Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

CHAPTER ONE

Fundamentals of G-quadruplex biology

F. Brad Johnson*
Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, United States
*Corresponding author: e-mail address: johnsonb@pennmedicine.upenn.edu

Contents

1. Introduction

It has been known since the 1960s that some DNA sequences with runs of adjacent guanines can assemble into helical arrangements of stacked G-quartets, with each quartet comprising four guanines linked by Hoogsteen hydrogen bonds.[1,2] G4s form under conditions of physiologic

Annual Reports in Medicinal Chemistry, Volume 54
ISSN 0065-7743
https://doi.org/10.1016/bs.armc.2020.06.004

pH and ionic strength, and some are as stable as, or even more stable than, duplex DNA. Their stability derives from the hydrogen bonds as well as from pi-orbital interactions among the quartets and from coordination by the quartets of centrally located cations (principally thought to be K + within cells). G4s are a polymorphic family of structures, which can be formed by guanines within a single nucleic acid strand (unimolecular) or from multiple strands (intermolecular). A minimum of two stacked quartets are required, but stability generally increases with additional quartets. The orientations of the phosphodiester backbones of G4s can vary, with all running in the same 5′ - 3′ direction (parallel) or two of them running opposite to the other two (antiparallel) or with one of the four running counter to the other three (hybrid). Furthermore, the glycosidic bond angles of the guanines can vary, as can the sequences, lengths, and connectivity of the intervening loops. There are also examples where a nucleotide interrupting what would otherwise be a run of Gs is looped-out to enable uninterrupted stacking of G-quartets to form a quadruplex. G4s can also form from RNA sequences, a key difference from G4 DNA being that G4 RNAs favor parallel folds. Thus rather than being a uniform structure, G4s are a family of interrelated nucleic acid folds.

Interest in the possibility that G-quadruplexes might play roles in biology grew as it became evident that sequences with the capacity to form G4s (at least *in vitro*) existed within telomeres and near the promoters of oncogenes.[3] The identification of proteins that can bind and process G4s with high affinity and some degree of selectivity, including transcription factors, components of telomere chromatin, helicases and nucleases, further stimulated interest, as did the discovery and design of small molecules that could bind and stabilize G4 folds (some early reviews include[4–6]). The earliest tests of G4 biology yielded intriguing, yet often inconclusive findings, and it is probably fair to characterize the view of most biologists until at least the mid-2000s as one of high skepticism concerning the physiologic importance of G4s. However, the discovery that sequences with the potential to form unimolecular G4s are abundant in the genomes of many eukaryotes,[3,7–10] coupled with the knowledge that individuals lacking DNA helicases that can remove G4s suffer from genome instability and transcriptional derangements (*e.g.* people with Werner and Bloom syndromes, who lack the WRN and BLM helicases, respectively; reviewed in[11]) convinced many investigators that at least G4s, if not prevented from forming, might in some way contribute to pathology. Supporting this view are the findings that bacterial transcriptomes are depleted of potential G4-forming sequences and that

artificial introduction of G4 RNA inhibits bacterial growth, whereas in mammalian cells RNAs with G4 forming potential appear to be predominantly unfolded.[12] Similarly, the most thermodynamically stable, and genetically unstable, G4-forming sequences were found to be depleted in most of 600 species analyzed.[13] In contrast, evidence that G4 forming sequences are evolutionarily conserved and overall are over-represented in many genomes, including that of humans, suggests that they must overall be of some benefit.[14] Regardless, investigation continued, and as described below, a large body of evidence now supports a variety of normal, as well as pathologic, roles for G4s in biology.

In this chapter I discuss (i) some basic experimental approaches to the investigation of biological G-quadruplexes, and (ii) highlight some of the evidence indicating their importance in particular biological processes. I also briefly discuss some of the more exciting and recent findings in the field, and it provides warnings about some of the many experimental difficulties that can accompany biological G4 investigations.

2. Experimental approaches to G-quadruplex biology

2.1 Some general principles and potential pitfalls

It is relatively challenging to demonstrate the existence and functions of G-quadruplexes in biological contexts, although as detailed later in this chapter, experimental findings have overall provided strong support for the importance of G4s in a remarkable range of biological processes. The challenge stems from several facts, chief among them that the presence or absence of a G-quadruplex reflects whether or not a particular nucleic acid has adopted a G4 conformation, rather than simply the absence or presence of the underlying nucleic acid itself. The fact that G4s are a family of related conformations, rather than a single structure, also adds to the difficulty of developing probes that can detect a range of G4 structures, while also maintaining high affinity and selectivity. Overall, it is important to understand that methods for detecting G4s unambiguously in biological contexts are not yet available, although the use of several independent methods can be used to build a strong case for G4 involvement in particular contexts.

One way to illustrate the challenge of accurately detecting G-quadruplex folds in biology is to compare the experimental approaches taken to explore G4 biology to those used to investigate the function of individual genes. One of the most revealing ways to understand the function of a gene is to examine the consequences of its selective inactivation, which has become

quite straightforward through the use of genome editing (*e.g. via* CRISPR). Any resulting changes in cell or organismal behavior reflect the normal function of the gene. Although some of the changes will be direct and others indirect, the fact remains that all of the changes are ultimately attributable to the gene itself. Furthermore, inactivation of a gene provides a critical control for testing the fidelity of probes for the RNA and protein products of the gene, *e.g.* antibodies and small molecules. Such probes are themselves essential for understanding the mechanisms by which the gene exerts its effects, for example by enabling the subcellular location and levels of a protein encoded by a gene to be monitored. Biological contexts are highly complex, and so a probe must distinguish with high selectivity among a large number of potential targets. For example, a typical human cell is estimated to contain 10^4–10^5 different proteins, and it is not uncommon for a probe of any given protein to react to some degree with additional proteins. Moreover, the relative abundance of different biomolecules can vary widely, *e.g.* concentrations of different proteins in human cells range over more than six orders of magnitude.[15] Thus even a probe with 1000-fold selectivity for a low abundance target *vs* a highly expressed and unintended target could conceivably bind primarily to the unintended target. Therefore, only if the reactivity of a probe is abolished by inactivation of the gene encoding its intended target, can it be concluded safely that off-target reactivity will not confound experimental findings; *i.e.* probe reactivity indicates the presence of the target.

The chain of inferences just described concerning the elucidation of gene function, and the confirmation of probe specificity for gene products, cannot be made in a similar fashion for G-quadruplexes in biological settings, for several reasons. First, although it is possible to reduce or eliminate the G4-forming potential of a particular nucleic acid *via* mutation of its sequence, any resulting changes in the function of that nucleic acid cannot necessarily be attributed to G4 loss *per se*, because the change in sequence could (and will likely) impact additional functions of the nucleic acid that are unrelated to G4 formation. For example, consider a DNA sequence near a gene promoter that forms stable G4s *in vitro*, and which has clear effects on the rate of transcription from the promoter *in vivo*. If mutation of the sequence to reduce its G4 forming potential alters the rate of transcription, this could reflect a change in binding by a sequence-specific duplex DNA-binding transcription factor protein, rather than a change in G4 stability. Compounding this problem is the common experimental practice of mutating several Gs within a G4-forming sequence to test the biological consequences, when in principle, mutation of a few guanines (in some cases

a single G) should significantly diminish G4 stability. Second, there is no known way to selectively remove all G4s from a biological context, and thus the specificity of G4 probes cannot currently be confirmed in a global fashion. Third, and perhaps most importantly, a probe for a G4 can be expected to stabilize the conformation that it detects (because its reactivity is based on having a surface energetically complementary to its target), and it can therefore be difficult to know to what degree the detected G4 existed prior to application of the probe. These and additional potential pitfalls are discussed in Sections 2.2–2.7, which provide commentary on some of the experimental approaches used to explore G4 biology.

2.2 *In silico* approaches

Algorithmic approaches to identifying sequences with the potential to form G4s have been essential to the investigation of G4 biology. Recently, these were reviewed comprehensively,[16] and only a few key points will be discussed here. The algorithms generally identify short stretches of nucleotides along one nucleic acid strand that contain four runs of at least two, but more typically three or more Gs, separated by one or more loop nucleotides of any type. In this chapter, such sequences are referred to as G4 "motifs," so as to distinguish them from actual G4 folds, because it is not yet possible to know precisely the propensity of an arbitrary such sequence to form a G4, *in vitro* and especially *in vivo*. G4 motifs are typically sought on one strand given the generally entropic favorability of intramolecular folds *vs* those requiring disparate strands to associate into intermolecular G4s, together with the reasonable assumption that G4 DNA is most likely to form when DNA is in a single-stranded form (*e.g.* during replication) and thus has diminished competition with duplex formation. However, it is worth mentioning that G-runs on different, nearby strands of a duplex could conceivably rearrange to form intermolecular G4s, for example under the influence of G4 binding proteins and torsional stresses driving duplex unwinding. Algorithms for identifying such duplex-derived intermolecular G4s have also been described, although to date they have not revealed much evidence of such G4 formation *in vivo*.[17,18] For intramolecular G4 searches, exactly how long the overall motif can be is not definite, although *in vitro* G4 stabilities are generally higher for shorter motifs, and if at least two of the loops are very short,[19] and parallel folds are favored when at least one of the loops is only one nucleotide long.[20] Some of the most commonly used open-source algorithms with web interfaces are G4 Hunter (http://bioinformatics.ibp.cz),

PQRSfinder (https://pqsfinder.fi.muni.cz/), QGRS (http://bioinformatics.ramapo.edu/QGRS/index.php) and G4RNA screener (http://scottgroup.med.usherbrooke.ca/G4RNA_screener/).

Several approaches have been used to improve the predictive power of *in silico* G4 algorithms. For example, G4Hunter incorporates data on the *in vitro* stability of 392 G4 motifs.[21] The G4RNA screener uses a machine learning derived model based on experimentally verified G4 and non G4 RNA structures in the G4 RNA database.[22] DNA and RNA G4-seq evaluate the genome and transcriptome-wide capacity of G4-motifs to impede the progression of a DNA or RNA polymerases selectively under G4-promoting conditions, typically the presence of K+ and the G4 small molecule ligand pyridostatin (PDS), and can thus be used to refine *in silico* predictions,[14,23] including incorporation of machine learning approaches.[24] Impressively, G4-seq has been used to compare the genome-wide G4-forming potential of 12 different species.[25] However, it is important to understand that although G4-seq helps reveal the G4-forming potential of different motifs *in vitro*, polymerases might also be impeded by other non-G4 secondary structures, and moreover that *bona fide* G4s identified by G4-seq and other *in vitro* approaches will not necessarily form *in vivo*. Similarly, the use of G4 stabilities as measured *in vitro* may not be highly predictive of those formed *in vivo*. Even though they employ physiologic pH and ionic strength, *in vitro* studies involve conditions that are otherwise quite different from those *in vivo*, and which may substantially impact G4s. In particular, *in vitro* studies of intramolecular G4s are typically carried out under dilute conditions, *e.g.* approximately 1 mM DNA, which is equivalent to 10 mg/mL for a 30-mer oligonucleotide and which is dramatically lower that the approximate overall intracellular macromolecular concentration of 400 mg/mL.[26] Agents that mimic intracellular conditions by enhancing molecular crowding and reducing water concentration (*e.g.* polyethylene glycol) have been found to stabilize G4s *vs* duplex DNA and to favor parallel *vs* other G4 folds.[26,27] Moreover, the presence *in vivo* of G4 binding proteins and unwinding helicases, histones and other chromatin proteins, phase-separation of intracellular components to generate varied biochemical subcellular chemical environments,[28] and regulators of topological DNA stresses are very likely to have major, but so far largely uncharacterized, impacts on G4 structure and stability.

Although *in silico* approaches are useful for identifying G4 motifs within particular genes of interest, they also enable genome-wide association studies to address the potential connection of particular factors of interest to G4

biology. For example, the WRN and BLM DNA helicases unwind G4s *in vitro*, and gene expression in cells lacking either of these helicases occur preferentially for genes with G4 motifs, consistent with the notion that G4 unwinding by these helicases impacts transcription.[29–32] Although this relationship does not hold for every gene, and thus cannot be predicted for an arbitrary gene, such genome-wide association studies enable the statistical significance of the general relationship to be evaluated by effectively testing it a large number of times. It is also important to consider *in silico* controls to further evaluate such relationships, for example by demonstrating that the relationship is weaker for motifs that imperfectly match G4 motifs, *e.g.* sequences that would qualify as G4 motifs except that they are missing one or two of the Gs required for G4 formation.[33]

2.3 Antibodies

Antibodies have relatively large and randomly generated protein surfaces that have been subjected to strong selective pressure to achieve high target affinity and specificity and are thus very useful probes of biomolecules. Indeed, anti-G-quadruplex antibodies have provided some of the clearest evidence of G4s in biology. Among the first, and strongest, such evidence came from antibodies raised against G4s formed *in vitro* by the single-stranded telomere DNA overhang sequence of the unicellular eukaryote *Stylonychia lemnae* ($T_4G_4T_4G_4$, which *in vitro* can form antiparallel G-G paired hairpins, two of which can pair to form a G4). *S. lemnae* possess macronuclei that contain several million chromosome fragments with telomeres at each end and are thus a rich source of whatever structures those telomeres might form. Immunofluorescence microscopy revealed that macronuclei of formaldehyde-crosslinked cells were stained using an antibody recognizing parallel as well as anti-parallel G4s, but not with an antibody that binds only parallel G4s, suggesting the telomeres might exist as the bimolecular G4s observed *in vitro*.[34] Interestingly, replicating chromatin was not stained, consistent with the expectation that G4s should be removed by an advancing replication fork. Nonetheless, there was uncertainty about the specificity of antibody binding within the context of a complex biological substrate and also about whether the staining might reflect G4s induced by the antibody, since molecular motion sufficient for G4 folding might occur even in formaldehyde crosslinked cells. However, RNAi-based genetic depletion of the telomere binding protein TEBPβ abolished antibody staining,[35] which was particularly remarkable because the homolog of TEBPβ in the related

protist *Oxytricha nova* had been previously shown to catalyze G4 formation *in vitro*.[36] This demonstration that antibody reactivity depended on the presence *in vivo* of a telomere binding protein that itself was expected to promote G4 folds, provided strong evidence for antibody specificity and that the telomeres form G4s naturally. Uncertainties remain, including what fraction of telomeres are in a G4 state at any given time (and this fraction could conceivably be low), and what exactly are the functions of the telomeric G4s. Nonetheless, these landmark studies encouraged further biological investigation of G4s.

