SURVEY AND SUMMARY

Modulating gene expression in breast cancer via DNA secondary structure and the CRISPR toolbox

Jessica A. Kretzmann [ID][1,2], **Kelly L. Irving** [ID][3], **Nicole M. Smith** [ID][3] **and Cameron W. Evans** [ID][3,*]

[1]Laboratory for Biomolecular Nanotechnology, Department of Physics, Technical University of Munich, Am Coulombwall 4a, 85748 Garching, Germany, [2]Munich Institute of Biomedical Engineering, Technical University of Munich, Boltzmannstraße 11, 85748 Garching, Germany and [3]School of Molecular Sciences, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia

ABSTRACT

Breast cancer is the most commonly diagnosed malignancy in women, and while the survival prognosis of patients with early-stage, non-metastatic disease is ~75%, recurrence poses a significant risk and advanced and/or metastatic breast cancer is incurable. A distinctive feature of advanced breast cancer is an unstable genome and altered gene expression patterns that result in disease heterogeneity. Transcription factors represent a unique therapeutic opportunity in breast cancer, since they are known regulators of gene expression, including gene expression involved in differentiation and cell death, which are themselves often mutated or dysregulated in cancer. While transcription factors have traditionally been viewed as 'undruggable', progress has been made in the development of small-molecule therapeutics to target relevant protein–protein, protein–DNA and enzymatic active sites, with varying levels of success. However, non-traditional approaches such as epigenetic editing, transcriptional control via CRISPR/dCas9 systems, and gene regulation through non-canonical nucleic acid secondary structures represent new directions yet to be fully explored. Here, we discuss these new approaches and current limitations in light of new therapeutic opportunities for breast cancers.

INTRODUCTION

Breast cancer recently surpassed lung cancer as the most commonly diagnosed cancer, with over 16 million people currently living with breast cancer worldwide (1,2). Breast cancer is a disease with many faces, owing to genomic instability and aberrant transcription patterns, which contribute to a high level of heterogeneity. Existing therapies target various points in the breast cancer pathway, from the initial stage of preventing oncogenesis in high-risk individuals to addressing therapeutic resistance in advanced metastatic scenarios. As it stands, there are no 'go-to' therapeutic options which are reliable and effective for treating breast cancer in all its variants. Instead, current clinical practice typically divides breast cancer into five intrinsic subtypes: Luminal A, Luminal B, HER2-enriched, Basal-like and Normal-like, on which basis different treatment strategies may be chosen. Key characteristics of the different breast cancer subtypes are shown in Figure 1A. Across all subtypes, while the initial response to therapy is often positive, recurrence is common; in fact, the majority of breast cancer-related deaths (~90%) are due to invasion and metastasis (3,4). For hormone receptor-positive subtypes in particular, there is a significantly elevated risk of recurrence in the following decades (5,6). Understanding how and why tumor cells leave the primary tumor site, acquire invasive properties, and colonize metastatic sites are therefore critical areas of research in breast cancer.

Epithelial–mesenchymal plasticity

Epithelial–mesenchymal plasticity (EMP) is a key mechanism that facilitates breast cancer recurrence and metastasis. The epithelial-to-mesenchymal transition (EMT) describes the process by which tumor cells adopt a mesenchymal phenotype, becoming more invasive and motile (Figure 1B). Often, the driving force fueling EMP is the aberrant expression of an array of developmental transcription factors (TFs), collectively termed EMT-TFs (7). These EMT-TFs are responsible for the maintenance of a proliferative, undifferentiated phenotype during development and tissue

*To whom correspondence should be addressed. Tel: +61 8 6488 4422; Email: cameron.evans@uwa.edu.au
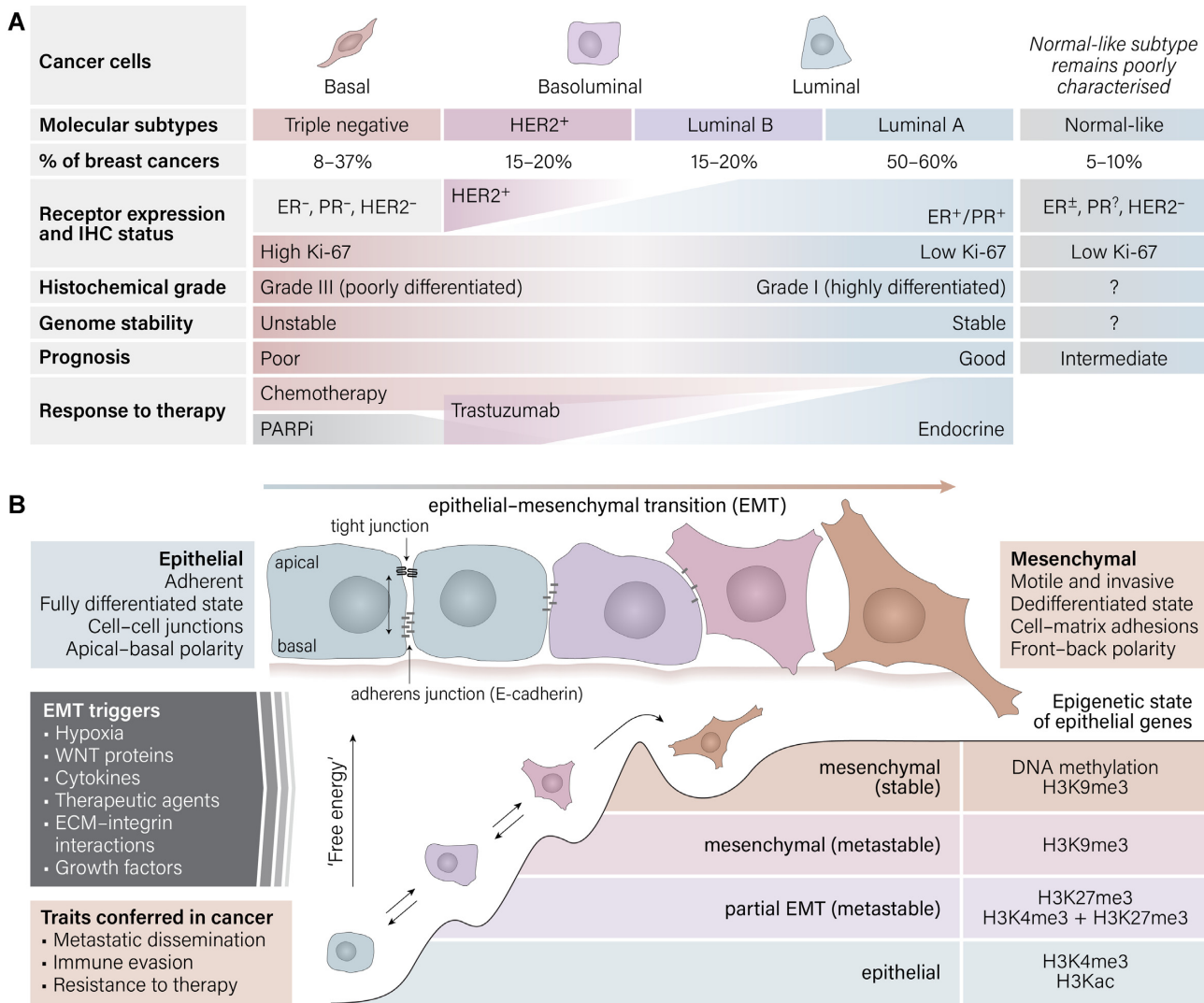
Figure 1. (**A**) Characteristics of breast cancer subtypes, based on collated data (241–245); ER, estrogen receptor, PR, progesterone receptor, PARPi, poly(ADP-ribose) polymerase (PARP) inhibitor. (**B**) Epithelial–mesenchymal transition showing key properties of epithelial and mesenchymal phenotypes, triggers and traits conferred by EMT in cancer. Adapted with permission from Springer Nature: Tam, W.L. and Weinberg, R.A. The epigenetics of epithelial–mesenchymal plasticity in cancer (2013). *Nat. Med.*, **19**, 1438–1449 (246).

regeneration, but their dysregulation in cancer can drive a gain of stemness, drug resistance, and increased invasive potential. Importantly, EMP is not regulated by an all-or-nothing switch, but rather a complex interplay of gene expression patterns, resulting in a continuum of intermediate cellular states (8–10). EMP is an epigenetically regulated and potentially reversible process (9,11,12), offering hope that with the correct therapeutic intervention, cells could transition back to a differentiated epithelial state or be re-sensitized to therapy.

Reducing the impact of EMT and re-sensitizing tumor cells to therapy relies on restoring a 'normal' profile of gene expression. Gene expression in breast cancer, just as in EMP, is highly plastic and affected by both mutations and epigenetic mechanisms. The regulation of gene expression is a circuitous process mediated by transcription factor activity and chromatin organization, which altogether modulate the ability of gene promoters to re-cruit RNA polymerases. Historically, TFs have been viewed as 'undruggable' due to the difficulty in designing small-molecule therapeutics to mediate specific protein–DNA or protein–protein interactions. More recent attempts have instead targeted the active sites of enzymes or have been designed for allosteric modulation of protein–protein interactions. Small-molecule drugs have also been developed to modulate the epigenetic state of the cancer cells by correcting DNA methylation, histone acetylation and histone methylation, for example. However, these agents all demonstrate low specificity and selectivity, and affect expression of a large proportion of genes, which ultimately leads to dose-limiting toxicity in patients (13). Furthermore, these approaches largely attempt to correct over-expression of oncogenic TFs, which is generally easier than attempting to upregulate dormant or silenced tumor-suppressor TFs, limiting the overall scope of potential targets.

The role of EMT in promoting cancer cell invasion, metastasis, and stem-like characteristics is well documented in several types of cancer, including breast cancer. Clinical samples of high-grade breast cancers have often been found to contain cells which express EMT program molecular signatures, particularly in tumors associated with poor patient prognosis [14,15]. The evidence suggests that these tumors contain cells which have progressed through 'partial EMT' with the acquisition of some mesenchymal markers, while also retaining particular epithelial characteristics. These highly heterogeneous intermediate phenotypic states arise from numerous interactions between epigenetic modifications and transcriptional regulators [9,16].

Epigenetic modifications are regulators of gene expression and EMP

Epigenetic modifications include histone acetylation, histone methylation and DNA methylation, which ultimately modulate chromatin accessibility, and together define a cell's epigenetic profile. Relationships between epigenetic modifications and transcriptional activity are complex and depend highly on the additional context in which they are placed, and have been described comprehensively elsewhere [17–19]. Epigenetic marks across the genome are highly altered in malignancies when compared to healthy tissue, and even differ between malignant cells across the EMP spectrum.

The progression of breast cancer through TF deregulation is complex, with many underlying shifts in epigenetic regulation and TFs acting in concert to direct the progression of cancerous cells through processes such as EMT. As demonstrated in Figure 1B, various histone modifications are gained during EMT, enabling a 'bivalent' epigenetic state of epithelial genes. However, the completion of EMT and stabilization of the mesenchymal phenotype requires epithelial genes to be repressed for extended periods of time. This is effected through highly stable DNA methylation (DNAme) near gene promoters, which causes gene silencing, and which is inherited with high fidelity over cell divisions [9]. Stable epigenetic repression drives malignant cells toward a mesenchymal phenotype and toward therapeutic resistance [9,11,12].

EMT is driven by an orchestra of EMT-TFs, including SNAIL, SLUG, ZEB1/2 and TWIST. These EMT-TFs are involved in the regulation of E-cadherin, N-cadherin and vimentin, as well as a range of tumor suppressor genes. Loss of *CDH1* (encoding E-cadherin) expression is a hallmark of EMT and can occur through either promoter methylation or transcriptional repression of *CDH1*. E-cadherin is both a tumor suppressor gene and a critical component in cell adhesion junctions [20–22]. EMT-TFs themselves drive widespread gene expression changes through cooperation with multiple epigenetic modifiers. For example, SNAIL (encoded by the *SNAI1* gene) cooperates with G9a, a H3K9 methyltransferase, to induce further recruitment of histone deacetylases (HDACs) and finally DNA methyltransferases (DNMTs), resulting in promoter DNAme and subsequent stable repression of E-cadherin [9,23]. Indeed, stable silencing of E-cadherin by DNAme is found in claudin-low breast cancers, one of the most mesenchymal subtypes. In a differ-

ent example, TWIST has been shown to recruit DNMT3B to the estrogen receptor 1 (*ESR1*) promoter, resulting in DNAme and loss of ER expression, and the progression of breast cancers that are ER-negative and hormone-resistant [24].

Hypermethylation of promoter CpG islands is a hallmark of cancer progression and typically correlates with transcriptional repression of the associated gene, as illustrated in Figure 2A [25]. However, it is important to note that the exact role and relationship between methylation and gene expression remains unresolved, and seems to depend highly on the specific context [26]. Promoter hypermethylation is commonly associated with a decrease in transcriptional activity and thought to alter the recruitment of regulatory proteins to the underlying DNA sequence, subsequently blocking transcriptional activation. Alternatively, methylation can provide binding sites for methyl-binding proteins which can act to mediate gene repression through their interactions with HDACs [27–29]. In breast cancers, genomic stability and epigenetic silencing of tumor suppressor genes is often associated with both DNA hypermethylation in combination with aberrant histone modification.

The relationships between epigenetic modifications and EMT-TFs are less well understood. For example, promoter hypermethylation of *TWIST1* has been observed but has not been directly correlated with gene repression in breast cancer, although the relationship has been established for other cancers [30]. There is evidence, however, that DNAme within regions of *SNAI1* (SNAIL) and *SNAI2* (SLUG) are correlated with gene repression, in addition to promoter hypermethylation of *ZEB2* [31]. Therefore, there is evidence to suggest that some EMT-TFs are themselves regulated through DNAme, as shown in Figure 2B.

EMT can be driven by different factors depending on the cellular context. For example, EMT in itself does not always lead to an increase in stemness; however, it can do so when driven by stable expression of pluripotency factors, such as *SOX2*. SOX2 is a developmentally important TF which impacts the progression of breast cancers. *SOX2* is not expressed in most normal adult tissues but is expressed in ∼43% of basal-like breast cancers [32]. *SOX2* expression is associated with an increased stem-like phenotype, increased invasiveness and metastatic potential, and therapy resistance [33,34]. Interestingly, Leis *et al.* suggest that the increased stemness through reactivation of *SOX2* in early stage breast cancer may explain the high tumor heterogeneity [35]. DNAme within the promoter and enhancer region functions as an epigenetic switch and forces cells into differentiation pathways. Stolzenburg *et al.* demonstrated that stable introduction of *de novo* DNAme in the *SOX2* promoter enabled long-lasting oncogenic repression in an MCF7 xenograft mouse model of breast cancer [7].

DNA methylation stabilizes the epigenetic profile of breast cancer

Aberrant DNAme patterns are implicated in the initiation, promotion, and progression of breast cancers. Breast cancers demonstrate DNA hypermethylation at promoter regions of tumor suppressor genes and hypomethylation in regulatory regions of oncogenes. Aberrant DNAme pat-
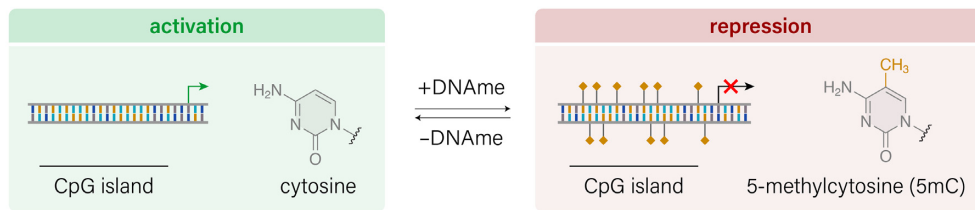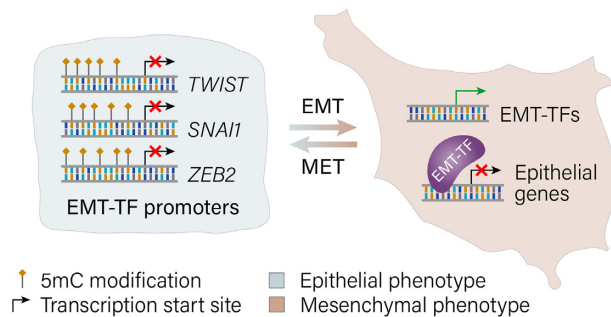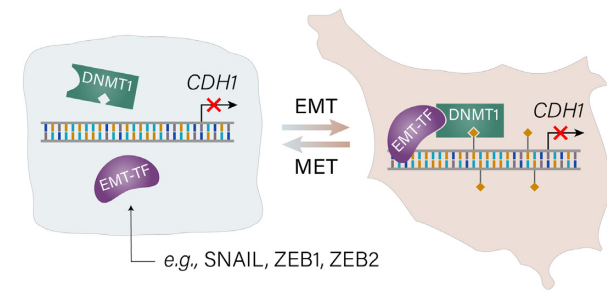
Figure 2. Key roles of DNAme in regulating gene expression in EMT. (**A**) DNA methylation of gene promoter CpG islands acts to repress gene expression. Methylation marks are deposited by DNA methyltransferase (DNMT), and removed by ten-eleven translocation (TET), families of enzymes. (**B**) Genes encoding EMT-TFs including TWIST, SNAIL (*SNAI1*) and ZEB2 are silenced by promoter methylation in epithelial cells. When these EMT-TFs are activated by EMT, they repress expression of epithelial genes, causing loss of cell adhesion and other epithelial phenotype characteristics. (**C**) EMT-TF expression can recruit epigenetic modifiers such as DNMT enzymes to methylate promoters of epithelial genes, resulting in stable and long-term repression. Adapted with permission from Elsevier: Skrypek, N., Goossens, S., De Smedt, E., Vandamme, N. and Berx, G. (2017) Epithelial-to-Mesenchymal Transition: Epigenetic Reprogramming Driving Cellular Plasticity. *Trends Genet. TIG*, **33**, 943–959 (31).

terns have also been associated with the acquisition of drug resistance, which is a major obstacle in breast cancer treatments and accounts for approximately 90% of unsuccessful treatments (36,37). Within human genomes, methylation predominantly occurs at cytosine (C) followed by a guanine (G) residue, termed CpG dinucleotides (38). CpG dinucleotides themselves are low in density throughout the genome, but cluster together in what are known as CpG islands within human gene promoter regions (38–40). Generally human promoter CpG islands are unmethylated, but can accumulate methylation patterns during development, or during other processes such as carcinogenesis (41,42).

