Differences in the DNA replication of unicellular eukaryotes and metazoans: known unknowns

Alessia Errico1 & Vincenzo Costanzo2+
London Research Institute, Hertfordshire, UK

Although the basic mechanisms of DNA synthesis are conserved across species, there are differences between simple and complex organisms. In contrast to lower eukaryotes, replication origins in complex eukaryotes lack DNA sequence specificity, can be activated in response to stressful conditions and require poorly conserved factors for replication firing. The response to replication fork damage is monitored by conserved proteins, such as the TIPIN–TIM–CLASPIN complex. The absence of this complex induces severe effects on yeast replication, whereas in higher eukaryotes it is only crucial when the availability of replication origins is limiting. Finally, the dependence of DNA replication on homologous recombination proteins such as RAD51 and the MRE11–RAD50–NBS1 complex is also different; they are dispensable for yeast S-phase but essential for accurate DNA replication in metazoans under unchallenged conditions. The reasons for these differences are not yet understood. Here, we focus on some of these known unknowns of DNA replication.

Keywords: DNA replication; checkpoint; homologous recombination; yeast; metazoans

EMBO reports (2010) 11, 270–278. doi:10.1038/embor.2010.27
See Glossary for abbreviations used in this article.

Introduction

All eukaryotes use similar machinery and regulatory mechanisms for DNA duplication and cell division. Indeed, the main players have been conserved throughout evolution from unicellular organisms to mammals (Table 1). However, despite this high level of conservation, important differences can be observed between higher and lower eukaryotes. Many of these variations have been described in the main DNA replication model systems, which are the *Xenopus laevis* egg extract and mammalian cell cultures for complex eukaryotes and metazoans, and yeast cells for unicellular eukaryotes. Here, we analyse the basic processes of DNA replication in which lower and higher eukaryotes differ, although the reasons for these differences remain obscure.

Establishment and selection of replication origins

DNA replication is a tightly controlled process. Several mechanisms have evolved to ensure that no regions of DNA are left unreplicated or are replicated more than once in every cell cycle. Prokaryotes use mostly a single origin to replicate their small genome, whereas eukaryotes—which have a high DNA content—have multiple replication origins distributed throughout the DNA. The first step in DNA replication is the assembly of a pre-replicative complex (pre-RC) at each origin (Diffley et al., 1995; Rowles et al., 1999), which consists in the binding of ORC1–6 (Bell & Stillman, 1992; Rabitsch et al., 2001), Cdc6 and Cdt1 (Gillespie et al., 2001) to chromatin, followed by the loading of the essential helicase activity, MCM2–7. In early S-phase, the pre-RC complex is converted into an initiation complex—which promotes DNA unwinding and polymerase loading (Diffley et al., 1995)—through the activity of S-phase kinases, the CDKs and DDK.

In *Saccharomyces cerevisiae*, replication origins are specified by a particular DNA sequence known as autonomously replicating sequence (ARS), which recruits the ORC. The ARS consists of an essential 11 bp consensus sequence (known as ACS) and several elements that contribute to promote initiation. Replication origins in the fission yeast *Schizosaccharomyces pombe* are much larger—500–1,000 bp compared with the 150 bp in *S. cerevisiae*—and do not display a clear consensus other than being extremely rich in A+T (Dai et al., 2005; Segurado et al., 2003). In contrast to yeast, replication origins in higher eukaryotes are defined less rigidly and apparently do not have a specific sequence requirement. An extreme case of a lack of ACS for origin specification can be found in embryonic systems—such as *X. laevis* and *Drosophila melanogaster*—in which DNA replication initiates at seemingly random sites spaced 10–15 kb apart (Blow et al., 2001; Shinomiy & Ina, 1991). A main difference between embryonic and somatic cells is the absence of transcription in embryos, which initiates at the midblastula transition—a stage with more defined initiation zones (Hyrien et al., 1995; Sasaki et al., 1999). In somatic cells replication origins are less frequent, being present approximately once every 150 kb. Until recently, only about 40 origins had been characterized in 10 metazoan organisms, from fruit flies to humans, with little evidence of a common consensus (Dimitrova et al., 1996). Two genome-wide approaches have led to the mapping of a greater number of initiation sites in the HeLa cell line genome. The first study characterized the presence of 283 origins using the HeLa S3 suspension cell line (Cadoret et al., 2008) and the second, more recent study identified 150 new origins in adherent HeLa cells (Karnani et al., 2008).
Although the two studies agree only partly about individual origins (Karnani et al., 2009), they both define specific features for the metazoan initiation sites, revealing a correlation between origins and transcription start sites (Cadoret et al., 2008; Karnani et al., 2009). These studies have started to clarify the connection between origin selection, gene regulation and chromatin structure. Future studies will probably lead to a clearer consensus that defines the metazoan origins of replication, although it seems unlikely to be related to the yeast origins.

