Genetic interactions of G-quadruplexes in humans

Katherine G Zyner[1†], Darcie S Mulhearn[1†], Santosh Adhikari[2],
Sergio Martínez Cuesta[1], Marco Di Antonio[2], Nicolas Erard[1], Gregory J Hannon[1],
David Tannahill[1], Shankar Balasubramanian[1,2,3]*

[1]Cancer Research United Kingdom Cambridge Institute, Cambridge, United Kingdom; [2]Department of Chemistry, University of Cambridge, Cambridge, United Kingdom; [3]School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom

Abstract G-quadruplexes (G4) are alternative nucleic acid structures involved in transcription, translation and replication. Aberrant G4 formation and stabilisation is linked to genome instability and cancer. G4 ligand treatment disrupts key biological processes leading to cell death. To discover genes and pathways involved with G4s and gain mechanistic insights into G4 biology, we present the first unbiased genome-wide study to systematically identify human genes that promote cell death when silenced by shRNA in the presence of G4-stabilising small molecules. Many novel genetic vulnerabilities were revealed opening up new therapeutic possibilities in cancer, which we exemplified by an orthogonal pharmacological inhibition approach that phenocopies gene silencing. We find that targeting the WEE1 cell cycle kinase or USP1 deubiquitinase in combination with G4 ligand treatment enhances cell killing. We also identify new genes and pathways regulating or interacting with G4s and demonstrate that the DDX42 DEAD-box helicase is a newly discovered G4-binding protein.
DOI: https://doi.org/10.7554/eLife.46793.001

*For correspondence:
sb10031@cam.ac.uk

†These authors contributed equally to this work

Introduction

G-quadruplex secondary structures (G4s) form in nucleic acids through the self-association of guanines (G) in G-rich sequences to form stacked tetrad structures (reviewed in *Bochman et al., 2012*; *Rhodes and Lipps, 2015*). In the human genome, over 700,000 G4s have been detected in vitro (*Chambers et al., 2015*). Sequences encoding G4s are enriched in regulatory regions consistent with roles in transcription and RNA regulation (*Huppert and Balasubramanian, 2007*; *Huppert, 2008*), and their over-representation in oncogene promoters, such as *MYC*, *KRAS* and *KIT*, suggests that they are important in cancer and are potential therapeutic targets (reviewed in *Balasubramanian et al., 2011*). Computationally predicted G4s have also been linked to replication origins (*Besnard et al., 2012*) and telomere homeostasis (reviewed in *Neidle, 2010*). In the transcriptome, more than 3000 mRNAs have been shown to contain G4 structures in vitro, particularly at 5' and 3' UTRs, suggestive of roles in posttranscriptional regulation (*Bugaut and Balasubramanian, 2012*; *Kwok et al., 2016*).

G4-specific antibodies have been used to visualise G4s in protozoa (*Schaffitzel et al., 2001*) and mammalian cells (*Biffi et al., 2013*; *Henderson et al., 2014*; *Liu et al., 2016*). More G4s are detected in transformed versus primary cells, and in human stomach and liver cancers compared to non-neoplastic tissues, supporting an association between G4 structures and cancer (*Biffi et al., 2014*; *Hänsel-Hertsch et al., 2016*). More recently, ChIP-seq was used to map endogenous G4 structure formation in chromatin revealing a link between G4s, promoters and transcription (*Hänsel-Hertsch et al., 2016*). G4s are found predominately in nucleosome-depleted chromatin within

promoters and 5' UTRs of highly transcribed genes, including cancer-related genes and regions of somatic copy number alteration. G4s may therefore be part of a regulatory mechanism to switch between different transcriptional states. At telomeres, tandem G4-repeat structures also may help protect chromosome ends by providing binding sites for shelterin complex components (reviewed in *Brázda et al., 2014*). As G4 structures can pause or stall polymerases, they must be resolved by helicases to allow replication and transcription to proceed. Several helicases, including WRN, BLM, PIF1, DHX36 and RTEL1, have been shown to unwind G4-structures in vitro (*Brosh, 2013*; *Mendoza et al., 2016*), and it is notable that fibroblasts from Werner (WRN) and Bloom (BLM) syndrome patients, who are predisposed to cancer, show altered gene expression that correlates with sites with potential to form G4s (*Damerla et al., 2012*).

Small molecules that selectively bind and stabilise G4 formation in vitro have been used to probe G4 biological function. G4 ligands, such as pyridostatin (PDS), PhenDC3 and TMPyP4, can reduce transcription of many genes harbouring a promoter G4, including oncogenes such as *MYC*, in multiple cancer cell lines (*Halder et al., 2012*; *McLuckie et al., 2013*; *Neidle, 2017*). G4-stabilising ligands also interfere with telomere homeostasis by inducing telomere uncapping/DNA damage through the inhibition of telomere extension by telomerase leading to senescence or apoptosis (reviewed in *Neidle, 2010*). 5' UTR RNA G4 structures may also be involved in eIF4A-dependent oncogene translation (*Wolfe et al., 2014*) and their stabilisation by G4-ligands can inhibit translation in vitro (*Bugaut and Balasubramanian, 2012*). Identification of several RNA G4-interacting proteins (reviewed in *Cammas and Millevoi, 2016*), including DEAD/DEAH helicases such as DDX3X, and DHX36 (*Chen et al., 2018*; *Herdy et al., 2018*) additionally suggests specific roles for G4 structures in RNA.

Some G4-stabilising ligands cause a DNA damage-response (DDR); for example, DNA damage sites induced by PDS in human lung fibroblasts mapped to genomic regions at G4s within several oncogenes including *SRC* (*Rodriguez et al., 2012*). Subsequent studies demonstrated that homologous recombination (HR) repair deficiencies can be exploited to selectively kill BRCA1/2-deficient cancer cells with G4 ligands (*McLuckie et al., 2013*; *Zimmer et al., 2016*). Recently, this concept has been applied to BRCA1/2-deficient breast cancers using CX-5461, a G4 ligand currently in clinical trials (*Xu et al., 2017*) (NCT02719977 ClinicalTrials.gov). Overall, these initial studies demonstrate that specific genotypes can be selectively vulnerable to G4-stabilisation and raises the question as to what other genotypes might provide further such opportunities.

We set out to address two main questions (*Figure 1*): 1) which human genes and cellular pathways interact with G4s and 2) what genetic backgrounds selectively lead to enhanced cell killing in the presence of G4 stabilising ligands? We employed PDS and PhenDC3 as representative G4 ligands as these are chemically and structurally dissimilar, but each shows a broad specificity for different G4 structural variants. Both ligands have been widely used as G4-targeting probes in biophysical (*De Cian et al., 2007b*; *Rodriguez et al., 2008*) and biological studies in which they have been shown to impart transcriptional inhibition, telomere dysfunction and replication stalling (*De Cian et al., 2007a*; *Halder et al., 2012*; *Mendoza et al., 2016*).

Results

Identification of genetic vulnerabilities to G4-ligands via genome-wide screening

An unbiased genome-wide shRNA screen was performed in A375 human melanoma cells to globally evaluate genetic vulnerabilities to G4-ligands and to identify genes and pathways involved with G4-structures (*Figure 2A*). For this, the pyridine-2,6-bis-quinolino-dicarboxamide derivative, PDS (*Rodriguez et al., 2012*), and bisquinolinium compound, PhenDC3 (*De Cian et al., 2007b*) were chosen (*Figure 2B*). We used the latest generation shERWOOD-Ultramir shRNA pLMN retroviral library, comprising 132,000 shRNAs across 12 randomised pools targeting the protein coding genome, with an average of five optimised hairpins per gene (*Figure 2C*) (*Knott et al., 2014*). A375 melanoma cells were used due to their rapid doubling, stable ploidy and success in other shRNA-dropout screens (*Sims et al., 2011*); they are TP53 wild-type and driven by oncogenic *BRAF* (V600E) and *CDKN2A* loss (*Forbes et al., 2015*). *Figure 2D* outlines our shRNA screening strategy. To identify shRNAs that are lost between the initial (t0) and final (fF) timepoints, unique 3'-antisense
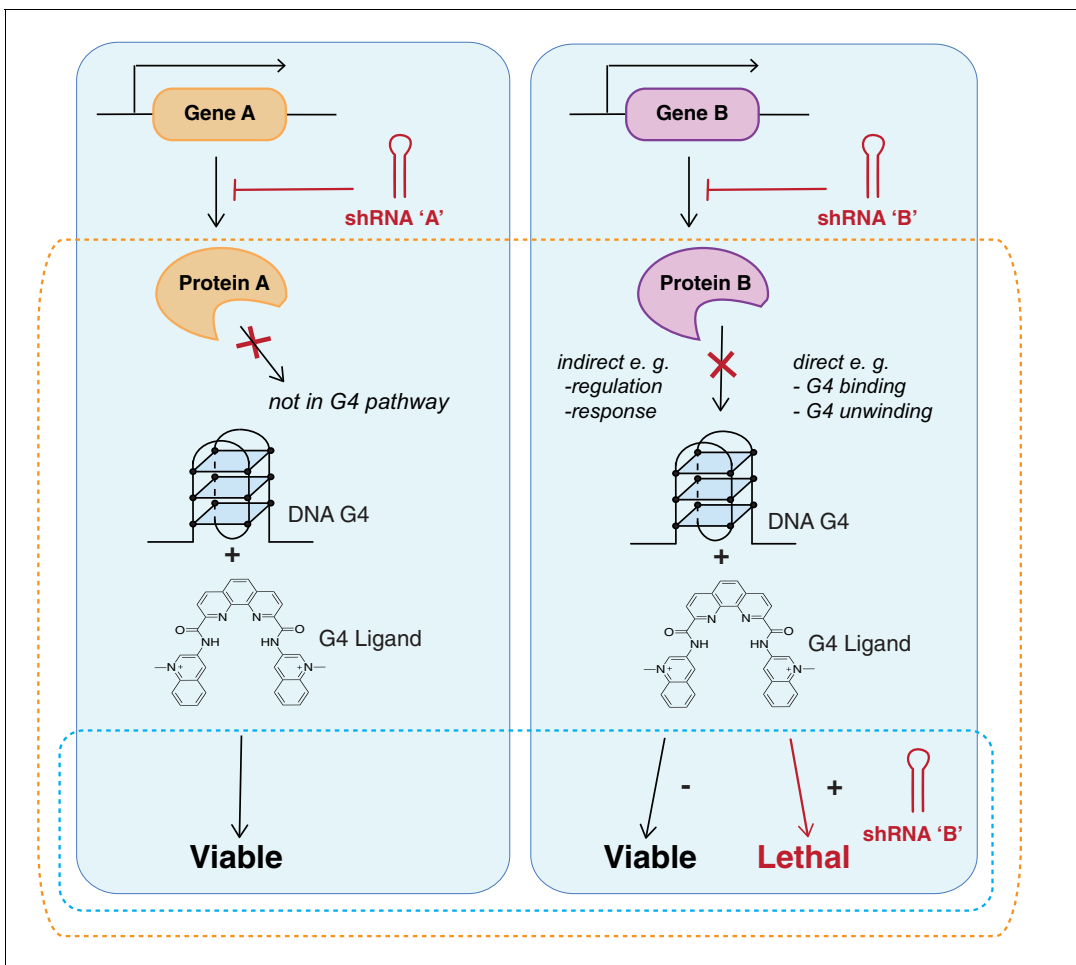
Figure 1. Strategy identifying genetic vulnerabilities involved with G4 biology. Genome-wide shRNA silencing combined with G4 structure stabilisation by small molecules identifies genes that when depleted compromise cell viability. Cells are infected with a genome-wide pool of shRNA lentiviruses targeting the protein coding genome followed by G4 ligand treatment to stabilise genomic and/or RNA G4 structures. Two general outcomes are possible: a gene is not required in a G4-dependent process so there is no effect on cell viability (left); or gene silencing results in cell death either due to loss of a direct G4 interaction (e.g. binding/unwinding) or indirectly through gene loss in a G4-dependent pathway (right). In absence of ligand, cells are viable in presence of the shRNA. Dotted boxes highlight genotypes of disease significance for possible G4-based therapies (blue) and genes and biological pathways that involve and/or interact with G4 structures (orange).

