The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery

Article (Accepted Version)

O'Driscoll, Mark (2017) The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery. *Journal of Pathology*, 241 (2). pp. 192-207. ISSN 0022-3417

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/65956/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery.

Mark O'Driscoll

Human DNA Damage Response Disorders Group,
Genome Damage & Stability Centre,
School of Life Sciences,
University of Sussex,
Brighton, BN1 9RQ,
United Kingdom.

Phone: 0044 (0) 1273 877 515

Fax: 0044 (0) 1273 678 121

Email: m.o-driscoll@sussex.ac.uk

Conflict of interest statement: No conflicts declared.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/path.4828

Abstract.

Accurate and efficient replication of the human genome occurs in the context of an array of constitutional barriers including regional topological constraints imposed by chromatin architecture and processes such as transcription, catenation of the helical polymer and spontaneously generated DNA lesions including base modifications and strand breaks. DNA replication is fundamentally important for tissue development and homeostasis; differentiation programmes are intimately linked with stem cell division. Unsurprisingly, impairments of the DNA replication machinery can have catastrophic consequences for genome stability and cell division. Functional impacts upon DNA replication and genome stability have long been known to play roles in malignant transformation through a variety of complex mechanisms, and significant further insight has been gained from studying model organisms in this context. Congenital hypomorphic defects in components of the DNA replication machinery have been and continue to be identified in humans. These disorders present with a wide range of clinical features. Indeed, in some instances, different mutations within the same gene underlie different clinical presentations. Understanding the origin and molecular basis of these features opens a window onto the range of developmental impacts of sub-optimal DNA replication and genome instability in humans. Here, I will briefly overview the basic steps involved in DNA replication and the key concepts that have emerged from this area of research before switching emphasis onto the pathological consequences of defects within the DNA replication network; the human disorders.

Keywords: Pre-replication complex, origins, replication stress, R-loops, growth retardation, genomic instability, mutagenesis, cancer.

Introduction.

The potential mutagenic and cytotoxic consequences of the plethora of different forms of DNA damage (e.g. chemical modifications, base loss, cross-links, strand breaks) has driven the consequent evolution of a plethora of DNA Damage Response (DDR) and DNA repair (DNA-R) pathways to mitigate against these threats. Collectively, these genome stability pathways are fundamentally important for accurate cell division and organismal development. Work over several decades using a range of model organisms from bacteria, nematode and yeasts to fruit flies, zebra fish and mice, has greatly informed our knowledge of the phenotypic consequences of defects in these pathways. There now exists several comprehensive online databases cataloguing these phenotypes (**Table 1**). Furthermore, congenital defects in various components of these pathways underlie a growing list of human syndromes exhibiting a wide range of clinical presentation (<http://www.ncbi.nlm.nih.gov/omim/>). Over the years, recurrent clinical features have emerged associated with congenital impairments of the networks governing genome stability (**Figure 1**).

This review will focus on describing the range of human disorders caused by defects in components of the DNA replication machinery, many of which have only recently received molecular-genetic diagnosis. Nonetheless, much work remains in elucidating the basis underlying genotype-phenotype associations. The molecular events involved in, and the problems encountered during, DNA replication that influence genome stability have been comprehensively reviewed and considered elsewhere [1-6]. Many of these reviews contain accessible supporting figures to aid the reader negotiate these complex topological processes. Before discussing some of these disorders I will briefly overview the basic mechanisms involved DNA replication and highlight some routes whereby defects herein contribute the genome instability.

2. The basic mechanics of DNA replication.

Each 'gap' phase (G1 and G2) of the cell cycle is separated by a 'synthesis' phase (S phase), where duplication of DNA must occur efficiently and accurately, coincident with histone-mediated packaging. This is followed by condensation into chromosomes (G2 phase), leading to equal segregation of the genetic material into the two daughter

cells during mitosis (M phase). Each phase of the cell cycle is tightly regulated to ensure an ordered and unidirectional progression from G1-S-G2-M.

Preparation for DNA replication actually occurs in G1 phase, when the replicative minichromosome maintenance (MCM) complex, composed of MCM2-7, localises to chromatin at sites referred to as 'replication origins' [7-10]. Two MCM2 complexes load onto each origin, and together with additional factors operate as replicative helicases, unwinding the double helix in opposing directions to allow duplication of the template following replication initiation [11] **(Figure 2A)**. Replicative origins become 'licensed' to replicate upon binding of the origin recognition complex (ORC), composed of ORC1-6, and its regulatory elements CDC6 (cell division cycle 6) and CDT1 (chromatin licensing and DNA replication factor 1) [12]. Together, these complexes constitute the 'pre-replicative complex' (pre-RC). Following pre-RC formation and at the G1-S transition, both CDC6 and CDT1 are ejected from the pre-RC and the origin is activated or 'fired' to replicate. Origin firing is initiated via two kinases; cyclin-dependent kinase (CDK) and Dbf4-dependent kinase (DDK), that activate the MCM2-7 complexes [5]. Replicative helicase activity involves recruitment of CDC45 and additional factors including the heterotetrameric GINS complex, MCM10, AND1 and the Clasp-Timeless-Tipin complex, amongst other proteins [2]. Within the bidirectional replisomes the Clasp-Timeless-Tipin complex coordinates DNA unwinding with DNA synthesis by the various DNA polymerases (Pol α , Pol δ and Pol ϵ) aided by other proteins including PCNA, TOPBP1, RECQL4 and Treslin [13].

Regulating origin firing.

During the G1 phase, the loading of MCM2-7 onto chromatin to form pre-RCs occurs approximately 3- to 10-fold in excess of subsequent licensing events [14] **(Figure 2B)**. Origins do not all fire simultaneously, and these additional or 'dormant' origins allow maximal flexibility to ensure effective replication of the genome. For example, if a replisome were to stall at a region of damaged DNA such as a strand break, another pre-RC close-by could then be rapidly licensed and fired to rescue the stalled replisome [15,16] **(Figure 2C)**.

Crucially, replication from an origin must occur only once during S phase. Re-replication from a fired origin has catastrophic consequences for genome stability [17] (**Figure 2B**). To ensure that replication occurs only once from an origin, licensing and firing are actively and temporally separated. High CDK activity triggers origin firing. Simultaneously, it also inhibits pre-RC assembly via multiple mechanisms. For example, CDK-mediated phosphorylation of CDT1 and ORC1 target these for ubiquitin-mediated degradation by the proteasome [18-20]. Indeed CDT1 is an important focal point for several regulatory mechanisms that collectively ensure origins are only fired once per cell cycle [21]. Specifically, CDT1 is subjected to replication coupled (via PCNA) ubiquitination and degradation during the S phase [22-24]. Additionally, Geminin, an inhibitor of CDT1, is expressed from S phase to M phase. It binds CDT1 and prevents its localisation to chromatin during DNA replication. Geminin is itself then targeted for degradation at the end of M phase through ubiquitination by the APC/cyclosome complex, thereby liberating CDT1 to licence pre-RCs in the subsequent G1 phase [25].

‘Replication stress’: a significant problem encountered during DNA replication.

‘Replication stress’ (RS), is a generic/collective term to describe aberrant events that can occur during DNA replication, such as replisome collision with DNA damage. This can cause replication fork stalling, coupled with helicase disassociation, generating stretches of single-stranded DNA (ssDNA) and/or replication fork collapse, ultimately generating ssDNA and double strand breaks (DSBs) [26] (**Figure 3A**). The maintenance of genome stability is fundamental to cell and organismal homeostasis [27]. The RS checkpoint, a signal transduction response involving activation of the protein kinases ATR and CHK1, is essential to ensure stabilisation of stalled replication forks and to prevent the cell’s entry into M phase with damaged and/or under-replicated DNA [28,29]. RS is a characteristic feature of most cancers; even detectable in precancerous lesions, suggesting it plays a direct role in transformation [30]. Oncogene activation is an important and well-established inducer of RS [31]. Importantly, supra-elevated levels of RS can have a detrimental impact on cell division. Manipulating RS in tumours, as a targeted therapeutic intervention, represents a very active area of research presently [6,32]. Two principal mechanisms based upon origin usage have been

proposed to explain the underlying basis of RS; insufficient origin usage and excessive origin usage and re-usage.

Insufficient origin usage.

Overexpression of the cyclin E oncogene can impair MCM2-7 chromatin binding in G1 resulting in a paucity of pre-RCs to allow effective completion of S phase [33,34]. Under-replicated regions of DNA can result in chromosome breakage in the subsequent M phase. This is frequently seen at common fragile sites (CFSs); specific genomic regions prone to breakage [35]. Lower densities of origin firing and impaired replisome progression have been documented in these regions [36]. CFSs are typically broken in tumours and several CFSs incorporate tumour suppressor genes. For example, *FHIT* (*fragile histidine triad*) within the FRA3B CFS on chromosome 3p14.2, is frequently disrupted and/or rearranged in solid tumours of stomach, colon and lung [37]. The concept of dormant origin firing has emerged as a counterweight to limiting availability of functional pre-RCs, allowing recovery during RS [16] **(Figure 2C)**. Nonetheless, when the pre-RC machinery itself is limiting, as has been demonstrated by hypomorphic alleles of *Mcm2* and *Mcm4* in mice, significant RS-mediated genome instability is observed along with markedly increased rates of tumorigenesis [38-41].

At a fundamental level, replication from a limited number of fired origins would likely result in replication forks having to travel greater distances, thereby increasing the risk of fork stalling. The associated reduced availability of nearby dormant origins to rescue these stalled forks would further increase the risk of fork collapse [36]. Replication fork stalling and collapse can stimulate template switching events and break-induced replication outcomes, situations that can result in the generation of genomic amplifications [42] **(Figure 3B)**. Persistent DSBs can also act as a platform for chromosomal translocations, breakage-fusion-bridge events and micronuclei formation [43].

Excessive origin usage and re-usage.

Excessive origin firing, as observed following over-expression of oncogenes including *RAS* and *MYC* also constitutes a threat to genome stability [44]. This can result in acute depletion of the cellular pools of dNTPs, ultimately inducing replication forking stalling and collapse [45]. Another related consequence of excessive origin firing is depletion of important genome stability factors. For example, the RPA (Replication Protein A) complex is a heterotrimeric (RPA1-3) ssDNA binding component required for most DNA strand transactions during replication and repair [46]. RPA coats newly generated ssDNA at on-going replication forks and at stalled forks, and is a fundamental mediator of the ATR-dependent DDR [28]. Depletion of the available pool of RPA due to increased levels of sequestration can increase the rate of ssDNA conversion into DSBs potentiating RS-mediated genomic instability [47].