Interestingly, DNA methylation can be directed through various stages of histone modifications. For example, histone methyltransferase EZH2 typically catalyzes trimethylation of histone H3 but can additionally control methylation through direct contact with DNA methyltransferases (43–46). In a similar way, EMT-TFs can also recruit DNA methyltransferases (Figure 2C). Over 100 genes have been observed to display promoter hypermethylation in breast cancers, with many of them involved in critical cell processes such as cell cycle regulation, genome integrity, apoptosis, metastasis and invasive potential (25,37,47,48).

Gene silencing through promoter hypermethylation is an important mechanism in the development of both carcinogenesis and tumor resistance. For example, promoter CpG island hypermethylation is observed in *BRCA1*; the wild-type *BRCA1* tumor suppressor gene is commonly silenced in nonfamilial breast and ovarian cancers (25) and is associated with increased cancer invasiveness and mortality (49). Additionally, *BRCA1* promoter methylation is found in 11–31% of sporadic breast cancers and 20–60% of sporadic TNBCs (50,51). In a further example, ER is a critical transcription factor regulating events important to mammary gland development, such as cell division (52). In general, high ER expression correlates positively with better clinical outcomes and recovery, but resistance is frequently acquired throughout the therapeutic course (36). ER expression may be lost through hypermethylation of the ER gene promoter (24) but can be reinduced in an ER-negative cell line (MDA-MB-231, TNBC) by inhibiting DNMT activity (53).

Summary

Gaining control over EMP and related epigenetic changes could lead to important new therapeutic approaches designed to limit resistance, metastasis and recurrence in breast cancer. Moving forward, researchers not only have an increased understanding of the complex regulatory roles and relationships between TFs but also new approaches that can be used for gene regulation. Both DNA secondary structure and an ever-expanding site-specific gene-modulation toolkit, through the Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated protein (CRISPR/Cas) genome engineering platform, can be used to modulate gene expression. Together, these approaches represent new ways to target the aberrant expression of TFs for breast cancer treatment.

NON-CANONICAL DNA SECONDARY STRUCTURES AS MODULATORS OF GENE EXPRESSION

Global chromatin architecture is tightly regulated at the epigenome level, where specific epigenetic marks promote relaxed or condensed chromatin states, in turn coordinating the transcriptional accessibility of genomic regions. The dysregulation of global epigenetic marks at both the DNA and histone level significantly disrupts this regulatory mechanism, promoting disease-associated aberrant transcript levels. Nucleosome-depleted regions can permit the formation of non-canonical DNA secondary structures which provide an additional layer of control over local DNA accessibility, modulating TF binding and transcriptional activity (54,55).

Non-canonical DNA conformations have revealed themselves as regulators of gene expression and disease progression, whether through direct or indirect effects on transcription. In this section, we will discuss non-canonical DNA secondary structures in the context of gene expression and their formation in cancer genomes, with a focus on modulation of TF expression and transcriptional activity.

G-quadruplexes and i-motifs are non-canonical DNA secondary structures

G-quadruplexes (G4s) and i-motifs are two highly dynamic non-canonical nucleic acid secondary structures which have emerged as important biological elements, with apparent regulatory roles over key processes such as replication and transcription (56–58). G4s and i-motifs can form at both the DNA and RNA levels, but the focus of this review will be at the DNA level. Since dysregulated TF–DNA interactions are key drivers in breast cancer progression, understanding the relationship between G4s, i-motifs, and TFs will help to uncover new therapeutic strategies for inhibiting EMP and modulating the epigenetic regulation of gene expression. Here, we discuss G4 and i-motif DNA secondary structures and their therapeutic potential within breast cancers.

G4 formation occurs within certain G-rich sequences containing several short tracts of contiguous guanine bases separated by intervening nucleotides (59–62). In contrast to the well-known Watson–Crick base pairing where a guanine base-pairs with cytosine (Figure 3A,B), four guanine nucleotides self-associate through Hoogsteen hydrogen bonding to form a G-tetrad (Figure 3C). Multiple G-tetrads π-stack upon each other to form a G4 with loops arising from intervening nucleotides between adjacent G-tracts (Figure 3D) (63,64). G4s can be intramolecular where they form from a single DNA strand, or intermolecular where they consist of two or four separate DNA strands (56,57). In contrast, i-motif structures form within certain C-rich sequences through hemi-protonated cytosine–cytosine, or C···C$^+$, base pairings (Figure 3E). These hemi-protonated cytosine–cytosine base pairs intercalate upon each other resulting in the i-motif structure (Figure 3F) (65–67). Like G4s, i-motifs can be intra- or inter-molecular, arising from single or two/four separate DNA strands, respectively. In most cases, G4s and i-motifs can be resolved by helicases during processes such as DNA replication and transcrip-

tion, and their stability varies depending on the extent of base stacking and intervening loop lengths. While G4 stability can be affected by type and concentrations of cations, i-motifs are more sensitive to pH alterations (68,69). In general, i-motifs have been much less studied within the cellular context than G4s (70,71).

Many oncogene promoters contain sequences capable of forming G4s or i-motifs. A greater abundance of G4s is observed in various cancer cell types compared to their non-tumorigenic counterparts (72,73), with G4 formation occurring at a higher density in promoters of actively transcribed genes, nucleosome depleted regions and being colocalized with RNA polymerase II and trimethylated histone (H3K4me3), which is a histone modification associated with active genes (74). Of significant interest to breast cancer are regulatory G4 and i-motif elements within the promoter regions of key oncogenes including *BCL-2* (56,75–79)*, c-MYC* (77,80–84), *KRAS* (85–88), *hTERT* (89–92), *PDGFR-β* (93,94), *c-KIT* (95–102), *RET* (103,104) and *VEGF* (105). Interest in the biological relevance of DNA secondary structures initially arose due to computational studies, which demonstrated sequences with the propensity to form G4s or i-motifs are overrepresented within regulatory regions of the human genome (59–62,69,106). Numerous studies have attempted to determine the mechanisms by which promoter G4 or i-motif structures shape gene expression, but no hard-and-fast rules have yet emerged. Multiple factors need to be considered such as positioning of the G4 on the template versus non-template strand and overlap with TF binding sites (107–111).

Promoter G4 structures were initially viewed as suppressive elements, physically blocking the accessibility and progression of transcriptional machinery (112,113). However, the first G4-ChIP-seq in human cells in 2013 found a strong association of endogenous G4s with the promoters of highly transcribed genes, linking G4s with increased transcriptional activity (74,114). Furthermore, a recent study demonstrated that endogenous G4s coincide with prominent TF binding sites in human chromatin, particularly in the promoters of highly transcribed genes (111). This finding was consistent across two different cell lines with distinct G4 landscapes, suggesting that G4 binding is a general property of certain TFs (115). However, a recent study using the G4-CUT&Tag method demonstrated that ligand-induced G4-stabilization actually reduced transcriptional activity by inhibiting binding of TFs (116). Thus, while endogenous G4s have been previously associated with enhanced transcription, G4-stabilizing ligands may actually downregulate or suppress transcription (116). Further investigation to see if this pattern is consistent over different ligand concentration, ligands, G4s and genes will be an interesting future direction.

Under Watson–Crick base pairing rules, G-rich regions of the genome are complemented by C-rich regions on the opposing DNA strand, and vice versa. Sites of G4 and i-motif formation should therefore exist at the same genomic loci on opposite strands, with an apparent potential for these two structural motifs to form simultaneously. However, increasing evidence indicates this does not occur (81,117–118). G4 and i-motif formation are altered during cell cycle progression, with G4 formation most prevalent in
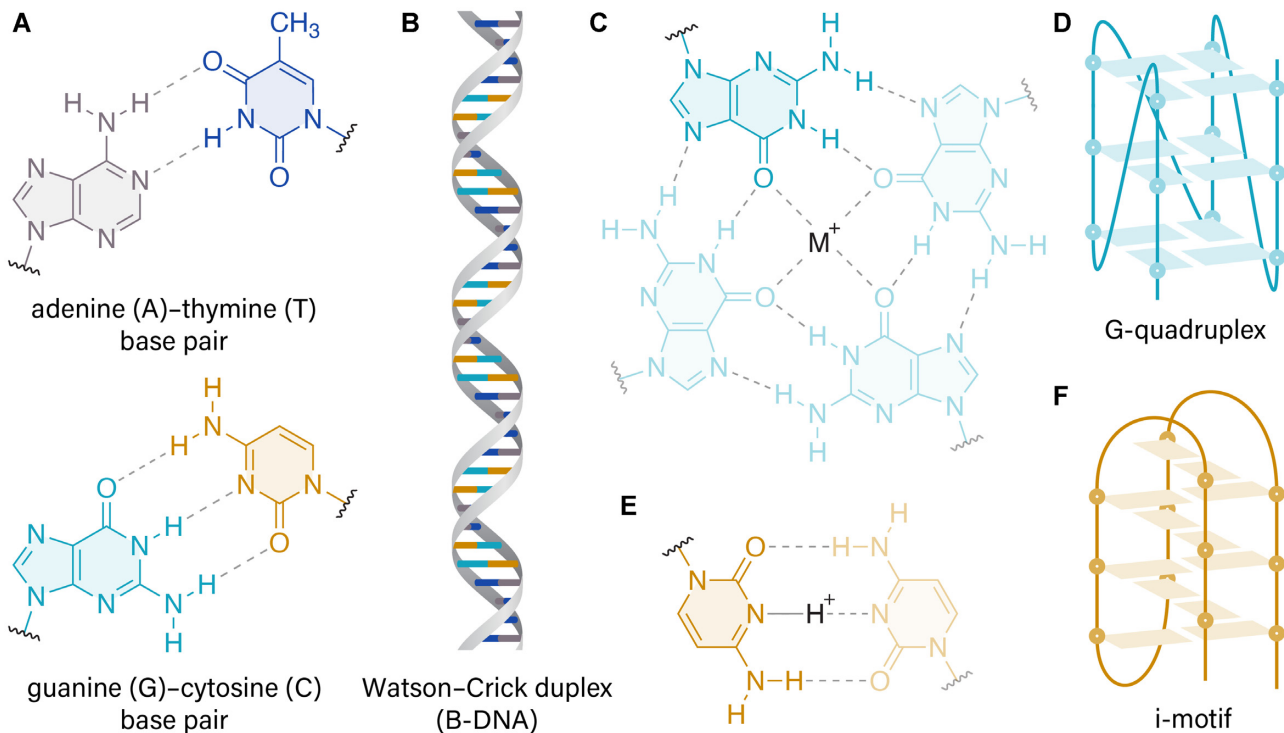
Figure 3. Schematic representation of DNA secondary structures. (**A**) Adenine–thymine and guanine–cytosine canonical Watson-Crick base pairing and (**B**) schematic representation of B-form DNA. (**C**) Planar G-tetrad structure formed through the self-association between four guanine bases via Hoogsteen hydrogen bonding, with a central monovalent cation shown within the center of the tetrad structure. (**D**) Schematic representation of an intramolecular G-quadruplex structure containing three stacked G-tetrads. (**E**) Hemi-protonated cytosine–cytosine base pairing, highlighting the N3 protonation of the cytosine ring. (**F**) Schematic representation of an intramolecular i-motif structure, through the self-association of several intercalated C···C^+ base pairings.

S phase, and i-motifs peaking at late G1 phase (70,114). Using molecular tweezers to measure torsional stress and extension forces, Sutherland *et al.* demonstrated that the *c-MYC* promoter sequence forms mutually exclusive G4 and i-motif structures in double-stranded DNA (81). It has been proposed that these complementary structures may act as a 'molecular switch' for gene expression (56,81,86). The observation that global stabilization of G4 structures in human cells reduces the abundance of i-motifs, and vice versa (71), indicates that the molecular switch mechanism may be a genome-wide phenomenon. Together, these findings point towards a biological 'on/off' switch which could potentially be exploited with future therapeutics.

DNA secondary structures contribute to genomic instability

Increasing evidence suggests that DNA secondary structures may contribute to aberrant oncogenic TF levels as well as influencing epigenetic disease states (54–55,119–121). The aberrant formation of G4 structures can result in epigenetic instability due to hindered biological processing of DNA. For example, unresolved G4 structures can physically block transcription, result in the accumulation of DNA replication errors, or cause double-strand breaks (122–125). Sites of highly stable DNA secondary structure formation may therefore act as mutational hotspots. Mutations localized at G4 formation sites within promoter regions could further influence cancer susceptibility by altering transcriptional activity, protein interac-

tions, DNA secondary structure formation and epigenetic patterns (89,123,126–130). Mapping the genome-wide occurrence of G4 structures using quantitative G4-ChIP-seq in 22 patient-derived tumor xenograft breast cancer models revealed differential G4 formation sites were significantly enriched in single-nucleotide variants (SNVs). SNV enrichment highlights a potential role for G4s in the development of breast cancer point mutations (131). Differential G4 formation sites were also enriched in the promoters of highly amplified genes and correlated with differential TF binding sites and differential TF expression levels. Interestingly, the 45 most common breast cancer driver regions characteristic for copy-number aberration-induced gene expression alterations were also linked to such G4-forming and TF-binding enriched sites. Further research into the link between DNA secondary structure formation and mutation may provide insights into how early cancer cells acquire disease-associated states.

G-quadruplex and i-motif promoter structures with therapeutic potential in breast cancer

It is interesting to note that each of the six hallmarks of cancer has been previously associated with oncogene promoter G4 or i-motif structures: self-sufficiency (*c-MYC* and *KRAS*), limitless replication (*hTERT*) (89–92,132), evasion of apoptosis (*BCL-2*) (56,77–79,133,134), sustained angiogenesis (*VEGF*) (105,135–139), invasive and metastatic potential (*PDGFA*) (140) and insensitivity to anti-growth sig-

nals (*pRb*) (66,141). For a more comprehensive review of this topic we direct the reader to Brooks *et al.* (66). The links between some important examples of promoter G4-/i-motif-mediated transcriptional control in the context of potential therapeutics for breast cancers are presented below.

Links between DNA secondary structure and cell proliferation genes. Breast cancer cells can acquire the ability to continuously proliferate through various molecular pathways. Genes including *hTERT*, *HER2* (*ERBB2*), *HRAS*, *KRAS* and *c-MYC* confer increased proliferative capacity in breast cancer (142,143). Each of these genes has been linked to silencing in cancer through DNA secondary structure formation, as discussed below. In particular, it is well established that the protein products of the *RAS* family (144) and *c-MYC* (145) are essentially undruggable due to a lack of targetable domains within the protein.

Telomerase, the enzyme essential for maintaining telomere length, is usually silenced in somatic cells but is expressed in ~90% of cancers, including breast cancers (146). The *hTERT* core promoter region contains two end-to-end stacked G4 structures connected by a hairpin loop that acts as a silencing element (89,90). The G4-forming region of *hTERT* is mutated in many cancers resulting in loss of G4 formation and subsequent *hTERT* activation. Kang *et al.* and Song *et al.* demonstrated that small molecule ligands which refold the hTERT promoter G4, even in the presence of these mutations, result in transcriptional repression of hTERT and cancer cell death (89,90).

HER2 is overexpressed in ~25% of human primary breast cancers across all subtypes. The HER2 promoter contains a G4-forming sequence which is the binding site for several TFs (147). When folded, the *HER2* promoter G4 element acts to block transcription resulting in repression of *HER2* (148). Downregulation of *HER2* levels is a potential therapeutic target and has been demonstrated within breast cancer cells via a luciferase reporter assay whereby stabilization of the *HER2* promoter G4 structure downregulated *HER2* expression at both the mRNA and protein levels (148).

Both the *KRAS* and *HRAS* promoter regions contain stable G4 and i-motif structures. G4 formation represses expression of each of these genes (85,86,149–152). In *KRAS* and *HRAS*, promoter G4s colocalize with binding sites for TFs such as MYC Associated Zinc Finger Protein (MAZ) and Specificity Protein 1 (SP1) (109,153). Binding of MAZ activates expression of *KRAS* and *HRAS* (152,154). Interestingly, the *HRAS* oncogene promoter contains two neighboring suppressive G4s, and both of which can be resolved by MAZ, restoring *HRAS* expression (150,151). Similarly, i-motif formation in *HRAS* and *KRAS* recruits heterogeneous nuclear ribonucleoprotein (hnRNP) family members hnRNP A1 and hnRNP K respectively, which unfold the i-motif structure resulting in increased oncogene expression (86,151). Breast cancers tend to lack the *RAS* promoter mutations which are commonly observed in other cancer types (155). However, the above TF-regulated G4/i-motif molecular switch mechanism influencing *KRAS* and *HRAS* expression can rationalize *RAS* activity in the absence of mutation. Accordingly, destabilization of hnRNP protein/i-

motif interactions or stabilization of promoter G4 structures has been reported to significantly inhibit *HRAS* and *KRAS* expression (85,154,156).

c-MYC is the primary oncogenic driver of cancer gene expression programs in a broad spectrum of cancer types, and is an inducer of EMT (145). The G4- and i-motif forming nuclease hypersensitive element III (1) (NHEIII$_1$) of the *c-MYC* promoter region is the principal regulator of *c-MYC* expression (80). The interplay between single-stranded DNA binding proteins and structure-specific resolving and/or stabilizing proteins determine structure formation capabilities pivotal for downstream *c-MYC* expression levels (81). Sutherland *et al.* demonstrated how the TF SP1 influences *c-MYC* expression through interactions between i-motif structure, hnRNP K and nucleolin (81). Ligand-based stabilization of the *c-MYC* G4 or i-motif has been shown to downregulate *c-MYC* expression (80,157).