DOI: https://doi.org/10.7554/eLife.46793.002

sequences were recovered by PCR and quantified by sequencing. If a gene knockdown compromises cell viability then the associated shRNA will be depleted compared to those targeting non-essential genes: the tF sequence count will be less than t0 thus $\log_2$ fold change (FC, tF/t0) is negative. A pilot using one shRNA pool established that a tF of 15 population doublings can be used to reveal significant G4-ligand-mediated changes [false discovery rate (FDR) ≤ 0.05] in shRNA levels using a ligand concentration resulting in 20% cell death (GI_{20}, see Materials and methods and *Figure 2—figure supplement 1* for details).

To understand the complete spectrum of G4 vulnerabilities, we first considered the combined set of sensitivities to PDS and PhenDC3 together. For the whole library, when individual shRNAs are considered 9509 (~7%) G4-ligand-specific hairpins (i.e. those not in DMSO) were found to be depleted (FDR ≤ 0.05; $\log_2$ FC <0, *Figure 3A*, *Supplementary file 1*). We then reasoned, for a gene knockdown to have compromised cell growth, that a minimum of either 50% or three shRNA hairpins should be significantly depleted for that gene (median $\log_2$ FC <0). This resulted in the identification of 843 G4 ligand-specific gene knockdowns not present in DMSO (*Figure 3B*). We then denoted a more stringent preliminary list of 758 G4 sensitisers as those having a median $\log_2$ FC $\leq$
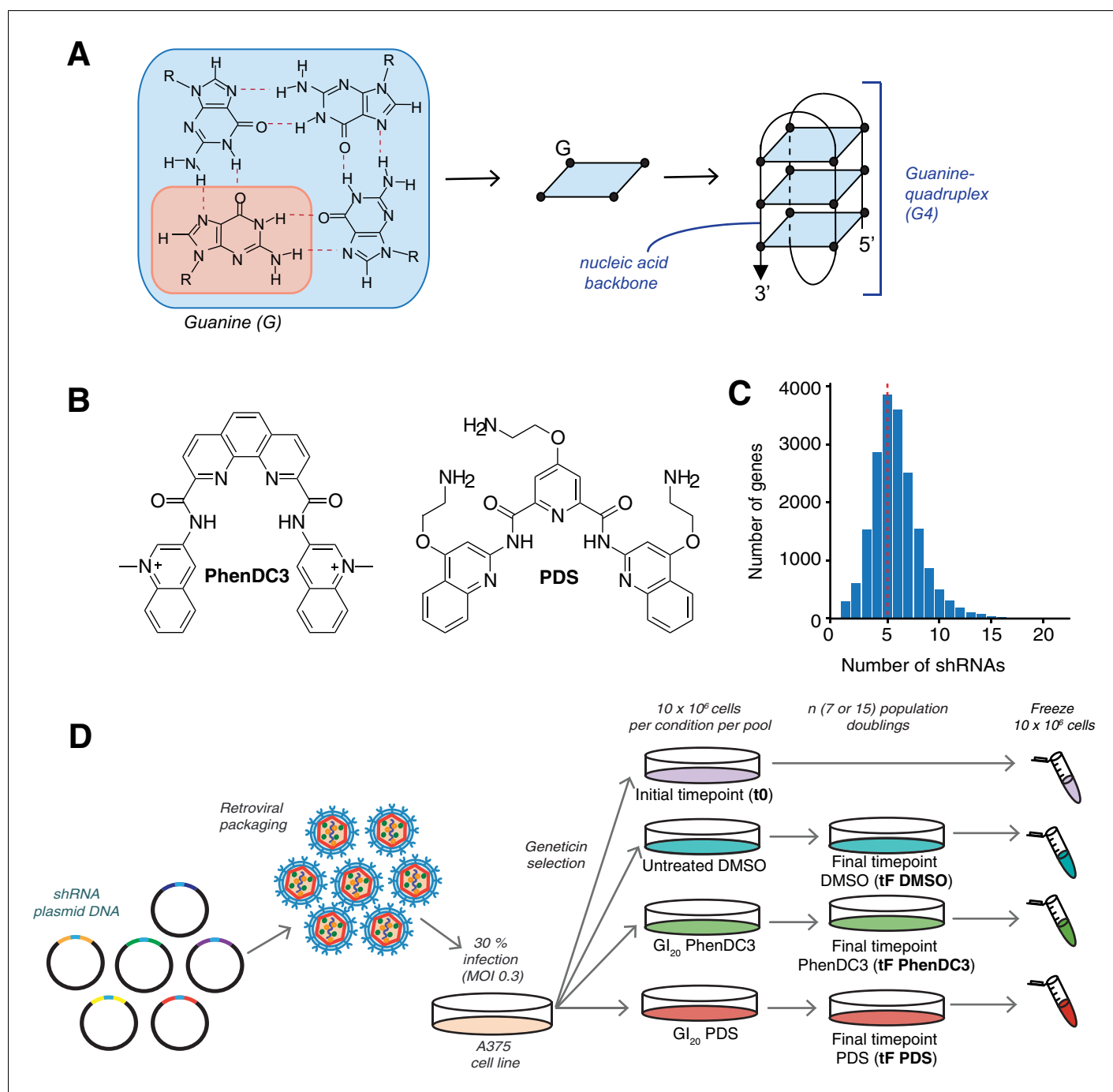
Figure 2. shRNA screening pipeline to uncover genetic vulnerabilities to G4 stabilisation. (A) A G-tetrad with four interacting guanines (left), which stack to form G4 structures (right). (B) Structures of the G4-stabilising small molecule ligands PDS and PhenDC3. (C) Distribution of the numbers of shRNAs targeting each gene, with the average indicated by a red dotted line. (D) Overall screening approach illustrated for one library pool. Plasmids are retrovirally packaged and A375 cells are infected at multiplicity of infection (MOI) of 0.3 (30%). Following antibiotic selection, an initial time point (t0) is harvested and then cells are cultured for 'n' population doublings in DMSO, PDS or PhenDC3 before the final time point was harvested (tF).

DOI: https://doi.org/10.7554/eLife.46793.003

The following figure supplement is available for figure 2:

Figure supplement 1. Genome-wide shRNA screen parameter optimisation.

DOI: https://doi.org/10.7554/eLife.46793.004

A Depleted shRNAs

B Depleted genes (log2FC < 0)

C G4 sensitiser genes (log2FC ≤ -1)

D Validated G4 ligand synthetic lethalities

Gene	Hairpins in t0	Depleted hairpins (PDS)	Median log2FC (PDS)	Depleted hairpins (PhenDC3)	Median log2FC (PhenDC3)	PUBMED/ BIORXIV (Rx) ID
ATRX	16	1	-1.6	5	-1.4	23563309; Rx 347542
BRCA1	5	3	-2.8	3	-2.1	28211448, 26748828
BRCA2	3	3	-2.2	2	-1.8	28211448, 26748828, 23782415
HERC2	4	0	-	2	-1.4	Rx 341420

E Additonal G4-related genes (from UniprotKB, GO & G4IPDB)

Gene	G4IPDB	GO term/ Uniprot KB G4 association	Hairpins in t0	Depleted hairpins (PDS)	Median log2FC (PDS)	Depleted hairpins (PhenDC3)	Median log2FC (PhenDC3)	PUBMED ID
ADAR1	YES	-	7	0	-	5	-3.9	24813121, 23381195
DHX36	YES	GO: G4 DNA binding, G4 RNA binding	5	1	-1.9	4	-4.6	29269411, 28069994, 25653156, 25611385, 24151078, 22238380, 21149580, 18842585, 16150737
DNA2	-	GO: G4 DNA unwinding	5	3	-2.2	0	-	23604072
FUS	YES	-	3	2	-3.2	-	-	23521792
MCRS1	-	GO: G4 RNA binding	2	0	-	1	-1.7	16571602
RECQL4	YES	-	6	3	-1.8	2	-2.6	25336622
SF3B3	YES	-	4	0	-	3	-2.9	23381195
XRN1	-	GO: G4 DNA binding, G4 RNA binding	4	3	-1.1	3	-2.7	9049243

F Manually curated top ranked G4-related genes (from Pubmed/PMC)