The antibodies used for the *S. lemnae* studies are so-called single chain antibodies (scFvs), in which variable regions of the heavy and light chains that compose the antigen recognition surface of a natural antibody are synthesized as a single polypeptide. This provides a convenient platform for random mutagenesis and selection to generate antibodies against a target of interest. Indeed, the Balasubramanian laboratory generated a set of anti-G4 scFvs against a intramolecular parallel G4 formed by a DNA sequence in the c-kit promoter ($CG_3CG_3CGCGAG_3AG_4$), including the scFvs hf1 and hf2.[37] These recognized several G4s and had approximately 1000-fold higher affinity for these than duplex DNA. Another advantage of scFvs is that they can be expressed in cells, and when hf1 was expressed in a human cell line, it altered the expression of genes with high G4-forming potential preferentially *vs* those without such potential, supporting its specificity in an *in vivo* context.[38] The scFv BG4, which is widely used as a probe of G4s in biology, was isolated in a similar fashion, and is remarkable for its capacity to recognize a broad array of DNA and RNA G4 structures with nanomolar affinity, while having substantially less affinity for some duplex and single-stranded DNA sequences.[39,40] The broad reactivity of BG4 for G4-forming sequences is arguably surprising given the diversity of different G4 folds, and it would be interesting to investigate precisely which G4 features are bound. Another anti-G4 scFv, called D1, was generated similarly against a G4 from the human *BCL2* gene, and proved to have high selectivity for parallel *vs* hybrid and anti-parallel G4s *in vitro* and to preferentially immunoprecipitate regions of chromatin from human cells having parallel G4 forming ability.[41] In addition to the above scFvs, a standard IgG type antibody, called 1H6, was generated using conventional monoclonal antibody techniques against the tetramolecular parallel G4 DNA formed by $(T_4G_4)_2$, and was found *in vitro* to bind several (but not all) DNA G4 structures with nanomolar affinity but not to several control RNA, ssDNA, and duplex DNA sequences.[42]

Several observations support the capacity of the above antibodies to detect G4s *in vivo*. For example, nuclear staining by BG4, D1 and 1H6 of cultured human cells, as detected by immunofluorescence microscopy, is augmented by small molecules that stabilize G4 formation.[39,41,42] Similarly nuclear 1H6 staining was enhanced by genetic deficiency for the G4-unwinding FANCJ helicase in human cells, and telomere staining by BG4 was increased by genetic deficiency for the G4-unwinding WRN and/or BLM helicases in mouse cells.[42,43] Furthermore, BG4 and D1 staining increases during S-phase as would be expected if G4s form from the single stranded DNA that is exposed during DNA replication (and at least in the case of D1, staining returns to baseline levels in the G2/M phase, demonstrating elevated S-phase staining is not simply caused by a doubling of DNA content).[39,41]

However, it must be stressed that within a biological context none of the above findings ensures that any of the antibodies will detect a given G4 in a sensitive fashion nor G4s in a selective fashion. First, an increase in reactivity following manipulations that promote G4s, for example observing additional nuclear foci upon antibody staining following treatment with a G4 binding ligand, does not prove that the reactivity under normal conditions is due to G4s. In other words, even though the manipulation may increase reactivity by enhancing the levels of G4s, this does not prove that the pre-manipulation reactivity was due to natural G4s (*e.g.* it could be cross-reactivity with a non-G4 target). Second, reactivity with the full range of potential targets present in a cell has not been assessed for any of the antibodies *in vitro*, but emerging data are already raising concerns. Indeed, it has been found that 1H6 has poor affinity for G4s lacking thymidines in their sequences, and moreover can bind with high affinity to poly-T single stranded DNA.[44] In the case of BG4, a recent DNA microarray survey of ~24,000 G4 and non-G4 forming sequences found that although the antibody binds to a wide array of G4s, it also binds to many ssDNA sequences incapable of forming G4s, to some even more strongly than to G4s.[45] Although these findings have not yet been validated by independent studies, it raises important questions about the many reports using BG4 reactivity to indicate G4s. Third, even if an antibody is shown to fully discriminate among different nucleic acids *in vitro*, it may yield misleading information in a biological context because it could bind a cross-reactive non-nucleic acid species or because it could induce a G4 fold artifactually. Demonstrating that antibody binding can be blocked with an excess of G4 competitor is encouraging, but it does not rule out the possibility of

binding to a cross-reactive target (*i.e.* if the G4 and cross-reactive molecule bind to overlapping surfaces on the antibody, they will compete with one another even though they are distinct molecules). The most convincing way to confirm specificity would be to demonstrate an absence of binding after all G4s have been selectively removed *in vivo*, *e.g.* by overexpressing a pan-G4-unwinding helicase (and DHX36 may be such a candidate[45]), but such findings have not been reported. Nonetheless, the specificity of the parallel G4-specific scFv D1 is suggested by the observations that nuclear and telomere staining are diminished when cells are treated with the berberine derivative Ber8, which binds anti-parallel G4[41]; presumably Ber8 converts the targeted parallel G4s to anti-parallel folds, but this has not yet been directly demonstrated. Overall these considerations indicate that binding of G4 antibodies to biological samples *per se* should not be assumed to reflect *bona fide* G4 structures. However, it is equally important to acknowledge that even imperfect probes for G4s can be very useful, provided appropriate controls are included and results are not overinterpreted.

2.4 Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) is a powerful approach to understanding the distribution across the genome of the various protein, RNA and DNA components that contribute to regulated gene expression, replication, and repair. The standard approach is to covalently cross-link proteins to each other and to nucleic acids within cells or tissues (typically with formaldehyde), and then isolate and fragment the chromatin into stretches of a few hundred DNA base pairs. The levels within each fragment of a feature of interest, *e.g.* a histone post-translational modification, can be determined by immunoprecipitating the fragments with an antibody that binds the feature, and then quantifying the relative efficiencies of fragment recovery, usually by using quantitative PCR (ChIP-qPCR) or next generation sequencing (ChIP-seq). ChIP using this approach and the G4 antibody BG4 has been reported,[46] and the results are consistent with the idea that naturally formed G4 folds are being detected. In particular, ChIP-seq with BG4 revealed that sequences with G4-forming potential, particularly in nucleosome depleted promoter regions, are more efficiently precipitated (*i.e.* "ChIPed") than those not predicted to form G4s.

There are several caveats that should be considered in interpreting G4 ChIP data. First, it is conceivable that during isolation and immunoprecipitation of chromatin natural G4 folds might be lost, or alternatively, formed

artifactually. The DNA within crosslinked chromatin will have some degree of conformational freedom, and although the crosslinking of proteins to the DNA might restrain changes in G4 folds to some degree, it is not clear that this is enough to prevent bulk changes in G4 formation or disassembly. An interesting approach that avoids potential losses and gains of G4s during chromatin isolation is to express anti-G4 scFvs within cells, which can thus be crosslinked *in vivo* to native chromatin prior to its isolation. Indeed, this approach was used for the hf1 and D1 scFvs, which were expressed in yeast and cultured human cervical carcinoma cells, respectively.[41,47] In the yeast case, telomeres were ChIPed by hf1 only under genetic conditions where G4s were expected to form (see Section 3.5 below[47]). In the human case, ChIPed chromatin was primarily gene-associated and enriched for DNA sequences that form parallel G4s *in vitro*, consistent with the demonstrated specificity of D1 for parallel G4s *in vitro*, and supporting the idea that D1 binds *bona fide* parallel G4s within chromatin.[41] However a second caveat is that anti-G4 antibodies are expected to induce the G4 folds to which they bind, as discussed above. Such artifactual G4 could be induced by scFvs expressed *in vivo*, as well as during antibody incubation with chromatin *in vitro* during standard ChIP (*e.g.* by BG4), and therefore ChIP evidence of G4 formation should be interpreted with caution. A third caveat involves so-called high-occupancy target (HOT) regions, which correspond to regions of chromatin, typically the promoters of highly expressed genes, that are recovered in ChIP experiments against a large number of different individual proteins, and thus have also been referred to as "hyper-ChIPable" regions. HOT regions were recently found to overlap with loci ChIPed by BG4.[41,48] Importantly, HOT ChIP peaks are often artefactual, as can be demonstrated by performing a control ChIP experiment on samples in which the protein targeted by the antibody has been removed by genetic deletion of the gene encoding it; any remaining ChIP signal cannot be due to the binding of the intended target, and such residual binding is typical for HOT regions. As discussed above, this control experiment is difficult to carry out for G4 targets, raising some uncertainty about G4 ChIP findings. Because HOT regions, like BG4-chipped loci, tend to be nucleosome-depleted, it is possible that the positively charged, G4-binding surface of the antibody is prone to bind DNA exposed by nucleosome depletion. Regardless, these considerations emphasize the need for studies to demonstrate whether G4 ChIP efficiency can be diminished using chromatin isolated from *in vivo* settings where natural G4s are likely to have been unwound (*e.g.* after overexpression of a pan-G4 unwinding helicase), thus

providing evidence that the baseline level of G4s are neither induced by the antibody nor are they explained by HOT features.

As an aside, the scFv hf2 was used to immunoprecipitate purified human genomic DNA, and the recovered sequences were enriched for G4 forming potential, indicating this scFv preferentially binds G4s within the context of genomic DNA.[49] Although the experiment described in this report has been frequently referred to as "ChIP" in subsequent literature, it is important to consider the differences between naked genomic DNA and DNA within the context of crosslinked chromatin, particularly with respect to the potential of an added antibody to induce G4 folds.

2.5 Small molecules

A large number of small molecule G4 ligands have been designed or identified, and can be explored in the G-quadruplex Ligands Database (G4LDB) (http://www.g4ldb.org/ci2/index.php). A few of the ones used commonly in biology are discussed briefly here, including the porphyrins 5,10,15,20-tetra(*N*-methyl-4-pyridyl)porphyrin (TmPyP4) and *N*-methyl mesoporphyrin IX (NMM), the bisquinolinium PhenDC3, the *N*,*N*′-bis(quinolinyl) pyridine-2,6-dicarboxamide pyridostatin (PDS), and the pentacyclic acridine RHPS4.[50–55] More detail about small molecule G4 ligands can be found in the other chapters of this volume, and in several outstanding reviews (*e.g.*[56,57]). All of the compounds just mentioned are capable of binding G4s with equilibrium dissociation constants in the submicromolar range, and of stabilizing G4 folds, but they differ in some important respects. First, TmPyP4 has only a few-fold higher affinity for G4s *vs* duplex DNA due to its ability to bind the minor groove of the latter.[51,58] This contrasts with the porphyrin NMM, which unlike TmPyP4 and many other G4 small molecule ligands is anionic and thus generally repelled by nucleic acids, making it a highly selective G4 ligand; it appears to derive its ability to stack on the quartet face of G4 DNA from a slight aplanarity caused by its central methyl group.[59] Remarkably, NMM is also capable of selectively binding and inducing parallel G4 folds.[60] Phen-DC3, PDS, and RHPS4 are also highly selective for G4 *vs* duplex DNA, and RHPS4 also has the interesting property of accumulating preferentially in mitochondria when added to cells at low doses.[61] Some small molecule ligands can fluoresce when bound to G4s, *e.g.* NMM and N-TASQ[62–64] making them potential visual reporters of intracellular G4 formation. Also, it is worth mentioning that although the impact of small molecule G4 ligands on cell biology is typically interpreted as a reflection of G4

stabilization, it is important to also consider that displacement of natural ligands from G4s, *e.g.* proteins or endogenous porphyrins, could be involved.

Goals for the design of small molecule G4 ligands include the selective targeting of G4s *vs* other nucleic acids, and selective targeting of particular G4 structures, for example to repress expression of particular oncogenes. These are formidable aims. The steady state level of G4 DNA *in vivo* is not known, but it seems likely to be less that 1% of total DNA. Thus even a ligand that has 100-fold selectivity for G4 *vs* duplex DNA would at best be equally distributed between G4 and duplex targets in aggregate; of course the level of the ligand would be lower for any particular duplex DNA target than for any particular G4, but depending on how the ligand functions, this might still be problematic. Furthermore, designing compounds that bind selectively to particular G4 folds is also challenging, although great progress is being made as described in other chapters in this volume. Among the clever approaches being taken are the design of multifunctional molecules that bind to G4s having particular loop sequences and topologies, to several neighboring G4s at once, or to G4s and flanking duplex DNA in a sequence-specific fashion (reviewed in[57]).

From a biological perspective, it is interesting that porphyrins are excellent ligands for G4s,[65] raising questions of whether natural porphyrins might impact G4 biology. In fact, recent study raised new suggestions for how G4s might impact cellular oxidative stress, in particular by sequestering heme.[66] Heme is an ancient molecule found in most forms of life and is an essential cofactor for many enzymes but is toxic when free, due to reactive oxygen species generated by the iron it carries. The study provided evidence that small molecule G4 ligands, particularly PhenDC3 and AQ1, which are unrelated to heme other than sharing an affinity for G4s, can displace heme from cellular G4s in human cells. Such displacement was inferred by the strong upregulation in cells treated with the ligands of genes involved in heme degradation, particularly *HMOX1*, and in iron transport including *FTH1, FTL1,* and genes involved in vacuolar acidification. Of note, these findings are consistent with studies in yeast, which found enrichment for G4 motifs in the promoters and ORFs of genes involved in iron metabolism, and that genes required for vacuolar acidification are the most important for providing resistance to growth inhibition by the G4 ligand NMM.[10] Thus it is possible that DNA or RNA G4s can serve to protect against heme-induced oxidative damage overall, but might also sacrifice themselves to such damage.[66] Because heme may be a natural ligand for G4s, these findings raise the possibility that G4 biology might be impacted by genetic

defects in heme synthesis, which cause a group of diseases called porphyrias, which are characterized by skin, nervous system, and other defects.[67] Environmental exposure to lead might also affect G4s, given that this element inhibits heme synthesis and can also bind within G4s itself.[68] Furthermore, given that heme catalytic activity is fine-tuned by the different proteins in which it binds, it is interesting to speculate about the possibility that G4s might similarly influence heme catalytic activity *in vivo*, as they can for hemin *in vitro*.[69]

2.6 Microscopy

The colocalization as determined by light microscopy of antibody or small molecule staining with subcellular features expected to form G4s, *e.g.* nuclear foci and nucleoli, or even simply entire nuclei or mitochondria, is often taken as evidence for G4 formation *in vivo*. However, it is important to acknowledge the resolution limits of this technique, at roughly half the wavelength of the light used to visualize samples, *i.e.* no less than approximately 0.2 μm, which is about the diameter of a typical *E. coli* cell. Clearly there are molecules not actually associated with one another that lie within this distance of one another *in vivo*, and so it is important not to overinterpret such colocalization data. Super resolution microscopy techniques can enhance resolution approximately 10-fold, but this approach does not yet appear to have been taken in studies of subcellular G4 localization. Electron microscopy, in contrast, can provide more definitive evidence of molecular associations and has begun to be used for G4 studies.[70,71]

2.7 Genetic approaches

Genetic approaches are critical to understanding the functions of G-quadruplexes in biology. Genetics can provide an avenue for directed, hypothesis-driven questions, and some general thoughts as well as specific examples of this are discussed throughout the chapter. However, one of the most powerful types of genetic approach is the so-called "genetic screen." These provide information about which parts of the genome impact a phenotype of interest, *e.g.* sensitivity to a G4 stabilizing ligand. They do this by showing that this phenotype is altered when a region of the genome, potentially as small as a single base pair, is mutated from the norm. The power of genetic screens derives precisely from the fact that they

require no *a priori* knowledge of what these regions of the genome might be. In a well-designed screen, important and often unanticipated clues to the underlying biology will emerge.

