pellicioli A, Liberi G, Haber J, Foiani M: **Checkpoint-mediated control of replisome-fork association and signalling in response to replication pausing**. *Oncogene* 2004, **23**:1206-1213.
13. Lopes M, Cotta-Ramusino C, Pellicioli A, Liberi G, Plevani P, Muzi-Falconi M, Newlon CS, Foiani M: **The DNA replication checkpoint response stabilizes stalled replication forks**. *Nature* 2001, **412**:557-561.
14. Cotta-Ramusino C, Fachinetti D, Lucca C, Doksani Y, Lopes M, Sogo J, Foiani M: **Exo1 processes stalled replication forks and counteracts fork reversal in checkpoint-defective cells**. *Molecular cell* 2005, **17**:153-159.
15. Hashash N, Johnson AL, Cha RS: **Regulation of fragile sites expression in budding yeast by MEC1, RRM3 and hydroxyurea**. *Journal of cell science* 2011, **124**:181-185.
16. Durkin SG, Glover TW: **Chromosome fragile sites**. *Annual review of genetics* 2007, **41**:169-192.
17. Hashash N, Johnson AL, Cha RS: **Topoisomerase II- and condensin-dependent breakage of MEC1ATR-sensitive fragile sites occurs independently of spindle tension, anaphase, or cytokinesis**. *PLoS genetics* 2012, **8**:e1002978.

18. Szilard RK, Jacques PE, Laramee L, Cheng B, Galicia S, Bataille AR, Yeung M, Mendez M, Bergeron M, Robert F, et al.: **Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX**. *Nature structural & molecular biology* 2010, **17**:299-305.
19. Ivessa AS, Lenzmeier BA, Bessler JB, Goudsouzian LK, Schnakenberg SL, Zakian VA: **The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes**. *Molecular cell* 2003, **12**:1525-1536.
20. Deshpande AM, Newlon CS: **DNA replication fork pause sites dependent on transcription**. *Science* 1996, **272**:1030-1033.
21. Aguilera A, Garcia-Muse T: **R loops: from transcription byproducts to threats to genome stability**. *Molecular cell* 2012, **46**:115-124.
22. Rozenzhak S, Mejia-Ramirez E, Williams JS, Schaffer L, Hammond JA, Head SR, Russell P: **Rad3 decorates critical chromosomal domains with gammaH2A to protect genome integrity during S-Phase in fission yeast**. *PLoS genetics* 2010, **6**:e1001032.
23. Lambert S, Froget B, Carr AM: **Arrested replication fork processing: interplay between checkpoints and recombination**. *DNA repair* 2007, **6**:1042-1061.
24. Branzei D, Foiani M: **Maintaining genome stability at the replication fork**. *Nature reviews. Molecular cell biology* 2010, **11**:208-219.
25. Ulrich HD: **Timing and spacing of ubiquitin-dependent DNA damage bypass**. *FEBS letters* 2011, **585**:2861-2867.
26. Zahn KE, Wallace SS, Doublie S: **DNA polymerases provide a canon of strategies for translesion synthesis past oxidatively generated lesions**. *Current opinion in structural biology* 2011, **21**:358-369.
27. Budzowska M, Kanaar R: **Mechanisms of dealing with DNA damage-induced replication problems**. *Cell biochemistry and biophysics* 2009, **53**:17-31.
28. Lopes M, Foiani M, Sogo JM: **Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions**. *Molecular cell* 2006, **21**:15-27.
29. Heller RC, Marians KJ: **Replication fork reactivation downstream of a blocked nascent leading strand**. *Nature* 2006, **439**:557-562.
30. Yeeles JT, Marians KJ: **The Escherichia coli replisome is inherently DNA damage tolerant**. *Science* 2011, **334**:235-238.
31. Minca EC, Kowalski D: **Multiple Rad5 activities mediate sister chromatid recombination to bypass DNA damage at stalled replication forks**. *Molecular cell* 2010, **38**:649-661.
32. Tercero JA, Diffley JF: **Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint**. *Nature* 2001, **412**:553-557.
33. Vazquez MV, Rojas V, Tercero JA: **Multiple pathways cooperate to facilitate DNA replication fork progression through alkylated DNA**. *DNA repair* 2008, **7**:1693-1704.
34. Cortes-Ledesma F, Aguilera A: **Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange**. *EMBO reports* 2006, **7**:919-926.
35. Roseaulin L, Yamada Y, Tsutsui Y, Russell P, Iwasaki H, Arcangioli B: **Mus81 is essential for sister chromatid recombination at broken replication forks**. *The EMBO journal* 2008, **27**:1378-1387.
36. Doksani Y, Bermejo R, Fiorani S, Haber JE, Foiani M: **Replicon dynamics, dormant origin firing, and terminal fork integrity after double-strand break formation**. *Cell* 2009, **137**:247-258.