Interplay between DNA secondary structures and EMT-TFs in breast cancer. Two crucial EMT-TFs, namely ZEB1 (*ZEB1*) and SNAIL (*SNAI1*), have the ability to form G4 structures within their respective gene promoter regions (158,159). Through recent independent studies, both the *ZEB1* and *SNAI1* promoter G4 elements have been shown to suppress transcription, whereas high expression levels of each of these genes were associated with a lack of G4 formation. Together, these studies provide some initial insight into relationships between EMT-TFs and promoter DNA secondary structures. Wang *et al*. demonstrated that G4 formation in the *SNAI1* promoter acts to repress expression and found a high number of mutations in this G4 region which affect G-tract lengths or distributions, potentially relating changes in *SNAI1* expression to changes in G4 formation (158). While the *SNAI1* promoter G4 has yet to be targeted with small molecules, stabilization of the *ZEB1* G4 structure was effective in downregulating *ZEB1* expression levels and inhibiting cell migration (159).

TFs such as hypoxia inducible factor 1 alpha (HIF1α) and MYB (encoded by the proto-oncogene *c-MYB*) are both strongly associated with the upregulation of EMT-TFs (160–162). The levels of both HIF1α and c-MYB TFs are mediated via suppressive G4 formation within their corresponding gene promoters (163,164). Stabilization of a *HIF1a* promoter G4 inhibits gene transcription by disrupting binding of the transcription factor AP2 to the promoter G4 (163), while G4 formation in the *c-MYB* promoter alters TF biding, where a complex between the G4 and the MAZ TF downregulates *c-MYB* expression (164). G4-mediated expression of both the c-MYB and HIF1α TFs therefore represent attractive targets for future research.

DNA secondary structure can modulate TF activity. Spiegel *et al.* and Lago *et al.* recently demonstrated an enrichment of TF binding at G4 forming sites within promoters across the human genome (111,115), which correlated with increased transcriptional activity when compared to promoter regions void of G4 structures. In fact, the interaction between DNA secondary structures and TFs is bidirectional. Just as G4 and i-motif formation affects TF binding, TFs themselves can also influence G4 and i-motif structures by unwinding, or by binding to DNA and preventing struc-

tures from forming (86,115,122,151,165). TF-governed formation of DNA secondary structures is, therefore, highly susceptible to aberrant TF levels associated with breast cancer progression. It is evident that a complex relationship exists between TF-mediated G4 stability and G4-dependent TF binding.

TF recognition of G4s or i-motifs may provide a unique opportunity for modulation of gene expression. For example, manipulation of TF binding has been demonstrated in relation to the i-motif-dependent hnRNP K regulation of *KRAS* expression, where the inhibition of i-motif formation prevented hnRNP K binding (86). Similarly, the stabilization of the *HIF1α* promoter G4 structure results in the displacement of AP-2 (activator protein 2) TF binding, leading to the subsequent reduction in *HIF1α* expression (163). The ability to modulate TF binding in such a precise manner is of interest when targeting promiscuous TFs, which often display both cancer promoting and cancer suppressing activities at different genomic loci. One example is SP1, which can act to dysregulate multiple target genes within breast cancer cells (83,115). Therefore, modulation of SP1 activity would therefore require locus-specific precision.

DNA–protein interactions play a crucial role in the coordination of chromatin states and distal interactions, an associated yet additional layer of genomic regulation on top of transcriptional activity (166). The observation that G4 and i-motif structures are found within nucleosome-depleted genomic regions suggests that these structures may influence chromatin dynamics (74). Therefore, understanding the mechanisms influencing the binding abilities of the previously mentioned TFs, in addition to chromatin regulating proteins, is an area of increasing interest. *ZEB1* repression via promoter G4 stabilization provides one recent example of the relationship between chromatin remodeling proteins and promoter G4 structures. Stabilization of the *ZEB1* promoter G4 displaces nucleolin, which impedes SP1 co-factor binding, preventing the further recruitment of the histone acetyltransferase P300. This cascade alters the promoter's architecture by shifting the chromatin into a closed (repressive) state, modulating transcriptional activity (159). A similar relationship has been observed for lysine-specific demethylase 1A (LSD1), which displays G4-mediated recruitment to the G4-bound TFs, telomeric repeat binding factor 2 (TRF2) and nucleoside diphosphate kinase 2 (NME2). G4-mediated LSD1 recruitment within the promoter regions of *p21* and *hTERT* promotes repressive histone marks resulting in gene silencing (167,168).

DNA sequences with the propensity to fold into secondary structures are implicated in DNAme density and deposition patterns, providing another mechanism by which G4 and i-motif structures impact the organization of chromatin (169,170). DNMT3A, DNMT3B and DNMT1 have high binding affinities for the 3D structure of folded G4s (171–173). Upon binding to a G4, DNMT1 loses enzymatic activity and physically blocks the surrounding DNA sequence from DNAme, thereby acting as local epigenetic regulatory element (171). Generally, G4 formation is negatively correlated with DNAme (173,174). Interestingly, however, breast cancer-associated hypermethylated sites display a threefold increase in G4-forming sequences. This observa-

tion suggests that while DNAme at sites of G4 formation is not biologically conserved, disease progression may act to alter this relationship (175). Comparison of DNAme patterns of 44 different i-motif forming sequences within a human breast cancer cell line (MCF7) and a non-tumorigenic mammary epithelial cell line (MCF10A) indicated 27% of i-motif sequences contained at least one 5mC, and of these, 83% were differentially methylated in MCF7 *vs* MCF10A, highlighting the differences in methylation between cancer and non-cancerous cell types (176).

At the molecular level, G4s tend to exhibit higher thermal stability when a DNAme modification is present within the sequence (177,178). Similarly, i-motif structures also tend to exhibit higher stability when a DNAme modification is present (176,179,180). However, the extent of methylation within an i-motif forming sequence has also been shown to be important, with hypermethylation destabilizing certain i-motif structures, such as the c-MYC i-motif (179). Furthermore, the position of the DNAme modification is also critical for the stability of certain i-motifs, such as the human telomeric hTelo i-motif (176). Unsurprisingly the change in structure stability is associated with altered binding capabilities of proteins. For example, aberrant DNAme modifications in the G4 forming sequence within the first exon of the hTERT oncogene enhances G4 stability, which in turn, impedes binding of the highly conserved CCCTC-binding factor (CTCF) TF, resulting in the oncogenic upregulation of hTERT expression (181).

In addition to chromatin state, G4-protein binding has been implicated in promoter-enhancer interactions, allowing chromatin looping either through G4-TF binding, or via long-range split G4 sequences that when facilitated by DNA looping can come together to form G4 structures. G4-dependent looping has been shown to be facilitated by the direct binding of the TF Yin Yang 1 (YY1) to promoter G4s, allowing DNA looping to occur between two bound YY1 sites (182). Disruption of G4 formation through helicase overexpression or CRISPR-gene editing diminished YY1-mediated DNA looping, highlighting the critical role of G4s in DNA looping. Disruption of YY1–G4 interactions via G4-stabilizing ligands resulted in altered gene expression not only of promoter G4-harboring genes but also of promoter G4-lacking genes through the dissociation of YY1 DNA looping with distal G4 structures. Interestingly, the YY1 promoter sequence itself is thought to be under the control of G4 formation, which is resolved within cancer cells due to elevated levels of the G4-resolving helicase G4R1 (165).

Summary

While the exact, mechanistic roles of G4s and i-motifs in breast cancers remain to be determined, it is clear that both secondary structures contribute to the progression of cancer and EMT. Regulatory regions of oncogenes are enriched in G4 and i-motif structures and are implicated in aberrant gene expression patterns and genomic instability through physical blocking of transcriptional machinery, and through their interactions with various TFs, epigenetic modifiers and modifications. Overall, the evidence suggests that G4 and i-motif structures control gene expression pat-

terns via a molecular switch mechanism, creating a potential new avenue for targeting therapeutics directly to DNA.

TAKING CONTROL OF GENE EXPRESSION FOR BREAST CANCER THERAPY

Previous attempts at targeting aberrant TF expression have involved targeting protein–protein and/or protein–DNA interactions, which has been a difficult task for small molecules where the protein of interest is undruggable (i.e. lacks binding sites for small molecule interactions). Targeting TF expression at the genetic level, however, has recently become possible due to developments in the field of noncanonical DNA secondary structure targeting and CRISPR technology. These two approaches will be discussed below, along with a discussion of their current limitations.

Modulating DNA secondary structure formation with small molecules

The unique structures of G4 and i-motifs enable the design of small-molecule therapeutics with preferential selectivity for G4s or i-motifs over the more ubiquitous B-DNA. Targeting G4s or i-motifs in gene regulatory regions compared to enzymes or proteins allows us to target genes regardless of the 'druggability' and/or copy number of the gene product present. Small molecule therapeutics targeting either G4s or i-motifs in breast cancer models are summarized in Table 1; however, it is important to note that there are other promising drugs available that modulate DNA structure that have not yet been tested in the context of breast cancer, such as RG260 (90), and are reviewed elsewhere in respect to different cancers (183). These small-molecule drugs are designed to either stabilize or destabilize the target structures, which modulates gene expression. To date, there have been many more small molecules developed for interaction with G4s than for i-motifs, owing to the more recent discovery of i-motifs as physiologically relevant structures (57,65,112,184–186), so we have focused the discussion towards small-molecules targeting G4s.

All G4s have a common core of stacked G-quartets with a central ion channel (often stabilized by monovalent cations, such as Na^+ or K^+) and four grooves with varying dimensions which are determined by overall topology and loop geometries. G4s can differ in their topologies and loop geometries. Topology is based on the directionality of strands that make up the G-quartet core, classified as parallel, antiparallel or hybrid (187,188) (Figure 4A). The size and sequence of the loops determine the loop geometries. Loop geometries are categorized as propeller or double-chain-reversal loops, edgewise or lateral loops, diagonal loops, or V-shaped loops, as illustrated in Figure 4B (188,189).

Ligands that only target the G-quartet core show poor discrimination between different G4 structures, but selectivity can be improved by additional targeting of loops and grooves (190). However, design of a G4 ligand that demonstrates selectivity toward a single G4 still remains a major challenge. Overall, the majority of G4-binding small molecules share a few basic characteristics, possessing a heteropolyaromatic chromophore that is usually planar to enable π–π stacking interactions with a terminal G-quartet

and one or more flexible side-chains with substituents featuring a cationic charge for interacting with the quadruplex grooves and loops (65,112,184,185). The majority of G4-targeting ligands stabilize the G4 structure and only a few ligands act to destabilize G4s (95,191–195). An example of a G4 structure interacting with PhenDC3 is depicted in Figure 4C,D; small-molecule therapeutics may stack upon the end of a G4 structure and/or interact with the distinctive loop structures of G4s (196).

There are several ways in which therapeutic targeting of DNA secondary structures can impact genomic stability and transcriptional activity. One of particular interest arises from the observation that genome-wide G4 or i-motif stabilization can result in the accumulation of DNA damage, leading to transient heterochromatin formation and transcriptional repression while damage is repaired (197,198). Breast cancer cells that have a compromised ability to repair DNA damage or properly unwind DNA during replication can be particularly sensitive to the stabilization of DNA secondary structures compared to their non-malignant counterparts. For example, cells with impaired homologous recombination repair pathways (which includes cells deficient in BRCA1 and BRCA2) display higher levels of genomic instability, cell cycle arrest and apoptosis following stabilization of G4 structures (199). In this scenario, stabilization of G4s blocks replication folds and induces single-stranded DNA gaps or breaks, where failure to repair this damage through the BRCA and non-homologous end joining pathways leads to cell death. Compound CX-5461 achieves specific toxicity against BRCA-deficient cancer cells through this mechanism. Importantly, CX-5461 resulted in dramatic tumor regression in polyclonal patient-derived xenograft models, including a TNBC tumor already pretreated heavily with other therapeutics (including platinum drugs), and a TNBC tumor derived from a patient who had been pretreated with anthracycline/taxane, and whose metastatic disease had minimal response to cisplatin. CX-5461 ultimately enables a new therapeutic option for aggressive cancers with BRCA deficiencies, including those resistant to PARP inhibition, and is in clinical trials (200,201).

Similarly, cells deficient in G4-unwinding helicases such as Fanconi anemia group J protein (FANCJ), which is commonly mutated in breast cancers, also prevents cells from resolving G4 structures, making G4-helicase deficient cells sensitive to secondary structure stabilizing treatments (122,202). Alternatively, G4-stabilizing molecules may increase susceptibility and enhance treatment efficiency when combined with traditional breast cancer chemotherapeutics. As an example, doxorubicin-resistant breast cancer cells can regain doxorubicin sensitivity after the addition of G4-stabilizing molecule 360A (167). Interestingly, Guilbaud *et al.* demonstrated that chromatin modifications which occur due to the stabilization of G4s can be inherited and persist even after the removal of the G4 stabilizing molecule (121), which suggests that transient DNA structure stabilization may be sufficient for a sustained therapeutic effect.

One of the major issues with traditional small molecule therapeutics for cancer treatments has been the severe off-target and side effects. Naturally, this also raises concerns about small-molecule therapeutics for targeting DNA sec-

Table 1. Examples of G4 and i-motif targeting using small molecules with therapeutic potential in breast cancer

Compound	Secondary structure	Impact on gene expression and/or cellular fate	Reference
360A	G4 stabilizing	Reactivation of *p21* expression, increased sensitivity to doxorubicin in breast cancer cells	(167)
Pyridostatin	G4 stabilizing	Activation of innate immune genes in breast cancer cells	(247)
PhenDC3	G4 stabilizing	Activation of innate immune genes in breast cancer cells	(247)
RHPS4 and derivatives	G4 stabilizing	Inhibits telomerase activity, induces damage foci at telomeres and apoptosis in breast cancer xenografts	(248–250)
GTC365	G4 stabilizing	Represses *hTERT* expression, decreased telomerase activity and telomere length, promoted cell death/senescence in breast cancer cells	(89)
BRACO-19	G4 stabilizing	Reduced telomerase activity and telomere length leading to cellular senescence in breast cancer cells	(89,251–254)
TGP18	G4 stabilizing	*BCL-2* downregulation, marginal inhibition of TNBC tumor growth in mouse xenograft breast cancer model	(255)
CX-5461	G4 stabilizing	Inhibits replication forks and induces DNA damage, resulting in cell death in BRCA-deficient TNBC patient derived xenograft models, currently in clinical trials	(200,201)
Cβ	G4 stabilizing	Represses *HER2* transcription and expression in breast cancer cells	(148)
QN-1	G4 stabilizing	*c-Myc* downregulation, cell cycle arrest and apoptosis, tumor suppression in TNBC mouse model	(256)
BTC f	G4 stabilizing	*c-Myc* downregulation demonstrated in hepatocellular carcinoma cells, inhibited growth in breast cancer cells	(257)
Pyridyl bis-prolinamide ligand 1	G4 stabilizing	Simultaneous inhibition of *c-Myc* and *BCL-2* expression, leading to cell cycle arrests, apoptosis and DNA damage in breast cancer cells	(258)
BMVC4	G4 stabilizing	*c-Myc* downregulation and telomerase inhibition in non-small cell lung carcinoma cells, inhibited growth in breast cancer cells	(259)
tanshinone IIA derivative 4	G4 stabilizing	Repression of c-*Myc*, *KRAS*, *VEGF-a/b*, *BCL-2* and *Tel-26* expression, suppressed metastasis and angiogenesis of TNBC cells	(260)
PBP2	G4 stabilizing	*BCL-2* downregulation, increased apoptosis in breast cancer cells	(261)
NiP	G4 stabilizing	Telomere uncapping, DNA damage and degradation of telomeric 3′ overhang, promoted apoptosis of breast cancer cell stem cells, reduced expression of stem cell markers	(262)
TMPyP4	G4 stabilizing	Decrease in *hTERT* expression, inhibition of telomerase activity, suppressed growth rate of breast cancer xenograft model	(263,264)
3,8,10-trisubstituted isoalloxazines	G4 stabilizing	Downregulation of *c-kit* gene expression in breast cancer cells	(95)
IMC-48	i-motif stabilizing	Upregulation of *BCL-2* in breast cancer cells	(56)
IMC-76	i-motif destabilizing	Downregulation of *BCL-2* in breast cancer cells	(56,76)
PBP1	i-motif stabilizing	*BCL-2* upregulation, decreased apoptosis in breast cancer cells	(261)

ondary structures. In general, the selectivity in targeting DNA secondary structures in cancer cells stems from the (i) the high metabolism and protein synthesis requirement in cancer cells (203), and (ii) the observation that the majority of small molecules do not induce G4s, but stabilize the ones which have formed (which occur at a higher level in cancerous cells, making these cells more susceptible). Together, these two general methods, along with the ability of healthy cells to unwind structures and repair DNA damage, seems to convey a level of specificity. Whether this is enough is yet to be answered. To reduce off-target effects, molecules targeted for G4 and i-motif interactions must have high selectivity over binding to duplex DNA and RNA, which would lead to nonspecific toxicity. This requirement seems to be readily achievable through molecular design, where selectivity for the larger surface area of terminal G-quartet increases with ligand size and overall number of side chain substituents. Further, the design of macrocyclic and crescent-shape molecules inherently show low duplex-binding ability due to the physical shape which prevents them from binding in duplex DNA intercalation sites (187). Selectivity for a specific secondary structure in relation to the control of a specific gene is more difficult, and while there has been some promise in demonstrating pref-

erence over one specific structure or conformation (such as preference of the Na^+-stabilized G4 over the K^+-stabilized G4 in the human telomeric sequence (204)), so far there are no predictive rules to govern molecular design. However, targeting a specific G4 may not be necessary, instead, it may be beneficial to target multiple G4-mediated pathways concurrently. The G4-stabilizing ligand triarylpyridine 20A affects several biological pathways with multiple G4s involved, which suggests multiple G4-dependent inhibitory effects (205). An alternative method of targeting at the gene level, which does enable controlled targeting of specific genes, is CRISPR/Cas technology.