As mentioned earlier, mammalian cells invest significant effort, especially involving the regulation of CDT1 function, to prevent the inappropriate re-licensing of fired origins, thus preventing re-replication [48] (**Figure 2B**). Significant re-replication would invoke similar depletion scenarios as described above (e.g. for dNTPs, RPA) and indeed cancer cells appear to have evolved unexpected routes to perpetuated re-replication-induced polyploidy [49]. Re-replication from an origin could also result in head-to-tail collisions of the newly fired fork with the original fork, resulting in all of the adverse issues inherent in fork stalling. Using an siRNA library-based screening approach, Vassilev and colleagues have recently characterised the landscape of genes involved in limiting genome duplication to once per cell cycle [50]. Along with factors controlling origin licensing inhibition, they described unanticipated roles here for components of the mitotic spindle assembly and checkpoint machinery, chromosome cohesion and segregation pathways and the cytokinesis network.

Replication colliding with transcription.

A type of collision event of intense recent focus is that of DNA replication fork collision with a transcription fork [1,51,52]. Replication and transcription fork collisions can generate significant torsional stress in the DNA molecule which can lead to replication fork reversal (**Figure 3A**). Regressed forks can constitute substrates for

inappropriate processing by structure-specific nucleases such as SLX4, MUS81, EXO1, CtIP and GEN1, generating DSBs and strand exchanges [1,53].

R-loops are DNA-RNA hybrid molecules containing displaced ssDNA which can form upon collision of a replisome with a transcription fork, when the nascent RNA from the transcription fork hybridizes to the complementary DNA strand [53,54] **(Figure 3C)**. Early replicating fragile sites (ERFSs), a class of fragile site found in highly transcribed, repetitive and CpG-rich genomic loci, are thought to involve R-loop formation. These structures are usually rapidly resolved through a combination of nuclease-mediated digestion (e.g. RNase H), RNA helicase action (e.g. senataxin) amongst other processes including those dependent upon RECQL5 [53-56] **(Fig 3C)**. If R-loops persist, the ssDNA intermediate is susceptible to breakage and/or inappropriate processing. Sollier and colleagues showed that when certain RNA processing factors are limiting (e.g. senataxin, RNA/DNA helicase Aquarius) or following topoisomerase I inhibition, R-loops can be processed into DSBs by nucleotide excision repair (NER) nucleases (XPF, XPG) [57].

Human disorders of the replication licensing machinery.

Prior to the genetic elucidation of these disorders, studies on various mouse models provided important insight into the mammalian phenotypes of hypomorphic defects in the licensing machinery. A prototypical example of this is *Cdc7* (cell division cycle 7) [58]. CDC7 kinase activity is essential for initiation of DNA replication and G1-S entry, interacting with pre-RC components (MCMs and ORCs) and DDK. *Cdc7*^{-/-} mice are early embryonic lethal and derived embryonic stem (ES) cells fail to cycle [59]. Whilst the ES cell phenotype was rescued via transgene expression of a single copy of *Cdc7*, viable progeny (born at sub-Mendelian frequency) generated from these cells exhibited multiple striking phenotypes. *Cdc7*^{-/-}tg mice displayed marked growth delay, testicular and ovarian atrophy (other organs were proportionately small), tail flexion anomalies and neonatal lethality [58]. The tail flexion phenotype was reminiscent of the *curl-tail* (*ct*) mouse model of neural tube defects [60]. *Cdc7*^{-/-}tg-derived mouse embryonic fibroblasts (MEFs) exhibited spontaneously reduced S phase capacity and reduced DNA synthesis following RS induction.

Disrupting the finely tuned balance between the temporal necessity for optimal stem and progenitor cell division and differentiation programmes are known to profoundly affect development, particularly body growth and brain development [61-63]. At least in mice, rapidly proliferating pluripotent embryonic stem cells exhibit constitutive RS and appear heavily reliant on replication-coupled pathways to maintain genome stability [64]. Severe congenital microcephaly with/without marked growth restriction is associated with defects in multiple components of the centrosome, microtubule spindle network and kinetochore [65,66]. It is thought that these defects negatively impinge upon normal mitotic cell division and spindle pole orientation, the latter being particularly important for normal brain development. Considering the developmental consequences of impaired mitotic cell division in this context, it seems logical to assume that human hypomorphic defects that impair DNA replicative capacity and/or maintenance of stem cell niches could also present with primordial dwarfism and/or neurodevelopmental deficits.

Seckel syndrome.

Seckel Syndrome (SS) represents a striking example of a microcephalic primordial dwarfism (MPD) [67,68]. Causative defects have been described in ATR, its interacting protein ATRIP, in certain nucleases and various centrosome proteins [69-74]. A humanised mouse model of ATR-defective SS showed that reduced *Atr* expression was associated with markedly elevated levels of RS and apoptosis throughout developing embryos, consistent with the model of intrauterine programming where adverse developmental impacts *in utero* underlie severe and non-progressive postnatal phenotypes [75]. Other mouse models of impaired *Atr* function further underscored the fundamental role for ATR in maintaining stem cells pools by limiting RS [76-79]. Impaired replication fork progression, stability and dormant origin firing have also been demonstrated in ATR-SS patient-derived cell lines [80]. These findings further reinforce the link between genome instability, DNA replication and limiting tissue regenerative capacity with systemic outcomes such as premature ageing, as previously highlighted by the laminopathy class of progeroid syndromes (e.g. Hutchinson-Gilford Progeria) [81]. Indeed progeria is a common feature of disorders associated with abnormalities of the genome stability, replication and transcription network [82]. Inhibition of ATR kinase activity using small molecule inhibitors has been shown to induce RS-mediated DNA

breakage and aberrant fork processing (e.g. excessive resection-induced ssDNA generation) through the inappropriate activities of specific nucleases including SLX4 and CtIP [83].

Polygenic defects in a growing number of genes encoding some components of the pre-RC licensing machinery have now been described. Pathogenic variant identification has been greatly facilitated by the widespread application of exome sequencing strategies. It is anticipated that novel defects in pre-RC components not yet attributable to a human disorder will be identified in the near future. Importantly, considering the essential nature of the DNA replication machinery, any anticipated pathogenic defects here are expected to be hypomorphic/neomorphic [58]. Therefore, molecular and cellular characterisation of these variants remains a crucial contributor to our understanding the landscape and complex etiopathology underlying these disorders.

Meier-Gorlin syndrome: origin licensing and S phase entry.

Multiple defects in the pre-RC machinery underlie Meier-Gorlin Syndrome (MGS), also known as ear, patella, short stature syndrome and/or microtia, absent patella, micrognathia syndrome; highlighting the core clinical phenotypes [84-87]. MGS is clinically variable and often associated with other features including urogenital abnormalities, congenital pulmonary emphysema and mammary hypoplasia. Longitudinal studies on MGS patients with defined molecular diagnoses have now been reported [88,89]. Recessive pathogenic variants have been identified in *ORC1*, *ORC4*, *ORC6*, *CDT1* and *CDC6* [90-93]. Functional analysis of cell lines derived from these patients identified impairments in origin licensing, S phase entry and progression, reduced dormant origin firing and impaired recovery following RS induction [91,93,94]. Reducing *orc1* levels in zebrafish embryos also recapitulated the growth restriction characteristic of MGS, and a *Drosophila* model of an MGS *Orc6* mutation demonstrated impaired replicative capacity at the developing larval stage [91,95]. MGS *ORC1* mutations localising to the BAH (bromo adjacent homology) domain of ORC1 likely impair its ability to bind specific histone marks that are normally enriched at origins

[96-98]. Modelling approaches have also shed mechanistic insight into how an MGS *ORC6* mutation impairs ORC formation [99].

Mutations in *CDC45L* have since been found to underlie MGS, MGS associated with craniosynostosis and in individuals with craniosynostosis alone [100]. *CDC45L* encodes CDC45, an essential protein within the replisome linking the MCM-helicase and GINS complex machinery to the replicative polymerases [101,102]. Although functional analysis of DNA replicative capacity, origin firing and stability or G1-S progression kinetics of the CDC45 patient-derived cells wasn't reported, it is expected that some aspects of DNA replication are compromised here [100].

Interestingly, *de novo* heterozygous and dominant mutations in *GMNN*, encoding Geminin, the inhibitor of CDT1, have also been described, representing a novel inheritance pattern underlying some MGS cases [103]. These findings have important clinical implications especially in MGS individuals not presenting with recessively inherited variants in components of the replicative machinery. Geminin is an important inhibitor of CDT1 and Geminin levels are tightly regulated during the cell cycle by ubiquitin-mediated proteosomal degradation. The pathogenic MGS Geminin variants were found to disrupt its destruction box motif thereby impairing APC/C mediated ubiquitination and subsequent Geminin degradation; prolonging Geminin stability and its consequent ability to inhibit CDT1 [103].

MCM2-7 complex disorders.

Mouse models have also lead the way in providing insight into the phenotypic presentation of hypomorphic defects in various MCM proteins of the pre-RC (i.e. MCM2-7); the majority of whom exhibit genomic instability, deficiencies in the proliferative cell compartments of most but not all tissues, elevated cancer incidence (particularly lymphoma) and depending on the allele, marked growth restriction. Key examples include mouse models for hypomorphism of *Mcm2*, *Mcm3* and *Mcm4* (*Chaos 3* allele: chromosome aberrations occurring spontaneously 3) [38-40,104].

To date, several human examples of pathogenic variants in some MCM2-7 complex components have been reported. Gao and colleagues described a single missense variant in *MCM2* as causative of nonsyndromic progressive sensorineural

hearing loss [105]. The mechanism underlying this particular clinical presentation is unclear, although interestingly, sensorineural hearing loss has been observed in related contexts (see section Disorders of the core DNA replication machinery, below). For *MCM4*, the clinical presentation better reflects the paradigm established by defective pre-RC components in MGS. Differing hypomorphic variants of *MCM4* have been identified in multiple individuals from different Irish Traveller families [106,107]. Affected individuals exhibit growth retardation along with selective Natural Killer (NK) cell deficiency and adrenal insufficiency. These patients' cells were also characterised by elevated chromosomal breakage rates, consistent with studies on the consequences of limiting pre-RC availability. Cancers have not been reported in these families although close surveillance has been advocated [106]. The basis underlying selective NK loss here is unclear, although *Mcm4^{Chaos3/-}/Mcm3^{+/-}* mice were found to exhibit abnormal adrenal morphology [106].

In another Irish Traveller family, Casey and colleagues reported an *MCM4* variant within the splice acceptor site of intron 1 that also overlaps with the 5' upstream region of *PRKDC* [108]. Here, the cells from the affected individuals showed elevated chromosome breakage, and the patients exhibited selective NK cell and glucocorticoid deficiency. *PRKDC* encodes DNA-PKcs, an important component of the non-homologous DNA end-joining (NHEJ) DSB-R pathway. The functional contribution of this *PRKDC* variant to DSB-R has not been investigated but potentially represents an additional confounding contributor to the chromosomal instability phenotype observed here. Hypomorphic variants in *MCM3,5,6* and *MCM7* have not been reported to date.