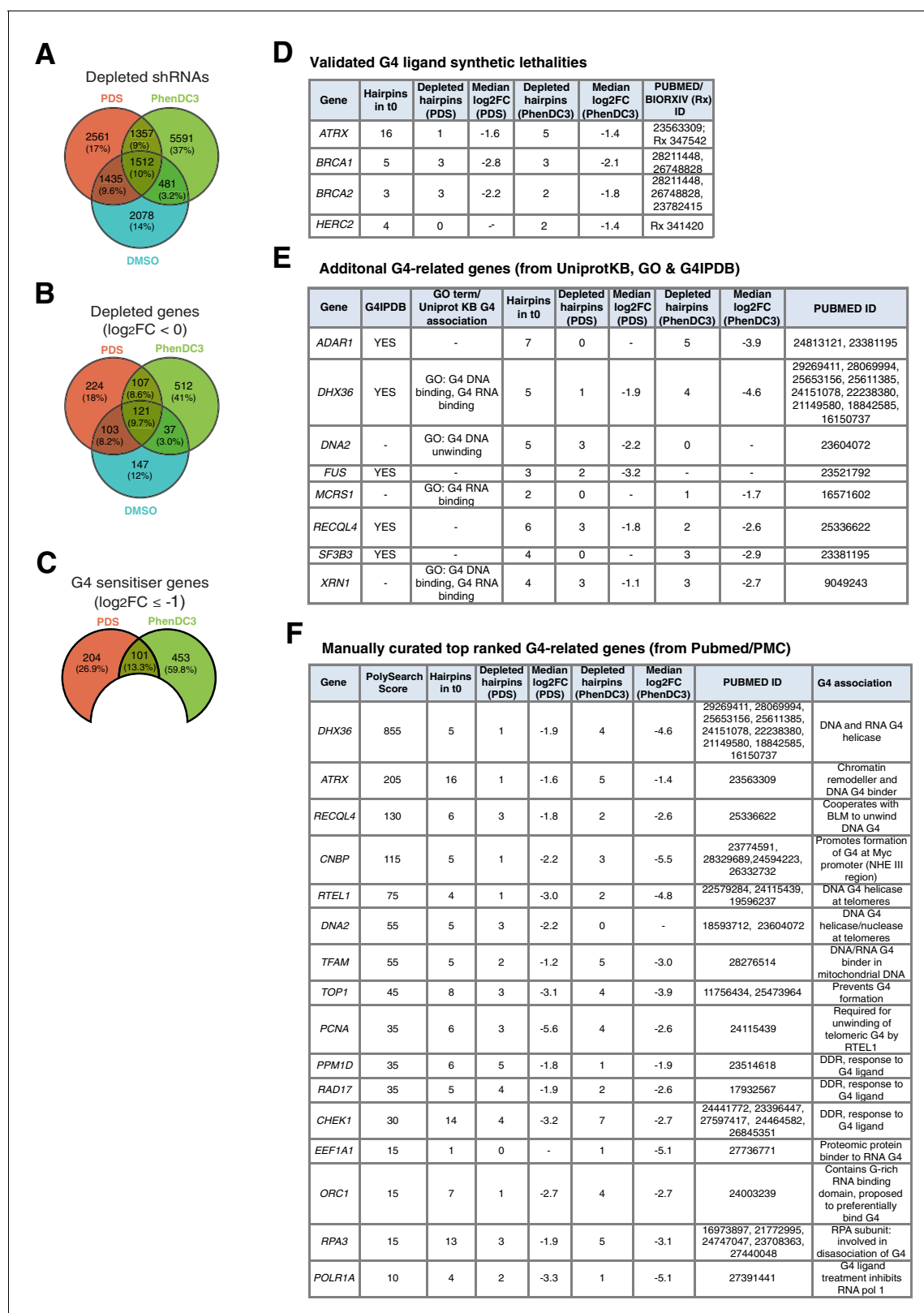| Gene	PolySearch Score	Hairpins in t0	Depleted hairpins (PDS)	Median log2FC (PDS)	Depleted hairpins (PhenDC3)	Median log2FC (PhenDC3)	PUBMED ID	G4 association
DHX36	855	5	1	-1.9	4	-4.6	29269411, 28069994, 25653156, 25611385, 24151078, 22238380, 21149580, 18842585, 16150737	DNA and RNA G4 helicase
ATRX	205	16	1	-1.6	5	-1.4	23563309	Chromatin remodeller and DNA G4 binder
RECQL4	130	6	3	-1.8	2	-2.6	25336622	Cooperates with BLM to unwind DNA G4
CNBP	115	5	1	-2.2	3	-5.5	23774591, 28329689, 24594223, 26332732	Promotes formation of G4 at Myc promoter (NHE III region)
RTEL1	75	4	1	-3.0	2	-4.8	22579284, 24115439, 19596237	DNA G4 helicase at telomeres
DNA2	55	5	3	-2.2	0	-	18593712, 23604072	DNA G4 helicase/nuclease at telomeres
TFAM	55	5	2	-1.2	5	-3.0	28276514	DNA/RNA G4 binder in mitochondrial DNA
TOP1	45	8	3	-3.1	4	-3.9	11756434, 25473964	Prevents G4 formation
PCNA	35	6	3	-5.6	4	-2.6	24115439	Required for unwinding of telomeric G4 by RTEL1
PPM1D	35	6	5	-1.8	1	-1.9	23514618	DDR, response to G4 ligand
RAD17	35	5	4	-1.9	2	-2.6	17932567	DDR, response to G4 ligand
CHEK1	30	14	4	-3.2	7	-2.7	24441772, 23396447, 27597417, 24464582, 26845351	DDR, response to G4 ligand
EEF1A1	15	1	0	-	1	-5.1	27736771	Proteomic protein binder to RNA G4
ORC1	15	7	1	-2.7	4	-2.7	24003239	Contains G-rich RNA binding domain, proposed to preferentially bind G4
RPA3	15	13	3	-1.9	5	-3.1	16973897, 21772995, 24747047, 23708363, 27440048	RPA subunit: involved in disassociation of G4
POLR1A	10	4	2	-3.3	1	-5.1	27391441	G4 ligand treatment inhibits RNA pol 1

Figure 3. Genome-wide screening in A375 cells reveals deficiencies in known G4-associated genes as sensitive to G4-stabilising small molecules. (A–C) Venn diagrams for: (A) significantly differentially expressed individual shRNAs (FDR ≤ 0.05); (B) significantly depleted genes (50% or three hairpins, FDR ≤ 0.05, median log2FC < 0) following DMSO, PDS and PhenDC3 treatment and (C) Significant PDS and PhenDC3 sensitiser genes not in DMSO and after applying a median log2FC ≤ −1 cut off. (D–F) Tables showing the number of depleted hairpins and median log2FC values for: (D) known G4

Figure 3 continued on next page

-1 (*Figure 3C*). It is reassuring that in this list we independently validated the known G4 sensitisers *BRCA1/2, ATRX* and *HERC2* (*McLuckie et al., 2013*; *Wang et al., 2019*; *Watson et al., 2013*; *Wu et al., 2018*; *Xu et al., 2017*; *Zimmer et al., 2016*; *Figure 3D*).

We next explored further genes already implicated in G4 biology, but whose deficiency has not yet been linked with any enhanced sensitivity to G4 ligands. For genes annotated with G4-related terms in the UniprotKB, Gene Ontology (GO) and G4IPDB databases (*Mishra et al., 2016*), an additional eight sensitisers (*ADAR, DHX36, DNA2, FUS, MCRS1, RECQL4, SF3B3 and XRN1*) were uncovered (*Figure 3E*). Text-mining with G4 search terms using PolySearch2 on PubMed abstracts and open access full texts (see Materials and methods; *Liu et al., 2015b*) revealed a further 12 sensitisers arising from our screen including helicases (*RTEL1*), DDR components (*CHEK1, RAD17*), transcriptional proteins (*POLR1A, CNBP*) and replication factors (*ORC1, RPA3, TOP1*) (*Figure 3F*).

Within the total 758 G4-sensitiser gene list, we uncovered five significant enriched KEGG pathway clusters (p<0.05): 'cell cycle', 'ribosome', 'spliceosome', 'ubiquitin-mediated proteolysis' and 'DNA replication' (*Figure 4A*, *Supplementary file 1*). Within each cluster are gene targets common to both G4 ligands, as well as genes unique to each ligand. To gain functional insights, enriched GO 'Biological Process' and 'Molecular Function' terms were determined (*Figure 4B*; *Supplementary file 1*) which showed 20 out of 45 of the former and all the latter terms into DNA or RNA classifications, consistent with PDS/PhenDC3 directly binding nucleic acid G4 targets. Furthermore, when protein domains were considered using GENE3D and PFAM databases (*Figure 4C*), we discovered enrichments in helicase C-terminal domains, RNA recognition motifs including RRM, RBD and RNP domains, and DNA-binding domains including zinc fingers, bZIP motifs and HMG boxes. Consistent with the ubiquitin-mediated proteolysis KEGG cluster, enrichments in multifunctional ATPase domains and in ubiquitin hydrolase domains, were also found. These latter findings suggest important areas of biology not previously known to be affected by G4 intervention in mammalian cells.

Cancer-associated gene depletion enhances sensitivity to G4-ligands

We next used the complete list of 758 genes, identified as stringent G4 ligand sensitisers above, to discover new cancer-associated gene vulnerabilities to G4-stabilising ligands. For this, we searched this list for any significant enrichment in the COSMIC database (v83) of genes causally implicated in cancer (*Forbes et al., 2015*). Of the 758 sensitisers, there was a two-fold enrichment (p=9.1×10^{-6}) for 50 cancer-associated genes, which increases to three-fold (p=2.5×10^{-3}) when considering only sensitisers common to both G4 ligands (*Figure 5A,B*, *Supplementary file 1*). Notably, when STRING network analysis (*Szklarczyk et al., 2017*) was used to investigate functional interactions, this revealed a DDR cluster that included *BRCA1* and *BRCA2*, as well as their interacting tumour suppressor partners *PALB2* and *BAP1*, two cancer-associated DDR genes not previously indicated as G4 ligand sensitisers. (*Figure 5C*). This analysis also identified as sensitisers a cluster consisting of several chromatin modifiers including *SMARCA4, SMARCB1* and *SMARCE1*.

Focused G4-sensitiser shRNA screening reveals robust G4-ligand genetic vulnerabilities and potential therapeutic targets