Targeted regulation of transcription factors using CRISPR/Cas systems

The ability to artificially modulate gene transcription could enable the reestablishment of normal gene regulation in diseased states, and program cell fate, behavior and tissue function. The ability to reprogram cells is fast becoming a relatively easy and accessible reality though CRISPR-based technology, and its quickly-expanding toolbox. CRISPR, standing for clustered regularly interspaced short palindromic repeat, together with CRISPR-associated proteins
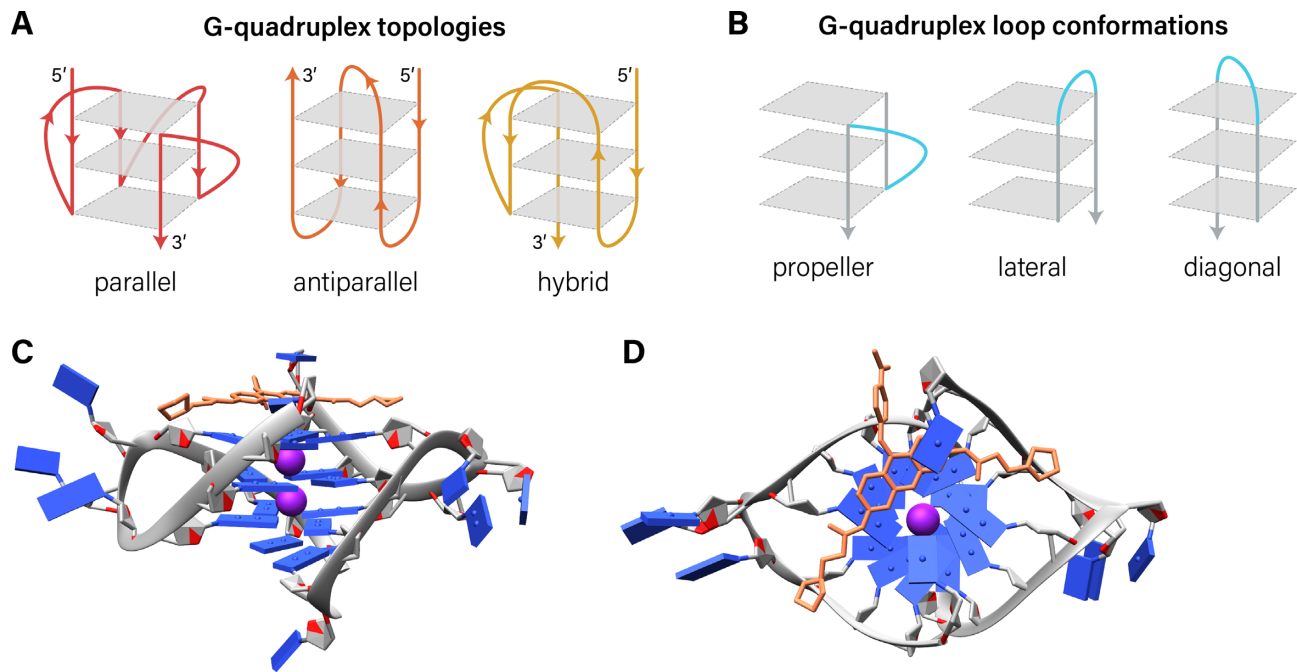
Figure 4. G-quadruplex topologies, loop conformations and example ligand binding modes. (**A**) G-quadruplex topologies for an intramolecular. (**B**) Loop conformations, highlighted in blue. (**C**) PhenDC3 ligand, shown in orange, interacts with a quadruplex via end-stacking and the grooves of this parallel G-quadruplex, PDB code: 3CE5. (**D**) Top-down view of the same structure as in (C).

(Cas), are effectively an RNA-guided 'homing' device which can be used to target nearly any sequence of DNA with high specificity. The most widely used CRISPR/Cas system is CRISPR/Cas9, specifically the *Streptococcus pyogenes* Cas9, however to date there are 6 different CRISPR/Cas types, and at least 29 subtypes (206). Wild-type CRISPR/Cas9 is an endonuclease, and can be used to cut DNA and facilitate gene editing through either non-homologous end joining (NHEJ) or homology-directed repair (HDR). Modification of the Cas9 protein to be catalytically dead results in deactivated Cas9 (dCas9), which can then be fused to an array of transcriptional regulators, or epimodifers, to directly modulate the expression and/or context of virtually any endogenous gene. For example, fusion of the CRISPR/Cas protein with DNA methyltransferase enzymes DNMT3A and DNMT3B can be used to catalyze *de novo* DNA methylation, while fusion with ten-eleven translocation (TET) proteins can be used to specifically achieve DNA demethylation. Additional fusions can be made for transactivation (fusion with VP64, VPR), transrepression (KRAB), and locus-specific histone modifications, such as acetylation (p300), deacetylation (HDAC3), methylation (PRDM9) and demethylation (LSD1), to name a few examples (207). Adding to this, while first generation systems involve the fusion of a single effector domain directly to the dCas9 protein, the development of second-generation systems has enabled amplified modification through recruitment of multiple effector copies, in addition to spatial and temporal control of the modifications (207–209). First and second-generation systems are illustrated in Figure 5.

One particularly advantageous feature of CRISPR/Cas technology for this type of application is the ability to multi-plex. 'Multiplexing' is the term given when numerous guide RNAs (gRNAs) or Cas enzymes are expressed simultaneously. Using multiple gRNAs to target one gene can often result in a synergistic effect, in addition to decreasing the level of off-target effects. While the use of multiple gRNAs to target different genes at once may enable more robust regulation of the complex gene expression patterns, and enable more efficient reprogramming. Currently, multiplexing has largely been used to identify and understand cell transformation processes (210,211), and has garnered only limited attention as a therapeutic strategy. Saunderson *et al.* demonstrated *de novo* DNAme by multiplexing to target a series of genes commonly methylated in breast cancer, using CRISPR/Cas9 fused to both DNMT3A and DNMT3L. This approach enabled long-term hypermethylation and gene silencing of *CDKN2A*, *RASSF1*, *HIC1* and *PTEN* in primary human myoepithelial cells. They were able to initiate aberrant cellular processes, increasing our understanding of DNAme-driven early changes in breast cancers (210). Finally, multiplexing by using several types of Cas enzymes can lead to robust epigenetic changes and gene regulation. For example, O'Geen *et al.* investigated requirements to achieve persistent repression via CRISPR/dCas9 transcriptional repression and epigenetic modifications. They demonstrated that while long-term repression (14 days) could be achieved using a combination of CRISPR/dCas9-KRAB and DNA methylation, locus-specific histone methylation through CRISPR/dCas9-EZH2 is required for robust, long-term silencing (212). Considering the complexity of aberrant gene expression networks, there has nevertheless been promising progress that may lead to future therapeutics for the treatment of breast cancers.
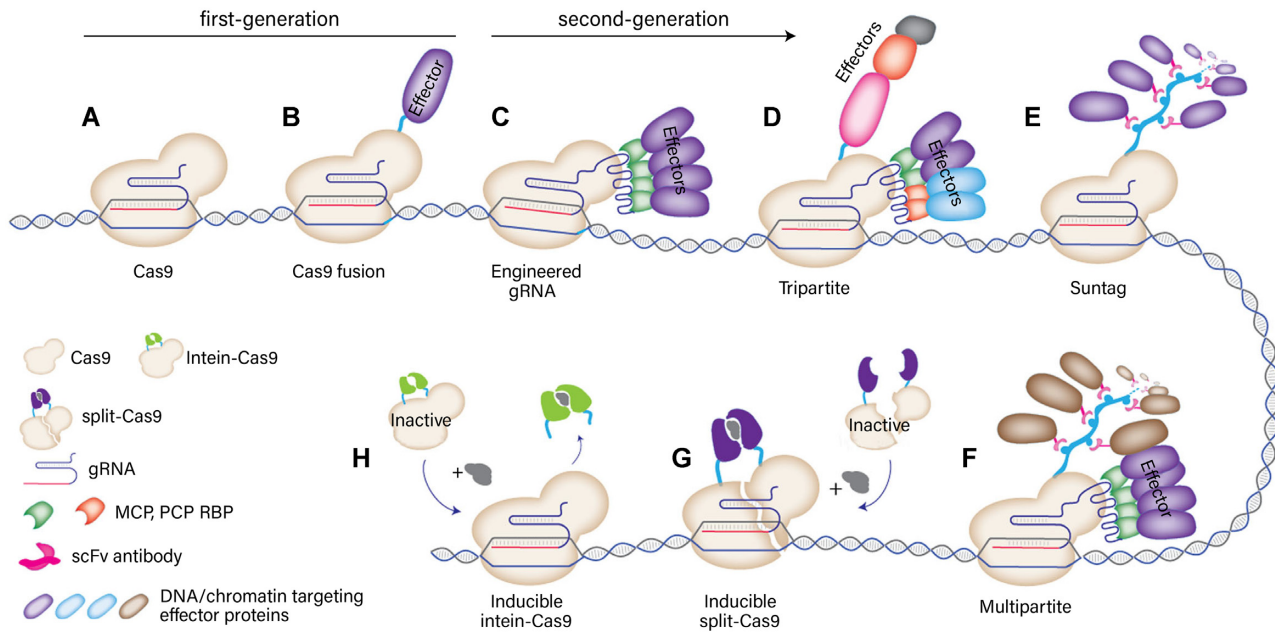
Figure 5. Schematic of the CRISPR/Cas9 toolbox. First-generation Cas9 systems include the (**A**) unmodified Cas9 system, and (**B**) the Cas9 enzyme fused to single effector proteins. Second-generation Cas9 systems (**C–H**) have evolved to enable either higher levels of modification, or more sophisticated mechanisms of controlling modification. The principle behind constructs aimed for higher levels of modification is to recruit multiple copies of effector proteins, which can be achieved through (C) engineering gRNA scaffold to include RNA aptamers (engineered sgRNA), (E) binding effectors along a repeating peptide array (Suntag), or (**D** and **F**) combinations thereof (tripartite and multipartite). Chemically inducible systems enable temporal control over the activity of the Cas9 fused effector proteins. (**G**) In the split-Cas9, each half of the Cas9 protein has to be induced to form the full, functional complex, while the intein-Cas9 approach (**H**) requires a protein segment to be excised after chemical induction. Adapted with permission from Adli *et al.* (2018) (207) under Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

Application of CRISPR/Cas therapeutics for targeted gene regulation in breast cancers. Interestingly, there are far fewer studies looking at the delivery of catalytically dead CRISPR/dCas technology as a potential therapeutic in comparison to the active nuclease. Perhaps this is due to the complexity of the relationships between epigenetics and gene expression, the intricate orchestra of TF expression relationships themselves, different dosage windows, or even due to gene editing being 'easier' to assess. Despite this, there have been several studies, both *in vitro* and *in vivo*, which do demonstrate the therapeutic potential of this approach. In 2016, Choudhury and coworkers achieved targeted demethylation of the *BRCA1* promoter with CRISPR/dCas9 fused to the catalytic domain of TET1. Their work demonstrated transcriptional upregulation of the gene after demethylation, in addition to a reduction in observed cell growth. The study was conducted using the MCF7 breast adenocarcinoma cell line *in vitro*, using Lipofectamine LTX as the delivery agent. This proof-of-concept study demonstrates that upregulation of BRCA1 expression exerts a significant inhibitory effect on cell viability, with or without additional chemotherapeutic agents (25). Huang *et al.* used a light-inducible genetic circuit using the dCas9 system fused to VPR a transactivation domain to activate the expression of either p53 or E-cadherin. They developed an AND logic gate, where the circuit requites presence of both the human telomerase reverse transcriptase (hTERT) and human uroplakin II (hUPII) promoters to be activated. In this example, activation of p53 and E-cadherin had a significant effect on reduction of cell

proliferation, invasion and apoptosis. While this particular study was conducted in a bladder cancer model, it serves as a good example of what can be achieved using dCas9 (213).

Rather than directly targeting TFs, which are notoriously difficult to drug, other studies have instead targeted upstream tumor suppressor genes. For example, *PTEN* (phosphatase and tensin homolog deleted from chromosome 10) is a critical tumor suppressor gene: *PTEN* expression negatively correlates with tumor size, pathological grade, and the expression of both ER and PR in breast cancer (47,214). PTEN expression is lost in up to 30% of sporadic breast tumors via DNA methylation of the PTEN promoter, and transcriptional repression of PTEN can lead to resistance to clinical treatments. Moses *et al.* used dCas9 fused to the transactivator VP64-p65pRta (VPR) to reactivate PTEN expression in triple negative breast cancer cell lines. This was achieved *in vitro* by first establishing cell lines stably expressing the dCas9-VPR system via lentiviral transduction, and then mixtures of gRNAs were transiently introduced using the commercial reagent Lipofectamine 2000. PTEN activation significantly repressed downstream oncogenic pathways, and suggested that combining CRISPR activation of PTEN with conventional small molecule inhibitors, such as PI3K/mTOR or B-Raf inhibitors, could increase the efficacy of tumor inhibition and limit potential of developing resistance (215). Liu *et al.* took a different approach and designed a polymeric nanoparticle for multistage delivery of CRISPR/dCas9 fused with a VP64 transactivator. The core construct of the nanoparticle includes phenylboronic acid (PBA) modified with low

molecular weight polyethyleneimine (PEI), which interacts and condenses with the CRISPR/dCas9 DNA. The shell was then formed by 2,3-dimethylmaleic anhydride-modified poly(ethylene glycol)-*b*-poly(lysine), which is designed to dissociate within the acidic tumor environment. miR-524 is a microRNA that has been found to be suppressed in several types of cancer, including breast cancers. Restoring endogenous expression of miR-524 can suppress the proliferation and metastatic potential of cancer cells. Liu *et al.* demonstrated suppressed tumor growth both *in vitro* and *in vivo,* in a TNBC model (216). Finally, work in our group has demonstrated tumor suppression and regression through the reactivation of silenced tumor suppressor genes *MASPIN* and *CCN6* in an MCF7 xenograft mouse model of breast cancer. Expression of *MASPIN* was reactivated using CRISPR/dCas9-VPR together with a synergistic activation mediator (SAM) complex multiplexed with 4 gRNAs, while *CCN6* expression was reactivated using CRISPR/dCas9-VPR targeted via 5 gRNAs. This study utilized a cationic polymer which had been designed for targeted delivery by way of a cyclic RGD peptide (217).

Disruption of secondary structures using Cas9 is also possible, albeit the field is in its infancy and currently aimed at understanding the regulatory role of structures, rather than application as a therapeutic. So far, the active Cas9 nuclease has been utilized to create mutations within G4-forming sequences. Huang *et al*. used this approach to disrupt the G4 structure involved in the regulation of the chloride intracellular channel 4 (*CLIC4*) gene, which has been found to have tumor-promoting properties. It was found that the transcription activity of CLIC4 decreased upon G4 disruption (218). Li *et al*. utilized Cas9 targeting G4 forming sequences to elicit the relationship between TF YY1-mediated DNA looping in a G4-dependent fashion. Disrupting the formation capability of specific G4 structures was found to result in the significant reduction of YY1 binding and DNA looping (182). Together, these examples begin to demonstrate the potential of combining the precise targeting capability of systems such as CRISPR and the unique genomic distribution and structure of G4/i-motif regulatory elements.

It is clear, however, that examples are somewhat limited, which may arise from several contributing reasons. Adding to the complexity of regulating gene expression patterns is that modulating gene expression is not always the only objective—the gene must be expressed in the correct context. For example, *OCT4*, a master transcription factor of pluripotency, can act as both a tumor suppressor gene or an oncogene, depending on the cellular context (219). MASPIN is another example, where localization of the expressed protein dictates its role as a tumor suppressor (220).

Hurdles in the administration of CRISPR/Cas.   Perhaps the biggest obstacle that CRISPR/Cas therapeutics face is the method of delivery (221). One of the main reasons for this issue is the size of the CRISPR/Cas construct, in addition to the requirement of simultaneous delivery of multiple components (222): the CRISPR/Cas, one or more short guide RNA (gRNA), and the effector domain (for epigenetic editing or transcriptional modulation) or template DNA, in the case of homology-directed editing. Transfecting large and multiple constructs results in a statistical distribution, where many cells do not receive all components (223). CRISPR/Cas therapeutics can be delivered in the form of RNA, DNA or protein, and can be delivered so as to achieve sustained expression, or as a transient 'hit and run' approach (224). These factors all affect the end dosage, the therapeutic window, and potential for off-target effects. Further, the target location within the body also impacts the choice of delivery method. For breast cancer, and other solid tumors, systemic delivery would be more clinically applicable than intracellular delivery methods such as microinjection or electroporation (222).

Gene delivery methods can be broadly divided into viral and non-viral delivery. Viral delivery involves the usage of engineered viruses, such as lentiviruses or adeno-associated viruses (AAVs), to deliver the therapeutic cargo. Viral delivery methods still face problems due to toxicity and immunological concerns (both initial and adaptive), limited intrinsic packaging capacity, and cost. Lentiviruses have a high carrying capacity of ~8 kb; however, they cause integration into the host genome which not only risks insertional mutagenesis but also causes long-term expression, which inherently increases the chances of off-target effects and toxicity when delivering CRISPR/Cas (225–227). AAVs are more favorable for gene delivery, as the chance of integration is low. However, AAVs struggle to encapsulate the CRISPR/Cas sequence as they are limited to a carrying capacity of ~4.7 kb. Smaller variants of the CRISPR/Cas have been developed (228), such as *Staphylococcus aureus* Cas9 (*Sa*Cas9). Nevertheless, for epigenetic or transcriptional applications, further space is required for the expression of the additional effectors, and so even the smaller orthologs can exceed the AAV carrying capacity (229). AAV-based delivery methods may be rendered ineffective by immune activation, as AAVs are common and 40–80% of adults will already have AAV antibodies. Therefore, patients often have to take immune suppressants in addition to the AAV therapeutic, and repeat dosing becomes a challenge, although there has been promising development of immune-orthogonal orthologues of AAV capsids to overcome this issue (230). Finally, AAV-based therapeutics hold the record for being the most expensive drugs available, and therefore are currently not applicable for treating something as widespread as breast cancer. Glybera, for example, an AAV-based therapeutic aimed at treating a rare inherited lipoprotein lipase deficiency, costs US$1.2 million per patient (231).