MCM-disorders outside of the MCM2-7 complex.

MCM8 and MCM9 are MCM protein family members that were originally identified as DNA replication factors interacting with the pre-RC, but who have since been ascribed roles in the homologous recombination repair (HRR) DSB-R pathway; most recently in regulating MRN (MRE11-RAD50-NBN)-mediated resection of DSBs [109-113]. Unexpectedly, *Mcm8^{-/-}* mice and *Mcm9^{-/-}* mice were found to be viable and with gonadal restricted aberrant phenotypes consistent with arrested meiosis [114]. HRR between homologous chromosomes during pachytene of meiosis I is a

programmed obligate step of meiosis. MEFs from these animals exhibited chromosomal instability, defective HRR and elevated replication fork collapse following induction of RS [114].

Using a combination of SNP arrays, comparative genomic hybridisation and exome sequencing strategies, pathogenic variants in *MCM8* and separately in *MCM9* were identified in multiple individuals exhibiting a syndrome of primary ovarian failure (POF) and growth restriction [115-117]. Patient-derived fibroblasts were found to exhibit elevated levels of chromosomal breakage when these were induced by mitomycin C (a DNA cross-linking agent). A homozygous truncating mutation in *MCM9* has since been identified in a family exhibiting hereditary mixed polyposis coli, colorectal carcinoma and POF; further expanding the clinical presentation of impaired MCM9 function [118].

Disorders of the core DNA replication machinery.

Proliferating cell nuclear antigen (PCNA)

PCNA functions as a homotrimeric sliding clamp that encircles the DNA double helix functioning as an interacting platform for many proteins. PCNA plays a central role in recruiting and retaining the DNA polymerases at the sites of DNA synthesis during replication, but also during the repair synthesis stage of multiple DNA-R pathways such as nucleotide excision repair (NER), DNA mismatch repair (MMR) and HRR [119]. Baple and colleagues identified a recurrent hypomorphic variant in *PCNA* in an extended Ohio Amish kindred. The affected individuals exhibited a fascinating phenotypic range reflective of a combination of several established genome instability conditions including Ataxia telangiectasia, Xeroderma pigmentosum and Cockayne syndrome [120]. The widely ranging clinical presentation, including short stature, photosensitivity, telangiectasia, sensorineural hearing loss, neurodegeneration, cerebellar atrophy and ataxia, likely reflects the breadth of functions of PCNA in multiple genome stability pathways. Analysis using a SILAC (stable isotope labelling with amino acids in cell culture)-mass spectroscopy approach showed that the pathogenic variant affected the constitution of the PCNA interactome [120]. Interestingly, patient-derived cells *did not*

exhibit impaired DNA replication, but they did exhibit elevated sensitivity to killing by, and impaired NER of, UV-photoproducts.

DNA polymerases Pol δ and Pol ϵ

The DNA template is copied in a semi-conservative fashion with remarkable fidelity through the coordinate action of three DNA polymerase multi-subunit holoenzyme complexes; Pol α , Pol δ and Pol ϵ [121]. Spontaneous error rates are estimated at around one incorrect base per 10^9 - 10^{10} nucleotides replicated [122]. This extremely high fidelity derives from the combined actions of the highly selective nucleotide incorporation and exonuclease proofreading activities of the replicative helicases Pol δ and Pol ϵ , coupled with post-replicative MMR. Germline defects in MMR components underlie Lynch syndrome/Hereditary Non-Polyposis Colorectal Carcinoma (HNPCC) [123,124]. Microsatellites are short repetitive sequences distributed throughout the genome, and are sites where replication is prone to DNA polymerase slippage [125]. MMR loss results in elevated spontaneous mutation frequency and a characteristic “microsatellite instability” (MSI) phenotype due to frameshifts within microsatellites [126-129].

Recently, germline mutations in *POLD1* and *POLE* were found to predispose carriers to colorectal carcinoma (CRC) [130,131]. *POLD1* encodes the catalytic subunit of Pol δ , whilst *POLE* encodes the catalytic subunit of the Pol ϵ holoenzyme. Interestingly, these germline mutations clustered within the region encoding the exonucleolytic proofreading domain, likely underlying the “ultramutation” phenotype characteristic of these malignancies [130,131]. This form of familial CRC has been termed “polymerase proofreading associated polyposis” (PPAP), and germline and somatic mutations in *POLD1* and *POLE* have since been found in other cancer types [132,133].

Weedon and colleagues reported an additional fascinating insight into the complex phenotypic presentation of germline *POLD1* mutation [134]. They found a recurrent heterozygous in-frame single codon deletion in *POLD1* (p.S605del) in several individuals exhibiting Mandibular hypoplasia, Deafness and Progeroid features (MDP)

syndrome, a multisystem disorder also characterised by lipodystrophy [134,135]. Functional analysis using engineered recombinant Pol δ showed that in contrast to the PPAP-associated variants, this pathogenic variant did not impact upon exonuclease proofreading activity of Pol δ . Rather, it impaired its catalytic activity. Subsequently, other individuals carrying the same deletion, along with an MDP case with a novel missense variant (p.R507C) were identified [136,137]. The p.R507C variant is worthy of functional interrogation as it localises to the exonuclease domain in contrast to p.S605del which localises to the POLD1 catalytic domain [137]. Nonetheless, caution is warranted in assuming that this class of *POLD1* variants universally presents exclusively as MDP. Lessel and colleagues found the p.S605del in several individuals originally classified as atypical Werner syndrome (WRN); that is, clinically presenting as the progeroid syndrome WRN, but lacking detectable mutations in *RECQL2*, the WRN causative gene [138]. These individuals also lacked the sensorineural hearing loss and mandibular hypoplasia characteristic of MDP. Interestingly, postnatal short stature appeared a common phenotype in this cohort.

The emerging picture appears to be that pathogenic variants of some component of the Pol δ and Pol ϵ DNA replicative polymerase holoenzymes can present as a wide range of clinical entities from familial cancer syndromes such as PPAP to developmental and metabolic conditions as typified by MDP, with and without growth delay [139].

Disorders of the DNA replication machinery involved in fork stabilisation and processing.

Bloom syndrome (BLM) and Werner syndrome (WRN) are well-known and long established multisystem disorders caused by defects in different members of the RECQ helicase family; BLM (*RECQL3*) and WRN (*RECQL2*). Both helicases play multiple accessory role in the resolution of topologically complex DNA structures during DNA-R and recovery from stalled replication forks [140-142]. BLM is characterised by growth retardation and microcephaly along with marked predisposition for most cancer types [143,144]. BLM patient cell lines present with a characteristic form of genomic instability: elevated spontaneous sister chromatid exchanges due to BLM's role in resolving certain HRR intermediates [145]. Additional clinical features include a

characteristic pattern of facial telangiectasia and diabetes mellitus. WRN individuals develop a prototypical segmental progeroid syndrome associated with premature greying, growth delay, diabetes, bilateral cataract, osteoporosis and cancer in multiple tissues [146]. Therefore, both conditions provide insight into the phenotypic consequences of elevated rates of replication fork stalling and collapse, with recurrent features include growth delay, metabolic dysregulation, genomic instability and cancer **(Figure 1)**. These core clinical features are typically common amongst defects in pathways playing fundamental roles in the repair and recovery of stalled replicative intermediates, as also exemplified by Fanconi Anaemia [147,148]. Indeed, the multi-system theme of the clinical features of these disorders is similarly reflected in the sole instance of hypomorphic mutation in *LIG1*, encoding DNA Ligase I. This is the main ligase activity required for DNA replication, joining Okazaki fragments during lagging strand synthesis [149,150]. Here, the affected female presented with a syndrome involving growth retardation, sunlight sensitivity, immunodeficiency, telangiectasia and absence of secondary sexual characteristics.

Defects in *RECQL4* encoding another member of the RECQ family are associated with a complex although overlapping range of human pathologies characterised by radial and growth defects along with predisposition to osteosarcoma in particular. *RECQL4* disorders include Rothmund-Thomson syndrome (RTS), Baller-Gerold syndrome (BGS) and RAPADILINO syndrome (*radial hypoplasia, patella hypoplasia and cleft or arched palate, diarrhoea and dislocated joints, little size and limb malformation, slender nose and normal intelligence*) [151,152]. *RECQL4* has long been known to be required for the initiation of DNA replication [153,154]. It is now thought that *RECQL4* plays essential enzymatic (helicase) and non-enzymatic roles during replication initiation along with MCM2-7 helicase in the formation of the CMG complex (CDC45-MCM-GINS) [155]. *RECQL5* represents yet another member of this family that plays multiple roles in maintaining genome stability in the context of DNA replication and its interface with transcription, although an associated human disorder remains to be identified [55,56]. The *Recql5* deletion mouse model may provide clues as to how this condition may present [156].

As detailed earlier, a replication fork stalling at a lesion or following nucleotide depletion generates a structure that must first be stabilised before attempts are made to

reinitiate synthesis. The complex enzymology of lesion by-pass, along with the clinical presentation of congenital defects in that machinery, represent topics that could easily occupy an entire journal issue and have recently been reviewed elsewhere [157,158]. Our knowledge of the interplay between and convergence of distinct pathways in processing stalled DNA replication intermediates continues to grow, as does our knowledge of how defects here present in humans. Sophisticated experimental approaches to capture the proteome at stalled forks, as pioneered by Cortez and colleagues, have proved to be invaluable aids for untangling this complexity [159-161]. An illustrative example of the depth of this complexity is BRCA2, which plays multifactorial roles at the stalled replication fork [162-164]. Interestingly, congenital defects in *BRCA2* have been associated with overlapping human pathologies; Fanconi anaemia and Seckel syndrome/MPD [73,165,166].

I will overview a short selection of disorders whose underlying defects have now been attributed to specific proteins that play different roles in the recovery of stalled forks. Although exhibiting a degree of clinical overlap, there are some clear differences between these disorders. The responsible proteins underlying these conditions usually act at the stalled fork preventing its aberrant processing and shunting the structure towards a favourable repair outcome (**Figure 3A**). For one of these conditions, the role in stalled replication fork recovery was identified *after* the molecular genetic association with a human syndrome. This represents another example of how functional characterisation of the protein helps inform our understanding the pathophysiology of the associated syndrome.