37. Ray Chaudhuri A, Hashimoto Y, Herrador R, Neelsen KJ, Fachinetti D, Bermejo R, Cocito A, Costanzo V, Lopes M: **Topoisomerase I poisoning results in PARP-mediated replication fork reversal**. *Nature structural & molecular biology* 2012, **19**:417-423.
38. Dronkert ML, Kanaar R: **Repair of DNA interstrand cross-links**. *Mutation research* 2001, **486**:217-247.
39. Lambert S, Mason SJ, Barber LJ, Hartley JA, Pearce JA, Carr AM, McHugh PJ: **Schizosaccharomyces pombe checkpoint response to DNA interstrand cross-links**. *Molecular and cellular biology* 2003, **23**:4728-4737.
40. Huang M, Kim JM, Shiotani B, Yang K, Zou L, D'Andrea AD: **The FANCM/FAAP24 complex is required for the DNA interstrand crosslink-induced checkpoint response**. *Molecular cell* 2010, **39**:259-268.
41. Hanada K, Budzowska M, Modesti M, Maas A, Wyman C, Essers J, Kanaar R: **The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks**. *The EMBO journal* 2006, **25**:4921-4932.
42. Long DT, Raschle M, Joukov V, Walter JC: **Mechanism of RAD51-dependent DNA interstrand cross-link repair**. *Science* 2011, **333**:84-87.
43. Azvolinsky A, Dunaway S, Torres JZ, Bessler JB, Zakian VA: **The S. cerevisiae Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes**. *Genes & development* 2006, **20**:3104-3116.
44. Labib K: **How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells?** *Genes & development* 2010, **24**:1208-1219.
45. Torres JZ, Bessler JB, Zakian VA: **Local chromatin structure at the ribosomal DNA causes replication fork pausing and genome instability in the absence of the S. cerevisiae DNA helicase Rrm3p**. *Genes & development* 2004, **18**:498-503.
46. Sabouri N, McDonald KR, Webb CJ, Cristea IM, Zakian VA: **DNA replication through hard-to-replicate sites, including both highly transcribed RNA Pol II and Pol III genes, requires the S. pombe Pfh1 helicase**. *Genes & development* 2012, **26**:581-593.
47. Torres JZ, Schnakenberg SL, Zakian VA: **Saccharomyces cerevisiae Rrm3p DNA helicase promotes genome integrity by preventing replication fork stalling: viability of rrm3 cells requires the intra-S-phase checkpoint and fork restart activities**. *Molecular and cellular biology* 2004, **24**:3198-3212.
48. Steinacher R, Osman F, Dalgaard JZ, Lorenz A, Whitby MC: **The DNA helicase Pfh1 promotes fork merging at replication termination sites to ensure genome stability**. *Genes & development* 2012, **26**:594-602.
49. Azvolinsky A, Giresi PG, Lieb JD, Zakian VA: **Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae**. *Molecular cell* 2009, **34**:722-734.
50. Sofueva S, Osman F, Lorenz A, Steinacher R, Castagnetti S, Ledesma J, Whitby MC: **Ultrafine anaphase bridges, broken DNA and illegitimate recombination induced by a replication fork barrier**. *Nucleic acids research* 2011, **39**:6568-6584.
51. Possoz C, Filipe SR, Grainge I, Sherratt DJ: **Tracking of controlled Escherichia coli replication fork stalling and restart at repressor-bound DNA in vivo**. *The EMBO journal* 2006, **25**:2596-2604.
52. Dubarry M, Loiodice I, Chen CL, Thermes C, Taddei A: **Tight protein-DNA interactions favor gene silencing**. *Genes & development* 2011, **25**:1365-1370.