The fact that metazoan origins occur at many sites in large initiation zones and tend to be organized in clusters is a further difference with respect to yeast. It has been proposed that these features confer a selective evolutionary advantage for complex organisms with large genomes, enabling them to easily replicate newly acquired DNA sequences (Hyrien & Mechali, 1993). These differences might also reflect the requirement for integrating the control of DNA replication with cell differentiation and organism development. In this case, origin specification could be dictated by the high-order structure of chromatin, which changes during cell differentiation.

Another intriguing feature of replication origin organization is that the number of MCM2–7 complexes loaded on DNA exceeds the number of ORC1–6 complexes. These extra MCM2–7 complexes have been proposed to be additional sites from which replication forks are stalled or slowed. This mechanism is potentially relevant to ensure the complete replication of the genome in the presence of obstacles to replication forks (Fig 1; Ge et al., 2007; Ibarra et al., 2008; Woodward et al., 2006). Whether dormant origins exist in yeast and what their role would be is not clear. A recent study has unexpectedly shown that the presence of double-strand breaks (DSBs) in yeast can trigger the firing of nearby dormant origins (Ibarra et al., 2008; Lei et al., 1996). Interestingly, these supplementary origins remain ‘dormant’ during S phase in X. laevis and mammals and only fire when replication forks stall or slow down. This mechanism is potentially relevant to ensure the complete replication of the genome in the presence of obstacles to replication forks (Fig 1; Ge et al., 2007; Ibarra et al., 2008; Woodward et al., 2006). Whether dormant origins exist in yeast and what their role would be is not clear. A recent study has unexpectedly shown that the presence of double-strand breaks (DSBs) in yeast can trigger the firing of nearby dormant origins (Doksani et al., 2009), suggesting that this feature is conserved across species.

The mechanism that leads to the firing of dormant origins is unclear. One hypothesis is that their firing is not due to an active mechanism, but to a kinetic and probabilistic process whereby, when forks stall, dormant origins have more time and a greater chance of being used before the region they occupy is replicated and inactivated by a fork coming from an adjacent active origin (Ge et al., 2007). There could also be an active process regulating dormant origins in response to replicative stress. The ATM/ATR-dependent intra-S-phase checkpoint regulates origin firing, thereby limiting the number of origins that actually fire in the presence of replication stress (Shechter et al., 2004). The intra-S-phase checkpoint needs to be downregulated transiently for the activation of dormant origins (Ge et al., 2007; Woodward et al., 2006) and Plx1 has been recently shown to have an important role in its suppression (Trenz et al., 2008). The ATM/ATR-dependent phosphorylation of MCM2 is essential for this Plx1-mediated function (Cortez et al., 2004; Yoo et al., 2004), as it promotes Plx1 binding to the MCM2–7 complex through its Polo box domain (Trenz et al., 2008). When this Plx1/MCM2–7 complex is in the proximity of stalled replication forks, it seems to be involved in the release of CHK1-mediated suppression of nearby dormant origins (Trenz et al., 2008). However, how Plx1 suppresses CHK1 activity in this process remains unclear. The phosphorylation of adaptors required for CHK1 activation might be involved in this pathway.