To enable more rigorous and further comparative analyses that focus solely on G4 sensitisers, we developed a custom shRNA screening panel encompassing the gene sensitisers identified above plus additional G4-associated genes noted from the literature (*Figure 6A*, *Figure 6—figure supplement 1*, see Materials and methods). This panel consisted of a single retroviral shRNA pool to allow all shRNAs to be screened simultaneously under standardised conditions and to minimise technical fluctuations. We first used this panel to recapitulate the findings of the genome-wide screen above
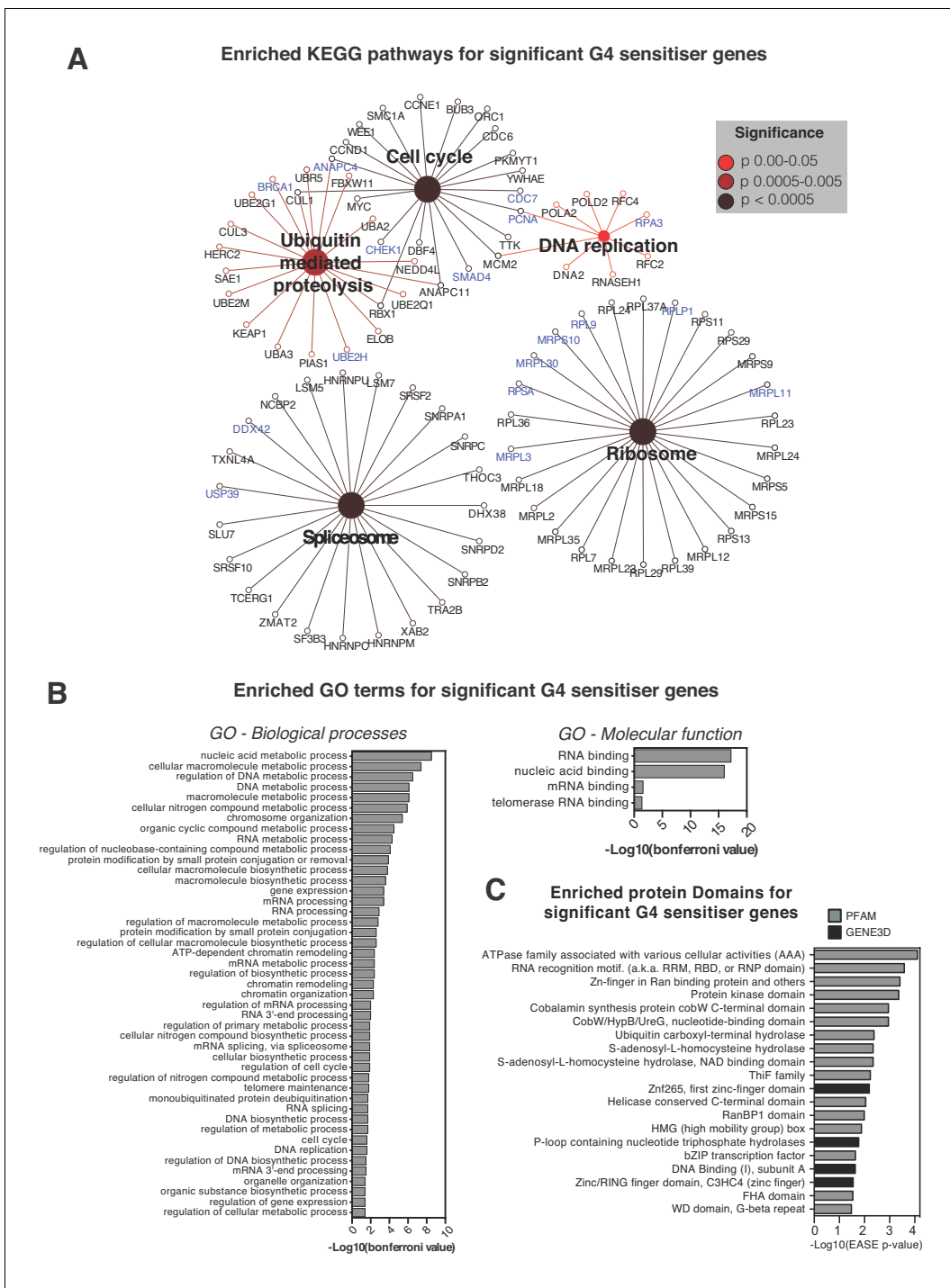
Figure 4. Pathways and processes showing sensitivity to G4-stabilising ligands. (**A**) Enriched KEGG pathways and (**B**) Gene Ontology terms, GO Biological Processes (BP) and Molecular Functions (MF), for the 758 genome-wide G4-sensitiser genes. Blue- genes common to both ligands; black-genes unique to either PDS or PhenDC3. A right-sided enrichment test with Bonferroni correction used (see Materials and methods). (**C**) Enriched protein domains (p≤0.05) within GENE3D (black) and PFAM databases (grey) ordered by -Log10 (EASE p-value). See also *Supplementary file 1*.
DOI: https://doi.org/10.7554/eLife.46793.006

and compare responses with different G4 ligands. Using A375 melanoma cells with PDS and PhenDC3, the custom panel recovered a total of 342 G4 sensitisers corresponding to 40.6% overlap (308 genes) with the complete genome-wide screen (*Figure 6B,C*). From this, we identified 290 G4

sensitisers with 89 and 161 unique for PDS and PhenDC3, respectively, and 40 genes common for both ligands (*Figure 6—figure supplement 1E*). Comparing PDS and PhenDC3 sensitisers by KEGG analysis shows that each ligand mostly interacts with different but related pathways (*Figure 6D,E*). Consistent with direct G4-targeting, nucleic-acid-related GO terms were enriched (*Figure 6—figure supplement 1F & G*, *Supplementary file 2*). We next considered that the 40 sensitiser genes common between PDS and PhenDC3 reflected the most robust sensitisers for G4 ligands in general and it is notable that 27 out of 40 associated with DNA or RNA binding processes, such as chromatin modification, replication transcription, and translation (*Figure 6F*). Again, the ubiquitin processes, which previously were not linked with G4 biology, were also uncovered as a significant sensitiser pathway. Overall, these results clearly show the spectrum of biological vulnerabilities that underpin the observed enhanced sensitivities for each G4-targeting ligand.

We next reasoned that the robust set of 290 G4 ligand sensitiser genes above provides a suitable test bed for exploring the arising therapeutic potential for combinatorial pharmacological inhibition and G4-ligands. We therefore looked for the presence of these sensitisers genes within the druggable genome interaction database (DGIdb) (*Griffith et al., 2013*). A total of 74 G4-sensitisers were found in the classifications 'Druggable Genome' (genes with known or predicted drug interactions) and 'Clinically Actionable' (genes used in targeted clinical cancer sequencing for precision medicine) with 13 being common to both classifications (*Figure 6G*, *Supplementary file 1*). Notably, this included KEAP1, an E3 ubiquitin ligase adapter protein and highlights a new therapeutic domain for the application of G4-based drugs. Performing a similar analysis on the 40 most robust sensitisers common to both G4 ligands gave 12 genes within DGIdb (*Figure 6H*, *Supplementary file 1*), including 5 (*BRCA1, CHEK1, CDK12, TOP1, PDKP1*) common to both druggable and clinically actionable classifications. These results therefore open up new possibilities for cancer therapies based on vulnerabilities to G4 ligands.

G4 sensitisers common to two independent cell lines

We next sought to extend the use of the custom shRNA lentiviral library to gain initial insights into possible commonalities and differences in the response to G4 ligands in cells from different lineages. We therefore applied the custom library to mesenchymal-derived HT1080 fibrosarcoma cells (wild-type *TP53*, driven by activated *NRAS* (Q61K) and *IDH1* mutation (R132C)) and compared the results to those from ectodermal A375 melanoma cells above (*Figure 7*, *Figure 7—figure supplement 1F & G*, *Supplementary files 1 & 2*). The custom HT1080 screen recovered a total of 121 G4 ligand sensitisers, with the majority (73 genes, 58%) shared with those seen for each ligand in the A375 genome-wide screen. Cytoscape network analysis (*Figure 7A*) revealed a core set of G4-associated genes/pathways for these genes in spliceosome, HR and ubiquitin-mediated proteolysis processes (p<0.0005). Overall, 29 PDS and 22 PhenDC3 gene sensitivities were found to be shared across all three screens (*Figure 7B,C*), and it is noteworthy that both G4 ligands targeted similar processes including transcription, splicing and ubiquitin-mediated proteolysis (*Figure 7D,E*).

BRCA1, TOP1, DDX42 and *GAR1* are key G4 ligand sensitiser genes

When we evaluated the data collectively from all screens, it was apparent that four genes were repeatedly found as G4 ligand sensitisers- *BRCA1, TOP1, DDX42* and *GAR1,* as they consistently appeared in both cell types and with both G4-ligands in all screens (*Figure 7F*, *Figure 7—figure supplement 1F*). To corroborate these genes as genuine G4 sensitisers, we developed an independent siRNA knockdown approach using a shorter timeframe (~6 days) to recapitulate ligand-induced growth inhibition (*Figure 8*). Both A375 and HT1080 cells were transfected with siRNAs targeting *BRCA1, TOP1, DDX42* or *GAR1* alongside non-targeting siRNA and non-transfected controls. Following 24 hr, cells were treated with two concentrations of PDS and PhenDC3 or vehicle control DMSO for 144 hr. Growth curves for non-transfected and non-targeting siRNA controls were similar across ligand treatments in both cell lines (*Figure 8—figure supplements 1* and *2*). For both HT1080 (*Figure 8A & B*) and A375 cells (*Figure 8—figure supplement 3A & B*), protein depletion following siRNA transfection was confirmed after 48 hr by immunoblotting cell lysates with the appropriate antibodies (average 76–92% knockdown for HT1080; 41–69% knockdown for A375 after 48 hr). The percentage difference in confluency compared to non-targeting siRNA control cells was
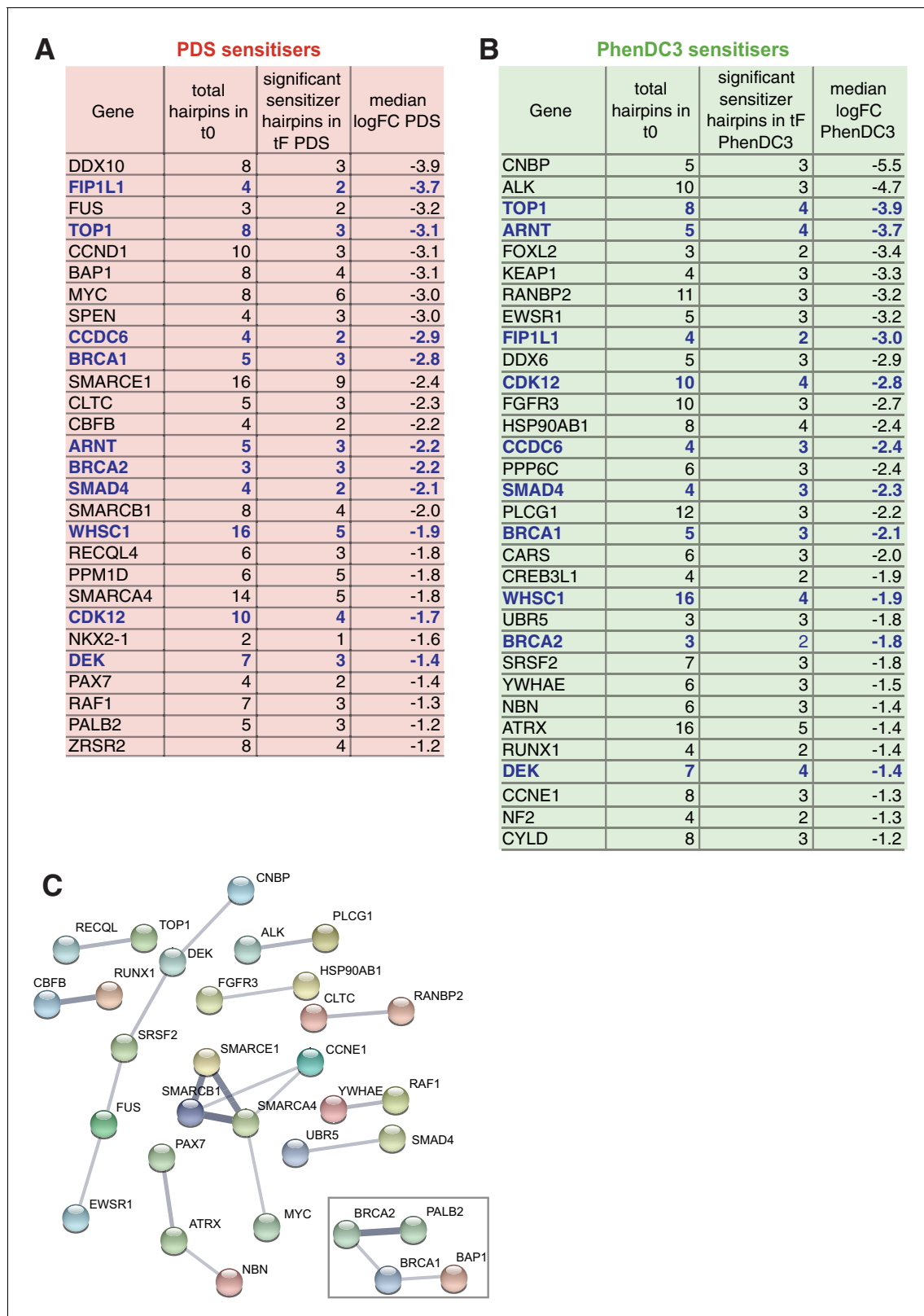
A   PDS sensitisers