The first genetic screen aimed at understanding G4 biology was conducted in the yeast *S. cerevisiae*.[10] It addressed which gene deletions cause yeast cells to be hypersensitive or resistant to growth inhibition by the small molecule G4 ligand NMM. As discussed above, NMM is a highly selective ligand for parallel G4s, including that formed by telomere repeats. Deletion of genes regulating telomere function, protein ubiquitination and chromatin structure were among those most strongly influencing NMM sensitivity, including those impacting the post-translational modification (PTM) of histone H3 on lysine 4 by methylation (of note, this histone PTM, H3K4me, has also been connected to G4 chromatin changes in vertebrates (see Section 3.1, below)). Recently, similar screens were reported in two cultured human cancer cell lines, which used genome-wide shRNA knockdowns of protein coding genes to identify those that normally guard against cell growth inhibition by the G4 ligands PDS and Phen-DC3.[10,72] Hypersensitivity to either or both compounds occurred preferentially when there were reductions in the activities of genes impacting gene expression, DNA replication, telomeres, mRNA splicing, ribosomes, and ubiquitin-mediated proteolysis, supporting roles for G4s in these processes. Because PDS and Phen-DC3 are structurally distinct compounds, and thus are likely to have different "off-target" (*i.e.* non-G4 related) effects, genes that modulate cell responses to both compounds are the most likely to be G4-specific, and interestingly, the screen yielded the highest level of enrichment for cancer-related genes when this subset was considered. Also of note, the four genes that are required for resistance to both compounds and in both cell lines, are *TOP1, BRCA1, DDX42* and *GAR1*, consistent with G4 roles in DNA replication and recombination, transcription, and translation. A caveat of this screen was that it was carried out in cancer cells that have various mutations that may perturb the impact of G4 ligands, *e.g.* they carry oncogenic mutations hyperactivating the RAS/BRAF/MAPK pathway, which is known to enhance replication fork collapse in telomeres.[73] Furthermore, to what degree the above screens provide information primarily about the impact of small molecules that stabilize G4s or rather about natural G4 biology *per se* is not yet clear. Screening for genes that modulate responses to genetic stabilization of G4s (*e.g.* deletion of G4 helicases) or to

G4 destabilization (by pharmacologic or genetic means) may be informative in this regard. Finally, it would be interesting to carry out parallel screens with structurally similar small molecules that have much less or no G4 stabilizing activity to further rule out potential G4-unrelated effects of the compounds, *e.g.* mesoporphyrin IX in place of NMM, or Bipy-DC3 in place of Phen-DC3.[54,65]

Genetic approaches can also enable hypothesis-directed investigation. For example, the Rap1 protein from the yeast *S. cerevisiae* was one of the first G4 binding proteins described.[74] Rap1 also binds in a sequence-specific fashion to G-rich duplex DNA, and it uses such binding to localize to telomeres and to the promoters of approximately 5% of all yeast genes, including those encoding ribosomal proteins.[75] At telomeres it functions to prevent end-resection by nucleases, regulate telomere lengthening by telomerase, and to recruit the Sir2/3/4 protein complex that forms heterochromatin and which can repress the transcription of nearby genes, whereas at most of its other genomic targets it cooperates with various other cofactor proteins to instead activate transcription. An obvious question thus arises: which of these functions require G4 binding, duplex DNA binding, or both? Simply deleting or overexpressing the gene encoding Rap1 cannot answer the question, because all Rap1 functions would be affected. However, structural studies have revealed that different faces of the Rap1 DNA binding domain engage each type of DNA target,[76] and thus it may be possible to test the *in vivo* consequences of mutant Rap1 proteins selectively deficient in G4 or duplex DNA binding (*i.e.* so-called "separation of function" mutations). Also of note, nucleosome displacement by Rap1 is coupled to its ability to activate transcription, and a small patch of amino acids in the Rap1 DNA binding domain required for direct binding of Rap1 to histone H3 and for nucleosome displacement was mapped recently.[77] Given the apparent connections between G4 DNA and nucleosome occupancy (see Section 3.1, below), it would be of interest to test whether G4-binding surfaces of Rap1 might also be involved in nucleosome displacement. The separation of function mutation approach could be of general use when investigating the functions of G4 interacting proteins, because it appears to be more often than not the case that such proteins have more than one cellular role. For example, G4 DNA-unwinding helicases in the RecQ family, such as human WRN and BLM and yeast Sgs1, appear to target G4s primarily *via* their so-called RQC domain (*e.g.* see Sections 3.2 and 3.4), but this domain is also important for the processing of other DNA structures.[78,79]

3. Biological processes impacted by G-quadruplexes

3.1 Chromatin

Genomic DNA is organized by a variety of proteins and RNAs into chromatin, the heart of which is the nucleosome, comprising two copies each of the core histone proteins H2A, H2B, H3 and H4 and the approximately 147 bp of DNA wrapped around them. The occupancy of nucleosomes within a region of DNA, as well as the organization of nucleosomes and post-translational modification (PTM) of their histones, can influence the recognition of DNA by the various factors that carry out its transcription, replication and repair. Euchromatin and heterochromatin refer to regions of relatively open and active chromatin *vs* more compact and silent chromatin, and whereas the transcriptional start sites of actively transcribed genes tend to be nucleosome-free, euchromatic and heterochromatic regions appear to have similar nucleosome occupancies overall.[80] Early evidence in yeast, *C. elegans*, and human cancer and helper T-cells indicated that gene promoters containing G4 motifs have reduced nucleosome occupancy, and subsequent studies have provided additional support for the idea that G-quadruplexes might generally tend to interfere with the wrapping of DNA into nucleosomes.[10,46,81–84] It is now evident that the interplay of G4s and chromatin is more complex and extends beyond nucleosome occupancy. For example, electron microscopy studies using the 1H6 antibody indicate that despite associations of G4s with active, and thus nucleosome free promoter regions, they might actually be more abundant within heterochromatin than euchromatin.[70] Moreover, there appear to be numerous interactions between G4s and the histone PTMs and chromatin remodeling enzymes that regulate nucleosome assembly, disassembly, and positioning, as first suggested by genetic studies in yeast.[10] For example, the chromatin remodeler and G4 binding protein ATRX deposits the H3.3 variant of histone H3 within telomere chromatin, and experimental reduction in H3.3 levels enhances telomere recombination (sister chromatid exchanges) in response to treatment with the G4 stabilizing ligand TmPyP4.[85,86] Plausibly connected to this H3.3 biology, during transcription of telomere DNA into TERRA RNA, ATRX suppresses the accumulation of telomere R-loops, which may enable G4s to form on the G-rich displaced telomeric DNA strand, and furthermore, ATRX deficiency itself confers greater sensitivity to the toxicity of the G4 small molecule ligand CX-3543.[87–89] Another example of G4-chromatin interactions involves

the LSD1-CoREST complex which can demethylate the lysine 4 and 9 residues on histone H3, and which may be recruited to telomeres by G4 RNA formed by TERRA and to the transcriptional promoters of the *CDKN1A* and *hTERT* loci by G4 DNA.[90–92] Furthermore, studies in chicken DT40 cells have revealed that G4 motifs can disturb the transmission of epigenetic information that normally maintains the repressed or active transcriptional state of genes through rounds of DNA replication by recycling H3K4 and H3K9 methylated histones from nucleosomes ahead of the replication fork to those reforming behind it.[93,94] That such chromatin changes are mediated by *bona fide* G4s is supported by the abilities of the WRN, BLM, REV1 and FANCJ G4 unwinding helicases to alleviate the impact of the G4 motifs, and by the capacity of the PDS-related small molecule PDC12 to stimulate this epigenetic reprogramming.[93,95]

Covalent modifications to DNA itself, particularly methylation of the 5′ position of cytosine (5mC), can also impact chromatin structure and function, including G4s. 5mC can stabilize G4 folds *via* enhancement of base pairing between cytosines, and the *de novo* human DNA methyltransferases DNMT3a and DNMT3b bind G4s avidly *in vitro*.[96–98] However genome-wide studies in normal human cells found lower levels of 5mC at G4 motifs,[99] suggesting G4s might generally limit 5mC modification of DNA. In addition, oxidative modification of G4s has been suggested to serve as an epigenetic modification.[100]

3.2 Transcription

Some of the earliest evidence for G4 biology came from studies of transcription, particularly those focused on G4 motifs in the promoter regions of oncogenes such as *MYC*.[101–103] These studies are discussed in detail in several of the other chapters of this volume, and have also been extensively reviewed (*e.g.*[104]). Overall, there is evidence that G4s, depending on context, can either enhance or inhibit transcription. Additional evidence for G4 involvement in transcription came from the observation that G4 motifs are over-represented in the upstream promoter regions of many organisms, and that treatment of yeast and human cancer cells with small molecule G4 ligands preferentially impacts the mRNA levels from genes with G4 motifs near their transcription start sites.[9,10,105] Similarly, selective modulation of transcript levels from genes containing G4 motifs has been observed in cells lacking the RecQ family G4 unwinding WRN, BLM or Sgs1 helicases.[10,29–32] Also in support, ChIP for the XPB and XPD helicases, which

bind, and bind and unwind, respectively, G4s *in vitro*, showed enrichment at G4 motifs especially near the transcriptional start sites of highly expressed genes.[106] For the WRN and BLM studies, there were several interesting differences in the findings reported, most notably that the first study showed that transcripts from genes with G4 motifs are upregulated by loss of WRN or BLM, whereas the later studies observed downregulation. This difference has not been explained, but it is interesting that the first study used cells grown in 3.5% oxygen, which is considered to be more physiological than the ambient oxygen levels used in the later studies.[107] Given the susceptibility of G4s to the oxidative modification 8-oxoG, which can impact G4 stability and conformation, and can interfere with transcription, along with the facts that there are defects in the repair of oxidatively damaged DNA in WRN and BLM mutant cells and that G4 oxidation has been linked to transcriptional regulation, it would be interesting to explore whether such G4 oxidation might contribute to the different findings.[108–112]

The exact mechanisms by which G4s regulate transcription are not well understood in most cases, but they can apparently affect RNA polymerase initiation, promoter-proximal pausing and elongation (reviewed by[113]). Some of the underlying mechanisms appear to involve recruitment of G4-binding transcription factors or exclusion of duplex binding transcription factors, regulation of histone PTMs, and nucleosome exclusion from the transcription start site and from transcription factor binding to regulatory sites, as discussed above. Furthermore, RNA polymerase progression may be impeded by G4s formed in front of the polymerase on the template strand, or by R-loops left behind by an advancing polymerase and stabilized by G4s on the non-template strand, thus interfering with subsequently initiated RNA polymerase. Conversely, G4s on the non-template strand might stimulate elongation by preventing re-annealing of the non-template with the template strand, as originally proposed for the ribosomal DNA which is generally highly transcribed and has numerous G4 motifs on the non-template strand.[87]

3.3 Messenger RNAs

The naturally single stranded nature of RNA, together with the fact that RNA G4s tend to have higher stability than DNA G4s *in vitro* has provided impetus to investigate the potential biological functions of RNA G4s. Molecular, biochemical and *in silico* approaches have identified numerous RNAs with G4 motifs that can form stable RNA G4s *in vitro*.[12,22,114–117]

However, the abundance in eukaryotic cells of single-stranded RNA binding proteins, *e.g.* the heterogeneous nuclear ribonucleoprotein proteins (hnRNPs) and many RNA helicases with G4-unwinding properties (reviewed in[118]), might prevent the accumulation of RNA G4s *in vivo*. Indeed, chemical probing of RNA within cultured human stem cells, mouse fibroblasts and yeast indicated that RNAs are almost entirely not in G4 conformations.[12] This held true even after ATP (required for helicase action generally) or the DHX36 G4-unwinding helicase had been depleted from cells. These findings are very important, although they still leave open the possibility that G4 RNA could form transiently or under special circumstances and thus impact biology. Furthermore they are compatible with the capacity of G4-stabilizing ligands to have significant impacts on biology.

Numerous studies have indicated potential roles for G4s in mRNA splicing, transport, stability and translation. For example, G4 motifs are associated with the use of alternative splice sites in the *TP53* and *BACE1* mRNAs, which can be modulated with the G4 ligand 360A or the hnRNP H protein, respectively.[119,120] The *PSD-95* and *CaMKIIa* mRNAs have G4 motifs in their 3' UTRs, which appears to mediate their intracellular localization to the neurites of neurons *via* binding to the Fragile-X mental retardation protein (FMRP).[121] Strongly supporting the functional importance of UTR G4 motifs is the recent demonstration that such motifs are under strong selective pressure and are enriched for RNA-binding protein interactions and for cis-eQTLs (expression quantitative trait loci), *i.e.* sequence variants that explain expression level differences of the mRNAs in which they reside.[122] With regard to translation, G4 motifs are often found in 5' UTRs, and such motifs in the *NRAS* 5' UTR are associated with inhibition of translation, which can be enhanced with G4 small molecule ligands.[123,124] A different type of G4-associated translational inhibition was revealed by transcriptome-wide ribosome profiling, which demonstrated that G4 motifs within 5' UTRs are associated with the translation of so-called upstream open reading frames (uORFs).[125] uORFs are short open reading frames upstream of the main ORFs within mRNAs, which inhibit translation of the main ORFs, and such inhibition was enhanced by deletion of the G4 unwinding helicases DHX9 and DHX36, supporting roles for G4s. G4 motifs are relatively underrepresented inside of ORFs, and when present tend to stall advancing ribosomes.[126–128] This connects with a newly discovered and active mechanism of canonical mRNA cleavage called ribothrypsis, which occurs in response to transient stalling of ribosome progression along the mRNA.[129] Ribothrypsis involves endonucleolytic

cleavage of the mRNA as it emerges from the ribosome channel, and is strongly associated with G4 motifs immediately ahead of the advancing ribosome.

3.4 Replication and genome stability

The unwinding of the DNA duplex and the generation of free DNA ends that are essential to replication puts the genome at risk for mutagenic events, and several lines of evidence implicate G4s in many aspects of these processes. With regard to replication, G4s appear to be involved in both initiation and the progression of replication forks. The replication of genomic DNA initiates at origins, which in metazoans are preferentially associated with OGREs (origin G-rich repeated elements) that lie approximately 250–300 bp 5′ from the initiation site. OGREs form G4s *in vitro*, apparently favoring parallel folds, and in murine cells were found to be required for full origin activity and to be sufficient to induce origin activity when inserted in an ectopic genomic location.[130] In support of G4 participation in OGRE function, treatment of cells with the PhenDC3 G4 ligand caused origins to form at novel sites containing G4 motifs, whereas it suppressed origin activity at sites lacking these motifs. How OGREs enhance origin firing is unclear, but they appear to function during the conversion of the pre-replication complex to the DNA synthesis elongation complex, which may relate to the apparent capacity of OGRE/G4s to exclude nucleosomes, facilitate the unwinding of duplex DNA, or to otherwise load replication factors.

In contrast to facilitating initiation, G4s may interfere with the progression of DNA replication and thus lead to genome instability. Early evidence came from *C. elegans* mutants lacking DOG-1, the homolog of the human FANCJ G4 DNA helicase. *dog-1* mutants accumulate 50–300 bp deletions beginning at the 3′ edges of G4 motifs comprising primarily uniform runs of Gs.[131,132] Remarkably the underlying G4 can apparently be inherited through cell division, together with un-replicated complementary strand across from the G4 (*i.e.* as a chromatid with a single stranded gap), and when the incompletely replicated strand is itself used as a replication template in the next cell cycle, this leads to double strand DNA breaks that are repaired by DNA polymerase theta-dependent end-joining to generate the deletion.[133,134] Similarly, human cells lacking *FANCJ*, which is a cause of Fanconi anemia, accumulate deletions associated with G4 motifs and are hypersensitive to the G4 small molecule ligand telomestatin.[135,136] Studies

in yeast lacking the Pif1 helicase also indicate that G4s can drive genome instability during replication, particularly those with short loops and thus high stability, and also indicated that the effect of *PIF1* deficiency can be mimicked by PhenDC3, which inhibits G4 unwinding by Pif1 *in vitro*.[137–139] Human cells lacking the BLM or WRN RecQ family G4 DNA helicases are characterized by genome instability, presumably explaining elevated cancer risk in these diseases, but remarkably, the elevated sister chromatid exchanges in BLM mutant cells is most pronounced at transcribed genes with abundant G4 motifs.[140] Therefore these RecQ family helicases might be more important in processing G4s during transcription than replication *per se*, which is consistent with G4-related transcriptional but not replication-associated defects in yeast lacking the RecQ family Sgs1 helicase.[10,141] It is conceivable that G4s formed on the non-template strand in the absence of these helicases enhances transcription-associated R-loop formation and persistence (or, alternatively, formation of hybrid RNA/DNA G4s, in cases where partial G4 motifs exist on the transcribed strand), which interferes with replication fork progression. This notion is consistent with the hypersensitivity of cells lacking topoisomerase I (TOP1) to G4 small molecule ligands (see Section 2.7, above), because TOP1 helps rewind untwisted DNA and thus dislodge RNA from R-loops.[142] It is also consistent with the recent demonstrations that R-loops underlie genome instability in human cancer cells treated with G4 stabilizing small molecules, as well as with the replication defects in cells lacking the FANCJ-related and G4 unwinding helicase RTEL1.[143,144] More broadly, these ideas also suggest that an underlying reason for organisms to possess so many different types of G4-unwinding helicases is that each may be tailored to processing distinct G4 folds or G4s in different molecular contexts.