### Table 1 | Yeast and vertebrate homologues of proteins involved in different aspects of DNA replication

<table>
<thead>
<tr>
<th>Vertebrates</th>
<th>Saccharomyces cerevisiae</th>
<th>Schizosaccharomyces pombe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-replicative complex components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORC1–6</td>
<td>Orc1–6</td>
<td>Orc1–6</td>
</tr>
<tr>
<td>CDC6</td>
<td>Cdc6</td>
<td>Cdc6</td>
</tr>
<tr>
<td>MCM2</td>
<td>Mcm2</td>
<td>Mcm2/Cdc19/Nda1</td>
</tr>
<tr>
<td>MCM3</td>
<td>Mcm3</td>
<td>Mcm3</td>
</tr>
<tr>
<td>MCM4</td>
<td>Mcm4/Cdc54</td>
<td>Cdc21</td>
</tr>
<tr>
<td>MCM5</td>
<td>Mcm5/Cdc46</td>
<td>Mcm5/Nda4</td>
</tr>
<tr>
<td>MCM6</td>
<td>Mcm6</td>
<td>Mcm6/Mix5</td>
</tr>
<tr>
<td>MCM7</td>
<td>Mcm7/Cdc47</td>
<td>Mcm7</td>
</tr>
<tr>
<td><strong>Pre-initiation complex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCM10</td>
<td>Mcm10</td>
<td>Cdc23</td>
</tr>
<tr>
<td>CDC45</td>
<td>Cdc45/Sld4</td>
<td>Cdc45/Sna41</td>
</tr>
<tr>
<td>TOPBP1</td>
<td>Dbp11</td>
<td>Cut5/Rad4</td>
</tr>
<tr>
<td>—</td>
<td>Sld2</td>
<td>Drc1</td>
</tr>
<tr>
<td>—</td>
<td>Sld3</td>
<td>Sld3</td>
</tr>
<tr>
<td>SLD5</td>
<td>Sld5</td>
<td>Sld5</td>
</tr>
<tr>
<td>PSF1</td>
<td>Psf1</td>
<td>Psf1</td>
</tr>
<tr>
<td>PSF2</td>
<td>Psf2</td>
<td>Psf2</td>
</tr>
<tr>
<td>PSF3</td>
<td>Psf3</td>
<td>Psf3</td>
</tr>
<tr>
<td><strong>Replication pausing complex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIPIN</td>
<td>Cam3</td>
<td>Swi3</td>
</tr>
<tr>
<td>TIM1</td>
<td>Tof1</td>
<td>Swi1</td>
</tr>
<tr>
<td>CLASPIN</td>
<td>Mrc1</td>
<td>Mrc1</td>
</tr>
<tr>
<td>AND1</td>
<td>Ctf4</td>
<td>Mc1</td>
</tr>
<tr>
<td><strong>S-phase and DNA damage checkpoint</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATR</td>
<td>Mec1</td>
<td>Rad3</td>
</tr>
<tr>
<td>ATRIP</td>
<td>Ddc2</td>
<td>Rad26</td>
</tr>
<tr>
<td>ATM</td>
<td>Tel1</td>
<td>Tel1</td>
</tr>
<tr>
<td>CHK1</td>
<td>Chk1</td>
<td>Chl1/Rad27</td>
</tr>
<tr>
<td>CHK2</td>
<td>Rad53</td>
<td>Cds1</td>
</tr>
<tr>
<td><strong>DNA repair (HR)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRE11</td>
<td>Mre11</td>
<td>Rad32</td>
</tr>
<tr>
<td>RAD50</td>
<td>Rad50</td>
<td>Rad50</td>
</tr>
<tr>
<td>NBS1</td>
<td>Xrs2</td>
<td>Nbs1</td>
</tr>
<tr>
<td>RAD51</td>
<td>Rad51</td>
<td>Rad51</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Rad9</td>
<td>Crb2</td>
</tr>
<tr>
<td>BRCA2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>RAD54</td>
<td>Rad54</td>
<td>Rad54</td>
</tr>
<tr>
<td>RAD52</td>
<td>Rad52</td>
<td>Rad52</td>
</tr>
<tr>
<td><strong>DNA repair (NHEJ)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIGIV</td>
<td>LigIV</td>
<td>LigIV</td>
</tr>
<tr>
<td>XRCC4</td>
<td>Lif1</td>
<td>—</td>
</tr>
<tr>
<td>XLF</td>
<td>Nej1</td>
<td>Xlf1/Nej1</td>
</tr>
<tr>
<td>Ku70</td>
<td>Ku70</td>
<td>Ku70</td>
</tr>
<tr>
<td>DNA-PK</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

HR, homologous recombination; NHEJ, non-homologous end-joining.
Overall, these data suggest that MCM2-recruited Plx1 promotes the progression of replication in the presence of replication stress. Consistent with this, DNA replication in the absence of Plx1 leads to the accumulation of DSBs, further supporting the role of Plx1 in promoting genome stability during S phase (Trenz et al., 2008). These findings have been recently confirmed in other vertebrates by showing that PLK1—a Plx1 orthologue—is required to promote DNA replication recovery after fork stalling by releasing the inhibition on origin firing in DT40 cells that lack FANCM (Schwab et al., 2010). This pathway was also shown to be dependent on MCM2 phosphorylation by ATR, demonstrating that this mechanism is highly conserved in metazoans. Neither the phosphorylation of MCM2 by ATM/ATR, nor a role for Plx1 orthologues in DNA replication have been described in yeast cells. Further work is required to understand in detail how dormant origins are regulated and their role in promoting genome replication in different organisms.