Gene	total hairpins in t0	significant sensitizer hairpins in tF PDS	median logFC PDS
DDX10	8	3	-3.9
FIP1L1	4	2	-3.7
FUS	3	2	-3.2
TOP1	8	3	-3.1
CCND1	10	3	-3.1
BAP1	8	4	-3.1
MYC	8	6	-3.0
SPEN	4	3	-3.0
CCDC6	4	2	-2.9
BRCA1	5	3	-2.8
SMARCE1	16	9	-2.4
CLTC	5	3	-2.3
CBFB	4	2	-2.2
ARNT	5	3	-2.2
BRCA2	3	3	-2.2
SMAD4	4	2	-2.1
SMARCB1	8	4	-2.0
WHSC1	16	5	-1.9
RECQL4	6	3	-1.8
PPM1D	6	5	-1.8
SMARCA4	14	5	-1.8
CDK12	10	4	-1.7
NKX2-1	2	1	-1.6
DEK	7	3	-1.4
PAX7	4	2	-1.4
RAF1	7	3	-1.3
PALB2	5	3	-1.2
ZRSR2	8	4	-1.2

B   PhenDC3 sensitisers

Gene	total hairpins in t0	significant sensitizer hairpins in tF PhenDC3	median logFC PhenDC3
CNBP	5	3	-5.5
ALK	10	3	-4.7
TOP1	8	4	-3.9
ARNT	5	4	-3.7
FOXL2	3	2	-3.4
KEAP1	4	3	-3.3
RANBP2	11	3	-3.2
EWSR1	5	3	-3.2
FIP1L1	4	2	-3.0
DDX6	5	3	-2.9
CDK12	10	4	-2.8
FGFR3	10	3	-2.7
HSP90AB1	8	4	-2.4
CCDC6	4	3	-2.4
PPP6C	6	3	-2.4
SMAD4	4	3	-2.3
PLCG1	12	3	-2.2
BRCA1	5	3	-2.1
CARS	6	3	-2.0
CREB3L1	4	2	-1.9
WHSC1	16	4	-1.9
UBR5	3	3	-1.8
BRCA2	3	2	-1.8
SRSF2	7	3	-1.8
YWHAE	6	3	-1.5
NBN	6	3	-1.4
ATRX	16	5	-1.4
RUNX1	4	2	-1.4
DEK	7	4	-1.4
CCNE1	8	3	-1.3
NF2	4	2	-1.3
CYLD	8	3	-1.2

C

Figure 5. Identification of cancer-associated genes whose loss promotes sensitivity to G4 ligands. (A, B) Median log$_2$FC and number of significantly depleted hairpins for G4 sensitisers overlapping the COSMIC database for PDS (A) and PhenDC3 (B). Genes common to both are indicated in blue. See also *Supplementary file 1*. (C) Functional interaction network analysis using STRING for the 50 COSMIC proteins indicated in A and B. Clusters are shown using confidence interactions > 0.4 from co-expression and experimental data. Box indicates the DDR cluster.

Figure 5 continued on next page

Figure 5 continued

plotted (*Figure 8—figure supplement 1B–E* and *Figure 8—figure supplement 3C–F*) and compared to DMSO treatment at 72, 96 and 120 hr (*Figure 8C–F*, *Figure 8—figure supplement 3G–J*).

Mirroring the shRNA screen findings, siRNA knockdown of all four genes in HT1080 cells imparted significant increases in sensitivity with PDS or PhenDC3 compared to DMSO. Some differences between the ligands and individual gene knockdowns were noted. For *BRCA1* and *TOP1* the lowest concentration of PDS resulted in the most sensitisation and this was evident early at 72 hr, whereas both PhenDC3 concentrations resulted in similar growth inhibition and was apparent later (*Figure 8C & D*, *Figure 8—figure supplement 1*). For DDX42 and GAR1, growth inhibition was mostly manifest from 96 hr, with both ligands and concentrations being broadly similar (*Figure 8E & F*, *Figure 8—figure supplement 1*). Results with the A375 cells also lend support to our observations, although there were some differences compared to HT1080 cells (*Figure 8—figure supplements 2* and *3*). While *GAR1* knockdown showed a similar sensitivity profile, *BRCA1* and *TOP1* deficiencies were sensitive to PDS but not PhenDC3. *DDX42* knockdown in A375 cells did not reflect the screens ligand sensitivities and this may in part be due to lower knockdown efficiency compared (~40%). Nonetheless, these independent siRNA short-term assays substantiate that *BRCA1*, *TOP1*, *DDX42* and *GAR1* are genetic vulnerabilities to G4 ligands and these may open up future possibilities for therapeutic development.

G4-targeting ligands plus pharmacological inhibitors of G4 sensitiser genes demonstrate synergistic cell killing

One of our aims was to identify potential cancer genotypes where G4-ligands could be therapeutically exploited. Cancers deficient in our newly discovered G4 sensitisers may be preferentially sensitive to G4-ligands as single agents. Alternatively, rather than exploiting a genetic deficiency per se, it may be possible to use pharmacological inhibition of a critical cancer gene product that phenocopies the deficiency in combination with G4 ligands as an orthogonal approach (*Figure 9A*). As proof-of-principle, we systematically evaluated cell death potentiation with the G4 ligand PDS in combination with pharmacological inhibitors for two new G4 sensitisers gene products, the WEE1 kinase or the deubiquitinase USP1 (*Figure 9B*). WEE1 is a crucial G2/M regulator overexpressed in several cancers (*Matheson et al., 2016*), and USP1 is involved in DDR regulation and is overexpressed in non-small cell lung and other cancers (reviewed in *García-Santisteban et al., 2013*). For our studies, we used MK1775 (AZD1775), a WEE1 kinase inhibitor that is being clinically evaluated in several cancers (*Richer et al., 2017*), and pimozide a potent USP1-targeting drug (*Chen et al., 2011a*). HT1080 and A375 cells were cultured in matrix combinations of PDS with MK1775 or pimozide at concentrations surrounding the GI_{50} values and cell viability measured after 96 hr using an end-point ATP luminescence-based assay (CellTiter-Glo, Promega). Combenefit software (*Di Veroli et al., 2016*) was then used to calculate synergy for different treatment combinations in which the percentage growth inhibition compared to single agent controls is used to plot a 3D-dose-response surface of synergy distribution in concentration space (*Figure 9C–F*). In HT1080 cells, synergy was found for both PDS and MK1775 or pimozide combinations (*Figure 9C,D*, *Figure 9—figure supplement 1*) with peak synergies of 21% and 24% at 156 nM PDS with 21 nM MK1775 or 6.25 μM pimozide, respectively (GI_{50} for PDS, MK1775 and pimozide alone = 322 nM, 59 nM and 8.4 μM, respectively). A375 cells showed lower synergy with PDS and MK1775 combination (*Figure 9E*, *Figure 9—figure supplement 1*), with peak synergy of 15% at 8 μM PDS, 444 nM MK1775 (GI_{50} for PDS, MK1775 and pimozide alone = 8.5 μM, 625 nM and 12.2 μM, respectively). The greatest synergy was seen in combinations of PDS and pimozide in A375 cells (*Figure 9F*, *Figure 9—figure supplement 1*) with a peak synergy of 61% at 5.33 μM PDS, 6.25 μM pimozide. Furthermore, long-term clonogenic survival assays revealed a similar potentiation of growth inhibition, albeit at lower compound concentrations, for PDS/MK1775 and PDS/pimozide drug combinations for both cell lines tested (*Figure 9—figure supplement 2*). Altogether, these results validate that appropriate drug combinations can synergistical act as a surrogate for gene deficiencies in the presence of G4 ligands and thus complements the findings uncovered by our genetic screening approach.
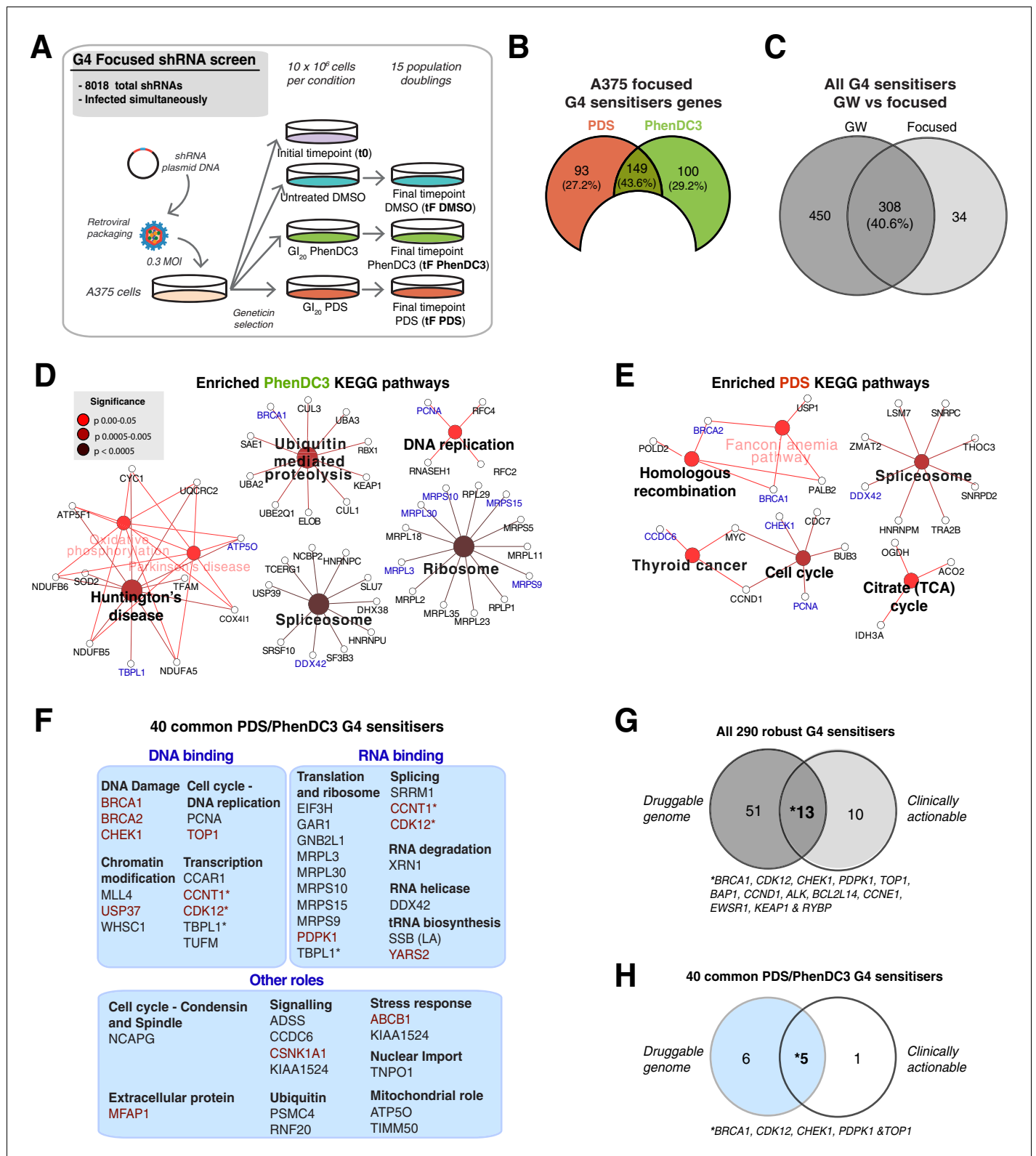
Figure 6. A custom G4 sensitiser shRNA panel reveals unique and common G4 ligand sensitivities. (**A**) A shRNAs custom retroviral pool (~8000 hairpins) was used to infect A375 cells. Following antibiotic selection, the reference time point (t0) was taken and then cells were cultured for 15 population doublings in DMSO, PDS or PhenDC3 before (tF). Three biological replicates were performed. (**B**) Significant sensitiser genes for the A375 focused screen (50% or three significantly depleted with median $\log_2$ FC ≤ -1). (**C**) Overlap of the genome-wide (GW) with A375 focused screen for PDS and PhenDC3 G4-sensitisers combined (see also *Figure 6—figure supplement 1*). (**D–E**) Enriched KEGG pathways for (**D**) PhenDC3 and (**E**) PDS sensitiser