G4s may also play important roles in programmed recombination events. For example, class switch recombination enables B cells to change the isotype of the antibody they produce, *e.g.* from IgM to IgG or IgA, and involves the generation of double strand breaks (DSBs) at G-rich switch regions, which are subsequently ligated to new sequences. Several studies together suggest that non-template strand G4s and R-loops which are formed during transcription of the switch regions (which together have been termed "G-loops"), and which are encouraged to form by the G4 helicase DDX1, provide a substrate for recognition by activation-induced deaminase (AID), an enzyme of the Apobec family. AID converts cytosines to uracils, leading to strand nicking by the base excision repair machinery, and such

nicks, when closely apposed and on opposite DNA strands, generate the DSBs required for class switch recombination.[87,145,146]

Another form of genome instability is that induced by transposons, which can move from one genomic site into another. G4s within the 3' UTRs of the human LINE1 family of transposons, which compose approximately 17% of the human genome, appear to enhance their transposition.[147] Similarly, the so-called "internal eliminated sequences" which are intentionally removed from the somatic genome of *Tetrahymena thermophila*, may do so *via* cooperation between G4s, the G4 binding protein Lia3 and the transposase Tpb2.[148] Interestingly, transposon activation occurs in response to various biological stresses, and there is recent evidence for broader involvement of G4s in cellular stress responses, including oxidative stress.[149] Furthermore, there is emerging evidence that G4s may be at elevated susceptibility to oxidative damage, and that 8-oxoG lesions may stabilize or destabilize G4s, depending on the particular type of G4, as recently reviewed.[150,151]

3.5 Telomeres

Telomeres are the chromatin structures at the termini of linear chromosomes, and are essential for normal chromosome stability (reviewed in[152]). They are distinguished from generic DNA ends, which are recognized by the DNA damage response (DDR) machinery. This distinction is mediated primarily by the shelterin protein complex, and can also involve the formation of "t-loops" wherein the single stranded 3′ overhang typically present at the telomere end forms a D-loop by invading internal telomere repeat duplex DNA. Telomeres are thus "capped" and do not activate the DDR. Because of an inability of the standard replication machinery to fully replicate DNA ends (the "end replication problem") and moreover the natural processing of telomeres to generate the 3′ overhangs that are essential to their capping, telomeres shorten with cell division until they reach the point where they activate the DDR unless maintained by lengthening mechanisms, such as that provided by the enzyme telomerase. In humans, telomerase is expressed at limiting levels in most normal cells, and this limitation appears to inhibit uncontrolled cell division that can lead to cancer; as such, telomerase is generally upregulated in the rare cells that form cancers, enabling their growth (see Section 3.7 below). The telomeres of most eukaryotes comprise DNA sequences with high G4 forming potential given that they contain repeated runs of Gs, and especially because these

G runs are on strand composing the 3′ overhang. There has thus been interest in exploring potential roles for G4s in telomere function (as noted above for *Stylonychia*), and indeed there is strong evidence that they impact telomere replication, capping, and length maintenance. Vertebrate telomeres comprise repeats of the sequence TTAGGG, which *in vitro* has the interesting property being able to readily interconvert between parallel, antiparallel or mixed hybrid G4 folds depending on experimental conditions,[153,154] which *in vivo* may enable opportunities for regulation based on fold-specific interactions, for example by interacting differentially with particular proteins. There is also emerging evidence for interactions among individual telomeric G4s *in vitro* that may reflect ways in which G4 impact the overall architecture of telomere chromatin (reviewed in[155]).

Telomeres are considered "fragile sites," in that they appear to be preferentially susceptible to breakage under conditions of replication stress, *e.g.* limiting dNTPs, inhibition of DNA polymerase, or oncogene-driven perturbation of normal replication.[73,156] Roles for the G4 unwinding helicases BLM, WRN, DNA2, RTEL1 in suppressing telomere fragility indicate that G4s may be contributory, which is also consistent with enhanced BG4 staining of replicating telomeres in Wrn and Blm-deficient mouse cells.[43,156–161] Similarly, evidence for replication-associated telomere fragility in cells treated with G4 small molecule ligands, such as RHPS4 and 360A, support this idea, as does the enhanced susceptibility of cells lacking the above helicases to small molecule G4 ligands.[158,162–165] Interestingly, using the STELA technique which can sensitively detect ultrashort, but rare, telomeres, it was recently demonstrated that the 360A G4 ligand, and particularly its dimeric form $(360A)_{2A}$,can induce the appearance of telomeres much shorter than the overall mean telomere length in the cells, consistent with telomere breakage during replication.[166] The CTC1-STN1-TEN1 (CST) complex, an RPA-like complex, can bind and melt G4s *in vitro* and apparently *in vivo*, and genetic depletion of CST also causes telomere fragility, which can be exacerbated by TmPyP4.[167,168] The TERRA RNA which is generated by transcription of the C-rich telomere strand, forms G4s; and it may also participate in G4 formation in telomere DNA by forming R-loops, thus also contributing to telomere replication difficulties, although G4 involvement in such TERRA related processes is not yet certain.[88,169–172] Furthermore, telomeric R-loops can contribute to the repair of oxidative telomere damage, and may thus not always be deleterious.[173]

Interference by G4s with telomere replication raises the question of why telomeres have not generally evolved away from such problematic sequences. This is an open and broad question, but G4s might provide clues. For example, genetic and biochemical experiments in yeast support the possibility that G4s can help cap telomeres, at least under conditions where the normal protein-based capping mechanisms have been perturbed.[47,174,175] In particular, when the function of the Cdc13 protein, which normally protects the telomere by binding the 3′ overhang, is diminished, G4 stabilization with G4 binding proteins or small molecule ligands, or by removal of G4 unwinding helicases, provided weak capping activity, associated with inhibition of both DDR responses and exonucleolytic degradation of the telomere end. Thus perhaps G4s had an ancestral role in telomere capping, which has been subsumed and improved upon by proteins that still recognize the underlying G-rich substrate. Whether G4s can ever contribute to natural telomere capping is unknown, but the capacity of G4 small molecule ligands to uncap telomeres is clear. This does not necessarily indicate that all G4s naturally uncap telomeres, as the ligands might favor particular G4 folds that interfere with capping or they may displace by steric hindrance the binding of natural capping factors that would not be displaced by the G4s themselves. Although the underlying mechanisms are not fully understood, there is indeed evidence that G4-binding ligands including BRACO-19, tetrasubstituted naphthalene diimides, and RHPS4 can displace shelterin proteins including TRF2 and POT1, which normally bind the duplex and 3′ overhang, respectively, and can also displace telomerase itself, which was recently found to have capping activity independent of its telomere lengthening role.[176–180]

Telomeres can be lengthened by telomerase, a reverse transcriptase comprising at its core the TERT catalytic subunit and a dedicated RNA template TERC, or can be lengthened by ALT (alternative lengthening of telomeres), which depends on homologous recombination (reviewed in[181]). In humans, telomerase is active primarily in dividing stem cells and cancer cells, whereas ALT is mainly restricted to cancer cells. Substantial evidence indicates that both mechanisms are impacted by G4s. For example, early *in vitro* studies indicated that telomeres in G4 conformations are poor substrates for telomerase, and *in vitro* and *in vivo* studies demonstrate that small molecule G4 ligands inhibit telomere lengthening by telomerase, whereas POT1 can reverse such inhibition apparently by unfolding G4s.[182,183] More recently, evidence has emerged that oxidation of telomere DNA to form 8-oxoG can

stimulate telomerase-mediated telomere lengthening by inhibiting G4 folds.[184] However, it must be mentioned that not all G4s may inhibit telomerase. Indeed, remarkable *in vitro* and *in vivo* studies have shown that both ciliate and human telomerase can apparently extend telomeres folded into bimolecular parallel G4s.[185,186] Potential G4 roles are not yet well understood in the case of ALT, but G4s may contribute to telomere recombination by causing replication-associated telomere breaks, and yet may also interfere with subsequent recombination events between telomere sequences.[89,187,188]

An additional type of connection between telomere biology and G4s concerns the role of telomere binding proteins in the regulation of gene expression at non-telomeric sites throughout the genome. For example, the localization of the human TRF2 protein to genes such as *CDKN1A* (which encodes the p21 CDK inhibitor) occurs in response to telomere shortening, and in this case binds to a promoter G4 in a fashion dependent on the G4-binding TRF2 basic domain, thus enabling the recruitment the repressive REST-coREST-LSD1 complex.[91,189] Perhaps similarly, when telomeres become critically shortened in yeast lacking telomerase, the G4-binding Rap1 protein re-localizes from telomeres to hundreds of new gene promoters in a DDR-dependent fashion, where it activates gene expression, although specific G4 connections in this case have not yet been reported.[190]

3.6 Mitochondria

Evidence is emerging that G4s play important roles in mitochondrial biology, including in replication, genome stability, and RNA transactions. This has been reviewed recently,[191] and only a few highlights will be mentioned here. Human mitochondrial DNA (mtDNA) has substantial asymmetry in base composition between its complementary strands, with twofold enrichment for Gs on the so-called heavy strand, yielding approximately threefold higher G4 motif density than the nuclear genome; similarly, the mtDNA of *S. cerevisiae* is 10-fold higher than that of its nuclear genome, despite its high AT content.[21,192] There are significant correlations between the location of G4 motifs and naturally occurring mtDNA deletion breakpoints, suggesting G4s may contribute to mtDNA instability.[193–195] Consistent with this, *in vitro* replication assays have demonstrated stalling of DNA synthesis at mtDNA regions associated with genetic variation, and that such pausing can be alleviated by the G4 helicase Pif1, which

has mitochondrial as well as nuclear forms.[196,197] With regard to mitochondrial G4 RNAs, the GRSF1 protein appears to bind and melt them, enabling their degradation by the degradosome and thus preventing their overaccumulation.[198] The G4 ligand BMVC localizes to the mitochondria preferentially in cancer cells, and its fluorescence properties when bound to G4s provided evidence of mitochondrial G4s in live cells.[199] BMVC inhibits transcription of mtDNA,[199] as does the G4 ligand RHPS4, which preferentially localizes mitochondria at low doses and also causes mtDNA levels to decline.[61] Finally, given the fact that disease variants of mtDNA are often present in mixed populations within cells together with wild type mtDNA (called heteroplasmy), a very clever approach was taken to use the higher G4-forming capacity of a particular disease variant (m.10191C), which is the cause of the neurological disorder Leigh syndrome (LS), to demonstrate a potential approach to therapy. Treatment of LS cells in culture with the G4 ligand berberine, which *in vitro* impedes mitochondrial DNA polymerase gamma progression, caused loss of the disease mtDNA variant, allowing for wild type mtDNA to accumulate and for mitochondrial respiration to return to healthy levels.[200]

3.7 Cancer

Cancer is fundamentally a disease driven by genome instability, and so it is reasonable to hypothesize that the genome-destabilizing roles that can be played by G4s, as described above, contributes to carcinogenesis. The association of DNA breakpoints in cancer cells with G4 motifs supports this notion.[201] Furthermore, studies with the BG4 antibody suggests that G4 levels may be higher in some types of cancer cells.[46,202] Although more evidence is needed, it would not be surprising if G4 levels were indeed higher in cancer cells given their generally unstable genomes which would tend to generate ssDNA regions that could form G4s, and because they tend to have generally more nucleosome free chromatin, and higher levels of transcription and R-loops.[203,204] Regardless, there is already substantial evidence that targeting G4s with small molecule ligands can selectively interfere with the growth and survival of cancer cells in culture and in animal experiments, as reviewed recently.[205]

Three main types of approaches are being taken to use G4 small molecule ligands to attack cancer. The first is to design ligands that selectively downregulate the expression of oncogenes such as *MYC* and *KRAS* by binding to G4s in their promoter regions. A second approach is to use G4 ligands to

interfere with telomere maintenance by telomerase or ALT, as discussed above briefly. Both of these approaches are the focus of other chapters in this volume. It is worth mentioning that telomerase inhibitors have so far not proved to be of major benefit in treating most cancers, although the fact that oncogenes tend to drive telomere fragility (see above) may make cancer cells particularly susceptible to replication associated damage by G4 ligands, in a fashion that is independent of telomerase inhibition.[73] Finally, a third approach, which appears promising in part because it is fundamentally easier to accomplish than the first two, is to use the capacity of G4 ligands to induce DNA damage to selectively inhibit cancer cell types that have deficiencies in DNA repair mechanisms. For example, people who carry mutations in *BRCA1*, and are thus at elevated risk for breast, ovarian and other cancers, typically inherit only one mutant allele, and it is the loss of the normal allele in a rare cell that puts that cell and its progeny at risk for becoming cancerous. This is because the BRCA1 protein is important for the repair of DNA, primarily by homologous recombination, and so a cell that has no functional copies of *BRCA1* is at increased risk for the accumulation of additional mutations that can drive carcinogenesis. Fortunately, this also means such cancer cells will be more susceptible to DNA damaging agents than are the normal cells that an individual needs to survive. This is the logic behind using inhibitors of polyADP ribose polymerase (PARP), *e.g.* olaparib, to treat cancer patients with inherited *BRCA1* mutations, because PARP or BRCA1 provide alternative pathways for the repair of DNA damage in normal cells, but only the PARP-dependent pathway is available in the cancer cells.[206] Because G4 ligands can induce DNA damage, as described above, they can selectively inhibit the growth and survival of cancer cells lacking pathways needed for repair of the damage. Indeed *BRCA1*, *BRCA2, or RAD51* mutant cells are hypersensitive to PDS, which correlates with elevated levels of telomere fragility, and holds true even for those that are resistant to olpaparib.[207] Similarly, cells deficient in BRCA1, BRCA2, or the non-homologous end joining factors LIG4 and DNA-PK, are hypersensitive to the G4 ligand CX-5461 (which is in clinical trials), and BRCA2 deficient tumor cells are sensitive to a combined G4 ligand/PARP inhibitor.[208,209] ALT tumors often lack ATRX, and because ATRX-deficient cells are hypersensitive to G4 ligands such tumors would be expected to be thus targeted, and recent evidence with malignant glioma cells supports this.[89,210,211] Finally, the genetic screen for gene knockdowns rendering human tumor cells sensitive to PDS or PhenDC3 mentioned above identified 50 genes known to be mutated in human cancers, including

PALB2, *BAP1*, *SMARCA4*, and *POLQ*, which encodes DNA polymerase theta mentioned above to be involved in the repair of G4 induced DSBs.[72] Thus cancers with these mutations may be susceptible to preferential inhibition by G4 ligands.

4. Closing remarks

The understanding of G4 biology is still at an early stage, but it is already evident that the area has an exciting future. Important fields impacted by G4s and not covered here, but which are the focus of other chapters in this volume include neurodegeneration and viral biology (this chapter was written during the SARS-CoV-2 pandemic, and it is thus noteworthy that coronaviral RNA replication depends on the DDX1 helicase mentioned above, suggesting potential inhibitory roles for G4s in the viral lifecycle.[212]) However, overall it remains challenging to know with certainty whether indications of G4 involvement in a biological process are fully attributable to *bona fide* G4s, and so experimentalists must continue to ask themselves "*G-Quad erat demonstrandum*?". Regardless, the capacity of G4 small molecule ligands to preferentially impact the functions of sequences with G4 forming potential is clear, and as such ligands with improved specificity and potency are developed they will provide additional tools to investigate the basic biological mechanisms of G4s, as well as provide important new therapeutic approaches to diseases.

Acknowledgments

The author is indebted to his many colleagues who have shared their interest, efforts, and ideas about G4 biology, and especially thanks Li-San Wang, Liliya Yatsunyk, Brett Kaufman, Peter Lansdorp, Daniella Rhodes, Jay Johnson, Zissimos Mourelatos, and Eric Brown for their critical insights and help. This work was supported by NIH grants R01GM110424, R01GM133154 and R01HL148821.

References

1. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Res.* **2006**, *34*(19), 5402–5415.
2. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The Regulation and Functions of DNA and RNA G-Quadruplexes. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*(8), 459–474. https://doi.org/10.1038/s41580-020-0236-x.
3. Eddy, J.; Maizels, N. Gene Function Correlates with Potential for G4 DNA Formation in the Human Genome. *Nucleic Acids Res.* **2006**, *34*(14), 3887–3896.
4. Johnson, J. E.; Smith, J. S.; Kozak, M. L.; Johnson, F. B. In Vivo Veritas: Using Yeast to Probe the Biological Functions of G-Quadruplexes. *Biochimie* **2008**, *90*(8), 1250–1263.