53. Tsang E, Carr AM: **Replication fork arrest, recombination and the maintenance of ribosomal DNA stability**. *DNA repair* 2008, **7**:1613-1623.
54. Wang JD, Berkmen MB, Grossman AD: **Genome-wide coorientation of replication and transcription reduces adverse effects on replication in Bacillus subtilis**. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**:5608-5613.
55. Srivatsan A, Tehranchi A, MacAlpine DM, Wang JD: **Co-orientation of replication and transcription preserves genome integrity**. *PLoS genetics* 2010, **6**:e1000810.
56. Mohanty BK, Bairwa NK, Bastia D: **The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of Saccharomyces cerevisiae**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**:897-902.
57. Biswas S, Bastia D: **Mechanistic insights into replication termination as revealed by investigations of the Reb1-Ter3 complex of Schizosaccharomyces pombe**. *Molecular and cellular biology* 2008, **28**:6844-6857.
58. Dalgaard JZ, Klar AJ: **A DNA replication-arrest site RTS1 regulates imprinting by determining the direction of replication at mat1 in S. pombe**. *Genes & development* 2001, **15**:2060-2068.
59. Arcangioli B, Klar AJ: **A novel switch-activating site (SAS1) and its cognate binding factor (SAP1) required for efficient mat1 switching in Schizosaccharomyces pombe**. *The EMBO journal* 1991, **10**:3025-3032.
60. Egel R: **Fission yeast mating-type switching: programmed damage and repair**. *DNA repair* 2005, **4**:525-536.
61. Eydmann T, Sommariva E, Inagawa T, Mian S, Klar AJ, Dalgaard JZ: **Rtf1-mediated eukaryotic site-specific replication termination**. *Genetics* 2008, **180**:27-39.
62. Inagawa T, Yamada-Inagawa T, Eydmann T, Mian IS, Wang TS, Dalgaard JZ: **Schizosaccharomyces pombe Rtf2 mediates site-specific replication termination by inhibiting replication restart**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**:7927-7932.
63. Lambert S, Mizuno K, Blaisonneau J, Martineau S, Chanet R, Freon K, Murray JM, Carr AM, Baldacci G: **Homologous recombination restarts blocked replication forks at the expense of genome rearrangements by template exchange**. *Molecular cell* 2010, **39**:346-359.
64. Lambert S, Watson A, Sheedy DM, Martin B, Carr AM: **Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier**. *Cell* 2005, **121**:689-702.
65. Mizuno K, Lambert S, Baldacci G, Murray JM, Carr AM: **Nearby inverted repeats fuse to generate acentric and dicentric palindromic chromosomes by a replication template exchange mechanism**. *Genes & development* 2009, **23**:2876-2886.
66. Pryce DW, Ramayah S, Jaendling A, McFarlane RJ: **Recombination at DNA replication fork barriers is not universal and is differentially regulated by Swi1**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**:4770-4775.
67. Ahn JS, Osman F, Whitby MC: **Replication fork blockage by RTS1 at an ectopic site promotes recombination in fission yeast**. *The EMBO journal* 2005, **24**:2011-2023.
68. Lorenz A, Osman F, Folkyte V, Sofueva S, Whitby MC: **Fbh1 limits Rad51-dependent recombination at blocked replication forks**. *Molecular and cellular biology* 2009, **29**:4742-4756.
69. Sun W, Nandi S, Osman F, Ahn JS, Jakovleska J, Lorenz A, Whitby MC: **The FANCM ortholog Fml1 promotes recombination at stalled replication forks and limits**