Figure 6 continued on next page

Identification of DDX42 as a new G4-binding protein

Another of our aims was to use the findings of our shRNA screen to identify proteins that may bind and/or regulate G4 structures in cells, such as G4 helicases. Indeed, DHX36 and DHX9, known G4 helicases (*Giri et al., 2011*; *Chen et al., 2018*; *Chakraborty and Grosse, 2011*; *Creacy et al., 2008* ; *Vaughn et al., 2005*) and the DEAD box protein DDX3X, that was recently shown to bind RNA G4s (*Herdy et al., 2018*), were identified as G4 sensitisers in our screen. Further members of the DDX/DHX helicase family also appeared as G4 sensitisers (*Figure 10A*), raising the question of whether these represent previously uncharacterized G4-binding proteins. To address this directly, we chose to investigate DDX42 as this was one of the four key G4 sensitisers identified above. DDX42 is a non-processive RNA helicase (*Uhlmann-Schiffler et al., 2006*) and has been associated with splicing (*Will et al., 2002*); however, this protein remains largely uncharacterised. By immuno-blotting of nuclear and cytoplasmic sub-cellular fractions (*Figure 10B–E*), we first established that DDX42 predominantly localises to the nucleus (~4 to 9-fold greater than cytoplasmic levels) in three independent cell lines, (HT1080, HEK293 and HeLa). As controls for fractionation, LaminB1 and GAPDH were found to partition as expected into nuclear and cytoplasmic fractions, respectively (*Figure 10C,D*).

As DDX42 is known to bind RNA, we next set out to demonstrate DDX42 affinity for a RNA-G4 structure as this has not previously been documented. For this, a G4 RNA oligonucleotide from the NRAS 5'UTR sequence, which forms a stable parallel G4 (*Kumari et al., 2007*), was used together with a mutated oligonucleotide unable to form a G4 structure and also a RNA hairpin as negative controls (*Herdy et al., 2018*). Oligonucleotides were folded in 100 mM KCl to promote G4 structure formation and the resultant structures confirmed by circular dichroism (CD) spectroscopy (*Figure 10—figure supplement 1*). The affinity of recombinant DDX42 was then investigated by Enzyme Linked Immunosorbent Assay (ELISA, *Figure 10F*) and binding parameters calculated using a non-linear regression model, assuming one-site-specific binding and saturation kinetics using Prism software. DDX42 bound the NRAS G4 folded in KCl with an apparent K_d of 71.1 ± 3.5 nM and did not bind detectably to the mutant oligonucleotide or RNA hairpin controls.

Given the nuclear localisation of DDX42 and as some DDX proteins also have DNA helicase activity (*Kikuma et al., 2004*), the DDX42 affinity for a DNA G4 structure was investigated. For this, an oligonucleotide corresponding to the stable parallel G4 structure in the promoter of *MYC* (*González and Hurley, 2010*; *Yang and Hurley, 2006*), and a non-G4 forming control, were used. The oligonucleotides were folded in 100 mM KCl and structures verified by CD spectroscopic analysis (*Figure 10—figure supplement 1B*). DDX42 affinity by ELISA (*Figure 10G*) showed that DDX42 binds to the MYC DNA G4 with an apparent K_d of 232.9 ± 23.5 nM with little binding to the mutant control. Thus, the G4 sensitiser screen has enabled us to identify and classify DDX42 as a G4-interacting protein as a new finding.

Discussion

G4 structures are emerging as promising clinical targets in cancer (*Xu et al., 2017*) but the range of disease-associated genetic backgrounds that potentiate G4 ligand effects has yet to be defined. Here, we have discovered many genes that when depleted enhance cell killing with the G4 ligands
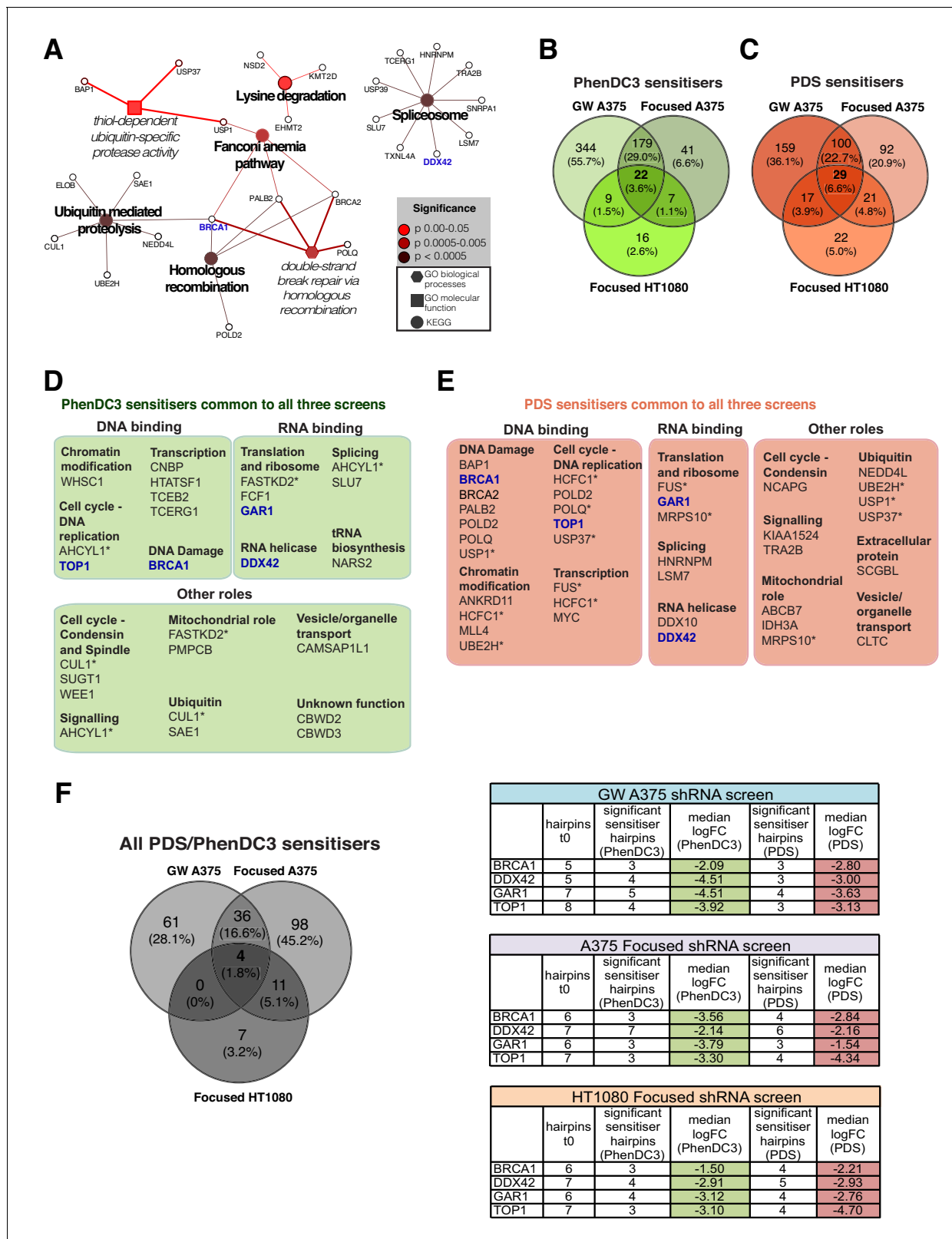
Figure 7. G4 sensitivities in two different cell lines. (A) Enriched KEGG and GO pathways for all G4 ligand-specific sensitisers (73 genes) shared between the genome-wide A375 and HT1080 screens. A right-sided enrichment test with Bonferroni correction used (see Materials and methods). (B–C) Comparison of G4 sensitisers across A375 focused, A375 genome-wide and HT1080 focused screens for (B) PhenDC3 and (C) PDS. (D–E) DAVID, STRING (experimental data, co-expression, medium confidence (≥0.4) interaction) and UniProtKB data analysis showing biochemical functions for

Figure 7 continued on next page

Figure 7 continued

common PhenDC3 (D) and PDS (E) sensitisers across all three screens. *=genes in multiple categories. Blue, four sensitisers common to both ligands. (F) Left, common sensitiser genes across all three screens. Right, number of depleted hairpins and median log$_2$FC values for four key genes found as both PDS and PhenDC3 sensitisers across all the three screens. See also *Figure 7—figure supplement 1*, *Supplementary file 1*.

The following figure supplement is available for figure 7:

Figure supplement 1. Focused HT1080 and A375 cell screening reveals shared PDS and PhenDC3 sensitivities.

PDS and/or PhenDC3. The majority of these have no documented link to G4 biology and the use of low ligand concentrations is likely to favour discovery of gene losses that are the most sensitive in imparting selective cell killing. Validating the success of our approach, we independently identified G4-associated protein coding genes known to be genetic vulnerabilities to G4 ligands including *BRCA1/2*, *HERC2* and *ATRX* (*McLuckie et al., 2013*; *Wang et al., 2019*; *Watson et al., 2013*; *Wu et al., 2018*; *Xu et al., 2017*; *Zimmer et al., 2016*). We now report for the first time genetic vulnerabilities in 20 other known G4-associated genes that promote sensitivity to G4-stabilising ligands. These include direct nucleic acid binders and/or unwinders, such as ADAR, DHX36, DNA2, FUS, MCRS1, RECQL4, SF3B3 and XRN1.

The clinical PARP inhibitor, olaparib has exemplified the concept of synthetic lethality in BRCA-deficient cells (*Bryant et al., 2005*; *Farmer et al., 2005*), and it is notable that BRCA deficiencies were isolated as one of the top genetic vulnerabilities for both G4 ligands in both A375 and HT1080 cells. While PDS and PhenDC3 have not been optimised by medicinal chemistry, the findings of Zimmer et al showing similar efficacy of PDS and olaparib in several BRCA-deficient models (*Zimmer et al., 2016*) lends further support that our screen detects robust, biologically relevant effects.