SMARCA1

Schimke immune-osseous dysplasia (SIOD) presents with the core clinical features of growth failure, T-cell immunodeficiency and spondyloepiphyseal dysplasia along with renal dysfunction [167-169]. In 2002, Boerkoel and colleagues identified multiple mutations in *SMARCA1* (*SWI/SNF2-related matrix associated, actin-dependent regulator of chromatin subfamily a-like 1*), as being causative of SIOD [170]. The SNF2-related family of proteins constitute an important group of chromatin re-modellers active during DNA replication, DNA-R and in controlling gene expression. Mutations in

genes encoding other SNF-2-related proteins have been identified in several syndromal contexts including *ERCC6* in Cockayne syndrome and *ATR* in α -thalassemia-mental retardation syndrome. Because of the pleiotropic nature of chromatin re-modelling it seems logical to assume that defects in components of this machinery could underlie a multi-system disorder. Nonetheless, a precise function of SMARCAL1 (also referred to as HARP: Hep-A related protein) in chromatin re-modelling was inferred based on homology to other SNF2-related proteins. Therefore, its physiological role(s) and how its impairment results in SIOD remained to be dissected.

In 2009 several groups showed that SMARCAL1 played an important function in maintaining genome stability, specifically through interaction with RPA at the replisome by limiting the emergence of RPA-coated ssDNA bubbles at stalled forks [171-176]. A SMARCAL1 interactome has since been described and has provided additional evidence of its complex activity at the replication fork [177]. Various models have been proposed to explain SMARCAL1's role at the replication fork; such as acting as an annealing/reverse helicase or translocase capable of rewinding ssDNA, or catalysing replication fork reversal into a chicken-foot structure [178,179] (**Figure 3A**). SMARCAL1 activity is tightly regulated at the stalled fork through direct phosphorylation by ATR, a key master regulator of replication fork stability [83,180]. Interestingly, SIOD patient-derived cells were found to exhibit spontaneously elevated levels of RS and impaired cell cycle progression, suggesting that these phenotypes may contribute to the complex pathophysiology of the disorder.

DNA2 & RBBP8 (CtIP)

FEN1 and DNA2 are nucleases active in removing 5'-flap structures generated as intermediates during long-patch base excision repair (BER) removal of modified (alkylation, oxidation) bases [181]. DNA 5'-ssDNA overhangs/flaps are also generated via helicase action and by strand displacement synthesis during DNA replication. These structures are rapidly processed through the action of several nucleases such as FEN1, which cleaves 5'-flaps of only a few nucleotides in length. Longer flaps (>25bp) become RPA coated and resistant to FEN1 action and are cleaved by DNA2. During Okazaki fragment maturation, 5'-flaps are generated by strand displacement

synthesis. Depending upon their length, they are subject to coordinate cleavage by FEN1 and DNA2 [182]. These nucleases also play important roles in processing stalled replication forks ensuring their stability, for example by limiting replication fork regression [183-185] (**Figure 3A**). Similarly, both nucleases function to limit 5'-flaps generated in the mitochondria through DNA polymerase γ -mediated replication of the mitochondrial genome. BER is active within mitochondria as a repair mechanism for the mitochondrial genome to counter the impacts of locally generated reactive oxygen species from the electron transport chain.

Depletion of DNA2 in human cells impairs mitochondrial DNA replication and induces nuclear instability [186]. Pathogenic variants in *DNA2* have been described as causative of two distinct pathologies. Ronchi and colleagues identified several distinct dominant *DNA2* variants in patients' exhibiting late-onset Progressive External Ophthalmoplegia (PEO), a mitochondrial myopathy [187]. These variants localised to both the nuclease and ATPase domains of DNA2, and functional analysis of recombinant proteins showed significant deficits in nuclease activity on model substrates *in vitro*. In contrast, Shaheen and colleagues found a homozygous truncating *DNA2* mutation in two affected individuals from a Seckel syndrome/MPD family [73]. Analysis of these patients' fibroblasts showed elevated spontaneous DNA breakage. Interestingly, the homozygous Seckel syndrome/MPD *DNA2* variant localised to the C-terminal end of the ATPase domain. Collectively, it is clear that differing hypomorphic defects in *DNA2* underlie a spectrum of human pathologies, likely reflecting its roles in both nuclear and mitochondrial genome processing.

Another nuclease associated with congenital human disease of relevance here is that encoded by *RBBP8* (*Retinoblastoma binding protein 8*), also known as CtIP (CTBP interacting protein). CtIP has endonuclease activity favouring 5'-flap structures, but also associates with the DSB exonuclease activity of the MRN complex [188]. Differing hypomorphic defects in *RBBP8* have been implicated in the MPD clinical spectrum, ranging from Seckel syndrome to Jawad syndrome (i.e. MPD with intellectual disability and digital abnormalities) [73,74,189]. *RBBP8*-defective patient cells exhibit ATR-dependent DDR defects [190].

Copy number variation involving DNA replication components and genomic disease.

Locus-specific rates of copy number variation (CNV) occur at orders of magnitude greater than single nucleotide variation, therefore representing a significant contributor to human genetic disease [191,192]. Complex chromosomal rearrangements are a class of CNVs whose formation has been proposed to rely on different replicative mechanisms involving fork stalling, template switching, DNA breakage and even mechanisms involving recombination-restart in the absence of DSBs [193-197].

Some well-known genomic diseases are caused by CNVs that also incorporate genes encoding proteins that function in DNA replication. Often these disorders exhibit growth delay and aberrant brain development as part of their clinical spectra. Whilst it can be challenging to ascribe precise genotype-phenotype relationships when CNVs involve multiple contiguous genes, a growing number of examples exists showing that copy number changes in some of these DNA replication genes are associated with DNA replicative and cell cycling abnormalities. For example, DiGeorge/Velocardiofacial syndrome (DG/VCS) patient-derived cells exhibit asynchronous DNA replication of certain loci and elevated RS-induced genome instability [198-200]. The latter phenotype can be induced by siRNA-mediated reduction of CDC45 in human cells [200]. *CDC45L* is amongst the approximately 30 genes haploinsufficient in DG/VCS, but is the only one encoding a product with an established role in DNA replication [201]. Haploinsufficiency of *RPA1* encoding the largest subunit of the Replication Protein A complex is a feature of Miller-Dieker syndrome and has been shown to underlie ATR-dependent DDR and cell cycle checkpoint activation defects in this context [202]. Increased *RPA1* copy number has also been implicated in genomic instability associated with genomic disease; in this instance causing impaired HRR [203]. Haploinsufficiency of *RFC2* encoding a component of the Replication Factor C complex is a feature of Williams-Beuren syndrome, and has been implicated in genomic instability in this context [202]. Finally, the characteristic haploinsufficiency of both *SLBP* and *NELF-A* in Wolf-Hirschhorn syndrome has been shown to underlie marked DNA replication and histone deposition abnormalities in cells from these individuals [204].

Compromised DNA replication and beyond. Could there be additional mechanistic contributors to phenotype?

Effective execution and termination of DNA replication is of course critical for tissue development through the coordinated expansion of stem cell niches and activation of differentiation programmes. It would therefore be anticipated that pathogenic variants in components of the DNA replication machinery would adversely impact upon the finely tuned balance between these pathways. Replication timing is also crucial for regulated expression of monoallelically expressed genes including X-linked genes, and some autosomal genes such as those encoding odorant receptors, immunoglobins, T-cell and NK-cell receptors [205]. The latter is particularly intriguing when considering the specific presentation of hypomorphic defects in *MCM4* [106,107].

As we endeavour to fully understand the basis of the often complex clinical presentation of congenital defects in the DNA replication machinery, it is important to consider the wealth of preceding work from model organisms showing that many of the DNA replication factors play other roles in addition to the central business of duplicating the genome. Specific components of the DNA replication machinery play roles in epigenetic inheritance through heterochromatin formation, gene silencing and transcriptional regulation, also influencing ribosomal biogenesis and cilia-mediated signal transduction [206-209]. This strongly argues that there can be wider consequences of defects in the DNA replication machinery, in addition to limiting stem cell expansion and RS-mediated stem cell attrition, which should be considered as potential pathomechanistic contributors to disorders in this context. A convincing illustration of this is provided by the impaired cilia signalling observed in specific MGS patient cells which demonstrates a functional connection between impaired origin licensing and reduced G1-S entry with inefficient hedgehog-pathway signalling and chondro-induction [93]. It's worth noting that the chondro-induction defects demonstrated in this study upon knockdown of the pre-RC machinery were derived from non-cycling G₀ cells [210]. In this context it is interesting to consider the overlapping cranial features of Sensenbrenner syndrome/cranioectodermal dysplasia, caused by defects in intraflagellar transport within cilia, and recent findings expanding the clinical presentation of MGS [100,211,212]. Cilia signalling is essential for chondrogenic differentiation and endochondral bone formation [213-217]. Indeed,

chondrogenic-related abnormalities are typical of ciliopathies such as Bardet-Biedl and Ellis-van Creveld syndrome [217,218]. Furthermore, it is intriguing to find several components of the core DNA replication machinery identified as constituents of the wider CPLANE (ciliogenesis and planar polarity effector) protein network [219]. The physiological relevance of these associations merit deeper investigation. Indeed functional deficits in cilia-mediated signalling have now also been documented in the context of reduced ATR-signalling which likely play some contribution to the clinical presentation of *ATR/ATRIP*-mutated Seckel syndrome [210].

Summary.

The expanding series of congenital defects in the DNA replication machinery highlights the importance of these proteins in maintaining genome stability and in normal human development. The clinical picture is extremely complex, indicating that less cell division is probably not the only phenotypic outcome of these pathogenic variants. Model organisms continue to offer important tools for understanding the wider impacts of these defects upon development. The zebrafish model in particular has remained an important and efficient additional, and/or alternative, to modelling gene defects in mice, and has lead the way in uncovering novel insight into the consequences of these kinds of defects [210]. Reliable and economic alternatives to the mouse system are required, particularly considering the genetic issues directly relevant to RS and DNA replication raised by some inbred mouse strains [220]. As pathogenic defects in the DNA replication machinery are likely hypomorphic/neomorphic, there exists a growing need to introduce the specific gene variants into the appropriate stem cell systems and/or model organisms, to better enable functional characterisation. These approaches are now greatly facilitated by the advances in genetic engineering such CRISPR/Cas9.

The wealth of genetic information generated by exome sequencing approaches has transformed the clinical genetics landscape. Consolidating this information, for example through collective approaches such as the Matchmaker Exchange portal (<http://www.matchmakerexchange.org/>) should further aid and advance pathogenic variant identification [221]. Importantly, functional characterisation of these variants

through cell and molecular-based approaches, needs to continue playing a central role in proving and/or interpreting pathogenicity. This is particularly important with regards variants of unknown significance, dominant variants and in situations where a specific variant doesn't result in the expected pathway defect (e.g. section PCNA above). The latter instances can often direct attention towards formulating novel pathomechanistic insights. Based on current trajectories, novel defect identification should hopefully continue at this accelerated rate enabling an ever deeper understanding of the consequences of pathogenic variants in components of the DNA replication machinery.

Acknowledgements.

The Human DNA Damage Response Disorders Group is programme funded (C24110/A15394) by Cancer Research UK to M O'D.

Statement of author contributions.

M O'D wrote the paper and drafted the figures.

Reference list.