5. Lipps, H. J.; Rhodes, D. G-Quadruplex Structures: In Vivo Evidence and Function. *Trends Cell Biol.* **2009**, *19*(8), 414–422.
6. Maizels, N.; Gray, L. T. The G4 Genome. *PLoS Genet.* **2013**, *9*(4), e1003468.
7. Todd, A. K.; Johnston, M.; Neidle, S. Highly Prevalent Putative Quadruplex Sequence Motifs in Human DNA. *Nucleic Acids Res.* **2005**, *33*(9), 2901–2907.
8. Huppert, J. L.; Balasubramanian, S. Prevalence of Quadruplexes in the Human Genome. *Nucleic Acids Res.* **2005**, *33*(9), 2908–2916.
9. Verma, A.; Halder, K.; Halder, R.; Yadav, V. K.; Rawal, P.; Thakur, R. K.; Mohd, F.; Sharma, A.; Chowdhury, S. Genome-Wide Computational and Expression Analyses Reveal G-Quadruplex DNA Motifs as Conserved Cis-Regulatory Elements in Human and Related Species. *J. Med. Chem.* **2008**, *51*(18), 5641–5649.
10. Hershman, S. G.; Chen, Q.; Lee, J. Y.; Kozak, M. L.; Yue, P.; Wang, L.-S.; Johnson, F. B. Genomic Distribution and Functional Analyses of Potential G-Quadruplex-Forming Sequences in Saccharomyces cerevisiae. *Nucleic Acids Res.* **2008**, *36*(1), 144–156.
11. Maizels, N. G4-Associated Human Diseases. *EMBO Rep.* **2015**, *16*(8), 910–922.
12. Guo, J. U.; Bartel, D. P. RNA G-Quadruplexes Are Globally Unfolded in Eukaryotic Cells and Depleted in Bacteria. *Science* **2016**, *353*, aaf5371.
13. Puig Lombardi, E.; Holmes, A.; Verga, D.; Teulade-Fichou, M.-P.; Nicolas, A.; Londoño-Vallejo, A. Thermodynamically Stable and Genetically Unstable G-Quadruplexes Are Depleted in Genomes across Species. *Nucleic Acids Res.* **2019**, *47*(12), 6098–6113.
14. Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. *Nat. Biotechnol.* **2015**, *33*(8), 877–881.
15. Wiśniewski, J. R.; Hein, M. Y.; Cox, J.; Mann, M. A "Proteomic Ruler" for Protein Copy Number and Concentration Estimation Without Spike-in Standards. *Mol. Cell. Proteomics* **2014**, *13*(12), 3497–3506.
16. Puig Lombardi, E.; Londoño-Vallejo, A. A Guide to Computational Methods for G-Quadruplex Prediction. *Nucleic Acids Res.* **2020**, *48*(1), 1–15.
17. Cao, K.; Ryvkin, P.; Johnson, F. B. Computational Detection and Analysis of Sequences with Duplex-Derived Interstrand G-Quadruplex Forming Potential. *Methods* **2012**, *57*(1), 3–10.
18. Kudlicki, A. S. G-Quadruplexes Involving Both Strands of Genomic DNA Are Highly Abundant and Colocalize with Functional Sites in the Human Genome. *PLoS One* **2016**, *11*(1), e0146174.
19. Guédin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. How Long Is Too Long? Effects of Loop Size on G-Quadruplex Stability. *Nucleic Acids Res.* **2010**, *38*(21), 7858–7868.
20. Bugaut, A.; Balasubramanian, S. A Sequence-Independent Study of the Influence of Short Loop Lengths on the Stability and Topology of Intramolecular DNA G-Quadruplexes. *Biochemistry* **2008**, *47*(2), 689–697.
21. Bedrat, A.; Lacroix, L.; Mergny, J.-L. Re-Evaluation of G-Quadruplex Propensity with G4Hunter. *Nucleic Acids Res.* **2016**, *44*(4), 1746–1759.
22. Garant, J.-M.; Perreault, J.-P.; Scott, M. S. Motif Independent Identification of Potential RNA G-Quadruplexes by G4RNA Screener. *Bioinformatics* **2017**, *33*(22), 3532–3537.
23. Kwok, C. K.; Marsico, G.; Sahakyan, A. B.; Chambers, V. S.; Balasubramanian, S. rG4-Seq Reveals Widespread Formation of G-Quadruplex Structures in the Human Transcriptome. *Nat. Methods* **2016**, *13*(10), 841–844.
24. Sahakyan, A. B.; Chambers, V. S.; Marsico, G.; Santner, T.; Di Antonio, M.; Balasubramanian, S. Machine Learning Model for Sequence-Driven DNA G-Quadruplex Formation. *Sci. Rep.* **2017**, *7*(1), 14535.

25. Marsico, G.; Chambers, V. S.; Sahakyan, A. B.; McCauley, P.; Boutell, J. M.; Antonio, M. D.; Balasubramanian, S. Whole Genome Experimental Maps of DNA G-Quadruplexes in Multiple Species. *Nucleic Acids Res.* **2019**, *47*(8), 3862–3874.
26. Ellis, R. J.; Minton, A. P. Cell Biology: Join the Crowd. *Nature* **2003**, *425*(6953), 27–28.
27. Zhou, J.; Tateishi-Karimata, H.; Mergny, J.-L.; Cheng, M.; Feng, Z.; Miyoshi, D.; Sugimoto, N.; Li, C. Reevaluation of the Stability of G-Quadruplex Structures Under Crowding Conditions. *Biochimie* **2016**, *121*, 204–208.
28. Hall, A. C.; Ostrowski, L. A.; Mekhail, K. Phase Separation as a Melting Pot for DNA Repeats. *Trends Genet.* **2019**, *35*(8), 589–600.
29. Johnson, J. E.; Cao, K.; Ryvkin, P.; Wang, L.-S.; Johnson, F. B. Altered Gene Expression in the Werner and Bloom Syndromes Is Associated with Sequences Having G-Quadruplex Forming Potential. *Nucleic Acids Res.* **2010**, *38*(4), 1114–1122.
30. Nguyen, G. H.; Tang, W.; Robles, A. I.; Beyer, R. P.; Gray, L. T.; Welsh, J. A.; Schetter, A. J.; Kumamoto, K.; Wang, X. W.; Hickson, I. D.; Maizels, N.; Monnat, R. J., Jr.; Harris, C. C. Regulation of Gene Expression by the BLM Helicase Correlates with the Presence of G-Quadruplex DNA Motifs. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*(27), 9905–9910.
31. Smestad, J. A.; Maher, L. J., 3rd Relationships Between Putative G-Quadruplex-Forming Sequences, RecQ Helicases, and Transcription. *BMC Med. Genet.* **2015**, *16*, 91.
32. Tang, W.; Robles, A. I.; Beyer, R. P.; Gray, L. T.; Nguyen, G. H.; Oshima, J.; Maizels, N.; Harris, C. C.; Monnat, R. J., Jr. The Werner Syndrome RECQ Helicase Targets G4 DNA in Human Cells to Modulate Transcription. *Hum. Mol. Genet.* **2016**, *25*(10), 2060–2069.
33. Ryvkin, P.; Hershman, S. G.; Wang, L.-S.; Johnson, F. B. Computational Approaches to the Detection and Analysis of Sequences with Intramolecular G-Quadruplex Forming Potential. *Methods Mol. Biol.* **2010**, *608*, 39–50.
34. Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Plückthun, A. In Vitro Generated Antibodies Specific for Telomeric Guanine-Quadruplex DNA React with Stylonychia lemnae Macronuclei. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*(15), 8572–8577.
35. Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. Telomere End-Binding Proteins Control the Formation of G-Quadruplex DNA Structures in Vivo. *Nat. Struct. Mol. Biol.* **2005**, *12*(10), 847–854.
36. Fang, G.; Cech, T. R. The Beta Subunit of Oxytricha Telomere-Binding Protein Promotes G-Quartet Formation by Telomeric DNA. *Cell* **1993**, *74*(5), 875–885.
37. Fernando, H.; Rodriguez, R.; Balasubramanian, S. Selective Recognition of a DNA G-Quadruplex by an Engineered Antibody. *Biochemistry* **2008**, *47*(36), 9365–9371.
38. Fernando, H.; Sewitz, S.; Darot, J.; Tavaré, S.; Huppert, J. L.; Balasubramanian, S. Genome-Wide Analysis of a G-Quadruplex-Specific Single-Chain Antibody That Regulates Gene Expression. *Nucleic Acids Res.* **2009**, *37*(20), 6716–6722.
39. Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. *Nat. Chem.* **2013**, *5*(3), 182–186.
40. Biffi, G.; Di Antonio, M.; Tannahill, D.; Balasubramanian, S. Visualization and Selective Chemical Targeting of RNA G-Quadruplex Structures in the Cytoplasm of Human Cells. *Nat. Chem.* **2014**, *6*(1), 75–80.
41. Liu, H.-Y.; Zhao, Q.; Zhang, T.-P.; Wu, Y.; Xiong, Y.-X.; Wang, S.-K.; Ge, Y.-L.; He, J.-H.; Lv, P.; Ou, T.-M.; Tan, J.-H.; Li, D.; Gu, L.-Q.; Ren, J.; Zhao, Y.; Huang, Z.-S. Conformation Selective Antibody Enables Genome Profiling and Leads to Discovery of Parallel G-Quadruplex in Human Telomeres. *Cell Chem. Biol.* **2016**, *23*(10), 1261–1270.

42. Henderson, A.; Wu, Y.; Huang, Y. C.; Chavez, E. A.; Platt, J.; Johnson, F. B.; Brosh, R. M., Jr.; Sen, D.; Lansdorp, P. M. Detection of G-Quadruplex DNA in Mammalian Cells. *Nucleic Acids Res.* **2014**, *42*(2), 860–869.
43. Drosopoulos, W. C.; Kosiyatrakul, S. T.; Schildkraut, C. L. BLM Helicase Facilitates Telomere Replication During Leading Strand Synthesis of Telomeres. *J. Cell Biol.* **2015**, *210*(2), 191–208.
44. Kazemier, H. G.; Paeschke, K.; Lansdorp, P. M. Guanine Quadruplex Monoclonal Antibody 1H6 Cross-Reacts with Restrained Thymidine-Rich Single Stranded DNA. *Nucleic Acids Res.* **2017**, *45*(10), 5913–5919.
45. Ray, S.; Tillo, D.; Boer, R. E.; Assad, N.; Barshai, M.; Wu, G.; Orenstein, Y.; Yang, D.; Schneekloth, J. S., Jr.; Vinson, C. Custom DNA Microarrays Reveal Diverse Binding Preferences of Proteins and Small Molecules to Thousands of G-Quadruplexes. *ACS Chem. Biol.* **2020**, *15*(4), 925–935.
46. Hänsel-Hertsch, R.; Beraldi, D.; Lensing, S. V.; Marsico, G.; Zyner, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M.; Tannahill, D.; Balasubramanian, S. G-Quadruplex Structures Mark Human Regulatory Chromatin. *Nat. Genet.* **2016**, *48*(10), 1267–1272.
47. Smith, J. S.; Chen, Q.; Yatsunyk, L. A.; Nicoludis, J. M.; Garcia, M. S.; Kranaster, R.; Balasubramanian, S.; Monchaud, D.; Teulade-Fichou, M.-P.; Abramowitz, L.; Schultz, D. C.; Johnson, F. B. Rudimentary G-Quadruplex-Based Telomere Capping in Saccharomyces cerevisiae. *Nat. Struct. Mol. Biol.* **2011**, *18*(4), 478–485.
48. Wreczycka, K.; Franke, V.; Uyar, B.; Wurmus, R.; Bulut, S.; Tursun, B.; Akalin, A. HOT or Not: Examining the Basis of High-Occupancy Target Regions. *Nucleic Acids Res.* **2019**, *47*(11), 5735–5745.
49. Lam, E. Y. N.; Beraldi, D.; Tannahill, D.; Balasubramanian, S. G-Quadruplex Structures Are Stable and Detectable in Human Genomic DNA. *Nat. Commun.* **2013**, *4*, 1796.
50. Anantha, N. V.; Azam, M.; Sheardy, R. D. Porphyrin Binding to Quadrupled T4G4. *Biochemistry* **1998**, *37*(9), 2709–2714.
51. Ren, J.; Chaires, J. B. Sequence and Structural Selectivity of Nucleic Acid Binding Ligands. *Biochemistry* **1999**, *38*(49), 16067–16075.
52. Gowan, S. M.; Heald, R.; Stevens, M. F.; Kelland, L. R. Potent Inhibition of Telomerase by Small-Molecule Pentacyclic Acridines Capable of Interacting with G-Quadruplexes. *Mol. Pharmacol.* **2001**, *60*(5), 981–988.
53. Parkinson, G. N.; Ghosh, R.; Neidle, S. Structural Basis for Binding of Porphyrin to Human Telomeres. *Biochemistry* **2007**, *46*(9), 2390–2397.
54. Cian, A. D.; De Cian, A.; DeLemos, E.; Mergny, J.-L.; Teulade-Fichou, M.-P.; Monchaud, D. Highly Efficient G-Quadruplex Recognition by Bisquinolinium Compounds. *J. Am. Chem. Soc.* **2007**, *129*(7), 1856–1857.
55. Rodriguez, R.; Müller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres. *J. Am. Chem. Soc.* **2008**, *130*(47), 15758–15759.
56. Monchaud, D.; Teulade-Fichou, M.-P. A Hitchhiker's Guide to G-Quadruplex Ligands. *Org. Biomol. Chem.* **2008**, *6*, 627–636.
57. Asamitsu, S.; Obata, S.; Yu, Z.; Bando, T.; Sugiyama, H. Recent Progress of Targeted G-Quadruplex-Preferred Ligands Toward Cancer Therapy. *Molecules* **2019**, *24*, 429.
58. Bennett, M.; Krah, A.; Wien, F.; Garman, E.; Mckenna, R.; Sanderson, M.; Neidle, S. A DNA-Porphyrin Minor-Groove Complex at Atomic Resolution: The Structural Consequences of Porphyrin Ruffling. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 9476–9481.