On the other hand, non-viral delivery encompasses all other methods of delivery, including physical methods such as electroporation, microinjection, hydrodynamic delivery and sonoporation, and synthetic delivery, such as with lipids, polymers and inorganic nanoparticles. While typically the delivery efficiency does decrease with large therapeutic cargoes, non-viral methods do not face a restricting size limitation, unlike viral delivery. All non-viral delivery methods are transient in nature, which lowers the chance of off-target effects (232). However, it also means that the therapy delivered must either be fast-acting, or create stable and inheritable changes within the cell. Some non-viral de-

livery methods such as highly cationic or PEGylated agents, have demonstrated potential toxicity or immunological responses (233,234), but the high variety in potential chemical design inherently enables enough flexibility to overcome these issues (227). The most limiting factor faced by nonviral delivery methods is simply that the delivery efficiency is not as high as their viral counterparts, especially when delivering to particular cell types, such as post-mitotic cells. Another challenge is achieving sufficient tumor targeting and penetration depth within highly dense and heterogeneous tumors, which has been a challenge even for promising delivery vehicles such as nanoparticles (235).

Summary

Targeting cancer progression at the genetic level has not been previously achievable but may be realized through the application of either small molecules that target DNA secondary structures or more flexible and advanced CRISPR/Cas9 technology. Small molecules capable of stabilizing G4 or i-motif structures show good promise as a therapeutic approach. However, the level of specificity that is attainable for these highly varied and ubiquitous structures, and only within target cells, is still uncertain. CRISPR/Cas9 technology may provide an alternative approach, enabling specific gene editing or regulation through transactivators/repressors and epigenetic enzymes. Offtarget delivery could potentially be minimized or avoided through sophisticated inducible Cas9 systems, and specific cell-type targeting. Since EMT-TFs act in concert, targeting just one gene at a time may not lead to robust gene expression and substantial phenotypic changes. Multiplexing to target multiple genes at once may solve this problem. Additionally, exploiting the heritability of epigenetic marks such as DNAme may assist in achieving a sustained effect even after transient treatment. One of the biggest limitations to the implementation of CRISPR/Cas9 technology is the difficulty of efficient and safe delivery.

CONCLUSION AND FUTURE OUTLOOK

Breast cancer is a highly heterogeneous and complex disease, with individual patients presenting unique profiles brought about by mutations, gene silencing or activation, and past therapy. Linking each of these profiles is anomalous gene expression, which can, in turn, alter TF interactions and in a feedback loop, alter the expression of other genes. This dysregulation of gene expression contributes to the hallmarks of cancer and processes such as EMT through alterations in TF binding, chromatin organization and epigenetic marks. Ultimately, disruption of normal patterns of gene expression contributes to recurrence and metastasis, which represent the principal cause of death in breast cancer patients today. Correcting and gaining control over aberrant gene expression therefore represents a highly attractive yet extremely complex therapeutic opportunity in breast cancer. However, as it stands for now, we still lack a coherent overview of the molecular and biochemical mechanisms inducing cells to enter various states along the epithelial–mesenchymal phenotypic spectrum and an improved understanding of the dynamic and plastic nature of the EMT program is required to fully utilize the therapeutic potential within this space.

As with all therapies, specificity will be key. Tight regulation of gene expression is crucial for normal cell processes and health, and so any therapeutic strategies to correct aberrant TF expression, disrupted TF binding, or abnormal secondary structure formation must not impact healthy cells and tissue. In this respect, small molecules designed to stabilize G4 and i-motif structures face a formidable challenge in terms of specificity and selectivity. With >700 000 potential G4 sites in the human genome and some 10 000 detected in chromatin (74), it will no doubt be challenging to design drugs that are highly specific for a given G4 or i-motif target. Furthermore, the complexity of TF signaling pathways complicates potential therapies designed to modulate DNA–protein interactions, as they may be highly interdependent. Although individual studies look at knockdown of individual TFs, further research to untangle the roles of individual TFs will be needed, which may require further development of multiplex or high-throughput methods such as CRISPR-based screening (236–239).

Targeting G4 stability through DNAme may present a potential alternative to small molecules. For example, regulating the formation of G4s or i-motifs may be possible through the utilization of CRISPR/Cas9 to methylate or demethylate particular loci. New therapeutic approaches will appear as our understanding of the relationship between DNA secondary structures, epigenetic modifications and gene expression increases, which may lead to robust gene regulation and EMP 'resetting' of carcinoma cells through CRISPR/Cas technology. However, before CRISPR/Cas technology can be utilized as a cancer therapeutic, the development of non-viral gene delivery is necessary and multidisciplinary collaborations between labs in these fields is crucial. Transient delivery of CRISPR remains a bottleneck associated with low therapeutic efficiency and hindered tumor penetration. Further research into the development of delivery agents, including in 3D culture to better model difficulties in delivery, will assist. In the future, personalized medicine will enable thorough profiling of each patients' mutations, epigenome and transcriptome, rather than diagnosis based on a few select markers. Having a complete tumour profile will allow for more accurate diagnoses, including a more precise determination of breast cancer subtype, and the identification of aberrant expression patterns that should be corrected. Such personalized approaches may also help to overcome bias in clinical, pathological and genetic cancer data, which are skewed toward European populations (240). Targeting particular genes or pathways through G4 or i-motif secondary structure, or with CRISPR-based technology, will offer highly specific approaches to therapy, including re-activation of silenced genes or successful interventions against undruggable targets.

Overall, there are many opportunities on the horizon that will enable gene-specific, or gene-pathway-specific, treatments for breast cancers. Targeting the aberrant expression of genes involved in cancer progression, metastasis, and development of resistance through EMT has the potential to yield novel therapies and improved management of high-grade malignancies.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Sung,H., Ferlay,J., Siegel,R.L., Laversanne,M., Soerjomataram,I., Jemal,A. and Bray,F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, **71**, 209–249.
2. Roser,M. and Ritchie,H. (2015) Cancer. *Our World in Data*. https://ourworldindata.org/cancer.
3. Riggio,A.I., Varley,K.E. and Welm,A.L. (2021) The lingering mysteries of metastatic recurrence in breast cancer. *Br. J. Cancer*, **124**, 13–26.
4. Pradella,D., Naro,C., Sette,C. and Ghigna,C. (2017) EMT and stemness: flexible processes tuned by alternative splicing in development and cancer progression. *Mol. Cancer*, **16**, 8.
5. Pan,H., Gray,R., Braybrooke,J., Davies,C., Taylor,C., McGale,P., Peto,R., Pritchard,K.I., Bergh,J., Dowsett,M. *et al.* (2017) 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. *N. Engl. J. Med.*, **377**, 1836–1846.
6. Colleoni,M., Sun,Z., Price,K.N., Karlsson,P., Forbes,J.F., Thürlimann,B., Gianni,L., Castiglione,M., Gelber,R.D., Coates,A.S. *et al.* (2016) Annual hazard rates of recurrence for breast cancer during 24 years of follow-up: results from the international breast cancer study group trials I to V. *J. Clin. Oncol.*, **34**, 927–935.
7. Stolzenburg,S., Beltran,A.S., Swift-Scanlan,T., Rivenbark,A.G., Rashwan,R. and Blancafort,P. (2015) Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. *Oncogene*, **34**, 5427–5435.
8. Casalino,L. and Verde,P. (2020) Multifaceted roles of DNA methylation in neoplastic transformation, from tumor suppressors to EMT and metastasis. *Genes*, **11**, 992.
9. Tam,W.L. and Weinberg,R.A. (2013) The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.*, **19**, 1438–1449.
10. Aban,C.E., Lombardi,A., Neiman,G., Biani,M.C., La Greca,A., Waisman,A., Moro,L.N., Sevlever,G., Miriuka,S. and Luzzani,C. (2021) Downregulation of E-cadherin in pluripotent stem cells triggers partial EMT. *Sci. Rep.*, **11**, 2048.
11. Dongre,A. and Weinberg,R.A. (2019) New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.*, **20**, 69–84.
12. Lambert,A.W. and Weinberg,R.A. (2021) Linking EMT programmes to normal and neoplastic epithelial stem cells. *Nat. Rev. Cancer*, **21**, 325–338.
13. Bushweller,J.H. (2019) Targeting transcription factors in cancer — from undruggable to reality. *Nat. Rev. Cancer*, **19**, 611–624.
14. Lien,H.C., Hsiao,Y.H., Lin,Y.S., Yao,Y.T., Juan,H.F., Kuo,W.H., Hung,M.-C., Chang,K.J. and Hsieh,F.J. (2007) Molecular signatures of metaplastic carcinoma of the breast by large-scale transcriptional profiling: identification of genes potentially related to epithelial–mesenchymal transition. *Oncogene*, **26**, 7859–7871.
15. Tomaskovic-Crook,E., Thompson,E.W. and Thiery,J.P. (2009) Epithelial to mesenchymal transition and breast cancer. *Breast Cancer Res.*, **11**, 213.
16. Kröger,C., Afeyan,A., Mraz,J., Eaton,E.N., Reinhardt,F., Khodor,Y.L., Thiru,P., Bierie,B., Ye,X., Burge,C.B. *et al.* (2019) Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proc. Natl. Acad. Sci.*, **116**, 7353–7362.
17. Carter,B. and Zhao,K. (2021) The epigenetic basis of cellular heterogeneity. *Nat. Rev. Genet.*, **22**, 235–250.
18. Lawrence,M., Daujat,S. and Schneider,R. (2016) Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.*, **32**, 42–56.
19. Zhao,Z. and Shilatifard,A. (2019) Epigenetic modifications of histones in cancer. *Genome Biol.*, **20**, 245.
20. Lee,J.-Y. and Kong,G. (2016) Roles and epigenetic regulation of epithelial–mesenchymal transition and its transcription factors in cancer initiation and progression. *Cell. Mol. Life Sci.*, **73**, 4643–4660.
21. Hu,J., Su,P., Jiao,M., Bai,X., Qi,M., Liu,H., Wu,Z., Sun,J., Zhou,G. and Han,B. (2018) TRPS1 Suppresses Breast Cancer Epithelial-mesenchymal Transition Program as a Negative Regulator of SUZ12. *Transl. Oncol.*, **11**, 416–425.
22. Lombaerts,M., van Wezel,T., Philippo,K., Dierssen,J.W.F., Zimmerman,R.M.E., Oosting,J., van Eijk,R., Eilers,P.H., van de Water,B., Cornelisse,C.J. *et al.* (2006) E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. *Br. J. Cancer*, **94**, 661–671.
23. Dong,C., Wu,Y., Yao,J., Wang,Y., Yu,Y., Rychahou,P.G., Evers,B.M. and Zhou,B.P. (2012) G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *J. Clin. Invest.*, **122**, 1469–1486.
24. Vesuna,F., Lisok,A., Kimble,B., Domek,J., Kato,Y., van der Groep,P., Artemov,D., Kowalski,J., Carraway,H., van Diest,P. *et al.* (2012) Twist contributes to hormone resistance in breast cancer by downregulating estrogen receptor-α. *Oncogene*, **31**, 3223–3234.
25. Choudhury,S.R., Cui,Y., Lubecka,K., Stefanska,B. and Irudayaraj,J. (2016) CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. *Oncotarget*, **7**, 46545–46556.
26. Sapozhnikov,D.M. and Szyf,M. (2021) Unraveling the functional role of DNA methylation using targeted DNA demethylation by steric blockage of DNA methyltransferase with CRISPR/dCas9. *Nat. Commun.*, **12**, 5711.
27. Jones,P.L., Veenstra,G.J.C., Wade,P.A., Vermaak,D., Kass,S.U., Landsberger,N., Strouboulis,J. and Wolffe,A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.*, **19**, 187–191.
28. Nan,X., Ng,H.-H., Johnson,C.A., Laherty,C.D., Turner,B.M., Eisenman,R.N. and Bird,A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386–389.
29. Dong,B., Qiu,Z. and Wu,Y. (2020) Tackle epithelial-mesenchymal transition with epigenetic drugs in cancer. *Front. Pharmacol.*, **11**, 596239.
30. Gort,E.H., Suijkerbuijk,K.P.M., Roothaan,S.M., Raman,V., Vooijs,M., van der Wall,E. and van Diest,P.J. (2008) Methylation of the TWIST1 promoter, TWIST1 mRNA levels, and immunohistochemical expression of TWIST1 in breast cancer. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.*, **17**, 3325–3330.
31. Skrypek,N., Goossens,S., Smedt,E.D., Vandamme,N. and Berx,G. (2017) Epithelial-to-mesenchymal transition: epigenetic reprogramming driving cellular plasticity. *Trends Genet.*, **33**, 943–959.
32. Rodriguez-Pinilla,S.M., Sarrio,D., Moreno-Bueno,G., Rodriguez-Gil,Y., Martinez,M.A., Hernandez,L., Hardisson,D., Reis-Filho,J.S. and Palacios,J. (2007) Sox2: a possible driver of the basal-like phenotype in sporadic breast cancer. *Mod. Pathol.*, **20**, 474–481.
33. Novak,D., Hüser,L., Elton,J.J., Umansky,V., Altevogt,P. and Utikal,J. (2020) SOX2 in development and cancer biology. *Semin. Cancer Biol.*, **67**, 74–82.

34. Liu,K., Xie,F., Gao,A., Zhang,R., Zhang,L., Xiao,Z., Hu,Q., Huang,W., Huang,Q., Lin,B. *et al.* (2017) SOX2 regulates multiple malignant processes of breast cancer development through the SOX2/miR-181a-5p, miR-30e-5p/TUSC3 axis. *Mol. Cancer*, **16**, 62.

35. Leis,O., Eguiara,A., Lopez-Arribillaga,E., Alberdi,M.J., Hernandez-Garcia,S., Elorriaga,K., Pandiella,A., Rezola,R. and Martin,A.G. (2012) Sox2 expression in breast tumours and activation in breast cancer stem cells. *Oncogene*, **31**, 1354–1365.

36. Chen,Y. and Zhang,Y. (2018) Application of the CRISPR/Cas9 system to drug resistance in breast cancer. *Adv. Sci.*, **5**, 1700964.

37. Romero-Garcia,S., Prado-Garcia,H. and Carlos-Reyes,A. (2020) Role of DNA methylation in the resistance to therapy in solid tumors. *Front. Oncol.*, **10**, 1152.

38. Babenko,V.N., Chadaeva,I.V. and Orlov,Y.L. (2017) Genomic landscape of CpG rich elements in human. *BMC Evol. Biol.*, **17**, 19.

39. Davuluri,R.V., Grosse,I. and Zhang,M.Q. (2001) Computational identification of promoters and first exons in the human genome. *Nat. Genet.*, **29**, 412–417.

40. Illingworth,R.S. and Bird,A.P. (2009) CpG islands – 'A rough guide'. *FEBS Lett.*, **583**, 1713–1720.

41. Suzuki,M.M. and Bird,A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.*, **9**, 465–476.

42. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.

43. Aravin,A.A., Sachidanandam,R., Bourc'his,D., Schaefer,C., Pezic,D., Fejes Toth,K., Bestor,T. and Hannon,G.J. (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell*, **31**, 785–799.

44. Viré,E., Brenner,C., Deplus,R., Blanchon,L., Fraga,M., Didelot,C., Morey,L., Van Eynde,A., Bernard,D., Vanderwinden,J.-M. *et al.* (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*, **439**, 871–874.

45. Linhart,H.G., Lin,H., Yamada,Y., Moran,E., Steine,E.J., Gokhale,S., Lo,G., Cantu,E., Ehrich,M., He,T. *et al.* (2007) Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. *Genes Dev.*, **21**, 3110–3122.

46. Hervouet,E., Vallette,F.M. and Cartron,P.-F. (2009) Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. *Epigenetics*, **4**, 487–499.

47. Wu,Y., Sarkissyan,M. and Vadgama,J.V. (2015) Epigenetics in breast and prostate cancer. *Methods Mol. Biol. Clifton NJ*, **1238**, 425–466.

48. Bediaga,N.G., Acha-Sagredo,A., Guerra,I., Viguri,A., Albaina,C., Ruiz Diaz,I., Rezola,R., Alberdi,M.J., Dopazo,J., Montaner,D. *et al.* (2010) DNA methylation epigenotypes in breast cancer molecular subtypes. *Breast Cancer Res.*, **12**, R77.

49. Xu,X., Gammon,M.D., Zhang,Y., Bestor,T.H., Zeisel,S.H., Wetmur,J.G., Wallenstein,S., Bradshaw,P.T., Garbowski,G., Teitelbaum,S.L. *et al.* (2008) BRCA1 promoter methylation is associated with increased mortality among women with breast cancer. *Breast Cancer Res. Treat.*, **115**, 397.

50. Sharma,P., Stecklein,S.R., Kimler,B.F., Sethi,G., Petroff,B.K., Phillips,T.A., Tawfik,O.W., Godwin,A.K. and Jensen,R.A. (2014) The prognostic value of BRCA1 promoter methylation in early stage triple negative breast cancer. *J. Cancer Ther. Res.*, **3**, 1–11.

51. Xu,J., Chen,Y. and Olopade,O.I. (2010) MYC and Breast Cancer. *Genes Cancer*, **1**, 629–640.

52. Carroll,J.S. (2016) Mechanisms of oestrogen receptor (ER) gene regulation in breast cancer. *Eur. J. Endocrinol.*, **175**, R41–R49.

53. Rahman,M.M., Brane,A.C. and Tollefsbol,T.O. (2019) MicroRNAs and epigenetics strategies to reverse breast cancer. *Cells*, **8**, 1214.