crossing over during DNA double-strand break repair. *Molecular cell* 2008, **32**:118-128.
70. Pomerantz RT, O'Donnell M: **What happens when replication and transcription complexes collide?** *Cell cycle* 2010, **9**:2537-2543.
71. Kaplan DL, Bastia D: **Mechanisms of polar arrest of a replication fork**. *Molecular microbiology* 2009, **72**:279-285.
72. Fachinetti D, Bermejo R, Cocito A, Minardi S, Katou Y, Kanoh Y, Shirahige K, Azvolinsky A, Zakian VA, Foiani M: **Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements**. *Molecular cell* 2010, **39**:595-605.
73. Singh SK, Sabatinos S, Forsburg S, Bastia D: **Regulation of replication termination by Reb1 protein-mediated action at a distance**. *Cell* 2010, **142**:868-878.
74. Chen JD, Pirrotta V: **Multimerization of the Drosophila zeste protein is required for efficient DNA binding**. *The EMBO journal* 1993, **12**:2075-2083.
75. Bastia D, Singh SK: **"Chromosome kissing" and modulation of replication termination**. *Bioarchitecture* 2011, **1**:24-28.
76. Prado F, Aguilera A: **Impairment of replication fork progression mediates RNA polII transcription-associated recombination**. *The EMBO journal* 2005, **24**:1267-1276.
77. Aguilera A, Gomez-Gonzalez B: **Genome instability: a mechanistic view of its causes and consequences**. *Nature reviews. Genetics* 2008, **9**:204-217.
78. Ansari A, Hampsey M: **A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping**. *Genes & development* 2005, **19**:2969-2978.
79. Blobel G: **Gene gating: a hypothesis**. *Proceedings of the National Academy of Sciences of the United States of America* 1985, **82**:8527-8529.
80. Cabal GG, Genovesio A, Rodriguez-Navarro S, Zimmer C, Gadal O, Lesne A, Buc H, Feuerbach-Fournier F, Olivo-Marin JC, Hurt EC, et al.: **SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope**. *Nature* 2006, **441**:770-773.
81. Bermejo R, Capra T, Jossen R, Colosio A, Frattini C, Carotenuto W, Cocito A, Doksani Y, Klein H, Gomez-Gonzalez B, et al.: **The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores**. *Cell* 2011, **146**:233-246.
82. Gomez-Gonzalez B, Garcia-Rubio M, Bermejo R, Gaillard H, Shirahige K, Marin A, Foiani M, Aguilera A: **Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles**. *The EMBO journal* 2011, **30**:3106-3119.
83. Bochman ML, Sabouri N, Zakian VA: **Unwinding the functions of the Pif1 family helicases**. *DNA Repair (Amst)* 2010, **9**:237-249.
84. McGlynn P: **Helicases that underpin replication of protein-bound DNA in Escherichia coli**. *Biochemical Society transactions* 2011, **39**:606-610.
85. Boubakri H, de Septenville AL, Viguera E, Michel B: **The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo**. *The EMBO journal* 2010, **29**:145-157.
86. Lobachev KS, Rattray A, Narayanan V: **Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells**. *Frontiers in bioscience : a journal and virtual library* 2007, **12**:4208-4220.
87. Voineagu I, Narayanan V, Lobachev KS, Mirkin SM: **Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**:9936-9941.