59. Nicoludis, J. M.; Miller, S. T.; Jeffrey, P. D.; Barrett, S. P.; Rablen, P. R.; Lawton, T. J.; Yatsunyk, L. A. Optimized End-Stacking Provides Specificity of N-Methyl Mesoporphyrin IX for Human Telomeric G-Quadruplex DNA. *J. Am. Chem. Soc.* **2012**, *134*(50), 20446–20456.
60. Nicoludis, J. M.; Barrett, S. P.; Mergny, J.-L.; Yatsunyk, L. A. Interaction of Human Telomeric DNA with N-Methyl Mesoporphyrin IX. *Nucleic Acids Res.* **2012**, *40*(12), 5432–5447.
61. Falabella, M.; Kolesar, J. E.; Wallace, C.; de Jesus, D.; Sun, L.; Taguchi, Y. V.; Wang, C.; Wang, T.; Xiang, I. M.; Alder, J. K.; Maheshan, R.; Horne, W.; Turek-Herman, J.; Pagano, P. J.; St Croix, C. M.; Sondheimer, N.; Yatsunyk, L. A.; Johnson, F. B.; Kaufman, B. A. G-Quadruplex Dynamics Contribute to Regulation of Mitochondrial Gene Expression. *Sci. Rep.* **2019**, *9*(1), 5605.
62. Arthanari, H.; Basu, S.; Kawano, T. L.; Bolton, P. H. Fluorescent Dyes Specific for Quadruplex DNA. *Nucleic Acids Res.* **1998**, *26*(16), 3724–3728.
63. Sabharwal, N. C.; Savikhin, V.; Turek-Herman, J. R.; Nicoludis, J. M.; Szalai, V. A.; Yatsunyk, L. A. N-Methylmesoporphyrin IX Fluorescence as a Reporter of Strand Orientation in Guanine Quadruplexes. *FEBS J.* **2014**, *281*(7), 1726–1737.
64. Laguerre, A.; Hukezalie, K.; Winckler, P.; Katranji, F.; Chanteloup, G.; Pirrotta, M.; Perrier-Cornet, J.-M.; Wong, J. M. Y.; Monchaud, D. Visualization of RNA-Quadruplexes in Live Cells. *J. Am. Chem. Soc.* **2015**, *137*(26), 8521–8525.
65. Li, Y.; Geyer, R.; Sen, D. Recognition of Anionic Porphyrins by DNA Aptamers. *Biochemistry* **1996**, *35*, 6911–6922.
66. Gray, L. T.; Lombardi, E. P.; Verga, D.; Guetta, C.; Nicolas, A.; Teulade-Fichou, M.-P.; Londoño-Vallejo, A.; Maizels, N. G-Quadruplexes Sequester Free Heme in Living Cells. *Cell Chem. Biol.* **2019**, *26*(12), 1681–1691.
67. Balwani, M.; Desnick, R. J. The Porphyrias: Advances in Diagnosis and Treatment. *Hematol. Am. Soc. Hematol. Educ. Program* **2012**, *2012*, 19–27.
68. Hardin, C. C.; Perry, A. G.; White, K. Thermodynamic and Kinetic Characterization of the Dissociation and Assembly of Quadruplex Nucleic Acids. *Biopolymers* **2000**, *56*(3), 147–194.
69. Golub, E.; Albada, H. B.; Liao, W.-C.; Biniuri, Y.; Willner, I. Nucleoapzymes: Hemin/G-Quadruplex DNAzyme-Aptamer Binding Site Conjugates with Superior Enzyme-like Catalytic Functions. *J. Am. Chem. Soc.* **2016**, *138*(1), 164–172.
70. Hoffmann, R. F.; Moshkin, Y. M.; Mouton, S.; Grzeschik, N. A.; Kalicharan, R. D.; Kuipers, J.; Wolters, A. H. G.; Nishida, K.; Romashchenko, A. V.; Postberg, J.; Lipps, H.; Berezikov, E.; Sibon, O. C. M.; Giepmans, B. N. G.; Lansdorp, P. M. Guanine Quadruplex Structures Localize to Heterochromatin. *Nucleic Acids Res.* **2016**, *44*(1), 152–163.
71. Artusi, S.; Perrone, R.; Lago, S.; Raffa, P.; Di Iorio, E.; Palù, G.; Richter, S. N. Visualization of DNA G-Quadruplexes in Herpes Simplex Virus 1-Infected Cells. *Nucleic Acids Res.* **2016**, *44*(21), 10343–10353.
72. Zyner, K. G.; Mulhearn, D. S.; Adhikari, S.; Martínez Cuesta, S.; Di Antonio, M.; Erard, N.; Hannon, G. J.; Tannahill, D.; Balasubramanian, S. Genetic Interactions of G-Quadruplexes in Humans. *Elife* **2019**, *8*, e46793.
73. Suram, A.; Kaplunov, J.; Patel, P. L.; Ruan, H.; Cerutti, A.; Boccardi, V.; Fumagalli, M.; Di Micco, R.; Mirani, N.; Gurung, R. L.; Hande, M. P.; d'Adda di Fagagna, F.; Herbig, U. Oncogene-Induced Telomere Dysfunction Enforces Cellular Senescence in Human Cancer Precursor Lesions. *EMBO J.* **2012**, *31*(13), 2839–2851.
74. Giraldo, R.; Rhodes, D. The Yeast Telomere-Binding Protein RAP1 Binds to and Promotes the Formation of DNA Quadruplexes in Telomeric DNA. *EMBO J.* **1994**, *13*(10), 2411–2420.

75. Morse, R. H. R. A. P. RAP, Open up! New Wrinkles for RAP1 in Yeast. *Trends Genet.* **2000**, *16*, 51–53.
76. Traczyk, A.; Liew, C. W.; Gill, D. J.; Rhodes, D. Structural Basis of G-Quadruplex DNA Recognition by the Yeast Telomeric Protein Rap1. *Nucleic Acids Res.* **2020**, *48*(8), 4562–4571.
77. Song, S.; Perez, J. V.; Svitko, W.; Ricketts, M. D.; Dean, E.; Schultz, D.; Marmorstein, R.; Johnson, F. B. Rap1-Mediated Nucleosome Displacement Can Regulate Gene Expression in Senescent Cells without Impacting the Pace of Senescence. *Aging Cell* **2020**, *19*(1)e13061.
78. Kitano, K.; Kim, S.-Y.; Hakoshima, T. Structural Basis for DNA Strand Separation by the Unconventional Winged-Helix Domain of RecQ Helicase WRN. *Structure* **2010**, *18*(2), 177–187.
79. Chatterjee, S.; Zagelbaum, J.; Savitsky, P.; Sturzenegger, A.; Huttner, D.; Janscak, P.; Hickson, I. D.; Gileadi, O.; Rothenberg, E. Mechanistic Insight into the Interaction of BLM Helicase with Intra-Strand G-Quadruplex Structures. *Nat. Commun.* **2014**, *5*, 5556.
80. Chereji, R. V.; Bryson, T. D.; Henikoff, S. Quantitative MNase-Seq Accurately Maps Nucleosome Occupancy Levels. *Genome Biol.* **2019**, *20*(1), 198.
81. Halder, K.; Halder, R.; Chowdhury, S. Genome-Wide Analysis Predicts DNA Structural Motifs as Nucleosome Exclusion Signals. *Mol. BioSyst.* **2009**, *5*(12), 1703–1712.
82. Wong, H. M.; Huppert, J. L. Stable G-Quadruplexes Are Found Outside Nucleosome-Bound Regions. *Mol. BioSyst.* **2009**, *5*(12), 1713–1719.
83. Kouzine, F.; Wojtowicz, D.; Baranello, L.; Yamane, A.; Nelson, S.; Resch, W.; Kieffer-Kwon, K.-R.; Benham, C. J.; Casellas, R.; Przytycka, T. M.; Levens, D. Permanganate/S1 Nuclease Footprinting Reveals Non-B DNA Structures with Regulatory Potential Across a Mammalian Genome. *Cell Syst.* **2017**, *4*(3), 344–356.e7.
84. Foulk, M. S.; Urban, J. M.; Casella, C.; Gerbi, S. A. Characterizing and Controlling Intrinsic Biases of Lambda Exonuclease in Nascent Strand Sequencing Reveals Phasing Between Nucleosomes and G-Quadruplex Motifs Around a Subset of Human Replication Origins. *Genome Res.* **2015**, *25*(5), 725–735.
85. Law, M. J.; Lower, K. M.; Voon, H. P. J.; Hughes, J. R.; Garrick, D.; Viprakasit, V.; Mitson, M.; De Gobbi, M.; Marra, M.; Morris, A.; Abbott, A.; Wilder, S. P.; Taylor, S.; Santos, G. M.; Cross, J.; Ayyub, H.; Jones, S.; Ragoussis, J.; Rhodes, D.; Dunham, I.; Higgs, D. R.; Gibbons, R. J. ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner. *Cell* **2010**, *143*, 367–378.
86. Udugama, M.; Chang, F. T. M.; Chan, F. L.; Tang, M. C.; Pickett, H. A.; McGhie, J. D. R.; Mayne, L.; Collas, P.; Mann, J. R.; Wong, L. H. Histone Variant H3.3 Provides the Heterochromatic H3 Lysine 9 Tri-Methylation Mark at Telomeres. *Nucleic Acids Res.* **2015**, *43*(21), 10227–10237.
87. Duquette, M. L.; Handa, P.; Vincent, J. A.; Taylor, A. F.; Maizels, N. Intracellular Transcription of G-Rich DNAs Induces Formation of G-Loops, Novel Structures Containing G4 DNA. *Genes Dev.* **2004**, *18*(13), 1618–1629.
88. Nguyen, D. T.; Voon, H. P. J.; Xella, B.; Scott, C.; Clynes, D.; Babbs, C.; Ayyub, H.; Kerry, J.; Sharpe, J. A.; Sloane-Stanley, J. A.; Butler, S.; Fisher, C. A.; Gray, N. E.; Jenuwein, T.; Higgs, D. R.; Gibbons, R. J. The Chromatin Remodelling Factor ATRX Suppresses R-Loops in Transcribed Telomeric Repeats. *EMBO Rep.* **2017**, *18*(6), 914–928.
89. Wang, Y.; Yang, J.; Wild, A. T.; Wu, W. H.; Shah, R.; Danussi, C.; Riggins, G. J.; Kannan, K.; Sulman, E. P.; Chan, T. A.; Huse, J. T. G-Quadruplex DNA Drives Genomic Instability and Represents a Targetable Molecular Abnormality in ATRX-Deficient Malignant Glioma. *Nat. Commun.* **2019**, *10*(1), 943.

90. Hirschi, A.; Martin, W. J.; Luka, Z.; Loukachevitch, L. V.; Reiter, N. J. G-Quadruplex RNA Binding and Recognition by the Lysine-Specific Histone Demethylase-1 Enzyme. *RNA* **2016**, *22*, 1250–1260.
91. Hussain, T.; Saha, D.; Purohit, G.; Kar, A.; Kishore Mukherjee, A.; Sharma, S.; Sengupta, S.; Dhapola, P.; Maji, B.; Vedagopuram, S.; Horikoshi, N. T.; Horikoshi, N.; Pandita, R. K.; Bhattacharya, S.; Bajaj, A.; Riou, J.-F.; Pandita, T. K.; Chowdhury, S. Transcription Regulation of CDKN1A (p21/CIP1/WAF1) by TRF2 Is Epigenetically Controlled Through the REST Repressor Complex. *Sci. Rep.* **2017**, *7*(1), 11541.
92. Saha, D.; Singh, A.; Hussain, T.; Srivastava, V.; Sengupta, S.; Kar, A.; Dhapola, P.; Dhople, V.; Ummanni, R.; Chowdhury, S. Epigenetic Suppression of Human Telomerase (hTERT) Is Mediated by the Metastasis Suppressor NME2 in a G-Quadruplex-Dependent Fashion. *J. Biol. Chem.* **2017**, *292*(37), 15205–15215.
93. Sarkies, P.; Murat, P.; Phillips, L. G.; Patel, K. J.; Balasubramanian, S.; Sale, J. E. FANCJ Coordinates Two Pathways That Maintain Epigenetic Stability at G-Quadruplex DNA. *Nucleic Acids Res.* **2012**, *40*(4), 1485–1498.
94. Schiavone, D.; Guilbaud, G.; Murat, P.; Papadopoulou, C.; Sarkies, P.; Prioleau, M.-N.; Balasubramanian, S.; Sale, J. E. Determinants of G Quadruplex-Induced Epigenetic Instability in REV1-Deficient Cells. *EMBO J.* **2014**, *33*(21), 2507–2520.
95. Guilbaud, G.; Murat, P.; Recolin, B.; Campbell, B. C.; Maiter, A.; Sale, J. E.; Balasubramanian, S. Local Epigenetic Reprogramming Induced by G-Quadruplex Ligands. *Nat. Chem.* **2017**, *9*(11), 1110–1117.
96. Hardin, C. C.; Corregan, M.; Brown, B. A., 2nd; Frederick, L. N. Cytosine-Cytosine+ Base Pairing Stabilizes DNA Quadruplexes and Cytosine Methylation Greatly Enhances the Effect. *Biochemistry* **1993**, *32*(22), 5870–5880.
97. Cree, S. L.; Fredericks, R.; Miller, A.; Pearce, F. G.; Filichev, V.; Fee, C.; Kennedy, M. A. DNA G-Quadruplexes Show Strong Interaction with DNA Methyltransferases in Vitro. *FEBS Lett.* **2016**, *590*(17), 2870–2883.
98. Smith, S. S.; Kan, J. L.; Baker, D. J.; Kaplan, B. E.; Dembek, P. Recognition of Unusual DNA Structures by Human DNA (cytosine-5)methyltransferase. *J. Mol. Biol.* **1991**, *217*(1), 39–51.
99. Halder, R.; Halder, K.; Sharma, P.; Garg, G.; Sengupta, S.; Chowdhury, S. Guanine Quadruplex DNA Structure Restricts Methylation of CpG Dinucleotides Genome-Wide. *Mol. BioSyst.* **2010**, *6*(12), 2439–2447.
100. Fleming, A. M.; Ding, Y.; Burrows, C. J. Oxidative DNA Damage Is Epigenetic by Regulating Gene Transcription via Base Excision Repair. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*(10), 2604–2609.
101. Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct Evidence for a G-Quadruplex in a Promoter Region and Its Targeting with a Small Molecule to Repress c-MYC Transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593–11598.
102. Simonsson, T.; Henriksson, M. C-Myc Suppression in Burkitt's Lymphoma Cells. *Biochem. Biophys. Res. Commun.* **2002**, *290*, 11–15.
103. Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. Small-Molecule Interaction with a Five-Guanine-Tract G-Quadruplex Structure from the Human MYC Promoter. *Nat. Chem. Biol.* **2005**, *1*(3), 167–173.
104. Francisco, A. P.; Paulo, A. Oncogene Expression Modulation in Cancer Cell Lines by DNA G-Quadruplex-Interactive Small Molecules. *Curr. Med. Chem.* **2017**, *24*(42), 4873–4904.
105. Halder, R.; Riou, J.-F.; Teulade-Fichou, M.-P.; Frickey, T.; Hartig, J. S. Bisquinolinium Compounds Induce Quadruplex-Specific Transcriptome Changes in HeLa S3 Cell Lines. *BMC Res. Notes* **2012**, *5*, 138.

106. Gray, L. T.; Vallur, A. C.; Eddy, J.; Maizels, N. G Quadruplexes Are Genomewide Targets of Transcriptional Helicases XPB and XPD. *Nat. Chem. Biol.* **2014**, *10*(4), 313–318.
107. Parrinello, S.; Samper, E.; Krtolica, A.; Goldstein, J.; Melov, S.; Campisi, J. Oxygen Sensitivity Severely Limits the Replicative Lifespan of Murine Fibroblasts. *Nat. Cell Biol.* **2003**, *5*(8), 741–747.
108. Gros, J.; da Silva, M. W.; De Cian, A.; Amrane, S.; Rosu, F.; Bourdoncle, A.; Saccà, B.; Alberti, P.; Lacroix, L.; Mergny, J.-L. Kinetics and Thermodynamics of G-Quadruplexes. *Nucleic Acids Symp. Ser.* **2005**, *49*(1), 61–62.
109. Zatterale, A.; Kelly, F. J.; Degan, P.; d'Ischia, M.; Pallardó, F. V.; Calzone, R.; Dunster, C.; Lloret, A.; Manini, P.; Coğulu, O.; Kavakli, K.; Pagano, G. Oxidative Stress Biomarkers in Four Bloom Syndrome (BS) Patients and in Their Parents Suggest in Vivo Redox Abnormalities in BS Phenotype. *Clin. Biochem.* **2007**, *40*(15), 1100–1103.
110. Kitsera, N.; Stathis, D.; Lühnsdorf, B.; Müller, H.; Carell, T.; Epe, B.; Khobta, A. 8-Oxo-7,8-Dihydroguanine in DNA Does Not Constitute a Barrier to Transcription, but Is Converted into Transcription-Blocking Damage by OGG1. *Nucleic Acids Res.* **2011**, *39*(14), 5926–5934.
111. Kanagaraj, R.; Parasuraman, P.; Mihaljevic, B.; van Loon, B.; Burdova, K.; König, C.; Furrer, A.; Bohr, V. A.; Hübscher, U.; Janscak, P. Involvement of Werner Syndrome Protein in MUTYH-Mediated Repair of Oxidative DNA Damage. *Nucleic Acids Res.* **2012**, *40*, 8449–8459.
112. Cheong, V. V.; Heddi, B.; Lech, C. J.; Phan, A. T. Xanthine and 8-Oxoguanine in G-Quadruplexes: Formation of a G·G·X·O Tetrad. *Nucleic Acids Res.* **2015**, *43*(21), 10506–10514.
113. Kim, N. The Interplay Between G-Quadruplex and Transcription. *Curr. Med. Chem.* **2019**, *26*(16), 2898–2917.
114. Lorenz, R.; Bernhart, S. H.; Qin, J.; Siederdissen, C. H. z.; Tanzer, A.; Amman, F.; Hofacker, I. L.; Stadler, P. F. 2D Meets 4G: G-Quadruplexes in RNA Secondary Structure Prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2013**, 832–844.
115. Beaudoin, J.-D.; Jodoin, R.; Perreault, J.-P. New Scoring System to Identify RNA G-Quadruplex Folding. *Nucleic Acids Res.* **2014**, *42*(2), 1209–1223.
116. Kwok, C. K.; Balasubramanian, S. Targeted Detection of G-Quadruplexes in Cellular RNAs. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 6751–6754.
117. Kwok, C. K.; Sahakyan, A. B.; Balasubramanian, S. Structural Analysis Using SHALiPE to Reveal RNA G-Quadruplex Formation in Human Precursor MicroRNA. *Angew. Chem. Int. Ed Engl.* **2016**, *55*(31), 8958–8961.
118. Sauer, M.; Paeschke, K. G-Quadruplex Unwinding Helicases and Their Function in Vivo. *Biochem. Soc. Trans.* **2017**, *45*(5), 1173–1182.
119. Marcel, V.; Tran, P. L. T.; Sagne, C.; Martel-Planche, G.; Vaslin, L.; Teulade-Fichou, M.-P.; Hall, J.; Mergny, J.-L.; Hainaut, P.; Van Dyck, E. G-Quadruplex Structures in TP53 Intron 3: Role in Alternative Splicing and in Production of p53 mRNA Isoforms. *Carcinogenesis* **2011**, *32*(3), 271–278.
120. Fisette, J.-F.; Montagna, D. R.; Mihailescu, M.-R.; Wolfe, M. S. A G-Rich Element Forms a G-Quadruplex and Regulates BACE1 mRNA Alternative Splicing. *J. Neurochem.* **2012**, *121*(5), 763–773.
121. Subramanian, M.; Rage, F.; Tabet, R.; Flatter, E.; Mandel, J.-L.; Moine, H. G-Quadruplex RNA Structure as a Signal for Neurite mRNA Targeting. *EMBO Rep.* **2011**, *12*(7), 697–704.
122. Lee, D. S. M.; Ghanem, L. R.; Barash, Y. Integrative Analysis Reveals RNA G-Quadruplexes in UTRs Are Selectively Constrained and Enriched for Functional Associations. *Nat. Commun.* **2020**, *11*(1), 527.

123. Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. An RNA G-Quadruplex in the 5' UTR of the NRAS Proto-Oncogene Modulates Translation. *Nat. Chem. Biol.* **2007**, *3*(4), 218–221.
124. Katsuda, Y.; Sato, S.-I.; Asano, L.; Morimura, Y.; Furuta, T.; Sugiyama, H.; Hagihara, M.; Uesugi, M. A Small Molecule That Represses Translation of G-Quadruplex-Containing mRNA. *J. Am. Chem. Soc.* **2016**, *138*, 9037–9040.
125. Murat, P.; Marsico, G.; Herdy, B.; Ghanbarian, A. T.; Portella, G.; Balasubramanian, S. RNA G-Quadruplexes at Upstream Open Reading Frames Cause DHX36- and DHX9-Dependent Translation of Human mRNAs. *Genome Biol.* **2018**, *19*(1), 229.
126. Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-Quadruplexes: The Beginning and End of UTRs. *Nucleic Acids Res.* **2008**, *36*(19), 6260–6268.
127. Endoh, T.; Kawasaki, Y.; Sugimoto, N. Suppression of Gene Expression by G-Quadruplexes in Open Reading Frames Depends on G-Quadruplex Stability. *Angew. Chem. Int. Ed Engl.* **2013**, *52*(21), 5522–5526.
128. Arachchilage, G. M.; Arachchilage, M. H.; Venkataraman, A.; Piontkivska, H.; Basu, S. Stable G-Quadruplex Enabling Sequences Are Selected against by the Context-Dependent Codon Bias. *Gene* **2019**, *696*, 149–161.
129. Ibrahim, F.; Maragkakis, M.; Alexiou, P.; Mourelatos, Z. Ribothrypsis, a Novel Process of Canonical mRNA Decay, Mediates Ribosome-Phased mRNA Endonucleolysis. *Nat. Struct. Mol. Biol.* **2018**, *25*(4), 302–310.
130. Prorok, P.; Artufel, M.; Aze, A.; Coulombe, P.; Peiffer, I.; Lacroix, L.; Guédin, A.; Mergny, J.-L.; Damaschke, J.; Schepers, A.; Ballester, B.; Méchali, M. Involvement of G-Quadruplex Regions in Mammalian Replication Origin Activity. *Nat. Commun.* **2019**, *10*(1), 3274.
131. Cheung, I.; Schertzer, M.; Rose, A.; Lansdorp, P. M. Disruption of Dog-1 in *Caenorhabditis Elegans* Triggers Deletions Upstream of Guanine-Rich DNA. *Nat. Genet.* **2002**, *31*, 405–409.
132. Kruisselbrink, E.; Guryev, V.; Brouwer, K.; Pontier, D. B.; Cuppen, E.; Tijsterman, M. Mutagenic Capacity of Endogenous G4 DNA Underlies Genome Instability in FANCJ-Defective *C. elegans*. *Curr. Biol.* **2008**, *18*(12), 900–905.
133. Lemmens, B.; van Schendel, R.; Tijsterman, M. Mutagenic Consequences of a Single G-Quadruplex Demonstrate Mitotic Inheritance of DNA Replication Fork Barriers. *Nat. Commun.* **2015**, *6*, 8909.
134. Koole, W.; van Schendel, R.; Karambelas, A. E.; van Heteren, J. T.; Okihara, K. L.; Tijsterman, M. A Polymerase Theta-Dependent Repair Pathway Suppresses Extensive Genomic Instability at Endogenous G4 DNA Sites. *Nat. Commun.* **2014**, *5*, 3216.
135. London, T. B. C.; Barber, L. J.; Mosedale, G.; Kelly, G. P.; Balasubramanian, S.; Hickson, I. D.; Boulton, S. J.; Hiom, K. FANCJ Is a Structure-Specific DNA Helicase Associated with the Maintenance of Genomic G/C Tracts. *J. Biol. Chem.* **2008**, *283*(52), 36132–36139.
136. Bharti, S. K.; Sommers, J. A.; Awate, S.; Bellani, M. A.; Khan, I.; Bradley, L.; King, G. A.; Seol, Y.; Vidhyasagar, V.; Wu, Y.; Abe, T.; Kobayashi, K.; Shin-ya, K.; Kitao, H.; Wold, M. S.; Branzei, D.; Neuman, K. C.; Brosh, R. M. A Minimal Threshold of FANCJ Helicase Activity Is Required for Its Response to Replication Stress or Double-Strand Break Repair. *Nucleic Acids Res.* **2018**, *46*, 6238–6256.
137. Lopes, J.; Piazza, A.; Bermejo, R.; Kriegsman, B.; Colosio, A.; Teulade-Fichou, M.-P.; Foiani, M.; Nicolas, A. G-Quadruplex-Induced Instability during Leading-Strand Replication. *EMBO J.* **2011**, *30*(19), 4033–4046.
138. Piazza, A.; Adrian, M.; Samazan, F.; Heddi, B.; Hamon, F.; Serero, A.; Lopes, J.; Teulade-Fichou, M.-P.; Phan, A. T.; Nicolas, A. Short Loop Length and High Thermal Stability Determine Genomic Instability Induced by G-Quadruplex-Forming Minisatellites. *EMBO J.* **2015**, *34*(12), 1718–1734.

139. Paeschke, K.; Capra, J. A.; Zakian, V. A. DNA Replication through G-Quadruplex Motifs Is Promoted by the Saccharomyces cerevisiae Pif1 DNA Helicase. *Cell* **2011**, *145*(5), 678–691.
140. van Wietmarschen, N.; Merzouk, S.; Halsema, N.; Spierings, D. C. J.; Guryev, V.; Lansdorp, P. M. BLM Helicase Suppresses Recombination at G-Quadruplex Motifs in Transcribed Genes. *Nat. Commun.* **2018**, *9*(1), 271.
141. Paeschke, K.; Bochman, M. L.; Garcia, P. D.; Cejka, P.; Friedman, K. L.; Kowalczykowski, S. C.; Zakian, V. A. Pif1 Family Helicases Suppress Genome Instability at G-Quadruplex Motifs. *Nature* **2013**, *497*(7450), 458–462.
142. Manzo, S. G.; Hartono, S. R.; Sanz, L. A.; Marinello, J.; De Biasi, S.; Cossarizza, A.; Capranico, G.; Chedin, F. DNA Topoisomerase I Differentially Modulates R-Loops across the Human Genome. *Genome Biol.* **2018**, *19*(1), 100.
143. Wu, W.; Bhowmick, R.; Vogel, I.; Özer, Ö.; Ghisays, F.; Thakur, R. S.; Sanchez de Leon, E.; Richter, P. H.; Ren, L.; Petrini, J. H.; Hickson, I. D.; Liu, Y. RTEL1 Suppresses G-Quadruplex-Associated R-Loops at Difficult-to-Replicate Loci in the Human Genome. *Nat. Struct. Mol. Biol.* **2020**, *27*(5), 424–437.
144. De Magis, A.; Manzo, S. G.; Russo, M.; Marinello, J.; Morigi, R.; Sordet, O.; Capranico, G. DNA Damage and Genome Instability by G-Quadruplex Ligands Are Mediated by R Loops in Human Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*(3), 816–825.
145. Qiao, Q.; Wang, L.; Meng, F.-L.; Hwang, J. K.; Alt, F. W.; Wu, H. AID Recognizes Structured DNA for Class Switch Recombination. *Mol. Cell* **2017**, *67*(3), 361–373.e4.
146. Ribeiro de Almeida, C.; Dhir, S.; Dhir, A.; Moghaddam, A. E.; Sattentau, Q.; Meinhart, A.; Proudfoot, N. J. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol. Cell* **2018**, *70*(4), 650–662.e8.
147. Sahakyan, A. B.; Murat, P.; Mayer, C.; Balasubramanian, S. G-Quadruplex Structures Within the 3' UTR of LINE-1 Elements Stimulate Retrotransposition. *Nat. Struct. Mol. Biol.* **2017**, *24*(3), 243–247.
148. Carle, C. M.; Zaher, H. S.; Chalker, D. L. A Parallel G Quadruplex-Binding Protein Regulates the Boundaries of DNA Elimination Events of Tetrahymena thermophila. *PLoS Genet.* **2016**, *12*(3), e1005842.
149. Byrd, A. K.; Zybailov, B. L.; Maddukuri, L.; Gao, J.; Marecki, J. C.; Jaiswal, M.; Bell, M. R.; Griffin, W. C.; Reed, M. R.; Chib, S.; Mackintosh, S. G.; MacNicol, A. M.; Baldini, G.; Eoff, R. L.; Raney, K. D. Evidence That G-Quadruplex DNA Accumulates in the Cytoplasm and Participates in Stress Granule Assembly in Response to Oxidative Stress. *J. Biol. Chem.* **2016**, *291*(34), 18041–18057.
150. Fleming, A. M.; Burrows, C. J. Interplay of Guanine Oxidation and G-Quadruplex Folding in Gene Promoters. *J. Am. Chem. Soc.* **2020**, *142*(3), 1115–1136.
151. Singh, A.; Kukreti, R.; Saso, L.; Kukreti, S. Oxidative Stress: Role and Response of Short Guanine Tracts at Genomic Locations. *Int. J. Mol. Sci.* **2019**, *20*, 4258.
152. de Lange, T. Shelterin-mediated telomere protection. *Annu. Rev. Genet.* **2018**, *52*, 223–247.
153. Gray, R. D.; Li, J.; Chaires, J. B. Energetics and Kinetics of a Conformational Switch in G-Quadruplex DNA. *J. Phys. Chem. B* **2009**, *113*, 2676–2683.
154. Bončina, M.; Vesnaver, G.; Chaires, J. B.; Lah, J. Unraveling the Thermodynamics of the Folding and Interconversion of Human Telomere G-Quadruplexes. *Angew. Chem. Int. Ed Engl.* **2016**, *55*(35), 10340–10344.
155. Kolesnikova, S.; Curtis, E. A. Structure and Function of Multimeric G-Quadruplexes. *Molecules* **2019**, *24*, 3074.

156. Sfeir, A.; Kosiyatrakul, S. T.; Hockemeyer, D.; MacRae, S. L.; Karlseder, J.; Schildkraut, C. L.; de Lange, T. Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. *Cell* **2009**, *138*(1), 90–103.
157. Zimmermann, M.; Kibe, T.; Kabir, S.; de Lange, T. TRF1 Negotiates TTAGGG Repeat-Associated Replication Problems by Recruiting the BLM Helicase and the TPP1/POT1 Repressor of ATR Signaling. *Genes Dev.* **2014**, *28*(22), 2477–2491.
158. Vannier, J.-B.; Pavicic-Kaltenbrunner, V.; Petalcorin, M. I. R.; Ding, H.; Boulton, S. J. RTEL1 Dismantles T Loops and Counteracts Telomeric G4-DNA to Maintain Telomere Integrity. *Cell* **2012**, *149*(4), 795–806.
159. Crabbe, L.; Verdun, R. E.; Haggblom, C. I.; Karlseder, J. Defective Telomere Lagging Strand Synthesis in Cells Lacking WRN Helicase Activity. *Science* **2004**, *306*(5703), 1951–1953.
160. Arnoult, N.; Saintome, C.; Ourliac-Garnier, I.; Riou, J.-F.; Londoño-Vallejo, A. Human POT1 Is Required for Efficient Telomere C-Rich Strand Replication in the Absence of WRN. *Genes Dev.* **2009**, *23*(24), 2915–2924.
161. Masuda-Sasa, T.; Polaczek, P.; Peng, X. P.; Chen, L.; Campbell, J. L. Processing of G4 DNA by Dna2 Helicase/nuclease and Replication Protein A (RPA) Provides Insights into the Mechanism of Dna2/RPA Substrate Recognition. *J. Biol. Chem.* **2008**, *283*(36), 24359–24373.
162. Rizzo, A.; Iachettini, S.; Zizza, P.; Cingolani, C.; Porru, M.; Artuso, S.; Stevens, M.; Hummersone, M.; Biroccio, A.; Salvati, E.; Leonetti, C. Identification of Novel RHPS4-Derivative Ligands with Improved Toxicological Profiles and Telomere-Targeting Activities. *J. Exp. Clin. Cancer Res.* **2014**, *33*, 81.
163. Gauthier, L. R.; Granotier, C.; Hoffschir, F.; Etienne, O.; Ayouaz, A.; Desmaze, C.; Mailliet, P.; Biard, D. S.; Boussin, F. D. Rad51 and DNA-PKcs Are Involved in the Generation of Specific Telomere Aberrations Induced by the Quadruplex Ligand 360A That Impair Mitotic Cell Progression and Lead to Cell Death. *Cell. Mol. Life Sci.* **2012**, *69*(4), 629–640.
164. Uringa, E.-J.; Lisaingo, K.; Pickett, H. A.; Brind'Amour, J.; Rohde, J.-H.; Zelensky, A.; Essers, J.; Lansdorp, P. M. RTEL1 Contributes to DNA Replication and Repair and Telomere Maintenance. *Mol. Biol. Cell* **2012**, *23*(14), 2782–2792.
165. Lin, W.; Sampathi, S.; Dai, H.; Liu, C.; Zhou, M.; Hu, J.; Huang, Q.; Campbell, J.; Shin-Ya, K.; Zheng, L.; Chai, W.; Shen, B. Mammalian DNA2 Helicase/nuclease Cleaves G-Quadruplex DNA and Is Required for Telomere Integrity. *EMBO J.* **2013**, *32*(10), 1425–1439.
166. Hwang, I. P.; Mailliet, P.; Hossard, V.; Riou, J.-F.; Bugaut, A.; Roger, L. Investigating the Effect of Mono- and Dimeric 360A G-Quadruplex Ligands on Telomere Stability by Single Telomere Length Analysis (STELA). *Molecules* **2019**, *24*, 577.
167. Zhang, M.; Wang, B.; Li, T.; Liu, R.; Xiao, Y.; Geng, X.; Li, G.; Liu, Q.; Price, C. M.; Liu, Y.; Wang, F. *Mammalian CST Averts Replication Failure by Preventing G-Quadruplex Accumulation*. *Nucleic Acids Res.* **2019**, 5243–5259. https://doi.org/10.1093/nar/gkz264.
168. Bhattacharjee, A.; Wang, Y.; Diao, J.; Price, C. M. Dynamic DNA Binding, Junction Recognition and G4 Melting Activity Underlie the Telomeric and Genome-Wide Roles of Human CST. *Nucleic Acids Res.* **2017**, *45*, 12311–12324.
169. Pérez-Martínez, L.; Öztürk, M.; Butter, F.; Luke, B. Npl3 Stabilizes R-Loops at Telomeres to Prevent Accelerated Replicative Senescence. *EMBO Rep.* **2020**, *21*(3), e49087.
170. Pfeiffer, V.; Crittin, J.; Grolimund, L.; Lingner, J. The THO Complex Component Thp2 Counteracts Telomeric R-Loops and Telomere Shortening. *EMBO J.* **2013**, *32*(21), 2861–2871.