Initiation of DNA replication

Two replication factors—Sld2 and Sld3—have recently emerged as crucial in the cell-cycle-dependent control of DNA replication initiation in yeast (Tanaka et al., 2007; Zegerman & Diffley, 2007). Sld2 and Sld3 represent the minimal set of substrates that need to be phosphorylated by CDKs to initiate DNA replication (Tanaka et al., 2007; Zegerman & Diffley, 2007). Their phosphorylation allows them to interact with a Dpb11—a BRCT-containing protein—which seems to facilitate the loading of Cdc45 and, therefore, origin firing. Orthologues of Sld2, Sld3 and Dpb11 can be found in fungi—although sequence conservation is low even among related species—but their presence in other organisms is uncertain. Most importantly, although TOPBP1/CUT5/MUS101 is the recognized orthologue of Dpb11 (Table 1; Garcia et al., 2005), no clear orthologues of Sld2 and Sld3 have been identified so far in animal cells. RECQL4, which interacts with TOPBP1 and is required for DNA replication, has been recently suggested to be the putative orthologue of Sld2. However, RECQL4 has a limited homology to Sld2 and its function does not seem to be regulated by CDK-dependent phosphorylation (Sangrithi et al., 2005). Therefore, although RECQL4 is necessary for origin firing, it is possibly not a crucial CDK target. An additional difference between yeast and mammals emerged from a recent study on the formation of the CMG complex, which is an association between CDC45, MCM2–7 and the GINS complex that requires the presence of RECQL4, Ctf4/AND1 and MCM10 to be assembled. The CMG complex seems to have a crucial role in the formation and progression of replication forks and, surprisingly, the depletion of TOPBP1—which has an essential role in the chromatin loading of CDC45 and GINS in yeast cells—does not significantly affect CMG complex formation in mammals (Im et al., 2009).

Sld3 seems to be even more divergent, as no putative orthologue has been identified according to primary sequence. The CDK-dependent regulation of the initiation of DNA replication is a conserved process and therefore it is likely that functional orthologues of both Sld2 and Sld3 will be found. It is tempting to speculate...
that owing to the variety of different cells in multicellular organisms, replication initiation requires a more complex regulation, probably through the phosphorylation of many substrates that fulfill the roles Sld2 and Sld3 have in yeast.

Another interesting distinction between higher and lower eukaryotes is that several proteins—such as geminin, which is a CDT1-regulatory protein (McGarry & Kirschner, 1998), and MCM9—have been identified as the main regulators of DNA replication factors only in higher eukaryotes. In all eukaryotic organisms, the assembly of a new origin is suppressed by a high concentration of CDKs. In addition to CDK activity, multicellular eukaryotes use geminin to regulate the assembly of replication origins and prevent re-replication. Geminin interacts tightly with CDT1, thereby preventing the binding of the MCM2–7 complex to origins (Wohlschlegel et al., 2000). The binding of geminin to CDT1 blocks licensing, whereas it is enabled by MCM9 binding to CDT1—which prevents the loading of geminin onto chromatin during licensing (Lutzmann & Mechali, 2008). Furthermore, yeast Cdt1 binds directly to the Mcm2–7 complex, whereas in multicellular organisms, a direct interaction between these proteins in the absence of chromatin has not been verified in vivo (Seo et al., 2005). The requirement of complex organisms for a more sophisticated regulation to integrate DNA licensing and replication within a development programme might explain the existence of additional factors in the metazoan pre-RC.

MCM8, which is an additional member of the MCM2–7 family, has also been found in higher eukaryotes (Maiorano et al., 2005). Studies performed in X. laevis egg extracts showed that MCM8 binds to chromatin after DNA synthesis is initiated and required for the efficient progression of replication forks. These data suggest that MCM8 is not involved in origin licensing but functions specifically as a DNA helicase in vivo, perhaps contributing to DNA unwinding during the elongation process of DNA replication (Maiorano et al., 2005). The requirement for MCM8 in higher eukaryotes might be related to the size and the complexity of the genome, associated with the need to ensure efficient processivity in replicating large genomes.