1. Gaillard H, Aguilera A. Transcription as a Threat to Genome Integrity. *Ann Rev of Biochem* 2016; **85**: 291-317.
2. Gaillard H, Garcia-Muse T, Aguilera A. Replication stress and cancer. *Nat Rev Cancer* 2015; **15**: 276-289.
3. Muñoz S, Méndez J. DNA replication stress: from molecular mechanisms to human disease. *Chromosoma* 2016: doi:10.1007/s00412-016-0573-x
4. Errico A, Costanzo V. Mechanisms of replication fork protection: a safeguard for genome stability. *Crit Rev Biochem Mol Biol* 2012; **47**: 222-235.
5. Hills Stephanie A, Diffley John FX. DNA Replication and Oncogene-Induced Replicative Stress. *Curr Biol* 2014; **24**: R435-R444.
6. Lecona E, Fernández-Capetillo O. Replication stress and cancer: It takes two to tango. *Exp Cell Res* 2014; **329**: 26-34.
7. Evrin C, Clarke P, Zech J, *et al.* A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc Natl Acad Scie U S A* 2009; **106**: 20240-20245.
8. Remus D, Beuron F, Tolun G, *et al.* Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. *Cell* 2009; **139**: 719-730.
9. Riera A, Tognetti S, Speck C. Helicase loading: How to build a MCM2-7 double-hexamer. *Semin Cell Dev Biol* 2014; **30**: 104-109.
10. Gambus A, Khoudoli GA, Jones RC, *et al.* MCM2-7 form double hexamers at licensed origins in xenopus egg extract. *J Biol Chem* 2011; **286**: 11855-11864.
11. Boos D, Frigola J, Diffley JFX. Activation of the replicative DNA helicase: breaking up is hard to do. *Curr Opin Cell Biol* 2012; **24**: 423-430.
12. Sclafani RA, Holzen TM. Cell Cycle Regulation of DNA Replication. *Annu Rev Genet* 2007; **41**: 237-280.
13. Masai H, Matsumoto S, You Z, *et al.* Eukaryotic Chromosome DNA Replication: Where, When, and How? *Annu Rev Biochem* 2010; **79**: 89-130.
14. Hyrien O, Marheineke K, Goldar A. Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *BioEssays* 2003; **25**: 116-125.

15. Blow JJ, Ge XQ. A model for DNA replication showing how dormant origins safeguard against replication fork failure. *EMBO Rep* 2009; **10**: 406-412.
16. Blow JJ, Ge XQ, Jackson DA. How dormant origins promote complete genome replication. *Trends Biochem Sci* 2011; **36**: 405-414.
17. Arias EE, Walter JC. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* 2007; **21**: 497 - 518.
18. Mendez J, Zou-Yang XH, Kim SY, *et al.* Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Mol Cell* 2002; **9**: 481 - 491.
19. Sugimoto N, Tatsumi Y, Tsurumi T, *et al.* Cdt1 phosphorylation by cyclin A-dependent kinases negatively regulates its function without affecting geminin binding. *J Biol Chem* 2004; **279**: 19691 - 19697.
20. Liu E, Li X, Yan F, *et al.* Cyclin-dependent kinases phosphorylate human Cdt1 and induce its degradation. *J Biol Chem* 2004; **279**: 17283 - 17288.
21. Tatsumi Y, Sugimoto N, Yugawa T, *et al.* Dereglulation of Cdt1 induces chromosomal damage without rereplication and leads to chromosomal instability. *J Cell Sci* 2006; **119**: 3128-3140.
22. Arias EE, Walter JC. PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nat Cell Biol* 2006; **8**: 84 - 90.
23. Senga T, Sivaprasad U, Zhu W, *et al.* PCNA is a co-factor for Cdt1 degradation by CUL4/DDB1 mediated N-terminal ubiquitination. *J Biol Chem* 2006; **281**: 6246 - 6252.
24. Zhong W, Feng H, Santiago FE, *et al.* CUL-4 ubiquitin ligase maintains genome stability by restraining DNA-replication licensing. *Nature* 2003; **423**: 885-889.
25. Fujita M. Cdt1 revisited: complex and tight regulation during the cell cycle and consequences of deregulation in mammalian cells. *Cell Division* 2006; **1**: 22.
26. Zou L, Elledge SJ. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* 2003; **300**: 1542-1548.
27. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. *Nature* 2009; **461**: 1071-1078.
28. Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol* 2008; **9**: 616-627.

29. López-Contreras AJ, Fernandez-Capetillo O. The ATR barrier to replication-born DNA damage. *DNA Repair* 2010; **9**: 1249-1255.
30. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability - an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* 2010; **11**: 220-228.
31. Halazonetis TD, Gorgoulis VG, Bartek J. An Oncogene-Induced DNA Damage Model for Cancer Development. *Science* 2008; **319**: 1352-1355.
32. Toledo LI, Murga M, Fernandez-Capetillo O. Targeting ATR and Chk1 kinases for cancer treatment: A new model for new (and old) drugs. *Molecular Oncology* 2011; **5**: 368-373.
33. Spruck CH, Won K-A, Reed SI. Deregulated cyclin E induces chromosome instability. *Nature* 1999; **401**: 297-300.
34. Ekholm-Reed S, Méndez J, Tedesco D, *et al.* Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J Cell Biol* 2004; **165**: 789-800.
35. Durkin SG, Glover TW. Chromosome Fragile Sites. *Annu Rev Genet* 2007; **41**: 169-192.
36. Ozeri-Galai E, Lebofsky R, Rahat A, *et al.* Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites. *Mol Cell* 2011; **43**: 122-131.
37. Letessier A, Millot GA, Koundrioukoff S, *et al.* Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* 2011; **470**: 120-123.
38. Kunnev D, Rusiniak ME, Kudla A, *et al.* DNA damage response and tumorigenesis in Mcm2-deficient mice. *Oncogene* 2010; **29**: 3630-3638.
39. Pruitt SC, Bailey KJ, Freeland A. Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. *Stem Cells* 2007; **25**: 3121-3132.
40. Shima N, Alcaraz A, Liachko I, *et al.* A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* 2007; **39**: 93-98.
41. Bagley BN, Keane TM, Maklakova VI, *et al.* A dominantly acting murine allele of Mcm4 causes chromosomal abnormalities and promotes tumorigenesis. *PLoS Genet* 2012; **8**: e1003034.
42. Arlt MF, Wilson TE, Glover TW. Replication stress and mechanisms of CNV formation. *Curr Opin Genet Dev* 2012; **22**: 204-210.

43. Malkova A, Ira G. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev* 2013; **23**: 271-279.
44. Lebofsky R, Walter JC. New Myc-anisms for DNA replication and tumorigenesis? *Cancer Cell* 2007; **12**: 102-103.
45. Bester Assaf C, Roniger M, Oren Yifat S, *et al.* Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* 2011; **145**: 435-446.
46. Wold MS. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem.* 1997; **66**: 61-92.
47. Toledo Luis I, Altmeyer M, Rask M-B, *et al.* ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell* 2013; **155**: 1088-1103.
48. Blow JJ, Gillespie PJ. Replication licensing and cancer [mdash] a fatal entanglement? *Nat Rev Cancer* 2008; **8**: 799-806.
49. Galanos P, Vougas K, Walter D, *et al.* Chronic p53-independent p21 expression causes genomic instability by deregulating replication licensing. *Nat Cell Biol* 2016; **18**: 777-789.
50. Vassilev A, Lee CY, Vassilev B, *et al.* Identification of genes that are essential to restrict genome duplication to once per cell division. *Oncotarget* 2016; **7**: 34956-34976.
51. Jones RM, Mortusewicz O, Afzal I, *et al.* Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress. *Oncogene* 2013; **32**: 3744-3753.
52. Helmrich A, Ballarino M, Nudler E, *et al.* Transcription-replication encounters, consequences and genomic instability. *Nat Struct Mol Biol* 2013; **20**: 412-418.
53. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat Rev Genet* 2015; **16**: 583-597.
54. Sollier J, Cimprich KA. Breaking bad: R-loops and genome integrity. *Trends Cell Biol* 2015; **25**: 514-522.
55. Saponaro M, Kantidakis T, Mitter R, *et al.* RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell* 2014; **157**: 1037-1049.

56. Urban V, Dobrovolna J, Hühn D, *et al.* RECQ5 helicase promotes resolution of conflicts between replication and transcription in human cells. *J Cell Biol* 2016; **214**: 401-415.
57. Sollier J, Stork Caroline T, García-Rubio María L, *et al.* Transcription-coupled nucleotide excision repair factors promote r-loop-induced genome instability. *Mol Cell* 2014; **56**: 777-785.
58. Kim JM, Takemoto N, Arai Ki, *et al.* Hypomorphic mutation in an essential cell-cycle kinase causes growth retardation and impaired spermatogenesis. *EMBO J* 2003; **22**: 5260-5272.
59. Kim JM, Nakao K, Nakamura K, *et al.* Inactivation of Cdc7 kinase in mouse ES cells results in S-phase arrest and p53-dependent cell death. *EMBO J* 2002; **21**: 2168-2179.
60. van Straaten HWM, Copp AJ. Curly tail: a 50 year history of the mouse spina bifida model. *Anat Embryol (Berl)* 2001; **203**: 225-237.
61. Arai Y, Pulvers JN, Haffner C, *et al.* Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat Commun* 2011; **2**: 154.
62. Dehay C, Kennedy H. Cell-cycle control and cortical development. *Nat Rev Neurosci* 2007; **8**: 438-450.
63. Encalada SE, Martin PR, Phillips JB, *et al.* DNA replication defects delay cell division and disrupt cell polarity in early *Caenorhabditis elegans* embryos. *Developmental Biology* 2000; **228**: 225-238.
64. Ahuja AK, Jodkowska K, Teloni F, *et al.* A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. *Nat Commun* 2016; **7**:10660.
65. Alcantara D, O'Driscoll M. Congenital microcephaly. *Am J Med Genet C Semin Med Genet* 2014; **166**: 124-139.
66. Verloes A, Drunat S, Gressens P, *et al.* Primary autosomal recessive microcephalies and Seckel syndrome spectrum disorders. In: GeneReviews(R). Pagon RA, Adam MP, Bird TD, *et al.*, (eds). University of Washington, Seattle: Seattle WA, 2013.
67. Seckel HPG. Bird-headed dwarfs: Studies in developmental anthropology including human proportions. In: Bird-headed dwarfs: Studies in developmental