88. Lobachev KS, Shor BM, Tran HT, Taylor W, Keen JD, Resnick MA, Gordenin DA: **Factors affecting inverted repeat stimulation of recombination and deletion in Saccharomyces cerevisiae**. *Genetics* 1998, **148**:1507-1524.
89. Lobachev KS, Stenger JE, Kozyreva OG, Jurka J, Gordenin DA, Resnick MA: **Inverted Alu repeats unstable in yeast are excluded from the human genome**. *The EMBO journal* 2000, **19**:3822-3830.
90. Paek AL, Kaochar S, Jones H, Elezaby A, Shanks L, Weinert T: **Fusion of nearby inverted repeats by a replication-based mechanism leads to formation of dicentric and acentric chromosomes that cause genome instability in budding yeast**. *Genes & development* 2009, **23**:2861-2875.
91. McMurray CT: **Mechanisms of trinucleotide repeat instability during human development**. *Nature reviews. Genetics* 2010, **11**:786-799.
92. Richard GF, Kerrest A, Dujon B: **Comparative genomics and molecular dynamics of DNA repeats in eukaryotes**. *Microbiology and molecular biology reviews : MMBR* 2008, **72**:686-727.
93. Mirkin SM: **DNA structures, repeat expansions and human hereditary disorders**. *Current opinion in structural biology* 2006, **16**:351-358.
94. Kim HM, Narayanan V, Mieczkowski PA, Petes TD, Krasilnikova MM, Mirkin SM, Lobachev KS: **Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair**. *The EMBO journal* 2008, **27**:2896-2906.
95. Wells RD: **DNA triplexes and Friedreich ataxia**. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2008, **22**:1625-1634.
96. Krasilnikova MM, Mirkin SM: **Replication stalling at Friedreich's ataxia (GAA)n repeats in vivo**. *Molecular and cellular biology* 2004, **24**:2286-2295.
97. Shishkin AA, Voineagu I, Matera R, Cherng N, Chernet BT, Krasilnikova MM, Narayanan V, Lobachev KS, Mirkin SM: **Large-scale expansions of Friedreich's ataxia GAA repeats in yeast**. *Molecular cell* 2009, **35**:82-92.
98. Goldfless SJ, Morag AS, Belisle KA, Sutera VA, Jr., Lovett ST: **DNA repeat rearrangements mediated by DnaK-dependent replication fork repair**. *Molecular cell* 2006, **21**:595-604.
99. Sinden RR: **DNA structure and function.** *San Diego: Academic Press* 1994:xxiii,398.
100. Bacolla A, Wojciechowska M, Kosmider B, Larson JE, Wells RD: **The involvement of non-B DNA structures in gross chromosomal rearrangements**. *DNA repair* 2006, **5**:1161-1170.
101. Zhang H, Freudenreich CH: **An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in S. cerevisiae**. *Molecular cell* 2007, **27**:367-379.
102. Ribeyre C, Lopes J, Boule JB, Piazza A, Guedin A, Zakian VA, Mergny JL, Nicolas A: **The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo**. *PLoS genetics* 2009, **5**:e1000475.
103. Paeschke K, Capra JA, Zakian VA: **DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase**. *Cell* 2011, **145**:678-691.
104. Lopes J, Piazza A, Bermejo R, Kriegsman B, Colosio A, Teulade-Fichou MP, Foiani M, Nicolas A: **G-quadruplex-induced instability during leading-strand replication**. *The EMBO journal* 2011, **30**:4033-4046.
105. Sogo JM, Lopes M, Foiani M: **Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects**. *Science* 2002, **297**:599-602.

106. Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA: **Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint**. *Genes & development* 2005, **19**:1040-1052.
107. Myung K, Datta A, Kolodner RD: **Suppression of spontaneous chromosomal rearrangements by S phase checkpoint functions in Saccharomyces cerevisiae**. *Cell* 2001, **104**:397-408.
108. Costanzo V: **Brca2, Rad51 and Mre11: performing balancing acts on replication forks**. *DNA repair* 2011, **10**:1060-1065.
109. Petermann E, Helleday T: **Pathways of mammalian replication fork restart**. *Nat Rev Mol Cell Biol* 2010, **11**:683-687.
110. Arlt MF, Wilson TE, Glover TW: **Replication stress and mechanisms of CNV formation**. *Current opinion in genetics & development* 2012, **22**:204-210.
111. Durkin SG, Ragland RL, Arlt MF, Mulle JG, Warren ST, Glover TW: **Replication stress induces tumor-like microdeletions in FHIT/FRA3B**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**:246-251.
112. Bester AC, Roniger M, Oren YS, Im MM, Sarni D, Chaoat M, Bensimon A, Zamir G, Shewach DS, Kerem B: **Nucleotide deficiency promotes genomic instability in early stages of cancer development**. *Cell* 2011, **145**:435-446.
113. Ozeri-Galai E, Bester AC, Kerem B: **The complex basis underlying common fragile site instability in cancer**. *Trends in genetics : TIG* 2012, **28**:295-302.
114. Blumrich A, Zapatka M, Brueckner LM, Zheglo D, Schwab M, Savelyeva L: **The FRA2C common fragile site maps to the borders of MYCN amplicons in neuroblastoma and is associated with gross chromosomal rearrangements in different cancers**. *Human molecular genetics* 2011, **20**:1488-1501.
115. Pelliccia F, Bosco N, Rocchi A: **Breakages at common fragile sites set boundaries of amplified regions in two leukemia cell lines K562 - Molecular characterization of FRA2H and localization of a new CFS FRA2S**. *Cancer letters* 2010, **299**:37-44.
116. Glover TW, Berger C, Coyle J, Echo B: **DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes**. *Human genetics* 1984, **67**:136-142.
117. Ozeri-Galai E, Lebofsky R, Rahat A, Bester AC, Bensimon A, Kerem B: **Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites**. *Mol Cell* 2011, **43**:122-131.
118. Le Tallec B, Dutrillaux B, Lachages AM, Millot GA, Brison O, Debatisse M: **Molecular profiling of common fragile sites in human fibroblasts**. *Nature structural & molecular biology* 2011, **18**:1421-1423.
119. Letessier A, Millot GA, Koundrioukoff S, Lachages AM, Vogt N, Hansen RS, Malfoy B, Brison O, Debatisse M: **Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site**. *Nature* 2011, **470**:120-123.
120. Palakodeti A, Lucas I, Jiang Y, Young DJ, Fernald AA, Karrison T, Le Beau MM: **Impaired replication dynamics at the FRA3B common fragile site**. *Human molecular genetics* 2010, **19**:99-110.
121. Petermann E, Orta ML, Issaeva N, Schultz N, Helleday T: **Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair**. *Molecular cell* 2010, **37**:492-502.
122. Kawabata T, Luebben SW, Yamaguchi S, Ilves I, Matise I, Buske T, Botchan MR, Shima N: **Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression**. *Molecular cell* 2011, **41**:543-553.

123. Ge XQ, Jackson DA, Blow JJ: **Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress**. *Genes & development* 2007, **21**:3331-3341.
124. Woodward AM, Gohler T, Luciani MG, Oehlmann M, Ge X, Gartner A, Jackson DA, Blow JJ: **Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress**. *The Journal of cell biology* 2006, **173**:673-683.
125. Zlotorynski E, Rahat A, Skaug J, Ben-Porat N, Ozeri E, Hershberg R, Levi A, Scherer SW, Margalit H, Kerem B: **Molecular basis for expression of common and rare fragile sites**. *Molecular and cellular biology* 2003, **23**:7143-7151.
126. Debatisse M, Le Tallec B, Letessier A, Dutrillaux B, Brison O: **Common fragile sites: mechanisms of instability revisited**. *Trends in genetics : TIG* 2012, **28**:22-32.
127. Smith DI, McAvoy S, Zhu Y, Perez DS: **Large common fragile site genes and cancer**. *Seminars in cancer biology* 2007, **17**:31-41.
128. Helmrich A, Stout-Weider K, Hermann K, Schrock E, Heiden T: **Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes**. *Genome research* 2006, **16**:1222-1230.
129. Helmrich A, Ballarino M, Tora L: **Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes**. *Molecular cell* 2011, **44**:966-977.
130. Jiang Y, Lucas I, Young DJ, Davis EM, Karrison T, Rest JS, Le Beau MM: **Common fragile sites are characterized by histone hypoacetylation**. *Human molecular genetics* 2009, **18**:4501-4512.