Together, these data suggest that the proteins and mechanisms involved in the initiation of DNA replication in higher eukaryotes differ from those in yeast systems (Fig 2).

**Stalled forks: a task for the replication pausing complex**

Once two adjacent origins have fired, the two converging forks progress until they meet, ensuring the complete replication of the DNA segment. However, forks can stall if they encounter DNA damage. To ensure that replication will resume after the obstacle is removed, it is important to stabilize the replication fork so that the replisome components do not dissociate. To this end, several proteins that are not essential for DNA synthesis are present at the replication fork through their interaction with members of the replisome (Branzei & Foiani, 2005; Gamburg et al., 2006). Among these, TIM1, TIPIN and CLASPIN have been identified—both in yeast and higher eukaryotes—as members of the ‘replication pausing complex’ that contributes both to fork stabilization and to checkpoint activation. A central role in this S-phase checkpoint response is carried out by the DNA-damage-sensing kinase ATR, its functional homologues in budding and fission yeast—which are Mec1 and Rad3, respectively—and the ATR downstream kinase CHK1. These kinases are required to promote fork stability both in the absence and in the presence of DNA damage (Branzei & Foiani, 2005). The mechanism that senses fork lesions has been studied in many systems, including the X. laevis egg extract. Work with this model suggested that when the polymerase encounters a lesion its progression is blocked, whereas the helicase keeps unwinding the DNA (Byun et al., 2005). The uncoupling between the stalled polymerase and the helicase generates a segment of single-strand (ss) DNA that constitutes the signal for recruiting the ATR–ATRIP complex through RPA binding. In budding yeast, Tof1 (TIM1), Csm3 (TIPIN) and Mrcl (CLASPIN) proteins are required for the Mec1/Rad53 (ATR/CHK2) checkpoint response that prevents the collapse of stalled replication forks and enables DNA replication to restart after recovery (Branzei & Foiani, 2005). In mammals, TIM1, TIPIN and CLASPIN seem to mediate the ATR–CHK1 signalling cascade (Chou & Elledge, 2006; Errico et al., 2007; Unsal-Kacmaz et al., 2007). Intriguingly, these proteins are also part of the replisome in the absence of DNA damage and travel with the replication fork (Errico et al., 2007; Katou et al., 2003; Tanaka et al., 2009). Consistent with this, yeast proteins Mrcl and Tof1 are important for the regulation of the normal progression of DNA replication (Hodgson et al., 2007; Katou et al., 2003; Tourriere et al., 2005), and a reduction in the expression levels of
components such as Csm3/TIPIN—couple the helicase to POL on the lagging strand template (Errico et al., 2009). The AND1–POL interaction when the dormant origins were suppressed (Errico et al., 2005; Zhou & Wang, 2004; Zhu et al., 2007), whereas the TIPIN–AND1 interaction has only been reported in X. laevis (Errico et al., 2009). A conserved mechanism might exist whereby Ctf4/AND1—and probably other replisome components such as Csm3/TIPIN—couple the helicase to POL on the lagging strand template (Errico et al., 2009; Gambus et al., 2009; Tanaka et al., 2009). Work in yeast has demonstrated that Mrc1 interacts with the catalytic subunit of DNA Polε, the leading strand polymerase, suggesting that Mrc1 is instead involved in coupling polymerization and unwinding on the leading strand at the replication fork (Lou et al., 2008).

Overall, these observations indicate that TIPIN, TIM1 and CLASPIN are structural components of the replication fork, representing a physical and functional link between the MCM2–7 helicase and other replication factors, such as DNA polymerases (Fig 3), and ensuring the stability of the replisome, which is a prerequisite for resuming DNA replication after stalling. The presence of orthologues of these proteins with similar functions in yeast indicates that the overall process of fork stabilization is conserved between complex and simple eukaryotes. However, these mechanisms seem to be partly redundant in higher eukaryotes to ensure that the disruption of the function of one gene is not detrimental to the whole process. This is consistent with a greater level of redundancy in relation to critical biological processes in higher eukaryotes.