54. Kim,N. (2019) The interplay between G-quadruplex and transcription. *Curr. Med. Chem.*, **26**, 2898–2917.

55. Wong,H.M. and Huppert,J.L. (2009) Stable G-quadruplexes are found outside nucleosome-bound regions. *Mol. Biosyst.*, **5**, 1713–1719.

56. Kang,H.-J., Kendrick,S., Hecht,S.M. and Hurley,L.H. (2014) The transcriptional complex between the BCL2 i-Motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.*, **136**, 4172–4185.

57. Brown,S.L. and Kendrick,S. (2021) The i-motif as a molecular target: more than a complementary DNA secondary structure. *Pharmaceuticals*, **14**, 96.

58. Varshney,D., Spiegel,J., Zyner,K., Tannahill,D. and Balasubramanian,S. (2020) The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.*, **21**, 459–474.

59. Huppert,J.L. and Balasubramanian,S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.

60. Bedrat,A., Lacroix,L. and Mergny,J.-L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.*, **44**, 1746–1759.

61. Belmonte-Reche,E. and Morales,J.C. (2020) G4-iM Grinder: when size and frequency matter. G-Quadruplex, i-Motif and higher order structure search and analysis tool. *NAR Genomics Bioinform.*, **2**, lqz005.

62. Sahakyan,A.B., Chambers,V.S., Marsico,G., Santner,T., Di Antonio,M. and Balasubramanian,S. (2017) Machine learning model for sequence-driven DNA G-quadruplex formation. *Sci. Rep.*, **7**, 14535.

63. Adrian,M., Heddi,B. and Phan,A.T. (2012) NMR spectroscopy of G-quadruplexes. *Methods*, **57**, 11–24.

64. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.

65. Day,H.A., Pavlou,P. and Waller,Z.A.E. (2014) i-Motif DNA: Structure, stability and targeting with ligands. *Bioorg. Med. Chem.*, **22**, 4407–4418.

66. Brooks,T.A., Kendrick,S. and Hurley,L. (2010) Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.*, **277**, 3459–3469.

67. Guéron,M. and Leroy,J.-L. (2000) The i-motif in nucleic acids. *Curr. Opin. Struct. Biol.*, **10**, 326–331.

68. Jana,J. and Weisz,K. (2021) Thermodynamic stability of G-quadruplexes: impact of sequence and environment. *Chembiochem. Eur. J. Chem. Biol.*, **22**, 2848–2856.

69. Wright,E.P., Huppert,J.L. and Waller,Z.A.E. (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.*, **45**, 2951–2959.

70. Zeraati,M., Langley,D.B., Schofield,P., Moye,A.L., Rouet,R., Hughes,W.E., Bryan,T.M., Dinger,M.E. and Christ,D. (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.*, **10**, 631–637.

71. King,J.J., Irving,K.L., Evans,C.W., Chikhale,R.V., Becker,R., Morris,C.J., Peña Martinez,C.D., Schofield,P., Christ,D., Hurley,L.H. *et al.* (2020) DNA G-quadruplex and i-motif structure formation is interdependent in human cells. *J. Am. Chem. Soc.*, **142**, 20600–20604.

72. Biffi,G., Tannahill,D., Miller,J., Howat,W.J. and Balasubramanian,S. (2014) Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues. *PLoS One*, **9**, e102711.

73. Bhatt,U., Kretzmann,A.L., Guédin,A., Ou,A., Kobelke,S., Bond,C.S., Evans,C.W., Hurley,L.H., Mergny,J.-L., Iyer,K.S. *et al.* (2021) The role of G-Quadruplex DNA in Paraspeckle formation in cancer. *Biochimie*, **190**, 124–131.

74. Hänsel-Hertsch,R., Beraldi,D., Lensing,S.V., Marsico,G., Zyner,K., Parry,A., Di Antonio,M., Pike,J., Kimura,H., Narita,M. *et al.* (2016) G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.*, **48**, 1267–1272.

75. Wang,X.-D, Ou,T.-M, Lu,Y.-J, Li,Z., Xu,Z., Xi,C., Tan,J.-H., Huang,S.-L., An,L.-K., Li,D. *et al.* (2010) Turning off transcription of the bcl-2 gene by stabilizing the bcl-2 promoter quadruplex with quindoline derivatives. *J. Med. Chem.*, **53**, 4390–4398.

76. Kendrick,S., Kang,H.-J., Alam,M.P., Madathil,M.M., Agrawal,P., Gokhale,V., Yang,D., Hecht,S.M. and Hurley,L.H. (2014) The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.*, **136**, 4161–4171.

77. Kendrick,S., Muranyi,A., Gokhale,V., Hurley,L.H. and Rimsza,L.M. (2017) Simultaneous drug targeting of the promoter MYC G-quadruplex and BCL2 i-motif in diffuse large B-cell lymphoma delays tumor growth. *J. Med. Chem.*, **60**, 6587–6597.

78. Dai,J., Chen,D., Jones,R.A., Hurley,L.H. and Yang,D. (2006) NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.*, **34**, 5133–5144.

79. Agrawal,P., Lin,C., Mathad,R.I., Carver,M. and Yang,D. (2014) The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in K+ solution. *J. Am. Chem. Soc.*, **136**, 1750–1753.

80. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) 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.*, **99**, 11593–11598.

81. Sutherland,C., Cui,Y., Mao,H. and Hurley,L.H. (2016) A mechanosensor mechanism controls the G-quadruplex/i-Motif molecular switch in the MYC promoter NHE III1. *J. Am. Chem. Soc.*, **138**, 14138–14151.

82. Phan,A.T., Kuryavyi,V., Gaw,H.Y. and Patel,D.J. (2005) Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nat. Chem. Biol.*, **1**, 167–173.

83. Sun,D. and Hurley,L.H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem.*, **52**, 2863–2874.

84. Brooks,T.A. and Hurley,L.H. (2010) Targeting MYC Expression through G-Quadruplexes. *Genes Cancer*, **1**, 641–649.

85. Cogoi,S. and Xodo,L.E. (2006) G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.*, **34**, 2536–2549.

86. Kaiser,C.E., Van Ert,N.A., Agrawal,P., Chawla,R., Yang,D. and Hurley,L.H. (2017) Insight into the complexity of the i-Motif and G-quadruplex DNA structures formed in the KRAS promoter and subsequent drug-induced gene repression. *J. Am. Chem. Soc.*, **139**, 8522–8536.

87. Ou,A., Schmidberger,J.W., Wilson,K.A., Evans,C.W., Hargreaves,J.A., Grigg,M., O'Mara,M.L., Iyer,K.S., Bond,C.S. and Smith,N.M. (2020) High resolution crystal structure of a KRAS promoter G-quadruplex reveals a dimer with extensive poly-A π-stacking interactions for small-molecule recognition. *Nucleic Acids Res.*, **48**, 5766–5776.

88. Marquevielle,J., Robert,C., Lagrabette,O., Wahid,M., Bourdoncle,A., Xodo,L.E., Mergny,J.-L. and Salgado,G.F. (2020) Structure of two G-quadruplexes in equilibrium in the KRAS promoter. *Nucleic Acids Res.*, **48**, 9336–9345.

89. Kang,H.-J., Cui,Y., Yin,H., Scheid,A., Hendricks,W.P.D., Schmidt,J., Sekulic,A., Kong,D., Trent,J.M., Gokhale,V. *et al.* (2016) A pharmacological chaperone molecule induces cancer cell death by restoring tertiary DNA structures in mutant hTERT promoters. *J. Am. Chem. Soc.*, **138**, 13673–13692.

90. Song,J.H., Kang,H.-J., Luevano,L.A., Gokhale,V., Wu,K., Pandey,R., Sherry Chow,H.-H., Hurley,L.H. and Kraft,A.S. (2019) Small-molecule-targeting hairpin loop of hTERT promoter G-quadruplex induces cancer cell death. *Cell Chem. Biol.*, **26**, 1110–1121.

91. Palumbo,S.L., Ebbinghaus,S.W. and Hurley,L.H. (2009) Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands. *J. Am. Chem. Soc.*, **131**, 10878–10891.

92. Monsen,R.C., DeLeeuw,L., Dean,W.L., Gray,R.D., Sabo,T.M., Chakravarthy,S., Chaires,J.B. and Trent,J.O. (2020) The hTERT core promoter forms three parallel G-quadruplexes. *Nucleic Acids Res.*, **48**, 5720–5734.

93. Qin,Y., Fortin,J.S., Tye,D., Gleason-Guzman,M., Brooks,T.A. and Hurley,L.H. (2010) Molecular cloning of the human platelet-derived growth factor receptor β (PDGFR-β) promoter and drug targeting of the G-quadruplex-forming region to repress PDGFR-β expression. *Biochemistry*, **49**, 4208–4219.

94. Brown,R.V., Wang,T., Chappeta,V.R., Wu,G., Onel,B., Chawla,R., Quijada,H., Camp,S.M., Chiang,E.T., Lassiter,Q.R. *et al.* (2017) The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β Core Promoter Nuclease Hypersensitive Element Can Explain the Unexpected Effects of Mutations and Provide Opportunities for Selective Targeting of Both Structures by Small Molecules To Downregulate Gene Expression. *J. Am. Chem. Soc.*, **139**, 7456–7475.

95. Bejugam,M., Sewitz,S., Shirude,P.S., Rodriguez,R., Shahid,R. and Balasubramanian,S. (2007) Trisubstituted Isoalloxazines as a New Class of G-Quadruplex Binding Ligands: Small Molecule Regulation of c-kit Oncogene Expression. *J. Am. Chem. Soc.*, **129**, 12926–12927.

96. Kuryavyi,V., Phan,A.T. and Patel,D.J. (2010) Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res.*, **38**, 6757–6773.

97. Todd,A.K., Haider,S.M., Parkinson,G.N. and Neidle,S. (2007) Sequence occurrence and structural uniqueness of a G-quadruplex in the human c-kit promoter. *Nucleic Acids Res.*, **35**, 5799–5808.

98. McLuckie,K.I.E., Waller,Z.A.E., Sanders,D.A., Alves,D., Rodriguez,R., Dash,J., McKenzie,G.J., Venkitaraman,A.R. and Balasubramanian,S. (2011) G-Quadruplex-Binding Benzo[a]phenoxazines Down-Regulate c-KIT Expression in Human Gastric Carcinoma Cells. *J. Am. Chem. Soc.*, **133**, 2658–2663.

99. Wei,D., Husby,J. and Neidle,S. (2015) Flexibility and structural conservation in a c-KIT G-quadruplex. *Nucleic Acids Res.*, **43**, 629–644.

100. Fernando,H., Reszka,A.P., Huppert,J., Ladame,S., Rankin,S., Venkitaraman,A.R., Neidle,S. and Balasubramanian,S. (2006) A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human c-kit Oncogene. *Biochemistry*, **45**, 7854–7860.

101. Hsu,S.-T.D., Varnai,P., Bugaut,A., Reszka,A.P., Neidle,S. and Balasubramanian,S. (2009) A G-Rich Sequence within the c-kit Oncogene Promoter Forms a Parallel G-Quadruplex Having Asymmetric G-Tetrad Dynamics. *J. Am. Chem. Soc.*, **131**, 13399–13409.

102. Phan,A.T., Kuryavyi,V., Burge,S., Neidle,S. and Patel,D.J. (2007) Structure of an Unprecedented G-Quadruplex Scaffold in the Human c-kit Promoter. *J. Am. Chem. Soc.*, **129**, 4386–4392.

103. Shin,Y.-J., Kumarasamy,V., Camacho,D. and Sun,D. (2015) Involvement of G-quadruplex structures in regulation of human RET gene expression by small molecules in human medullary thyroid carcinoma TT cells. *Oncogene*, **34**, 1292–1299.

104. Guo,K., Pourpak,A., Beetz-Rogers,K., Gokhale,V., Sun,D. and Hurley,L.H. (2007) Formation of Pseudosymmetrical G-Quadruplex and i-Motif Structures in the Proximal Promoter Region of the RET Oncogene. *J. Am. Chem. Soc.*, **129**, 10220–10228.

105. Guo,K., Gokhale,V., Hurley,L.H. and Sun,D. (2008) Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.*, **36**, 4598–4608.

106. Abdelhamid,M.A.S. and Waller,Z.A.E. (2020) Tricky topology: persistence of folded human telomeric i-motif DNA at ambient temperature and neutral pH. *Front. Chem.*, **8**, 40.

107. Agarwal,T., Roy,S., Kumar,S., Chakraborty,T.K. and Maiti,S. (2014) In the sense of transcription regulation by G-quadruplexes: asymmetric effects in sense and antisense strands. *Biochemistry*, **53**, 3711–3718.

108. Du,Z., Zhao,Y. and Li,N. (2008) Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription. *Genome Res.*, **18**, 233–241.

109. Kumar,P., Yadav,V.K., Baral,A., Kumar,P., Saha,D. and Chowdhury,S. (2011) Zinc-finger transcription factors are associated with guanine quadruplex motifs in human, chimpanzee, mouse and rat promoters genome-wide. *Nucleic Acids Res.*, **39**, 8005–8016.

110. Hou,Y., Li,F., Zhang,R., Li,S., Liu,H., Qin,Z.S. and Sun,X. (2019) Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure. *Epigenetics*, **14**, 894–911.

111. Spiegel,J., Cuesta,S.M., Adhikari,S., Hänsel-Hertsch,R., Tannahill,D. and Balasubramanian,S. (2021) G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.*, **22**, 117.

112. Balasubramanian,S., Hurley,L.H. and Neidle,S. (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.*, **10**, 261–275.

113. Eddy,J., Vallur,A.C., Varma,S., Liu,H., Reinhold,W.C., Pommier,Y. and Maizels,N. (2011) G4 motifs correlate with promoter-proximal transcriptional pausing in human genes. *Nucleic Acids Res.*, **39**, 4975–4983.

114. Biffi,G., Tannahill,D., McCafferty,J. and Balasubramanian,S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, **5**, 182–186.

115. Lago,S., Nadai,M., Cernilogar,F.M., Kazerani,M., Domíniguez Moreno,H., Schotta,G. and Richter,S.N. (2021) Promoter

G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nat. Commun.*, **12**, 3885.

116. Li,C., Wang,H., Yin,Z., Fang,P., Xiao,R., Xiang,Y., Wang,W., Li,Q., Huang,B., Huang,J. *et al.* (2021) Ligand-induced native G-quadruplex stabilization impairs transcription initiation. *Genome Res.*, **31**, 1546–1560.

117. Cui,Y., Kong,D., Ghimire,C., Xu,C. and Mao,H. (2016) Mutually Exclusive Formation of G-Quadruplex and i-Motif Is a General Phenomenon Governed by Steric Hindrance in Duplex DNA. *Biochemistry*, **55**, 2291–2299.

118. Dhakal,S., Yu,Z., Konik,R., Cui,Y., Koirala,D. and Mao,H. (2012) G-Quadruplex and i-Motif Are Mutually Exclusive in ILPR Double-Stranded DNA. *Biophys. J.*, **102**, 2575–2584.

119. De,S. and Michor,F. (2011) DNA secondary structures and epigenetic determinants of cancer genome evolution. *Nat. Struct. Mol. Biol.*, **18**, 950–955.

120. Halder,K., Halder,R. and Chowdhury,S. (2009) Genome-wide analysis predicts DNA structural motifs as nucleosome exclusion signals. *Mol. Biosyst.*, **5**, 1703–1712.

121. Guilbaud,G., Murat,P., Recolin,B., Campbell,B.C., Maiter,A., Sale,J.E. and Balasubramanian,S. (2017) Local epigenetic reprogramming induced by G-quadruplex ligands. *Nat. Chem.*, **9**, 1110–1117.

122. Wu,Y., Shin-ya,K. and Brosh,R.M. (2008) FANCJ helicase defective in Fanconia anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol. Cell. Biol.*, **28**, 4116–4128.

123. Xu,Y.-Z., Jenjaroenpun,P., Wongsurawat,T., Byrum,S.D., Shponka,V., Tannahill,D., Chavez,E.A., Hung,S.S., Steidl,C., Balasubramanian,S. *et al.* (2020) Activation-induced cytidine deaminase localizes to G-quadruplex motifs at mutation hotspots in lymphoma. *Nar Can.*, **2**, zcaa029.

124. Magis,A.D., Manzo,S.G., Russo,M., Marinello,J., Morigi,R., Sordet,O. and Capranico,G. (2019) DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci.*, **116**, 816–825.

125. Amparo,C., Clark,J., Bedell,V., Murata-Collins,J.L., Martella,M., Pichiorri,F., Warner,E.F., Abdelhamid,M.a.S., Waller,Z.a.E. and Smith,S.S. (2020) Duplex DNA from Sites of Helicase-Polymerase Uncoupling Links Non-B DNA Structure Formation to Replicative Stress. *Cancer Genomics Proteomics*, **17**, 101–115.

126. Šviković,S. and Sale,J.E. (2017) The Effects of Replication Stress on S Phase Histone Management and Epigenetic Memory. *J. Mol. Biol.*, **429**, 2011–2029.

127. Sarkies,P., Reams,C., Simpson,L.J. and Sale,J.E. (2010) Epigenetic Instability due to Defective Replication of Structured DNA. *Mol. Cell*, **40**, 703–713.

128. Schiavone,D., Guilbaud,G., Murat,P., Papadopoulou,C., Sarkies,P., Prioleau,M.-N., Balasubramanian,S. and Sale,J.E. (2014) Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. *EMBO J.*, **33**, 2507–2520.

129. De,S. and Michor,F. (2011) DNA replication timing and long-range DNA interactions predict mutational landscapes of cancer genomes. *Nat. Biotechnol.*, **29**, 1103–1108.

130. Katapadi,V.K., Nambiar,M. and Raghavan,S.C. (2012) Potential G-quadruplex formation at breakpoint regions of chromosomal translocations in cancer may explain their fragility. *Genomics*, **100**, 72–80.

131. Hänsel-Hertsch,R., Simeone,A., Shea,A., Hui,W.W.I., Zyner,K.G., Marsico,G., Rueda,O.M., Bruna,A., Martin,A., Zhang,X. *et al.* (2020) Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat. Genet.*, **52**, 878–883.