Figure legends

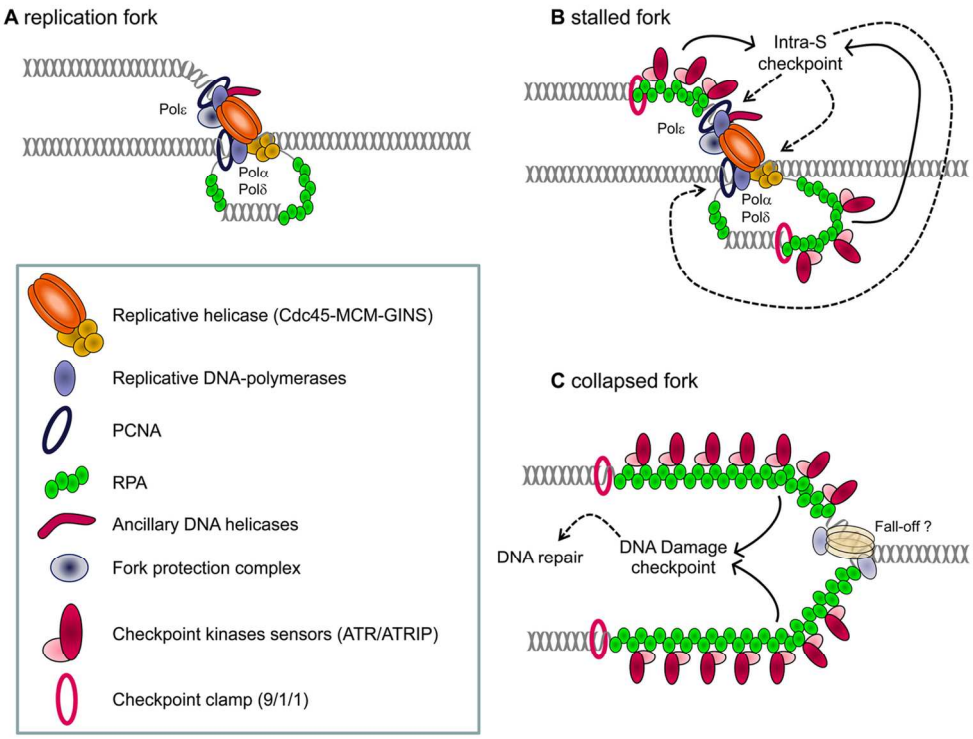
Figure 1: Maintaining replication fork integrity

A. Schematic model of a moving replication fork, showing of the replicative helicase (Cdc45-MCM-GINS) associated with the DNA polymerases Polα/δ and Polε on the lagging and leading strand, respectively. The Fork protection complex, composed of at least Mrc1 and Tof1 in budding yeast, and ancillary DNA helicases such as Rrm3 are thought to travel with the replication fork via their interaction with Polε. B. The slowing down or the transient pause of fork progression leads to the exposure of a limited amount of ssDNA (around 300 nucleotides) at the fork. This triggers the activation of the intra-S checkpoint which, in turn, regulates the activity of DNA polymerases and replicative helicases to limit uncoupling. Once the cause of impediment to fork progression is resolved, the fork resumes its progression without the assistance of additional mechanisms. C. At collapsed forks, a large amount of ssDNA is exposed, activating the DNA-damage checkpoint. The replisome losses functionality and might even dissociate from the fork. The restart of the fork thus requires the assistance of additional mechanisms, such as homologous recombination, to rebuild a replisome. Collapsed forks may arise upon certain forms of fork-arrest or when the intra-S checkpoint fails.