Dealing with DSBs during DNA replication

DSBs are a significant threat to genome integrity and can be generated by genotoxic agents. However, the most common cause of these lesions in proliferating cells is aberrant DNA replication (Costanzo et al., 2001; Haber, 1998; Kuzminov, 2001). Eukaryotic cells repair DSBs through two main DNA repair pathways: homologous recombination (HR) and non-homologous end joining (NHEJ; Fig 4; Valerie & Povirk, 2003). HR uses an undamaged template—a sister chromatid or homologous chromosome—to restore chromosome integrity without any loss of genetic information. It occurs only during the S and G2 phases of the cell cycle, when sister chromatids are available, and relies on several proteins including Rad51/52/54/50, Mre11, Nbs1, RPA and Brc1/2 (Li & Heyer, 2008). By contrast, NHEJ is potentially mutagenic, as broken ends are processed and directly religated in the absence of homologous pairing (Lewis & Resnick, 2000; Lieber et al., 2003). The Mre11–Rad50–Nbs1 (MRN) complex has a crucial role in the DNA damage response and, together with ATM, is the primary sensor of DSBs. The MRN complex is also important in the initial steps of both HR and NHEJ (Mimitou & Symington, 2009).

The relative contribution of NHEJ and HR to DSB repair varies substantially between budding yeast and mammalian cells. In S. cerevisiae, DSBs are repaired mainly through the HR pathway, whereas NHEJ has only a minor role (Aylon & Kupiec, 2004; Lee et al., 1999). By contrast, NHEJ seems to be of greater importance in mammalian cells (Crichtlow & Jackson, 1998), as it is responsible for the repair of more than 60% of the exogenously induced DSBs in mouse embryonic stem cells (Liang et al., 1998). Surprisingly, mammalian cells that are deficient in NHEJ exhibit few spontaneous chromosome breaks and are viable, although this is the pathway that is preferentially used in vertebrates (Sonoda et al., 2006). This could be associated with the fact that NHEJ is dispensable and probably actively suppressed during S phase.

When sister chromatids are available, HR becomes the repair mechanism of choice for DSBs arising from collapsed replication forks. Intriguingly, despite the high degree of conservation of the single proteins, the requirement of each protein orthologue for cell survival and their contribution to the HR reaction differs profoundly across different species for many of the HR proteins. For example, yeast mutants lacking Rad51, Rad52 or Rad54 exhibit similar mild
lethality in vertebrate cells, whereas mice carrying disrupted proficient in meiosis (Essers et al. 2004; Llorente & Symington, 2004). In summary, yeast mutants in knockout of MRE11 whereas mutations in the nuclease domain of MRE11 in S. cerevisiae have a limited effect, and Mre11-null cells are mostly viable (Bressan et al., 1998; Krogh et al., 2005; Lewis et al., 2004; Llorente & Symington, 2004). In summary, yeast mutants in many of the key HR proteins are viable, whereas the loss of the same proteins in higher eukaryotes results in cell or embryonic lethality.

The reasons behind this discrepancy are largely unclear. One possible explanation is the greater requirement for HR proteins—such as MRE11, RAD51 or BRCA2—for repairing and restarting stalled and collapsed forks in higher eukaryotes. The larger size of the genome might indeed lead to a higher percentage of stalled and collapsed forks in metazoans compared with yeast cells. For example, about 1 in 12 yeast cells lacking Rad52 gives rise to one dead and one living sister cell—as it would be expected if there were a DSB on one sister chromatid requiring repair (Jim Haber, personal communication). If this lesion frequency is scaled up to the vertebrate genome—which is 400 times larger—one would expect perhaps 30 lesions in the absence of RAD51, which would probably be sufficient to compromise the survival of a vertebrate cell. Therefore, these proteins are possibly just as necessary when measured in any defined region undergoing replication, and the same argument might apply to MRN proteins, which have so many different tasks. Other factors besides the genome size could contribute to this increased occurrence of corrupted forks, such as DNA sequence complexity, higher metabolic requirement, oxidative status or chromatin organization in higher organisms. Therefore, the replication machinery might rely more heavily on HR proteins to fix replication errors.