In dropout screens, dissociating minor from robust growth effects is important and is highly dependent on parameters such as compound dose, genotype and cell line selected. Our screen was designed with stringent parameters to detect genes deficiencies worthy of further exploration. Indeed, we demonstrate potent growth inhibition of up to 80% of the four top G4 sensitisers genes in a parallel siRNA approach.

The gene sensitivities uncovered here have potential to be exploited chemotherapeutically in cancer by deploying a G4-stabilising drug as a single-agent therapy. Alternatively, in the absence of a particular gene deficiency, pharmacological inhibition of a critical oncogene could phenocopy the genetic sensitivities described here and be used in combinatorial treatments with G4-stabilising drugs. This may be attractive as cells are less likely to simultaneously develop resistance against two drugs (reviewed in *Chan and Giaccia, 2011*). Furthermore, as lower drug doses are used, this increases the therapeutic window and has less adverse side effects. As proof-as-principle for this, we selected the WEE1 cell cycle kinase and the deubiquitinase USP1, and demonstrated that their pharmacological inhibition, with MK1775 and pimozide, respectively, leads to the potentiation of cell death in conjunction with the G4 ligand PDS. For example, 5.3 µM PDS or 6.25 µM pimozide alone impart little growth inhibition (14% and 6% respectively), but together they lead to strong growth inhibition (79%). *Table 1* highlights further potential combinatorial opportunities for cancer-associated genes with clinical and/or experimental drugs. Additional therapeutic possibilities for other gene sensitivities that are largely still to be explored from a pharmacological perspective are illustrated in *Table 2*.

While the custom HT1080 screen recovered 58% of sensitisers seen for each ligand in the A375 genome-wide screen, it is striking that this increases to 93% (i.e. 112 out of 121) when considering all screens irrespective of G4 ligand, suggesting remarkable consistency when comparing G4 ligand effects globally. Differences in individual ligand sensitives may arise from variances in cellular uptake and dose, for example, the GI20 dose of PhenDC3 is ten-fold higher for A375 compared to HT1080; G4 ligand-dependent molecular preference for G-tetrad end binding (*Le et al., 2015*) and/or the accessibility of G4s in the chromatin of individual cell lines (*Hänsel-Hertsch et al., 2016*). These points plus differences in protein knockdown efficiency, especially in A375 cells, may contribute to the differences in G4 ligand growth inhibition in our siRNA experiments. In the siRNA experiments, the G4 ligand-induced growth inhibition of both A375 and HT1080 appear not to follow a 'typical'
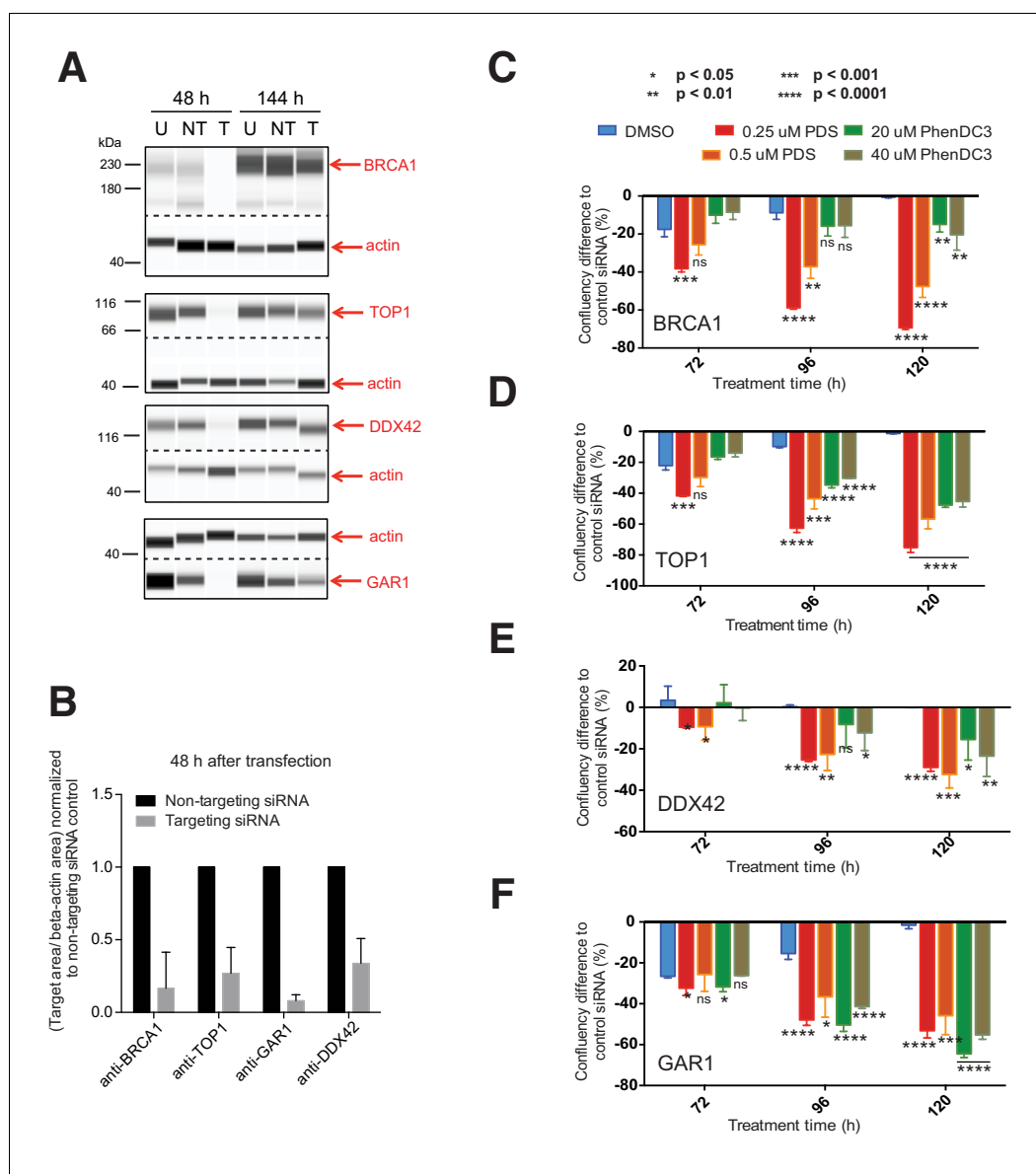
Figure 8. siRNA knockdowns validate *BRCA1*, *TOP1*, *DDX42* and *GAR1 as* key G4 ligand sensitiser genes. (**A**) HT1080 cells were treated with non-targeting (NT) or targeting (T) siRNAs for *BRCA1*, *TOP1*, *DDX42* and *GAR1*. 48 hr and 144 hr after transfection, cell lysates and a non-transfected cell lysate (U) were probed with appropriate antibodies and actin control by western blotting. (**B**) Protein levels for targeting (T) and non-targeting (NT) 48 hr lysates were normalised to the internal actin control and then normalised to NT levels for three biological replicates (mean ± standard deviation). (**C–F**) HT1080 cells were transfected with targeting siRNAs for 24 hr before PDS, PhenDC3 or DMSO treatment. Comparative box plots of confluency differences and significance (unpaired parametric t-test) at selected timepoints for (**C**) *BRCA1*, (**D**) *TOP1*, (**E**) *DDX42*, (**F**) *GAR* (ns = not significant) for three separate siRNA transfections. See also *Figure 8—figure supplements 1*, *2* and *3*.

DOI: https://doi.org/10.7554/eLife.46793.012

The following source data and figure supplements are available for figure 8:

Source data 1. Source files for western blots.
DOI: https://doi.org/10.7554/eLife.46793.017

Figure supplement 1. Short-term siRNA knockdowns of four key sensitisers in HT1080 cells show dose-dependent growth inhibition with G4-ligands.
DOI: https://doi.org/10.7554/eLife.46793.013

Figure supplement 2. Short-term siRNA knockdowns of four key G4-sensitisers in A375 cells show dose-dependent growth inhibition with G4-ligands.

Figure 8 continued on next page

Figure 8 continued

Figure supplement 3. Short-term siRNA knockdowns validate four key sensitivities from the shRNA screening in A375 cells.
Figure supplement 3—source data 1. Source files for western blots.

dose response where higher concentrations lead to greater effects. This may in part be due to there being an optimum G4 ligand dose for a particular gene loss leading to enhanced cell death. Indeed, it is thought that lower drug concentrations better fall within a 'synthetic lethality window' (*Nijman, 2011*). Higher doses may mask these effects, by targeting more G4s that are not dependent on the particular gene lost and/or be due to other off-target effects. This is also supported by the experiments in *Figure 9* that show synergy is only apparent at defined concentrations.

Our data additionally provides insights into the possible functions of the identified G4 sensitisers and indicates roles in DNA damage response (DDR), transcription/chromatin remodelling, nucleic acid unwinding, splicing and ubiquitin-mediated proteolysis. Our findings substantially advance our knowledge of G4 interactions with DDR beyond *BRCA1/2* as several key HR genes were identified as novel G4-sensitisers including *PALB2*, *BAP1* and the deubiquitinase *USP1*. Importantly, this highlights that such HR repair mechanisms are an integral and important cellular response in preventing cell death induced through the increased persistence of G4s. Persistent G4 structures are also inhibitory to DNA replication/cell cycle progression (reviewed in *Valton and Prioleau, 2016*), and it is of note that we also uncovered many cell cycle/DNA replication sensitivities such as *PCNA*, *CHEK1*, *CCND1*, *CDC7*, *RFC2* and *RFC4.* Taken together these suggest that G4 stabilisation with small molecules could be an attractive therapeutic strategy to inhibit cell growth.

Deficits in G4-unwinding helicases are predicted to increase the persistence of G4 structures resulting in heightened sensitivity to G4 ligands. Several known G4-associated helicase deficiencies were recovered, including *RECQL4*, *RTEL1* and *DHX36*, alongside many others with no known G4 link (see *Figure 10A*). Here, we demonstrate for the first time that the DDX42 DEAD/DEAH helicase is in fact a previously unidentified structure-specific G4-binding protein. On a wider level, this acts as proof-of-principle that other specific G4 interacting proteins exist within the sensitiser list of over 700 proteins. Other known G4-helicases such as BLM, WRN, PIF1 and FANCJ (reviewed in *Wu and Brosh, 2010*) were not identified as sensitisers, which may reflect functional redundancy (*Spillare et al., 2006*), or a low ligand concentrations and/or cell type effects.