- anthropology including human proportions. Thomas CT, (ed). S. Karger: Basel, 1960.
68. Majewski F, Goecke T, Opitz JM. Studies of microcephalic primordial dwarfism I: Approach to a delineation of the Seckel syndrome. *Am J Med Genet* 1982; **12**: 7-21.
69. O'Driscoll M, Ruiz-Perez VL, Woods CG, *et al.* A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. *Nature Genetics* 2003; **33**: 497-501.
70. Ogi T, Walker S, Stiff T, *et al.* Identification of the first ATRIP-deficient patient and novel mutations in ATR define a clinical spectrum for ATR-ATRIP Seckel syndrome. *PLoS Genet* 2012; **8**: e1002945.
71. Al-Dosari MS, Shaheen R, Colak D, *et al.* Novel CENPJ mutation causes Seckel syndrome. *J Med Genet* 2010; **47**: 411-414.
72. Kalay E, Yigit G, Aslan Y, *et al.* CEP152 is a genome maintenance protein disrupted in Seckel syndrome. *Nat Genet* 2011; **43**: 23-26.
73. Shaheen R, Faqeih E, Ansari S, *et al.* Genomic analysis of primordial dwarfism reveals novel disease genes. *Genome Res* 2014; **24**: 291-299.
74. Qvist P, Huertas P, Jimeno S, *et al.* CtIP mutations cause Seckel and Jawad syndromes. *PLoS Genet* 2011; **7**: e1002310.
75. Murga M, Bunting S, Montana MF, *et al.* A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. *Nat Genet* 2009; **41**: 891-898.
76. O'Driscoll M. Mouse models for ATR deficiency. *DNA Repair* 2009; **8**: 1333-1337.
77. Lee Y, Shull ER, Frappart PO, *et al.* ATR maintains select progenitors during nervous system development. *EMBO Journal* 2012; **31**: 1177-1189.
78. Ruzankina Y, Pinzon-Guzman C, Asare A, *et al.* Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell* 2007; **1**: 113-126.
79. Schoppy DW, Ragland RL, Gilad O, *et al.* Oncogenic stress sensitizes murine cancers to hypomorphic suppression of ATR. *J Clin Invest* 2012; **122**: 241-252.
80. Mokrani-Benhelli H, Gaillard L, Biasutto P, *et al.* Primary microcephaly, impaired dna replication, and genomic instability caused by compound heterozygous ATR mutations. *Hum Mutat* 2012; **34**: 374-384.

81. Gonzalo S, Kreienkamp R. DNA repair defects and genome instability in Hutchinson–Gilford Progeria Syndrome. *Curr Opin Cell Biol* 2015; **34**: 75-83.
82. Wolters S, Schumacher B. Genome maintenance and transcription integrity in aging and disease. *Front Genet* 2013; **4**: 19
83. Couch FB, Bansbach CE, Driscoll R, *et al.* ATR phosphorylates SMARCA1 to prevent replication fork collapse. *Genes Dev* 2013; **27**: 1610-1623.
84. Gorlin R, Cervenka J, Moller K, *et al.* Malformation syndromes. A selected miscellany. *Birth Defects Orig Artic Ser* 1975; **11**: 39-50.
85. Bongers E, Van Kampen A, Van Bokhoven H, *et al.* Human syndromes with congenital patellar anomalies and the underlying gene defects. *Clin Genet* 2005; **68**: 302-319.
86. Bongers EMHF, Opitz JM, Fryer A, *et al.* Meier-Gorlin syndrome: Report of eight additional cases and review. *Am J Med Genet* 2001; **102**: 115-124.
87. Bongers EMHF, Van Bokhoven H, Van Thienen M-N, *et al.* The small patella syndrome: description of five cases from three families and examination of possible allelism with familial patella aplasia-hypoplasia and nail-patella syndrome. *J Med Genet* 2001; **38**: 209-214.
88. de Munnik SA, Bicknell LS, Aftimos S, *et al.* Meier-Gorlin syndrome genotype-phenotype studies: 35 individuals with pre-replication complex gene mutations and 10 without molecular diagnosis. *Eur J Hum Genet* 2012; **20**: 598-606.
89. de Munnik SA, Otten BJ, Schoots J, *et al.* Meier-Gorlin syndrome: Growth and secondary sexual development of a microcephalic primordial dwarfism disorder. *Am J Med Genet A* 2012; **158A**: 2733-2742.
90. Bicknell LS, Bongers EMHF, Leitch A, *et al.* Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat Genet* 2011; **43**: 356-359.
91. Bicknell LS, Walker S, Klingseisen A, *et al.* Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome. *Nat Genet* 2011; **43**: 350-355.
92. Guernsey DL, Matsuoka M, Jiang H, *et al.* Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nat Genet* 2011; **43**: 360-364.

93. Stiff T, Alagoz M, Alcantara A, *et al.* Deficiency in origin licensing proteins impairs cilia formation: implications for the aetiology of Meier-Gorlin syndrome. *PLoS Genet* 2013; **9**: e1003360.
94. Kerzendorfer C, Colnaghi R, Abramowicz I, *et al.* Meier–Gorlin syndrome and Wolf–Hirschhorn syndrome: Two developmental disorders highlighting the importance of efficient DNA replication for normal development and neurogenesis. *DNA Repair* 2013; **12**: 637-644.
95. Balasov M, Akhmetova K, Chesnokov I. Drosophila model of Meier-Gorlin syndrome based on the mutation in a conserved C-Terminal domain of Orc6. *Am J Med Genet A* 2015; **167**: 2533-2540.
96. Kuo AJ, Song J, Cheung P, *et al.* The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* 2012; **484**: 115-119.
97. Noguchi K, Vassilev A, Ghosh S, *et al.* The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. *EMBO J* 2006; **25**: 5372-5382.
98. Zhang W, Sankaran S, Gozani O, *et al.* A Meier-Gorlin syndrome mutation impairs the ORC1-nucleosome association. *ACS Chem Biol* 2015; **10**: 1176-1180.
99. Bleichert F, Balasov M, Chesnokov I, *et al.* A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *eLife* 2013; **2**: e00882.
100. Fenwick Aimee L, Kliszczak M, Cooper F, *et al.* Mutations in CDC45, encoding an essential component of the pre-initiation complex, cause Meier-Gorlin Syndrome and craniosynostosis. *Am J Hum Genet* 2016; **99**: 125-138.
101. Zou L SB. Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin. *Science* 1998; **280**: 593-596.
102. Moyer SE, Lewis PW, Botchan MR. Isolation of the Cdc45/Mcm2-7/GINS(CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci USA* 2006; **103**: 10236-10241.
103. Burrage Lindsay C, Charng W-L, Eldomery Mohammad K, *et al.* De novo GMNN mutations cause autosomal-dominant primordial dwarfism associated with Meier-Gorlin syndrome. *Am J Hum Genet* 2015; **97**: 904-913.

104. Alvarez S, Diaz M, Flach J, *et al.* Replication stress caused by low MCM expression limits fetal erythropoiesis and hematopoietic stem cell functionality. *Nat Commun* 2015; **6**: 8548.
105. Gao J, Wang Q, Dong C, *et al.* Whole exome sequencing identified *MCM2* as a novel causative gene for autosomal dominant nonsyndromic deafness in a Chinese family. *PLoS One* 2015; **10**: e0133522.
106. Hughes CR, Guasti L, Meimaridou E, *et al.* MCM4 mutation causes adrenal failure, short stature, and natural killer cell deficiency in humans. *J Clin Invest* 2012; **122**: 814-820.
107. Gineau L, Cognet C, Kara N, *et al.* Partial MCM4 deficiency in patients with growth retardation, adrenal insufficiency, and natural killer cell deficiency. *J Clin Invest* 2012; **122**: 821-832.
108. Casey JP, Nobbs M, McGettigan P, *et al.* Recessive mutations in MCM4/PRKDC cause a novel syndrome involving a primary immunodeficiency and a disorder of DNA repair. *J Med Genet* 2012; **49**: 242-245.
109. Volkening M, Hoffmann I. Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. *Mol Cell Biol* 2005; **25**: 1560-1568.
110. Maiorano D, Cuvier O, Danis E, *et al.* MCM8 as an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* 2005; **120**: 315-328.
111. Lee KY, Im J-S, Shibata E, *et al.* MCM8-9 complex promotes resection of double-strand break ends by MRE11-RAD50-NBS1 complex. *Nat Commun* 2015; **6**: 7744.
112. Park J, Long DT, Lee KY, *et al.* The MCM8-MCM9 complex promotes RAD51 recruitment at DNA damage sites to facilitate homologous recombination. *Mol Cell Biol* 2013; **33**: 1632-1644.
113. Nishimura K, Ishiai M, Horikawa K, *et al.* Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. *Mol Cell* 2012; **47**: 511-522.
114. Lutzmann M, Grey C, Traver S, *et al.* MCM8- and MCM9-deficient mice reveal gametogenesis defects and genome instability due to impaired homologous recombination. *Mol Cell* 2012; **47**: 523-534.

115. Wood-Trageser Michelle A, Gurbuz F, Yatsenko Svetlana A, *et al.* MCM9 mutations are associated with ovarian failure, short stature, and chromosomal instability. *Am J Hum Genet* 2014; **95**: 754-762.
116. AlAsiri S, Basit S, Wood-Trageser MA, *et al.* Exome sequencing reveals MCM8 mutation underlies ovarian failure and chromosomal instability. *J Clin Invest* 2015; **125**: 258-262.
117. Tenenbaum-Rakover Y, Weinberg-Shukron A, Renbaum P, *et al.* Minichromosome maintenance complex component 8 (MCM8) gene mutations result in primary gonadal failure. *J Med Genet* 2015; **52**: 391-399.
118. Goldberg Y, Halpern N, Hubert A, *et al.* Mutated MCM9 is associated with predisposition to hereditary mixed polyposis and colorectal cancer in addition to primary ovarian failure. *Cancer Genetics* 2015; **208**: 621-624.
119. Moldovan G-L, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. *Cell* 2007; **129**: 665-679.
120. Baple EL, Chambers H, Cross HE, *et al.* Hypomorphic PCNA mutation underlies a human DNA repair disorder. *J Clin Invest* 2014; **124**: 3137-3146.
121. Kunkel TA, Burgers PM. Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol* 2008; **18**: 521-527.
122. Johansson E, Dixon N. Replicative DNA polymerases. *Cold Spring Harb Perspect Biol* 2013; **5**: pii: a012799.
123. Fishel R, Lescoe MK, Rao M, *et al.* The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. *Cell* 1993; **75**: 1027-1038.
124. Leach FS, Nicolaides NC, Papadopoulos N, *et al.* Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell* 1993; **75**: 1215-1225.
125. Karran P. Microsatellite instability and DNA mismatch repair in human cancer. *Semin Cancer Biol* 1996; **7**: 15-24.
126. Thibodeau S, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science* 1993; **260**: 816-819.
127. Aaltonen L, Peltomaki P, Leach F, *et al.* Clues to the pathogenesis of familial colorectal cancer. *Science* 1993; **260**: 812-816.
128. Peltomaki P, Aaltonen L, Sistonen P, *et al.* Genetic mapping of a locus predisposing to human colorectal cancer. *Science* 1993; **260**: 810-812.