If HR is the main pathway to repair DSBs arising at replication forks, it should be noted that the collapse of a replication fork can generate a one-ended DSB (Fig 4), which is not the classical HR substrate. DSBs with only one free end are thought to be repaired by a sub-pathway of HR called break-induced replication (BIR; Poser et al., 2008; McEachern & Haber, 2006). The first step of this sub-pathway is similar to HR—the steps of which have been recently clarified
Sidebar A | In need of answers

(i) What consensus defines the metazoan origins of replication?
(ii) Are there Sld2 and Sld3 orthologues in higher eukaryotes? Do they represent the minimal set of CDK targets necessary for origin firing?
(iii) How many proteins that are important for cell cycle processes are present only in higher vertebrates?
(iv) Which DNA replication mechanisms have diverged and why?
(v) Is HR required for DNA replication in metazoans? If yes, at which stage?
(vi) What is the impact of BIR and MMBIR on mammalian DNA replication?

(Deem & Symington, 2009)—in that the 5’ end of the broken arm is resected in a highly regulated fashion by a set of nucleases. In HR, this involves the sequential action of the Mre11 and DNA2 nucleases coupled to the Sgs1 helicase, which act redundantly with Exo1 to produce a 3’ ssDNA filament (Deem & Symington, 2009) that is used to prime DNA synthesis on a new template (Fig 4). Once formed, this strand invades DNA templates in repeated attempts to find a suitable region of homology downstream or upstream from the point of fork collapse (Llorente et al., 2008; Smith et al., 2007). BIR could participate in replication fork recovery in yeast and, as such, it has been suggested to be the underlying mechanism of some chromosomal structural changes (Deem et al., 2008; Payen et al., 2008; Schmidt et al., 2006). However, the extent of BIR involvement in replication fork recovery in higher eukaryotes is unknown. It is tempting to speculate that BIR is a more important pathway to restart collapsed forks in higher eukaryotes, as it would be favoured by the presence of highly repetitive sequences that would facilitate homology-driven invasion. However, this remains to be established.

In principle, BIR is an accurate process that depends on recombination proteins and requires extensive homology for strand invasion. Nevertheless, it can lead to loss of heterozygosity and chromosomal rearrangements if the invading strand is paired with homologous allelic and non-allelic sequences (Deem et al., 2008; Payen et al., 2008; Smith et al., 2007). Indeed, BIR-based mechanisms can explain the complexity of the chromosomal structural changes that occur in cancer cells (Hastings et al., 2009; Lydeard et al., 2007; Smith et al., 2007). This is particularly relevant for a BIR-related pathway, microhomology-mediated BIR (MMBIR; Fig 4), that has been recently elucidated; this seems to be involved in the repair of one-ended DSBs that pair with stretches of non-related ssDNA molecules, which share microhomology with the invading 3’ ssDNA. MMBIR probably accounts for only a small fraction of DSB repair in yeast, whereas in mammalian cells it seems to be more efficient (Bentley et al., 2004). Genome-wide DNA sequencing studies of different cancer cell lines and primary tumours indicate that many rearrangements might derive from BIR and MMBIR-mediated events (Pleasance et al., 2009; Stephens et al., 2009).

Conclusions

The molecules and mechanisms that ensure a faithful DNA replication have been highly conserved throughout evolution. However, there are important differences between simple and complex organisms, just a few of which we have highlighted here. Among the differences that we have not considered there is an important class of genes, known as Fanconi anaemia (FA) proteins, that are involved in HR and DNA replication control in mammalian cells; except for a few members, FA proteins do not have clear homologues in unicellular eukaryotes.

A clear conclusion from this type of analysis is that the extrapolation from one organism to another cannot be considered universally reliable, although it has been extremely useful in studying DNA replication and its regulatory mechanisms. The challenge for the future is to understand the differences by taking advantage of more sophisticated approaches and innovative methods such as direct visualization of replication intermediates with advanced microscopy-based techniques. In addition, the further development of existing model systems that are capable of recapitulating DNA replication and repair will be useful for these studies (Sidebar A).

ACKNOWLEDGEMENTS

We thank members of Clare Hall Laboratories and of the Genome Stability Unit for their comments. V. Costanzo is supported by Cancer Research UK, the Lister Institute of Preventive Medicine, the European Research Council start-up grant (206281) and the EMBO Young Investigator Programme.

REFERENCES

Blow JJ, Gillespie PJ, Francis D, Jackson DA (2001) Replication origins in Xenopus egg extract are 5–15 kilobases apart and are activated in clusters that fire at different times. J Cell Biol 152: 15–25


Maiorano D, Cuvier O, Danis E, Mechali M (2005) MCM8 is an MCM2–7–related protein that functions as a DNA helicase during replication elongation and not initiation. Cell 120: 315–328


Alessia Errico  Vincenzo Costanzo