132. Yu,Z., Gaerig,V., Cui,Y., Kang,H., Gokhale,V., Zhao,Y., Hurley,L.H. and Mao,H. (2012) Tertiary DNA Structure in the Single-Stranded hTERT Promoter Fragment Unfolds and Refolds by Parallel Pathways via Cooperative or Sequential Events. *J. Am. Chem. Soc.*, **134**, 5157–5164.

133. Kendrick,S., Akiyama,Y., Hecht,S.M. and Hurley,L.H. (2009) The i-Motif in the bcl-2 P1 Promoter Forms an Unexpectedly Stable Structure with a Unique 8:5:7 Loop Folding Pattern. *J. Am. Chem. Soc.*, **131**, 17667–17676.

134. Roy,B., Talukder,P., Kang,H.-J., Tsuen,S.S., Alam,M.P., Hurley,L.H. and Hecht,S.M. (2016) Interaction of Individual Structural Domains of hnRNP LL with the BCL2 Promoter i-Motif DNA. *J. Am. Chem. Soc.*, **138**, 10950–10962.

135. Salvati,E., Zizza,P., Rizzo,A., Iachettini,S., Cingolani,C., D'Angelo,C., Porru,M., Randazzo,A., Pagano,B., Novellino,E. *et al.* (2014) Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res.*, **42**, 2945–2957.

136. Agrawal,P., Hatzakis,E., Guo,K., Carver,M. and Yang,D. (2013) Solution structure of the major G-quadruplex formed in the human VEGF promoter in K+: insights into loop interactions of the parallel G-quadruplexes. *Nucleic Acids Res.*, **41**, 10584–10592.

137. Wu,Y., Zan,L.-P., Wang,X.-D., Lu,Y.-J., Ou,T.-M., Lin,J., Huang,Z.-S. and Gu,L.-Q. (2014) Stabilization of VEGF G-quadruplex and inhibition of angiogenesis by quindoline derivatives. *Biochim. Biophys. Acta BBA - Gen. Subj.*, **1840**, 2970–2977.

138. Muench,D., Rezzoug,F., Thomas,S.D., Xiao,J., Islam,A., Miller,D.M. and Sedoris,K.C. (2019) Quadruplex-forming oligonucleotide targeted to the VEGF promoter inhibits growth of non-small cell lung cancer cells. *PLoS One*, **14**, e0211046.

139. Sun,D., Guo,K. and Shin,Y.-J. (2011) Evidence of the formation of G-quadruplex structures in the promoter region of the human vascular endothelial growth factor gene. *Nucleic Acids Res.*, **39**, 1256–1265.

140. Qin,Y., Rezler,E.M., Gokhale,V., Sun,D. and Hurley,L.H. (2007) Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. *Nucleic Acids Res.*, **35**, 7698–7713.

141. Xu,Y. and Sugiyama,H. (2006) Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). *Nucleic Acids Res.*, **34**, 949–954.

142. Cui,X., Chen,H., Zhang,Q., Xu,M., Yuan,G. and Zhou,J. (2019) Exploration of the Structure and Recognition of a G-quadruplex in the her2 Proto-oncogene Promoter and Its Transcriptional Regulation. *Sci. Rep.*, **9**, 3966.

143. Liu,M., Casimiro,M.C., Wang,C., Shirley,L.A., Jiao,X., Katiyar,S., Ju,X., Li,Z., Yu,Z., Zhou,J. *et al.* (2009) p21CIP1 attenuates Ras- and c-Myc-dependent breast tumor epithelial mesenchymal transition and cancer stem cell-like gene expression in vivo. *Proc. Natl. Acad. Sci.*, **106**, 19035–19039.

144. Moore,A.R., Rosenberg,S.C., McCormick,F. and Malek,S. (2020) Author Correction: RAS-targeted therapies: is the undruggable drugged? *Nat. Rev. Drug Discov.*, **19**, 902–902.

145. Dang,C.V., O'Donnell,K.A., Zeller,K.I., Nguyen,T., Osthus,R.C. and Li,F. (2006) The c-Myc target gene network. *Semin. Cancer Biol.*, **16**, 253–264.

146. Holysz,H., Lipinska,N., Paszel-Jaworska,A. and Rubis,B. (2013) Telomerase as a useful target in cancer fighting—the breast cancer case. *Tumor Biol.*, **34**, 1371–1380.

147. Zhang,T., Zhang,H., Wang,Y. and McGown,L.B. (2012) Capture and identification of proteins that bind to a GGA-rich sequence from the ERBB2 gene promoter region. *Anal. Bioanal. Chem.*, **404**, 1867–1876.

148. Cui,X., Chen,H., Zhang,Q., Xu,M., Yuan,G. and Zhou,J. (2019) Exploration of the Structure and Recognition of a G-quadruplex in the her2 Proto-oncogene Promoter and Its Transcriptional Regulation. *Sci. Rep.*, **9**, 3966.

149. Morgan,R.K., Batra,H., Gaerig,V.C., Hockings,J. and Brooks,T.A. (2016) Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1859**, 235–245.

150. Cogoi,S., Shchekotikhin,A.E. and Xodo,L.E. (2014) HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property. *Nucleic Acids Res.*, **42**, 8379–8388.

151. Miglietta,G., Cogoi,S., Pedersen,E.B. and Xodo,L.E. (2016) GC-elements controlling HRAS transcription form i-motif structures unfolded by heterogeneous ribonucleoprotein particle A1. *Sci. Rep.*, **5**, 18097.

152. Membrino,A., Cogoi,S., Pedersen,E.B. and Xodo,L.E. (2011) G4-DNA Formation in the HRAS Promoter and Rational Design of Decoy Oligonucleotides for Cancer Therapy. *PLoS One*, **6**, e24421.

153. Todd,A.K. and Neidle,S. (2008) The relationship of potential G-quadruplex sequences in cis -upstream regions of the human genome to SP1-binding elements. *Nucleic Acids Res.*, **36**, 2700–2704.

154. Cogoi,S., Paramasivam,M., Membrino,A., Yokoyama,K.K. and Xodo,L.E. (2010) The KRAS Promoter Responds to Myc-associated Zinc Finger and Poly(ADP-ribose) Polymerase 1 Proteins, Which Recognize a Critical Quadruplex-forming GA-element. *J. Biol. Chem.*, **285**, 22003–22016.

155. Prior,I.A., Hood,F.E. and Hartley,J.L. (2020) The Frequency of Ras Mutations in Cancer. *Cancer Res.*, **80**, 2969–2974.

156. Cogoi,S., Zorzet,S., Rapozzi,V., Géci,I., Pedersen,E.B. and Xodo,L.E. (2013) MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.*, **41**, 4049–4064.

157. Shu,B., Cao,J., Kuang,G., Qiu,J., Zhang,M., Zhang,Y., Wang,M., Li,X., Kang,S., Ou,T.-M. *et al.* (2018) Syntheses and evaluation of new acridone derivatives for selective binding of oncogene c-myc promoter i-motifs in gene transcriptional regulation. *Chem. Commun.*, **54**, 2036–2039.

158. Wang,Y., Yan,T., Li,J., Wang,W., Wu,Q., Shi,J., Li,D., Stovall,D.B., Zhang,Q. and Sui,G. (2020) The SNAIL1 promoter contains G-quadruplex structures regulating its gene expression and DNA replication. *Exp. Cell Res.*, **394**, 112158.

159. Dutta,A., Maji,N., Sengupta,P., Banerjee,N., Kar,S., Mukherjee,G., Chatterjee,S. and Basu,M. (2021) Promoter G-quadruplex favours epigenetic reprogramming-induced atypical expression of ZEB1 in cancer cells. *Biochim. Biophys. Acta Gen. Subj.*, **1865**, 129899.

160. Li,Y., Jin,K., Pelt,G.W. van, Dam,H. van, Yu,X., Mesker,W.E., Dijke,P. ten, Zhou,F. and Zhang,L. (2016) c-Myb Enhances Breast Cancer Invasion and Metastasis through the Wnt/β-Catenin/Axin2 Pathway. *Cancer Res.*, **76**, 3364–3375.

161. Gao,Y., Zhang,W., Liu,C. and Li,G. (2019) miR-200 affects tamoxifen resistance in breast cancer cells through regulation of MYB. *Sci. Rep.*, **9**, 18844

162. Yang,M.-H., Wu,M.-Z., Chiou,S.-H., Chen,P.-M., Chang,S.-Y., Liu,C.-J., Teng,S.-C. and Wu,K.-J. (2008) Direct regulation of TWIST by HIF-1α promotes metastasis. *Nat. Cell Biol.*, **10**, 295–305.

163. Chen,H., Long,H., Cui,X., Zhou,J., Xu,M. and Yuan,G. (2014) Exploring the Formation and Recognition of an Important G-Quadruplex in a HIF1α Promoter and Its Transcriptional Inhibition by a Benzo[c]phenanthridine Derivative. *J. Am. Chem. Soc.*, **136**, 2583–2591.

164. Palumbo,S.L., Memmott,R.M., Uribe,D.J., Krotova-Khan,Y., Hurley,L.H. and Ebbinghaus,S.W. (2008) A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity. *Nucleic. Acids. Res.*, **36**, 1755–1769.

165. Huang,W., Smaldino,P.J., Zhang,Q., Miller,L.D., Cao,P., Stadelman,K., Wan,M., Giri,B., Lei,M., Nagamine,Y. *et al.* (2012) Yin Yang 1 contains G-quadruplex structures in its promoter and 5′-UTR and its expression is modulated by G4 resolvase 1. *Nucleic Acids Res.*, **40**, 1033–1049.

166. Zhang,X., Spiegel,J., Martínez Cuesta,S., Adhikari,S. and Balasubramanian,S. (2021) Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat. Chem.*, **13**, 626–633.

167. Hussain,T., Saha,D., Purohit,G., Kar,A., Kishore Mukherjee,A., Sharma,S., Sengupta,S., Dhapola,P., Maji,B., Vedagopuram,S. *et al.* (2017) Transcription regulation of CDKN1A (p21/CIP1/WAF1) by TRF2 is epigenetically controlled through the REST repressor complex. *Sci. Rep.*, **7**, 11541.

168. Saha,D., Singh,A., Hussain,T., Srivastava,V., Sengupta,S., Kar,A., Dhapola,P., Dhople,V., Ummanni,R. and Chowdhury,S. (2017) Epigenetic suppression of human telomerase (hTERT) is mediated by the metastasis suppressor NME2 in a G-quadruplex–dependent fashion. *J. Biol. Chem.*, **292**, 15205–15215.

169. Clark,J. and Smith,S.S. (2008) Secondary Structure at a Hot Spot for DNA Methylation in DNA from Human Breast Cancers. *Cancer Genomics Proteomics*, **5**, 241–251.

170. Smith,S.S., Baker,D.J. and Jardines,L.A. (1989) A G4-DNA/B-DNA junction at codon 12 of c-Ha-ras is actively and asymmetrically methylated by DNA(cytosine-5)methyltransferase. *Biochem. Biophys. Res. Commun.*, **160**, 1397–1402.

171. Mao,S.-Q., Ghanbarian,A.T., Spiegel,J., Martínez Cuesta,S., Beraldi,D., Di Antonio,M., Marsico,G., Hänsel-Hertsch,R., Tannahill,D. and Balasubramanian,S. (2018) DNA G-quadruplex structures mold the DNA methylome. *Nat. Struct. Mol. Biol.*, **25**, 951–957.

172. Cree,S.L., Fredericks,R., Miller,A., Pearce,F.G., Filichev,V., Fee,C. and Kennedy,M.A. (2016) DNA G-quadruplexes show strong interaction with DNA methyltransferases in vitro. *FEBS Lett.*, **590**, 2870–2883.

173. Halder,R., Halder,K., Sharma,P., Garg,G., Sengupta,S. and Chowdhury,S. (2010) Guanine quadruplex DNA structure restricts methylation of CpG dinucleotides genome-wide. *Mol. Biosyst.*, **6**, 2439–2447.

174. Jara-Espejo,M. and Line,S.R. (2020) DNA G-quadruplex stability, position and chromatin accessibility are associated with CpG island methylation. *FEBS J.*, **287**, 483–495.

175. Malousi,A., Kouidou,S., Tsagiopoulou,M., Papakonstantinou,N., Bouras,E., Georgiou,E., Tzimagiorgis,G. and Stamatopoulos,K. (2019) MeinteR: A framework to prioritize DNA methylation aberrations based on conformational and cis-regulatory element enrichment. *Sci. Rep.*, **9**, 19148.

176. Wright,E.P., Abdelhamid,M.A.S., Ehiabor,M.O., Grigg,M.C., Irving,K., Smith,N.M. and Waller,Z.A.E. (2020) Epigenetic modification of cytosines fine tunes the stability of i-motif DNA. *Nucleic Acids Res.*, **48**, 55–62.

177. Lin,J., Hou,J., Xiang,H., Yan,Y., Gu,Y., Tan,J., Li,D., Gu,L., Ou,T. and Huang,Z. (2013) Stabilization of G-quadruplex DNA by C-5-methyl-cytosine in bcl-2 promoter: Implications for epigenetic regulation. *Biochem. Biophys. Res. Commun.*, **433**, 368–373.

178. Tsukakoshi,K., Saito,S., Yoshida,W., Goto,S. and Ikebukuro,K. (2018) CpG Methylation Changes G-Quadruplex Structures Derived from Gene Promoters and Interaction with VEGF and SP1. *Molecules*, **23**, 944.

179. Xu,B., Devi,G. and Shao,F. (2015) Regulation of telomeric i-motif stability by 5-methylcytosine and 5-hydroxymethylcytosine modification. *Org. Biomol. Chem.*, **13**, 5646–5651.

180. Bhavsar-Jog,Y.P., Van Dornshuld,E., Brooks,T.A., Tschumper,G.S. and Wadkins,R.M. (2014) Epigenetic Modification, Dehydration, and Molecular Crowding Effects on the Thermodynamics of i-Motif Structure Formation from C-Rich DNA. *Biochemistry*, **53**, 1586–1594.

181. Li,P.-T., Wang,Z.-F., Chu,I.-T., Kuan,Y.-M., Li,M.-H., Huang,M.-C., Chiang,P.-C., Chang,T.-C. and Chen,C.-T. (2017) Expression of the human telomerase reverse transcriptase gene is modulated by quadruplex formation in its first exon due to DNA methylation. *J. Biol. Chem.*, **292**, 20859–20870.

182. Li,L., Williams,P., Ren,W., Wang,M.Y., Gao,Z., Miao,W., Huang,M., Song,J. and Wang,Y. (2021) YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat. Chem. Biol.*, **17**, 161–168.

183. Kosiol,N., Juranek,S., Brossart,P., Heine,A. and Paeschke,K. (2021) G-quadruplexes: a promising target for cancer therapy. *Mol. Cancer*, **20**, 40.

184. Monchaud,D. and Teulade-Fichou,M.-P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.*, **6**, 627–636.

185. Carvalho,J., Mergny,J.-L., Salgado,G.F., Queiroz,J.A. and Cruz,C. (2020) G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies. *Trends Mol. Med.*, **26**, 848–861.

186. Abou Assi,H., Garavís,M., González,C. and Damha,M.J. (2018) i-Motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.*, **46**, 8038–8056.

187. Neidle,S. (2016) Quadruplex Nucleic Acids as Novel Therapeutic Targets. *J. Med. Chem.*, **59**, 5987–6011.

188. Patel,D.J., Phan,A.T. and Kuryavyi,V. (2007) Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.

189. Nishio,M., Tsukakoshi,K. and Ikebukuro,K. (2021) G-quadruplex: Flexible conformational changes by cations, pH, crowding and its applications to biosensing. *Biosens. Bioelectron.*, **178**, 113030.

190. Cao,Q., Li,Y., Freisinger,E., Qin,P.Z., Sigel,R.K.O. and Mao,Z.-W. (2017) G-quadruplex DNA targeted metal complexes acting as potential anticancer drugs. *Inorg. Chem. Front.*, **4**, 10–32.

191. Waller,Z.A.E., Sewitz,S.A., Hsu,S.-T.D. and Balasubramanian,S. (2009) A Small Molecule That Disrupts G-Quadruplex DNA

Structure and Enhances Gene Expression. *J. Am. Chem. Soc.*, **131**, 12628–12633.

192. Zamiri,B., Reddy,K., Macgregor,R.B. and Pearson,C.E. (2014) TMPyP4 Porphyrin Distorts RNA G-quadruplex Structures of the Disease-associated r(GGGGCC)n Repeat of the C9orf72 Gene and Blocks Interaction of RNA-binding Proteins *. *J. Biol. Chem.*, **289**, 4653–4659.

193. Kaluzhny,D., Ilyinsky,N., Shchekotikhin,A., Sinkevich,Y., Tsvetkov,P.O., Tsvetkov,V., Veselovsky,A., Livshits,M., Borisova,O., Shtil,A. *et al.* (2011) Disordering of Human Telomeric G-Quadruplex with Novel Antiproliferative Anthrathiophenedione. *PLoS One*, **6**, e27151.

194. O'Hagan,M.P., Haldar,S., Duchi,M., Oliver,T.A.A., Mulholland,A.J., Morales,J.C. and Galan,M.C. (2019) A Photoresponsive Stiff-Stilbene Ligand Fuels the Reversible Unfolding of G-Quadruplex DNA. *Angew. Chem.*, **131**, 4378–4382.

195. Lejault,P., Mitteaux,J., Sperti,F.R. and Monchaud,D. (2021) How to untie G-quadruplex knots and why? *Cell Chem. Biol.*, **28**, 436–455.

196. Parkinson,G.N., Ghosh,R. and Neidle,S. (2007) Structural Basis for Binding of Porphyrin to Human Telomeres. *Biochemistry*, **46**, 2390–2397.