129. Peltomäki P, Lothe RA, Aaltonen LA, *et al.* microsatellite instability is associated with tumors that characterize the Hereditary Non-Polyposis Colorectal Carcinoma syndrome. *Cancer Res* 1993; **53**: 5853-5855.
130. Rayner E, van Gool IC, Palles C, *et al.* A panoply of errors: polymerase proofreading domain mutations in cancer. *Nat Rev Cancer* 2016; **16**: 71-81.
131. Palles C, Cazier J-B, Howarth KM, *et al.* Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* 2013; **45**: 136-144.
132. Church DN, Briggs SEW, Palles C, *et al.* DNA polymerase ϵ and δ exonuclease domain mutations in endometrial cancer. *Human Molecular Genetics* 2013; **22**: 2820-2828.
133. Briggs S, Tomlinson I. Germline and somatic polymerase ϵ and δ mutations define a new class of hypermutated colorectal and endometrial cancers. *J Pathol* 2013; **230**: 148-153.
134. Weedon MN, Ellard S, Prindle MJ, *et al.* An in-frame deletion at the polymerase active site of POLD1 causes a multisystem disorder with lipodystrophy. *Nat Genet* 2013; **45**: 947-950.
135. Shastri S, Simha V, Godbole K, *et al.* A novel syndrome of mandibular hypoplasia, deafness, and progeroid features associated with lipodystrophy, undescended testes, and male hypogonadism. *J Clin Endocrinol Metabol* 2010; **95**: E192-E197.
136. Reinier F, Zoledziewska M, Hanna D, *et al.* Mandibular hypoplasia, deafness, progeroid features and lipodystrophy (MDPL) syndrome in the context of inherited lipodystrophies. *Metabolism* 2015; **64**: 1530-1540.
137. Pelosini C, Martinelli S, Ceccarini G, *et al.* Identification of a novel mutation in the polymerase delta 1 (POLD1) gene in a lipodystrophic patient affected by mandibular hypoplasia, deafness, progeroid features (MDPL) syndrome. *Metabolism* 2014; **63**: 1385-1389.
138. Lessel D, Hisama FM, Szakszon K, *et al.* POLD1 Germline mutations in patients initially diagnosed with Werner syndrome. *Hum Mutat* 2015; **36**: 1070-1079.
139. Nicolas E, Golemis EA, Arora S. POLD1: Central mediator of DNA replication and repair, and implication in cancer and other pathologies. *Gene* 2016; **590**: 128-141.

140. Chu WK, Hickson ID. RecQ helicases: multifunctional genome caretakers. *Nat Rev Cancer* 2009; **9**: 644-654.
141. Croteau DL, Popuri V, Opresko PL, *et al.* Human RecQ helicases in DNA repair, recombination, and replication. *Annu Rev Biochem* 2014; **83**: 519-552.
142. Larsen NB, Hickson ID. RecQ Helicases: Conserved guardians of genomic integrity. In: DNA Helicases and DNA Motor Proteins. Spies M, (ed). Springer New York: New York, NY, 2013; 161-184.
143. German J. Bloom syndrome: A Mendelian prototype of somatic mutational disease. *Medicine* 1993; **72**: 393-406.
144. Ellis NA, Groden J, Ye T-Z, *et al.* The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 1995; **83**: 655-666.
145. Ellis NA, Lennon DJ, Proytcheva M, *et al.* Somatic intragenic recombination within the mutated locus BLM can correct the high sister-chromatid exchange phenotype of Bloom syndrome cells. *Am J Hum Genet* 1995; **57**: 1019-1027.
146. Oshima J, Sidorova JM, Monnat RJ Jr. Werner syndrome: Clinical features, pathogenesis and potential therapeutic interventions. *Ageing Res Rev* 2016 doi: 10.1016/j.arr.2016.03.002. [Epub ahead of print].
147. Shimamura A, Alter BP. Pathophysiology and management of inherited bone marrow failure syndromes. *Blood Rev* 2010; **24**: 101-122.
148. Clauson C, Schärer OD, Niedernhofer L. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harb Perspect Biol* 2013; **5**: a012732.
149. Webster ADB, Barnes DE, Arlett CF, *et al.* Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene. *Lancet* 1992; **339**: 1508-1509.
150. Barnes DE, Tomkinson AE, Lehmann AR, *et al.* Mutations in the DNA ligase 1 gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA damaging agents. *Cell* 1992; **69**: 495-503.
151. Siitonen HA, Sotkasiira J, Biervliet M, *et al.* The mutation spectrum in RECQL4 diseases. *Eur J Hum Genet* 2008; **17**: 151-158.
152. Larizza L, Roversi G, Volpi L. Rothmund-Thomson syndrome. *Orphanet J Rare Dis* 2010; **5**: 1-16.

153. Sangrithi MN, Bernal JA, Madine M, *et al.* Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell* 2005; **121**: 887-898.
154. Matsuno K, Kumano M, Kubota Y, *et al.* The N-terminal noncatalytic region of *Xenopus* RecQ4 is required for chromatin binding of DNA polymerase α in the initiation of DNA replication. *Mol Cell Biol* 2006; **26**: 4843-4852.
155. Capp C, Wu J, Hsieh T-s. RecQ4: the second replicative helicase? *Crit Rev Biochem Mol Biol* 2010; **45**: 233-242.
156. Hu Y, Raynard S, Sehorn MG, *et al.* RECQL5/Recql5 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments. *Genes Dev* 2007; **21**: 3073-3084.
157. Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harb Perspect Biol* 2013; **5**: a012708.
158. Chatterjee N, Siede W. Replicating damaged DNA in eukaryotes. *Cold Spring Harb Perspect Biol* 2013; **5**: a019836.
159. Sirbu BM, McDonald WH, Dungrawala H, *et al.* identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. *J Biol Chem* 2013; **288**: 31458-31467.
160. Dungrawala H, Cortez D. Purification of proteins on newly synthesized DNA using iPOND. In: The Nucleus. Hancock R, (ed). Springer New York: New York, NY, 2015; 123-131.
161. Dungrawala H, Rose Kristie L, Bhat Kamakoti P, *et al.* The replication checkpoint prevents two types of fork collapse without regulating replisome stability. *Mol Cell* 2015; **59**: 998-1010.
162. Schlacher K, Christ N, Siaud N, *et al.* Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* 2011; **145**: 529-542.
163. Schlacher K, Wu H, Jasin M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell* 2012; **22**: 106-116.

164. Trego Kelly S, Groesser T, Davalos Albert R, *et al.* Non-catalytic roles for XPG with BRCA1 and BRCA2 in homologous recombination and genome stability. *Mol Cell* 2016; **61**: 535-546.
165. Howlett NG, Taniguchi T, Olson S, *et al.* Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* 2002; **297**: 606-609.
166. Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev* 2012; **26**: 1393-1408.
167. Boerkoel FC, O'Neill S, André LJ, *et al.* Manifestations and treatment of Schimke immuno-osseous dysplasia: 14 new cases and a review of the literature. *Eur J Pediatr* 2000; **159**: 1-7.
168. Spranger J, Hinkel GK, Stöss H, *et al.* Schimke immuno-osseous dysplasia: A newly recognized multisystem disease. *J Pediatr* 1991; **119**: 64-72.
169. Schimke RN, Horton W, King C. Chondroitin-6-sulphaturia, defective cellular immunity, and nephrotic syndrome. *Lancet* 1971; **2**: 1088-1089.
170. Boerkoel CF, Takashima H, John J, *et al.* Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. *Nat Genet* 2002; **30**: 215-220.
171. Bansbach CE, Bétous R, Lovejoy CA, *et al.* The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes Dev* 2009; **23**: 2405-2414.
172. Ciccio A, Bredemeyer AL, Sowa ME, *et al.* The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes Dev* 2009; **23**: 2415-2425.
173. Postow L, Woo EM, Chait BT, *et al.* Identification of SMARCAL1 as a component of the DNA damage response. *J Biol Chem* 2009; **284**: 35951-35961.
174. Yusufzai T, Kong X, Yokomori K, *et al.* The annealing helicase HARP is recruited to DNA repair sites via an interaction with RPA. *Genes Dev* 2009; **23**: 2400-2404.
175. Yuan J, Ghosal G, Chen J. The annealing helicase HARP protects stalled replication forks. *Genes Dev* 2009; **23**: 2394-2399.
176. Driscoll R, Cimprich KA. HARPing on about the DNA damage response during replication. *Genes Dev* 2009; **23**: 2359-2365.
177. Bétous R, Glick GG, Zhao R, *et al.* Identification and characterization of SMARCAL1 protein complexes. *PLoS One* 2013; **8**: e63149.

178. Bansbach CE, Boerkoel CF, Cortez D. SMARCAL1 and replication stress: An explanation for SIOD? *Nucleus* 2010; **1**: 245-248.
179. Bétous R, Mason AC, Rambo RP, *et al.* SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes Dev* 2012; **26**: 151-162.
180. Couch FB, Cortez D. Fork reversal, too much of a good thing. *Cell Cycle* 2014; **13**: 1049-1050.
181. Krokan HE, Bjørås M. Base Excision Repair. *Cold Spring Harb Perspect Biol* 2013; **5**: a012583.
182. Kang Y-H, Lee C-H, Seo Y-S. Dna2 on the road to Okazaki fragment processing and genome stability in eukaryotes. *Crit Rev Biochem Mol Biol* 2010; **45**: 71-96.
183. Kumar S, Burgers PM. Lagging strand maturation factor Dna2 is a component of the replication checkpoint initiation machinery. *Genes Dev* 2013; **27**: 313-321.
184. Lai Mong S, Foiani M. Dna2 Offers Support for Stalled Forks. *Cell* 2012; **149**: 1181-1183.
185. Hu J, Sun L, Shen F, *et al.* The intra-S phase checkpoint targets Dna2 to prevent stalled replication forks from reversing. *Cell* 2012; **149**: 1221-1232.
186. Duxin JP, Dao B, Martinsson P, *et al.* Human Dna2 is a nuclear and mitochondrial DNA maintenance protein. *Mol Cell Biol* 2009; **29**: 4274-4282.
187. Ronchi D, Di Fonzo A, Lin W, *et al.* Mutations in DNA2 link progressive myopathy to mitochondrial DNA instability. *Am J Hum Genet* 2013; **92**: 293-300.
188. Makharashvili N, Paull TT. CtIP: A DNA damage response protein at the intersection of DNA metabolism. *DNA Repair* 2015; **32**: 75-81.
189. Borglum AD, Balslev T, Haagerup A, *et al.* A new locus for Seckel syndrome on chromosome 18p11.31-q11.2. *Eur J Hum Genet* 2001; **9**: 753-757.
190. Alderton GK, Joenje H, Varon R, *et al.* Seckel syndrome exhibits cellular features demonstrating defects in the ATR signalling pathway. *Hum Mol Genet* 2004; **13**: 3127-3138.
191. Lupski JR. Genomic rearrangements and sporadic disease. *Nat Genet* 2007; **39**: S43-47.
192. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010; **61**: 437-455.