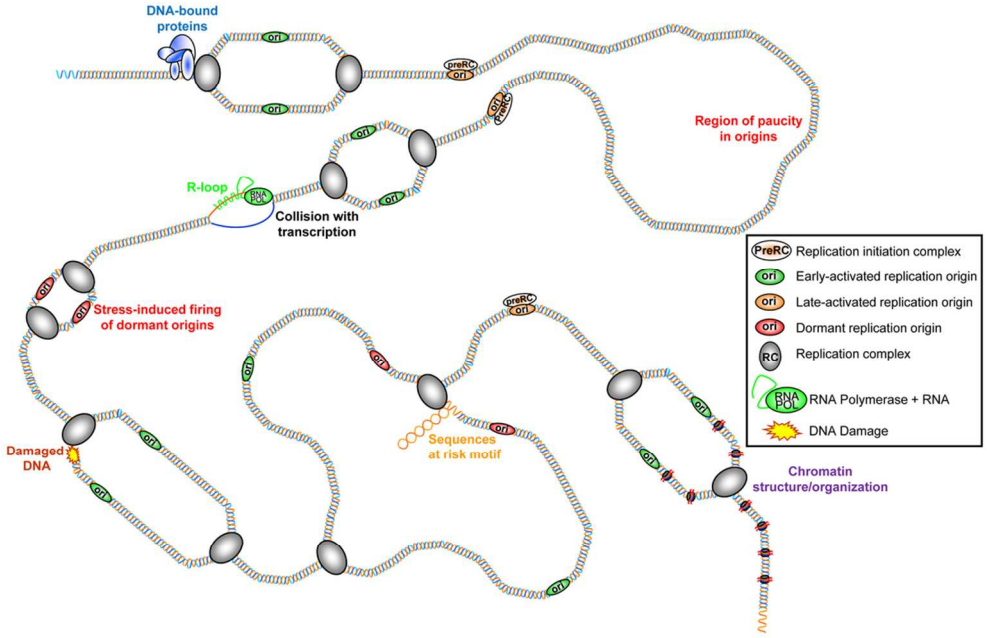
Figure 2: Overview of obstacles causing RFB.

Typical obstacles causing RFBs (DNA-bound proteins, chromatin organisation, at risk sequences motif, DNA-damage) and interference between replication and transcription all have the potential to interfere with the progression of replication forks and its associated DNA synthesis. Global replication-stress and impediments to fork progression result in the activation of dormant origins that helps to complete DNA replication. In region of paucity in replication origins, completion of replication relies on long-traveling forks which are thus particularly sensitive to fork arrest and replication inhibition.

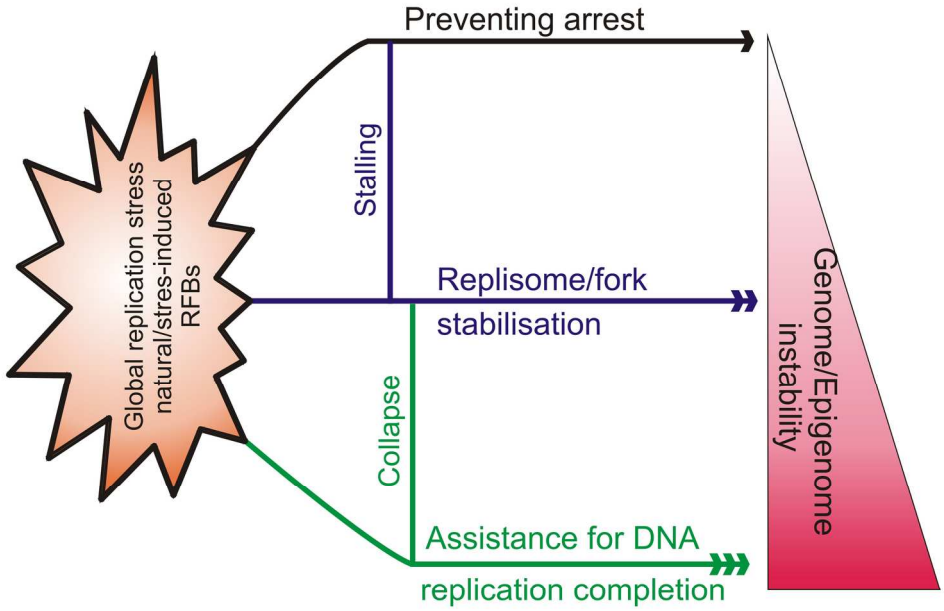
Figure 3: Three distinct strategies that prevent genome stability at arrested forks.



132x101mm (300 x 300 DPI)



103x66mm (300 x 300 DPI)



Cell lines of defence:

DNA repair/Helicases/Translocases

Checkpoint pathways

Fork-restart mechanisms

170x155mm (300 x 300 DPI)