197. Shanbhag,N.M., Rafalska-Metcalf,I.U., Balane-Bolivar,C., Janicki,S.M. and Greenberg,R.A. (2010) ATM-Dependent Chromatin Changes Silence Transcription In cis to DNA Double-Strand Breaks. *Cell*, **141**, 970–981.

198. Ayrapetov,M.K., Gursoy-Yuzugullu,O., Xu,C., Xu,Y. and Price,B.D. (2014) DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin. *Proc. Natl. Acad. Sci.*, **111**, 9169–9174.

199. Zimmer,J., Tacconi,E.M.C., Folio,C., Badie,S., Porru,M., Klare,K., Tumiati,M., Markkanen,E., Halder,S., Ryan,A. *et al.* (2016) Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Mol. Cell*, **61**, 449–460.

200. Xu,H., Di Antonio,M., McKinney,S., Mathew,V., Ho,B., O'Neil,N.J., Santos,N.D., Silvester,J., Wei,V., Garcia,J. *et al.* (2017) CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat. Commun.*, **8**, 14432.

201. Canadian Cancer Trials Group (2021) A Phase I Study of CX5461 clinicaltrials.gov.

202. Odermatt,D.C., Lee,W.T.C., Wild,S., Jozwiakowski,S.K., Rothenberg,E. and Gari,K. (2020) Cancer-associated mutations in the iron-sulfur domain of FANCJ affect G-quadruplex metabolism. *PLos Genet.*, **16**, e1008740.

203. Derenzini,M., Trerè,D., Pession,A., Govoni,M., Sirri,V. and Chieco,P. (2000) Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. *J. Pathol.*, **191**, 181–186.

204. Carvalho,J., Quintela,T., Gueddouda,N.M., Bourdoncle,A., Mergny,J.-L., Salgado,G.F., Queiroz,J.A. and Cruz,C. (2018) Phenanthroline polyazamacrocycles as G-quadruplex DNA binders. *Org. Biomol. Chem.*, **16**, 2776–2786.

205. Beauvarlet,J., Bensadoun,P., Darbo,E., Labrunie,G., Rousseau,B., Richard,E., Draskovic,I., Londono-Vallejo,A., Dupuy,J.-W., Nath Das,R. *et al.* (2019) Modulation of the ATM/autophagy pathway by a G-quadruplex ligand tips the balance between senescence and apoptosis in cancer cells. *Nucleic Acids Res.*, **47**, 2739–2756.

206. Pickar-Oliver,A. and Gersbach,C.A. (2019) The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.*, **20**, 490–507.

207. Adli,M. (2018) The CRISPR tool kit for genome editing and beyond. *Nat. Commun.*, **9**, 1911.

208. Shakirova,K.M., Ovchinnikova,V.Y. and Dashinimaev,E.B. (2020) Cell reprogramming with CRISPR/Cas9 based transcriptional regulation systems. *Front. Bioeng. Biotechnol.*, **8**, 882.

209. Pandelakis,M., Delgado,E. and Ebrahimkhani,M.R. (2020) CRISPR-based synthetic transcription factors in vivo: the future of therapeutic cellular programming. *Cell Syst.*, **10**, 1–14.

210. Saunderson,E.A., Stepper,P., Gomm,J.J., Hoa,L., Morgan,A., Allen,M.D., Jones,J.L., Gribben,J.G., Jurkowski,T.P. and Ficz,G. (2017) Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat. Commun.*, **8**, 1450.

211. Mei,Y., Cai,D. and Dai,X. (2020) Modulating cancer stemness provides luminal a breast cancer cells with HER2 positive-like features. *J. Cancer*, **11**, 1162–1169.

212. O'Geen,H., Ren,C., Nicolet,C.M., Perez,A.A., Halmai,J., Le,V.M., Mackay,J.P., Farnham,P.J. and Segal,D.J. (2017) dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res.*, **45**, 9901–9916.

213. Huang,X., Zhou,Q., Wang,M., Cao,C., Ma,Q., Ye,J. and Gui,Y. (2021) A Light-Inducible Split-dCas9 System for Inhibiting the Progression of Bladder Cancer Cells by Activating p53 and E-cadherin. *Front. Mol. Biosci.*, **7**, 627848.

214. Zhang,H.-Y., Liang,F., Jia,Z.-L., Song,S.-T. and Jiang,Z.-F. (2013) PTEN mutation, methylation and expression in breast cancer patients. *Oncol. Lett.*, **6**, 161–168.

215. Moses,C., Nugent,F., Waryah,C.B., Garcia-Bloj,B., Harvey,A.R. and Blancafort,P. (2019) Activating PTEN Tumor Suppressor Expression with the CRISPR/dCas9 System. *Mol. Ther. - Nucleic Acids*, **14**, 287–300.

216. Liu,Q., Zhao,K., Wang,C., Zhang,Z., Zheng,C., Zhao,Y., Zheng,Y., Liu,C., An,Y., Shi,L. *et al.* (2019) Multistage Delivery Nanoparticle Facilitates Efficient CRISPR/dCas9 Activation and Tumor Growth Suppression In Vivo. *Adv. Sci.*, **6**, 1801423.

217. Kretzmann,J.A., Evans,C.W., Moses,C., Sorolla,A., Kretzmann,A.L., Wang,E., Ho,D., Hackett,M.J., Dessauvagie,B.F., Smith,N.M. *et al.* (2019) Tumour suppression by targeted intravenous non-viral CRISPRa using dendritic polymers. *Chem. Sci.*, **10**, 7718–7727.

218. Huang,M.-C., Chu,I.-T., Wang,Z.-F., Lin,S., Chang,T.-C. and Chen,C.-T. (2018) A G-Quadruplex Structure in the Promoter Region of CLIC4 Functions as a Regulatory Element for Gene Expression. *Int. J. Mol. Sci.*, **19**, 2678.

219. Jin,X., Li,Y., Guo,Y., Jia,Y., Qu,H., Lu,Y., Song,P., Zhang,X., Shao,Y., Qi,D. *et al.* (2019) ERα is required for suppressing OCT4-induced proliferation of breast cancer cells via DNMT1/ISL1/ERK axis. *Cell Prolif.*, **52**, e12612.

220. Goulet,B., Kennette,W., Ablack,A., Postenka,C.O., Hague,M.N., Mymryk,J.S., Tuck,A.B., Giguère,V., Chambers,A.F. and Lewis,J.D. (2011) Nuclear localization of maspin is essential for its inhibition of tumor growth and metastasis. *Lab. Invest.*, **91**, 1181–1187.

221. Wilbie,D., Walther,J. and Mastrobattista,E. (2019) Delivery Aspects of CRISPR/Cas for in Vivo Genome Editing. *Acc. Chem. Res.*, **52**, 1555–1564.

222. Mout,R., Ray,M., Lee,Y.-W., Scaletti,F. and Rotello,V.M. (2017) In Vivo Delivery of CRISPR/Cas9 for Therapeutic Gene Editing: Progress and Challenges. *Bioconjug. Chem.*, **28**, 880–884.

223. Kretzmann,J.A., Ho,D., Evans,C.W., Plani-Lam,J.H.C., Garcia-Bloj,B., Mohamed,A.E., O'Mara,M.L., Ford,E., Tan,D.E.K., Lister,R. *et al.* (2017) Synthetically controlling dendrimer flexibility improves delivery of large plasmid DNA. *Chem. Sci.*, **8**, 2923–2930.

224. Amabile,A., Migliara,A., Capasso,P., Biffi,M., Cittaro,D., Naldini,L. and Lombardo,A. (2016) Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell*, **167**, 219–232.

225. LaFountaine,J.S., Fathe,K. and Smyth,H.D.C. (2015) Delivery and therapeutic applications of gene editing technologies ZFNs, TALENs, and CRISPR/Cas9. *Int. J. Pharm.*, **494**, 180–194.

226. Yin,H., Kanasty,R.L., Eltoukhy,A.A., Vegas,A.J., Dorkin,J.R. and Anderson,D.G. (2014) Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.*, **15**, 541–555.

227. Yin,H., Kauffman,K.J. and Anderson,D.G. (2017) Delivery technologies for genome editing. *Nat. Rev. Drug Discov.*, **16**, 387–399.

228. Ran,F.A., Cong,L., Yan,W.X., Scott,D.A., Gootenberg,J.S., Kriz,A.J., Zetsche,B., Shalem,O., Wu,X., Makarova,K.S. *et al.* (2015) In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, **520**, 186–191.

229. Kretzmann,J.A., Evans,C.W., Norret,M., Blancafort,P. and Swaminathan Iyer,K. (2018) Non-viral Methodology for Efficient Co-transfection. In: Jeltsch,A. and Rots,M.G. (eds). *Epigenome Editing: Methods and Protocols, Methods in Molecular Biology*. Springer, New York, NY, pp. 241–254.

230. Moreno,A.M., Palmer,N., Alemán,F., Chen,G., Pla,A., Jiang,N., Leong Chew,W., Law,M. and Mali,P. (2019) Immune-orthogonal orthologues of AAV capsids and of Cas9 circumvent the immune

response to the administration of gene therapy. *Nat. Biomed. Eng.*, **3**, 806–816.

231. Wang,D., Tai,P.W.L. and Gao,G. (2019) Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.*, **18**, 358–378.

232. Shen,C.-C., Hsu,M.-N., Chang,C.-W., Lin,M.-W., Hwu,J.-R., Tu,Y. and Hu,Y.-C. (2019) Synthetic switch to minimize CRISPR off-target effects by self-restricting Cas9 transcription and translation. *Nucleic Acids Res.*, **47**, e13.

233. Li,S.-D. and Huang,L. (2006) Gene therapy progress and prospects: non-viral gene therapy by systemic delivery. *Gene Ther.*, **13**, 1313–1319.

234. Judge,A., McClintock,K., Phelps,J.R. and MacLachlan,I. (2006) Hypersensitivity and Loss of Disease Site Targeting Caused by Antibody Responses to PEGylated Liposomes. *Mol. Ther.*, **13**, 328–337.

235. Wilhelm,S., Tavares,A.J., Dai,Q., Ohta,S., Audet,J., Dvorak,H.F. and Chan,W.C.W. (2016) Analysis of nanoparticle delivery to tumours. *Nat. Rev. Mater.*, **1**, 1–12.

236. le Sage,C., Lawo,S., Panicker,P., Scales,T.M.E., Rahman,S.A., Little,A.S., McCarthy,N.J., Moore,J.D. and Cross,B.C.S. (2017) Dual direction CRISPR transcriptional regulation screening uncovers gene networks driving drug resistance. *Sci. Rep.*, **7**, 17693.

237. Shalem,O., Sanjana,N.E. and Zhang,F. (2015) High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.*, **16**, 299–311.

238. Wang,T., Wei,J.J., Sabatini,D.M. and Lander,E.S. (2014) Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science*, **343**, 80–84.

239. Joung,J., Konermann,S., Gootenberg,J.S., Abudayyeh,O.O., Platt,R.J., Brigham,M.D., Sanjana,N.E. and Zhang,F. (2017) Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.*, **12**, 828–863.

240. Zavala,V.A., Bracci,P.M., Carethers,J.M., Carvajal-Carmona,L., Coggins,N.B., Cruz-Correa,M.R., Davis,M., de Smith,A.J., Dutil,J., Figueiredo,J.C. *et al.* (2021) Cancer health disparities in racial/ethnic minorities in the United States. *Br. J. Cancer*, **124**, 315–332.

241. Sims,A.H., Howell,A., Howell,S.J. and Clarke,R.B. (2007) Origins of breast cancer subtypes and therapeutic implications. *Nat. Clin. Pract. Oncol.*, **4**, 516–525.

242. Johansson,H.J., Socciarelli,F., Vacanti,N.M., Haugen,M.H., Zhu,Y., Siavelis,I., Fernandez-Woodbridge,A., Aure,M.R., Sennblad,B., Vesterlund,M. *et al.* (2019) Breast cancer quantitative proteome and proteogenomic landscape. *Nat. Commun.*, **10**, 1600.

243. Yersal,O. and Barutca,S. (2014) Biological subtypes of breast cancer: prognostic and therapeutic implications. *World J. Clin. Oncol.*, **5**, 412.

244. Dai,X., Li,T., Bai,Z., Yang,Y., Liu,X., Zhan,J. and Shi,B. (2015) Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.*, **5**, 2929–2943.

245. Weigelt,B., Geyer,F.C. and Reis-Filho,J.S. (2010) Histological types of breast cancer: how special are they? *Mol. Oncol.*, **4**, 192–208.

246. Tam,W.L. and Weinberg,R.A. (2013) The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.*, **19**, 1438–1449.

247. Miglietta,G., Russo,M., Duardo,R.C. and Capranico,G. (2021) G-quadruplex binders as cytostatic modulators of innate immune genes in cancer cells. *Nucleic Acids Res.*, **49**, 6673–6686.

248. Salvati,E., Leonetti,C., Rizzo,A., Scarsella,M., Mottolese,M., Galati,R., Sperduti,I., Stevens,M.F.G., D'Incalci,M., Blasco,M. *et al.* (2007) Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J. Clin. Invest.*, **117**, 3236–3247.

249. Phatak,P., Cookson,J.C., Dai,F., Smith,V., Gartenhaus,R.B., Stevens,M.F.G. and Burger,A.M. (2007) Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism. *Br. J. Cancer*, **96**, 1223–1233.

250. Rizzo,A., Iachettini,S., Zizza,P., Cingolani,C., Porru,M., Artuso,S., Stevens,M., Hummersone,M., Biroccio,A., Salvati,E. *et al.* (2014) Identification of novel RHPS4-derivative ligands with improved toxicological profiles and telomere-targeting activities. *J. Exp. Clin. Cancer Res.*, **33**, 81.

251. Perry,P.J., Reszka,A.P., Wood,A.A., Read,M.A., Gowan,S.M., Dosanjh,H.S., Trent,J.O., Jenkins,T.C., Kelland,L.R. and Neidle,S. (1998) Human Telomerase Inhibition by Regioisomeric Disubstituted Amidoanthracene-9,10-diones. *J. Med. Chem.*, **41**, 4873–4884.

252. Sun,D., Thompson,B., Cathers,B.E., Salazar,M., Kerwin,S.M., Trent,J.O., Jenkins,T.C., Neidle,S. and Hurley,L.H. (1997) Inhibition of Human Telomerase by a G-Quadruplex-Interactive Compound. *J. Med. Chem.*, **40**, 2113–2116.

253. Burger,A.M., Dai,F., Schultes,C.M., Reszka,A.P., Moore,M.J., Double,J.A. and Neidle,S. (2005) The G-Quadruplex-Interactive Molecule BRACO-19 Inhibits Tumor Growth, Consistent with Telomere Targeting and Interference with Telomerase Function. *Cancer Res.*, **65**, 1489–1496.

254. Gowan,S.M., Harrison,J.R., Patterson,L., Valenti,M., Read,M.A., Neidle,S. and Kelland,L.R. (2002) A G-Quadruplex-Interactive Potent Small-Molecule Inhibitor of Telomerase Exhibiting in Vitro and in Vivo Antitumor Activity. *Mol. Pharmacol.*, **61**, 1154–1162.

255. Suseela,Y.V., Satha,P., Murugan,N.A. and Govindaraju,T. (2020) Recognition of G-quadruplex topology through hybrid binding with implications in cancer theranostics. *Theranostics*, **10**, 10394–10414.

256. Hu,M.-H., Wu,T.-Y., Huang,Q. and Jin,G. (2019) New substituted quinoxalines inhibit triple-negative breast cancer by specifically downregulating the c-MYC transcription. *Nucleic Acids Res.*, **47**, 10529–10542.

257. Panda,D., Debnath,M., Mandal,S., Bessi,I., Schwalbe,H. and Dash,J. (2015) A Nucleus-Imaging Probe That Selectively Stabilizes a Minor Conformation of c-MYC G-quadruplex and Down-regulates c-MYC Transcription in Human Cancer Cells. *Sci. Rep.*, **5**, 13183.

258. Paul,R., Das,T., Debnath,M., Chauhan,A. and Dash,J. (2020) G-Quadruplex-Binding Small Molecule Induces Synthetic Lethality in Breast Cancer Cells by Inhibiting c-MYC and BCL2 Expression. *ChemBioChem*, **21**, 963–970.

259. Huang,F.-C., Chang,C.-C., Wang,J.-M., Chang,T.-C. and Lin,J.-J. (2012) Induction of senescence in cancer cells by the G-quadruplex stabilizer, BMVC4, is independent of its telomerase inhibitory activity. *Br. J. Pharmacol.*, **167**, 393–406.

260. Zeng,L., Wu,Q., Wang,T., Li,L.-P., Zhao,X., Chen,K., Qian,J., Yuan,L., Xu,H. and Mei,W.-J. (2021) Selective stabilization of multiple promoter G-quadruplex DNA by using 2-phenyl-1H-imidazole-based tanshinone IIA derivatives and their potential suppressing function in the metastatic breast cancer. *Bioorg. Chem.*, **106**, 104433.

261. Debnath,M., Ghosh,S., Chauhan,A., Paul,R., Bhattacharyya,K. and Dash,J. (2017) Preferential targeting of i-motifs and G-quadruplexes by small molecules. *Chem. Sci.*, **8**, 7448–7456.

262. Qin,H., Zhao,C., Sun,Y., Ren,J. and Qu,X. (2017) Metallo-supramolecular complexes enantioselectively eradicate cancer stem cells in vivo. *J. Am. Chem. Soc.*, **139**, 16201–16209.

263. Grand,C.L., Han,H., Muñoz,R.M., Weitman,S., Von Hoff,D.D., Hurley,L.H. and Bearss,D.J. (2002) The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. *Mol. Cancer Ther.*, **1**, 565–573.

264. Konieczna,N., Romaniuk-Drapała,A., Lisiak,N., Totoń,E., Paszel-Jaworska,A., Kaczmarek,M. and Rubiś,B. (2019) Telomerase Inhibitor TMPyP4 Alters Adhesion and Migration of Breast-Cancer Cells MCF7 and MDA-MB-231. *Int. J. Mol. Sci.*, **20**, 2670.