193. Hastings PJ, Ira G, Lupski JR. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* 2009; **5**: e1000327.
194. Hastings PJ, Lupski JR, Rosenberg SM, *et al.* Mechanisms of change in gene copy number. *Nat Rev Genet* 2009; **10**: 551-564.
195. Hart L, O'Driscoll M. Causes and consequences of structural genomic alterations in the human genome. In: eLS. John Wiley & Sons Ltd, Chichester, 2013. <http://www.els.net> [doi: 10.1002/9780470015902.a0024976].
196. Mizuno KI, Miyabe I, Schalbetter SA, *et al.* Recombination-restarted replication makes inverted chromosome fusions at inverted repeats. *Nature* 2013; **493**: 246-249.
197. Lambert S, Mizuno Ki, Blaisonneau J, *et al.* Homologous recombination restarts blocked replication forks at the expense of genome rearrangements by template exchange. *Mol Cell* 2010; **39**: 346-359.
198. Amiel A, Korenstein A, Gaber E, *et al.* Asynchronous replication of alleles in genomes carrying an extra autosome. *Eur J Hum Genet* 1999; **7**: 223-230.
199. Amiel A, Reish O, Gaber E, *et al.* Asynchronous replication of alleles in genomes carrying a microdeletion. *Isr Med Assoc J* 2002; **4**: 702-705.
200. Colnaghi R, Carpenter G, Volker M, *et al.* The consequences of structural genomic alterations in humans: Genomic Disorders, genomic instability and cancer. *Semin Cell Dev Biol* 2011; **22**: 875-885.
201. Vitelli. F, Lindsay. E.A, A. B. Genetic dissection of the DiGeorge syndrome phenotype. *Cold Spring Harb Symp Quant Biol* 2002; **67**: 327-332.
202. O'Driscoll M, Dobyns WB, van Hagen JM, *et al.* Cellular and Clinical Impact of Haploinsufficiency for Genes Involved in ATR Signaling. *Am J Hum Genet* 2007; **81**: 77-86.
203. Outwin E, Carpenter G, Bi W, *et al.* Increased RPA1 gene dosage affects genomic stability potentially contributing to 17p13.3 duplication syndrome. *PLoS Genet* 2011; **7**: e1002247.
204. Kerzendorfer C, Hannes F, Colnaghi R, *et al.* Characterizing the functional consequences of haploinsufficiency of NELF-A (WHSC2) and SLBP identifies novel cellular phenotypes in Wolf-Hirschhorn syndrome. *Hum Mol Genet* 2012; **21**: 2181-2193.

205. Ensminger AW, Chess A. Coordinated replication timing of monoallelically expressed genes along human autosomes. *Hum Mol Genet* 2004; **13**: 651-658.
206. Gilbert DM. Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* 2002; **14**: 377-383.
207. Sasaki T, Gilbert DM. The many faces of the origin recognition complex. *Curr Opin Cell Biol* 2007; **19**: 337-343.
208. Wallace JA, Orr-Weaver TL. Replication of heterochromatin: insights into mechanisms of epigenetic inheritance. *Chromosoma* 2005; **114**: 389-402.
209. Nakayama J-i, Allshire RC, Klar AJS, *et al.* A role for DNA polymerase [alpha] in epigenetic control of transcriptional silencing in fission yeast. *EMBO J* 2001; **20**: 2857-2866.
210. Stiff T, Casar Tena T, O'Driscoll M, *et al.* ATR promotes cilia signalling: links to developmental impacts. *Hum Mol Genet* 2016; **25**: 1574-1587.
211. Arts HH, Bongers EMHF, Mans DA, *et al.* C14ORF179 encoding IFT43 is mutated in Sensenbrenner syndrome. *J Med Genet* 2011; **48**: 390-395.
212. Gilissen C, Arts HH, Hoischen A, *et al.* Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am J Hum Genet* 2010; **87**: 418-423.
213. Haycraft CJ, Serra R. Cilia involvement in patterning and maintenance of the skeleton. *Curr Top Dev Biol* 2008; **85**: 303-332.
214. Serra R. Role of intraflagellar transport and primary cilia in skeletal development. *Anat Rec (Hoboken)* 2008; **291**: 1049-1061.
215. Haycraft CJ, Zhang Q, Song B, *et al.* Intraflagellar transport is essential for endochondral bone formation. *Development* 2007; **134**: 307-316.
216. Song B, Haycraft CJ, Seo H-s, *et al.* Development of the post-natal growth plate requires intraflagellar transport proteins. *Dev Biol* 2007; **305**: 202-216.
217. Kaushik AP, Martin JA, Zhang Q, *et al.* Cartilage abnormalities associated with defects of chondrocytic primary cilia in Bardet-Biedl syndrome mutant mice. *J Orthop Res* 2009; **27**: 1093-1099.
218. Pacheco M, Valencia M, Caparrós-Martín JA, *et al.* Evc works in chondrocytes and osteoblasts to regulate multiple aspects of growth plate development in the appendicular skeleton and cranial base. *Bone* 2012; **50**: 28-41.

- 219. Toriyama M, Lee C, Taylor SP, *et al.* The ciliopathy-associated CPLANE proteins direct basal body recruitment of intraflagellar transport machinery. *Nat Genet* 2016; **48**: 648-656.
- 220. Coulombe P, Grégoire D, Tsanov N, *et al.* A spontaneous Cdt1 mutation in 129 mouse strains reveals a regulatory domain restraining replication licensing. *Nat Commun* 2013; **4**: 2065.
- 221. Philippakis AA, Azzariti DR, Beltran S, *et al.* The Matchmaker Exchange: A platform for rare disease gene discovery. *Hum Mutat* 2015; **36**: 915-921.

Table 1. Some commonly used model organism phenotype-cataloguing databases

Model organism	Database
Mouse	Mouse Genome Informatics (MGI) http://www.informatics.jax.org/ Knockout Mouse Project (KOMP) https://www.komp.org/
Zebrafish	https://zfin.org/
Yeast	<i>Saccharomyces cerevisiae</i> http://www.yeastgenome.org/ <i>Schizosaccharomyces pombe</i> http://www.pombase.org/
Nematode	http://www.wormbase.org/#012-34-5
Bacteria	http://www.wormbase.org/#012-34-5

Figure Legends.

Figure 1.

The landscape of the clinical features commonly exhibited by disorders with defects in the pathways governing genome stability.

Body size is profoundly reduced in several genome instability conditions (IUGR: intra-uterine growth retardation. PNGR: post-natal growth retardation). Examples include DDR disorders such as Seckel syndrome (e.g. those defective in *ATR*, *ATRIP*, *DNA2*, *BRCA2*) and DSB-R defective conditions such as those compromised in NHEJ (e.g. defects in *XRCC4*, *LIG4*) or HRR (e.g. BLM syndrome, Fanconi anaemia). Haematological features can range from pancytopenia, SCID (severe combined immunodeficiency), CID (combined immunodeficiency) and even sub-clinical immune impacts as seen in some NHEJ-defective individuals (e.g. Artemis-SCID, *LIG4*, *XLF/NEHJ1*) to frank anaemia (e.g. Fanconi anaemia). Cancers of the haematopoietic system are also observed in some conditions (e.g. Ataxia telangiectasia, Nijmegen breakage syndrome, Fanconi anaemia). Disorders of the genome stability networks can present with complex neurological features including microcephaly (*LIG4*, *XLF/NEHJ1*, Seckel syndrome, BLM syndrome,

PNKP-Microcephaly and seizure syndrome) and neurodegeneration (Purkinje loss in Ataxia telangiectasia) which can also result in gait abnormalities. A range of other features are also typical of these types of disorders including solid tumours such as skin cancers (basal and squamous cell carcinoma) in nucleotide excision repair disorders (e.g. Xeroderma pigmentosum), head and neck cancers (e.g. Fanconi anaemia) and even sarcomas (BLM syndrome, *RECQL4* syndromes). Metabolic dysfunctions such as lipodystrophy (e.g. Cockayne syndrome, Seckel syndrome), but specifically insulin-resistant diabetes (e.g. Ataxia telangiectasia, BLM syndrome), also appear to frequently occur along with marked progeroid features (e.g. WRN, laminopathies).

Figure 2.

Origin firing, re-replication and dormant origins.

(A) Upon firing of the licensed origins the MCM2-7 containing helicase complexes move in opposite directions coincident with the DNA synthesis and accessory machinery as indicated by the relative positioning of the idealised replisomes.

(B) Origins are indicated in black and licensed origins in grey. Not all licensed origins are fired and not all origins fire simultaneously, giving rise to replication bubbles of various sizes. Origin under-firing increases the distance that fired origins have to travel thereby increasing the risk of them stalling. On the other hand, over-firing origins can deplete dNTPs and the availability of the replicative machinery, increasing the risk of

fork stalling. Re-replication from a fired origin represents a serious threat to genome stability.

(C) Dormant origin firing can rescue stalled replication forks. If the progress of a fork is stalled (red symbol), an adjacent licensed origin can be fired to replicate the intervening region. Once the reason underlying the stalling of the origin fork has been successfully addressed, replication can proceed bi-directionally.

Figure 3.

Replication fork stalling, processing stalled forks and R-loops.

(A) Stalling of a replication fork at a lesion (e.g. strand break/acute dNTP depletion: red symbol) can result in uncoupling of the helicase activity of the replisome from the polymerases, generating tracks of ssDNA. These become RPA bound and this structure then activates the ATR-dependent DDR to stabilise the stalled fork. If the fork is not stabilised it can collapse generating DSBs. Stabilised forks can be processed in many ways, typically involving the controlled engagement of specific nucleases. These can generate flap structures as intermediates in the repair process, whilst fork regression into a 'chicken foot' structure represents another topological route to recovery from fork stalling. A regressed fork can be subjected to nuclease processing to enable re-initiation of synthesis from the fork.

(B) Stalled forks can also induce template switching events where an adjacent fork containing some degree of sequence homology can be exploited as a template to bypass a replication blocking lesion. This form of microhomology mediated break-induced repair (MMBIR) results in amplification of the DNA sequence.

(C) R-loops are DNA-RNA hybrid structures (grey & red strands respectively) generated upon collision of a replication fork with a transcription fork. R-loop resolution involves the coordinate action of DNA-RNA helicase activity to unwind the structure (SETX: senataxin) and RNA digestion via RNase H. A growing list of additional factors have also been implicated in R-loop processing (e.g. RECQL5, XPF, XPG, CS-B). The exposed ssDNA region of the R-loop (black strand) is susceptible to breakage if the R-loops are not properly resolved.