171. Martadinata, H.; Phan, A. T. Structure of Human Telomeric RNA (TERRA): Stacking of Two G-Quadruplex Blocks in K(+) Solution. *Biochemistry* **2013**, *52*, 2176–2183.
172. Randall, A.; Griffith, J. D. Structure of Long Telomeric RNA Transcripts: The G-Rich RNA Forms a Compact Repeating Structure Containing G-Quartets. *J. Biol. Chem.* **2009**, *284*(21), 13980–13986.
173. Tan, J.; Duan, M.; Yadav, T.; Phoon, L.; Wang, X.; Zhang, J.-M.; Zou, L.; Lan, L. An R-Loop-Initiated CSB-RAD52-POLD3 Pathway Suppresses ROS-Induced Telomeric DNA Breaks. *Nucleic Acids Res.* **2020**, *48*(3), 1285–1300.
174. Hayashi, N.; Murakami, S. STM1, a Gene Which Encodes a Guanine Quadruplex Binding Protein, Interacts with CDC13 in Saccharomyces cerevisiae. *Mol. Genet. Genomics* **2002**, *267*(6), 806–813.
175. Jurikova, K.; Gajarsky, M.; Hajikazemi, M.; Nosek, J.; Prochazkova, K.; Paeschke, K.; Trantirek, L.; Tomaska, L. Role of Folding Kinetics of Secondary Structures in Telomeric G-Overhangs in the Regulation of Telomere Maintenance in Saccharomyces cerevisiae. *J. Biol. Chem.* **2020**, *295*, 8958–8971.
176. Zhou, G.; Liu, X.; Li, Y.; Xu, S.; Ma, C.; Wu, X.; Cheng, Y.; Yu, Z.; Zhao, G.; Chen, Y. Telomere Targeting with a Novel G-Quadruplex-Interactive Ligand BRACO-19 Induces T-Loop Disassembly and Telomerase Displacement in Human Glioblastoma Cells. *Oncotarget* **2016**, *7*(12), 14925–14939.
177. Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J. A.; Neidle, S. The G-Quadruplex-Interactive Molecule BRACO-19 Inhibits Tumor Growth, Consistent With Telomere Targeting and Interference with Telomerase Function. *Cancer Res.* **2005**, *65*(4), 1489–1496.
178. Hampel, S. M.; Sidibe, A.; Gunaratnam, M.; Riou, J.-F.; Neidle, S. Tetrasubstituted Naphthalene Diimide Ligands with Selectivity for Telomeric G-Quadruplexes and Cancer Cells. *Bioorg. Med. Chem. Lett.* **2010**, *20*(22), 6459–6463.
179. Salvati, E.; Rizzo, A.; Iachettini, S.; Zizza, P.; Cingolani, C.; D'Angelo, C.; Porru, M.; Mondello, C.; Aiello, A.; Farsetti, A.; Gilson, E.; Leonetti, C.; Biroccio, A. A Basal Level of DNA Damage and Telomere Deprotection Increases the Sensitivity of Cancer Cells to G-Quadruplex Interactive Compounds. *Nucleic Acids Res.* **2015**, *43*(3), 1759–1769.
180. Perera, O. N.; Sobinoff, A. P.; Teber, E. T.; Harman, A.; Maritz, M. F.; Yang, S. F.; Pickett, H. A.; Cesare, A. J.; Arthur, J. W.; MacKenzie, K. L.; Bryan, T. M. Telomerase Promotes Formation of a Telomere Protective Complex in Cancer Cells. *Sci. Adv.* **2019**, *5*(10), eaav4409.
181. Shay, J. W.; Wright, W. E. Telomeres and Telomerase: Three Decades of Progress. *Nat. Rev. Genet.* **2019**, *20*(5), 299–309.
182. Zaug, A. J.; Podell, E. R.; Cech, T. R. Human POT1 Disrupts Telomeric G-Quadruplexes Allowing Telomerase Extension in Vitro. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*(31), 10864–10869.
183. Gomez, D. L. M.; Armando, R. G.; Cerrudo, C. S.; Ghiringhelli, P. D.; Gomez, D. E. Telomerase as a Cancer Target. Development of New Molecules. *Curr. Top. Med. Chem.* **2016**, *16*(22), 2432–2440.
184. Fouquerel, E.; Lormand, J.; Bose, A.; Lee, H.-T.; Kim, G. S.; Li, J.; Sobol, R. W.; Freudenthal, B. D.; Myong, S.; Opresko, P. L. Oxidative Guanine Base Damage Regulates Human Telomerase Activity. *Nat. Struct. Mol. Biol.* **2016**, *23*(12), 1092–1100.
185. Oganesian, L.; Moon, I. K.; Bryan, T. M.; Jarstfer, M. B. Extension of G-Quadruplex DNA by Ciliate Telomerase. *EMBO J.* **2006**, *25*(5), 1148–1159.
186. Moye, A. L.; Porter, K. C.; Cohen, S. B.; Phan, T.; Zyner, K. G.; Sasaki, N.; Lovrecz, G. O.; Beck, J. L.; Bryan, T. M. Telomeric G-Quadruplexes Are a Substrate and Site of Localization for Human Telomerase. *Nat. Commun.* **2015**, *6*, 7643.

187. Temime-Smaali, N.; Guittat, L.; Sidibe, A.; Shin-ya, K.; Trentesaux, C.; Riou, J.-F. The G-Quadruplex Ligand Telomestatin Impairs Binding of Topoisomerase IIIα to G-Quadruplex-Forming Oligonucleotides and Uncaps Telomeres in ALT Cells. *PLoS One* **2009**, *4,* e6919.
188. Amato, R.; Valenzuela, M.; Berardinelli, F.; Salvati, E.; Maresca, C.; Leone, S.; Antoccia, A.; Sgura, A. G-Quadruplex Stabilization Fuels the ALT Pathway in ALT-Positive Osteosarcoma Cells. *Genes* **2020**, *11*, 304.
189. Mukherjee, A. K.; Sharma, S.; Sengupta, S.; Saha, D.; Kumar, P.; Hussain, T.; Srivastava, V.; Roy, S. D.; Shay, J. W.; Chowdhury, S. Telomere Length-Dependent Transcription and Epigenetic Modifications in Promoters Remote from Telomere Ends. *PLOS Genetics* **2018**, *14,* e1007782.
190. Platt, J. M.; Ryvkin, P.; Wanat, J. J.; Donahue, G.; Ricketts, M. D.; Barrett, S. P.; Waters, H. J.; Song, S.; Chavez, A.; Abdallah, K. O.; Master, S. R.; Wang, L.-S.; Johnson, F. B. Rap1 Relocalization Contributes to the Chromatin-Mediated Gene Expression Profile and Pace of Cell Senescence. *Genes Dev.* **2013**, *27*(12), 1406–1420.
191. Falabella, M.; Fernandez, R. J.; Johnson, F. B.; Kaufman, B. A. Potential Roles for G-Quadruplexes in Mitochondria. *Curr. Med. Chem.* **2019**, *26*(16), 2918–2932.
192. Capra, J. A.; Paeschke, K.; Singh, M.; Zakian, V. A. G-Quadruplex DNA Sequences Are Evolutionarily Conserved and Associated with Distinct Genomic Features in Saccharomyces cerevisiae. *PLoS Comput. Biol.* **2010**, *6*(7), e1000861.
193. Dong, D. W.; Pereira, F.; Barrett, S. P.; Kolesar, J. E.; Cao, K.; Damas, J.; Yatsunyk, L. A.; Johnson, F. B.; Kaufman, B. A. Association of G-Quadruplex Forming Sequences with Human mtDNA Deletion Breakpoints. *BMC Genomics* **2014**, *15*, 677.
194. Bharti, S. K.; Sommers, J. A.; Zhou, J.; Kaplan, D. L.; Spelbrink, J. N.; Mergny, J.-L.; Brosh, R. M., Jr. DNA Sequences Proximal to Human Mitochondrial DNA Deletion Breakpoints Prevalent in Human Disease Form G-Quadruplexes, a Class of DNA Structures Inefficiently Unwound by the Mitochondrial Replicative Twinkle Helicase. *J. Biol. Chem.* **2014**, *289*(43), 29975–29993.
195. Oliveira, P. H.; da Silva, C. L.; Cabral, J. M. S. An Appraisal of Human Mitochondrial DNA Instability: New Insights into the Role of Non-Canonical DNA Structures and Sequence Motifs. *PLoS One* **2013**, *8*(3), e59907.
196. Butler, T. J.; Estep, K. N.; Sommers, J. A.; Maul, R. W.; Moore, A. Z.; Bandinelli, S.; Cucca, F.; Tuke, M. A.; Wood, A. R.; Bharti, S. K.; Bogenhagen, D. F.; Yakubovskaya, E.; Garcia-Diaz, M.; Guilliam, T. A.; Byrd, A. K.; Raney, K. D.; Doherty, A. J.; Ferrucci, L.; Schlessinger, D.; Ding, J.; Brosh, R. M. Mitochondrial Genetic Variation Is Enriched in G-Quadruplex Regions That Stall DNA Synthesis in Vitro. *Hum. Mol. Genet.* **2020**, *29*(8), 1292–1309.
197. Sparks, M. A.; Singh, S. P.; Burgers, P. M.; Galletto, R. Complementary Roles of Pif1 Helicase and Single Stranded DNA Binding Proteins in Stimulating DNA Replication through G-Quadruplexes. *Nucleic Acids Res.* **2019**, *47*(16), 8595–8605.
198. Pietras, Z.; Wojcik, M. A.; Borowski, L. S.; Szewczyk, M.; Kulinski, T. M.; Cysewski, D.; Stepien, P. P.; Dziembowski, A.; Szczesny, R. J. Dedicated Surveillance Mechanism Controls G-Quadruplex Forming Non-Coding RNAs in Human Mitochondria. *Nat. Commun.* **2018**, *9*(1), 2558.
199. Huang, W.-C.; Tseng, T.-Y.; Chen, Y.-T.; Chang, C.-C.; Wang, Z.-F.; Wang, C.-L.; Hsu, T.-N.; Li, P.-T.; Chen, C.-T.; Lin, J.-J.; Lou, P.-J.; Chang, T.-C. Direct Evidence of Mitochondrial G-Quadruplex DNA by Using Fluorescent Anti-Cancer Agents. *Nucleic Acids Res.* **2015**, *43*, 10,102–10,113.
200. Naeem, M. M.; Maheshan, R.; Costford, S. R.; Wahedi, A.; Trajkovski, M.; Plavec, J.; Yatsunyk, L. A.; Ciesielski, G. L.; Kaufman, B. A.; Sondheimer, N. G-Quadruplex-Mediated Reduction of a Pathogenic Mitochondrial Heteroplasmy. *Hum. Mol. Genet.* **2019**, *28*(19), 3163–3174.

201. De, S.; Michor, F. DNA Secondary Structures and Epigenetic Determinants of Cancer Genome Evolution. *Nat. Struct. Mol. Biol.* **2011**, *18*(8), 950–955.
202. Biffi, G.; Tannahill, D.; Miller, J.; Howat, W. J.; Balasubramanian, S. Elevated Levels of G-Quadruplex Formation in Human Stomach and Liver Cancer Tissues. *PLoS One* **2014**, *9*(7), e102711.
203. Kotsantis, P.; Silva, L. M.; Irmscher, S.; Jones, R. M.; Folkes, L.; Gromak, N.; Petermann, E. Increased Global Transcription Activity as a Mechanism of Replication Stress in Cancer. *Nat. Commun.* **2016**, *7*, 13087.
204. Corces, M. R.; Granja, J. M.; Shams, S.; Louie, B. H.; Seoane, J. A.; Zhou, W.; Silva, T. C.; Groeneveld, C.; Wong, C. K.; Cho, S. W.; Satpathy, A. T.; Mumbach, M. R.; Hoadley, K. A.; Robertson, A. G.; Sheffield, N. C.; Felau, I.; Castro, M. A. A.; Berman, B. P.; Staudt, L. M.; Zenklusen, J. C.; Laird, P. W.; Curtis, C.; Cancer Genome Atlas Analysis Network; Greenleaf, W. J.; Chang, H. Y. The Chromatin Accessibility Landscape of Primary Human Cancers. *Science* **2018**, 362, eaaav1898.
205. Nakanishi, C.; Seimiya, H. G-Quadruplex in Cancer Biology and Drug Discovery. *Biochem. Biophys. Res. Commun.* **2020**, S0006-291X(20)30676-30678.
206. Lord, C. J.; Ashworth, A. PARP Inhibitors: Synthetic Lethality in the Clinic. *Science* **2017**, *355*(6330), 1152–1158.
207. Zimmer, J.; Tacconi, E. M. C.; Folio, C.; Badie, S.; Porru, M.; Klare, K.; Tumiati, M.; Markkanen, E.; Halder, S.; Ryan, A.; Jackson, S. P.; Ramadan, K.; Kuznetsov, S. G.; Biroccio, A.; Sale, J. E.; Tarsounas, M. Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Mol. Cell* **2016**, *61*(3), 449–460.
208. Xu, H.; Di Antonio, M.; McKinney, S.; Mathew, V.; Ho, B.; O'Neil, N. J.; Santos, N. D.; Silvester, J.; Wei, V.; Garcia, J.; Kabeer, F.; Lai, D.; Soriano, P.; Banáth, J.; Chiu, D. S.; Yap, D.; Le, D. D. D.; Ye, F. B.; Zhang, A.; Thu, K.; Soong, J.; Lin, S.-C.; Tsai, A. H. C.; Osako, T.; Algara, T.; Saunders, D. N.; Wong, J.; Xian, J.; Bally, M. B.; Brenton, J. D.; Brown, G. W.; Shah, S. P.; Cescon, D.; Mak, T. W.; Caldas, C.; Stirling, P. C.; Hieter, P.; Balasubramanian, S.; Aparicio, S. CX-5461 Is a DNA G-Quadruplex Stabilizer with Selective Lethality in BRCA1/2 Deficient Tumours. *Nat. Commun.* **2017**, *8*, 14432.
209. Salvati, E.; Botta, L.; Amato, J.; Di Leva, F. S.; Zizza, P.; Gioiello, A.; Pagano, B.; Graziani, G.; Tarsounas, M.; Randazzo, A.; Novellino, E.; Biroccio, A.; Cosconati, S. Lead Discovery of Dual G-Quadruplex Stabilizers and Poly(ADP-Ribose) Polymerases (PARPs) Inhibitors: A New Avenue in Anticancer Treatment. *J. Med. Chem.* **2017**, *60*(9), 3626–3635.
210. Watson, L. A.; Ashley Watson, L.; Solomon, L. A.; Li, J. R.; Jiang, Y.; Edwards, M.; Shin-ya, K.; Beier, F.; Bérubé, N. G. Atrx Deficiency Induces Telomere Dysfunction, Endocrine Defects, and Reduced Life Span. *J. Clin. Invest.* **2013**, *123*, 2049–2063.
211. Leung, J. W.-C.; Ghosal, G.; Wang, W.; Shen, X.; Wang, J.; Li, L.; Chen, J. Alpha Thalassemia/Mental Retardation Syndrome X-Linked Gene Product ATRX Is Required for Proper Replication Restart and Cellular Resistance to Replication Stress. *J. Biol. Chem.* **2013**, *288*, 6342–6350.
212. Wu, C.-H.; Chen, P.-J.; Yeh, S.-H. Nucleocapsid Phosphorylation and RNA Helicase DDX1 Recruitment Enables Coronavirus Transition from Discontinuous to Continuous Transcription. *Cell Host Microbe* **2014**, *16*(4), 462–472.