Our findings highlight the ubiquitin-protesome pathway and modifications such, as neddylation as unexplored areas with respect to G4s. The only documented ubiquitin-G4 relationship in human cells is with HERC2, an E3 ubiquitin ligase that is implicated in G4 resolution whose loss sensitises cells to G4 ligands (*Wu et al., 2018*). We also independently validate *HERC2* as a G4 sensitiser in our screen and extend our observations to cover the full breadth of the proteosomal degradation pathway, including members of E1 ligase (*UBA3*, *UBA2*, *SAE1*), E2 ligase (*UBE2H*), E3 ligase (*NEDD4L*, *RBX1*, *CUL1*, *RNF20*), deubiquitinating enzyme (*USP1* and *USP37*) and proteosome (*PSMC2)* families (see *Table 2*) (*Senft et al., 2018*; *Wei and Lin, 2012*) Given the involvement of ubiquitin-proteasomal regulation in pathways, such as DDR and cell cycle, that are generally deregulated in cancer (*Harrigan et al., 2018*), this opens up an interesting intersection between ubiquitin regulation and G4s. As ubiquitin components are being targeted for anticancer therapies (*Huang and Dixit, 2016*), their efficacy might be enhanced through simultaneous G4 targeting and here we have provided strong proof-of-principle of this using synergistic combinations of pimozide (targeting UPS1) and the G4 ligand PDS.

In contrast to other genetic screens identifying sensitiser genes that enhance the efficacy of anti-cancer agents (*Azorsa et al., 2009*; *Martens-de Kemp et al., 2017*), our work suggests that persistent G4s are problematic for splicing. We identified several cancer-associated splicing factors as G4 sensitisers, including SRSF10, HNRNPM and the known G4-interactor FUS, which is overexpressed in several cancers (*Crozat et al., 1993*; *Dvinge et al., 2016*; *Takahama et al., 2013*). For the latter, a drug inhibiting general spliceosome assembly (*Table 1*) has been pharmacologically explored
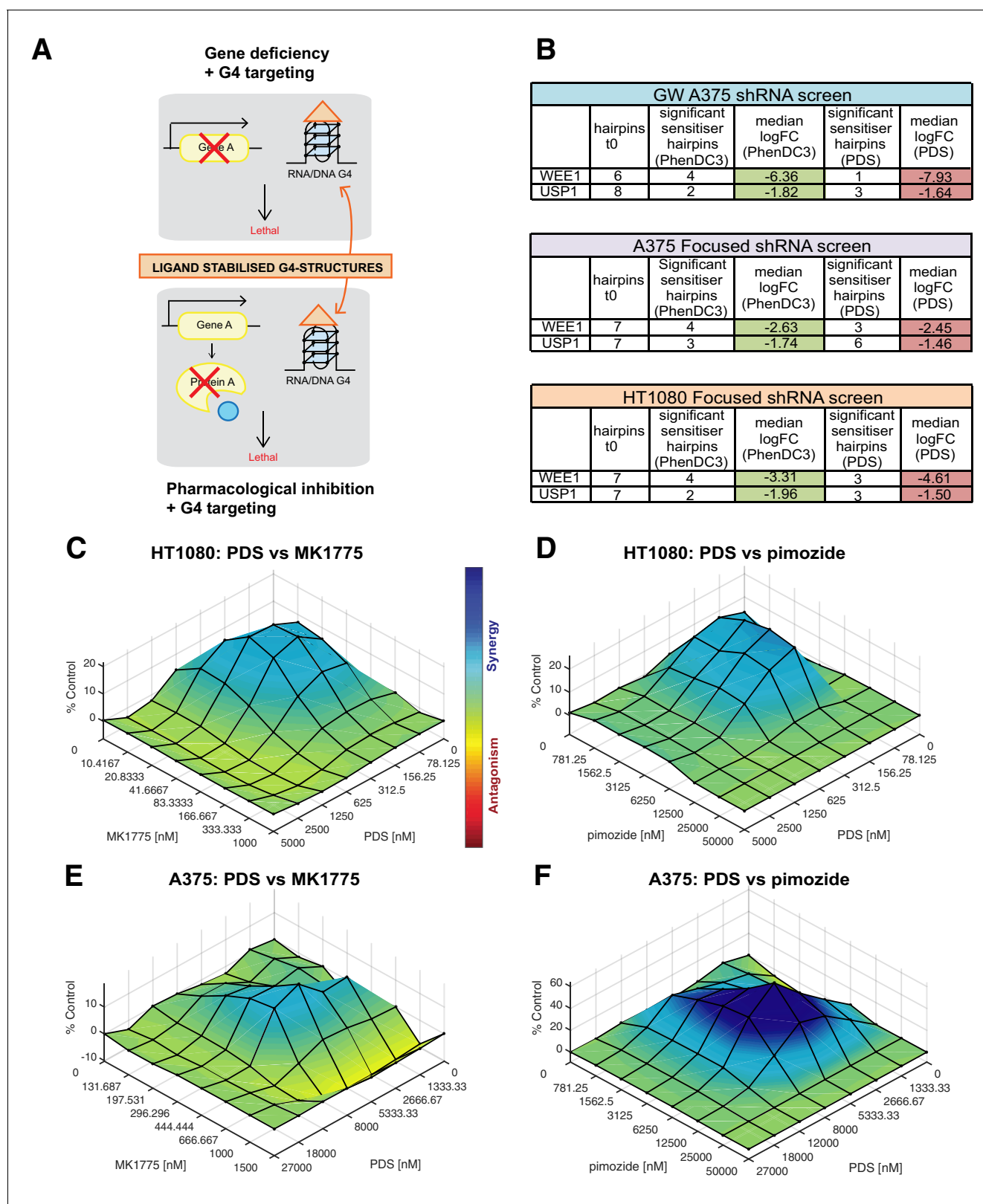
Figure 9. Cell death potentiation mediated by pharmacological inhibition of WEE1 or USP1 with the G4-stabilising ligand PDS. (A) Cell death potentiation with G4-stabilising ligands in combination with either gene deficiencies, such as shRNA-mediated knockdown (top), or pharmacological inhibition of a protein (bottom). (B) Numbers of depleted shRNA hairpins and median $\log_2$FC values for WEE1 and USP1 in the genome-wide and focused screens. (C–F) Synergy plots for HT1080 (C, D) and A375 (E, F) cells treated with PDS in combination with MK1775 (C, E) or pimozide (D, F). To

Figure 9 continued on next page

Figure 9 continued

determine any synergy in cell killing, 3D response surface plots were calculated using Combenefit software with the BLISS model for an average of three biological replicas. Heat bar- blue shading indicates synergy combinations, red indicates antagonism (see also *Figure 9—figure supplements 1* and *2*).

DOI: https://doi.org/10.7554/eLife.46793.018

The following figure supplements are available for figure 9:

Figure supplement 1. Synergy calculations for PDS with MK1775 and pimozide.

DOI: https://doi.org/10.7554/eLife.46793.019

Figure supplement 2. Clonogenic cell survival assay demonstrates enhanced cell death upon treatment with PDS in combination MK1775 or pimozide.

DOI: https://doi.org/10.7554/eLife.46793.020

(*Kotake et al., 2007*) raising the possibility of potentiation by G4-stabilising ligand combinatorial treatment.

We designated four of the genetic vulnerabilities as 'key' genes (*BRCA1*, *TOP1*, *DDX42*, and *GAR1)* whose deficiencies stood out with respect to consistent sensitivity to PDS and PhenDC3 in both cell lines tested. Given this, we postulate that deficiencies in any of these four genes will impart significant G4 ligand sensitivity for a range of cell types and/or with other G4 ligands. As GAR1-deficiencies are implicated in chronic lymphocytic leukaemia and contribute to telomere dysfunction (*Dos Santos et al., 2017*), we suggest that this cancer may be acutely sensitive to G4-stabilisation by small molecules.

In conclusion, we have revealed genes and pathways that interact with stabilised G4 structures. This information provides new insights into G4-related biology, especially into the functional pathways and roles as G4-interacting proteins. Furthermore, this work reveals novel disease-related genetic vulnerabilities for G4-ligands. Overall, these data provide a unique and comprehensive resource that can be further explored to understand biology that may involve G4s and also inspire new therapeutic possibilities.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Cell line (*H. Sapiens*)	A375	ATCC	Cat# CRL-1619, RRID:CVCL_0132	
Cell line (*H. Sapiens*)	HT1080	ATCC	Cat# CCL-121, RRID:CVCL_0317	
Cell line (*H. Sapiens*)	Plat-A	Cell Biolabs Inc	Cat# RV-102, RRID:CVCL_B489	
Antibody	Mouse monoclonal anti-Beta Actin	Merck	Cat# A5441, RRID:AB_476744	WB (1:250)
Antibody	Mouse polyclonal anti-DDX42	Abcam	Cat# ab80975, RRID:AB_2041042	WB (1:250)
Antibody	Rabbit monoclonal anti-Beta Actin	Cell Signalling Technology	Cat# 4970, RRID:AB_2223172	WB (1:500)
Antibody	Rabbit polyclonal anti-BRCA1	Cell Signalling Technology	Cat# 9010, RRID:AB_2228244	WB (1:50)
Antibody	Rabbit monoclonal anti-GAPDH	Cell Signalling Technology	Cat# 5174, RRID:AB_10622025	WB (1:50)
Antibody	Rabbit polyclonal anti-GAR1	NovusBio	Cat# NBP2-31742, RRID:AB_2801566	WB (1:100)
Antibody	Rabbit polyclonal anti-GST, HRP-conjugated	Abcam	Cat# ab3416, RRID:AB_30378	ELISA (1:10,000)

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	Rabbit monoclonal anti-LaminB1	Cell Signalling Technology	Cat# 12586, RRID:AB_2650517	WB (1:250)
Antibody	Rabbit monoclonal anti-TOP1	Abcam	Cat# ab109374, RRID:AB_10861978	WB (1:250)
Recombinant DNA reagent	pCMV-VSV-G plasmid	Addgene	Cat # 8454, RRID:Addgene_8454	plasmid
Recombinant DNA reagent	G-quadruplex focused shRNA plasmid library	transOMIC technologies, this paper		supplied as a glycerol stock, Materials and methods subsection: 'Composition and recombinant DNA reproduction of shRNA libraries'
Recombinant DNA reagent	transOMIC LMN genome-wide shRNA plasmid library	transOMIC technologies		supplied as multiple glycerol stocks
Sequence-based reagent	Biotinylated oligonucleotides	*Biffi et al. (2013)*, *Herdy et al. (2018)*, this paper		Materials and methods subsection 'Oligonucleotide annealing'
Sequence-based reagent	Genomic qPCR primers	this paper		Materials and methods subsection 'Barcode recovery, adapter ligation and sequencing'
Sequence-based reagent	Pasha/DGCR8 siRNA	Qiagen	Cat# 1027423	
Sequence-based reagent	siRNAs	this paper		Materials and methods subsection 'siRNA validation experiments – transfection, experimental outline, immunoblotting'
Peptide, recombinant protein	Recombinant human DDX42	NovusBio	Cat# H0001325-P01	
Commercial assay or kit	BluePippin 2% Internal Standard Marker Kit	Sage Science	Cat# BDF2010	
Commercial assay or kit	CellTitre-Glo One Solution Assay Reagent	Promega	Cat# G8461	
Commercial assay or kit	KAPA library quantification kit for Illumina platforms	Kapa Biosystems	Cat# 07960140001	
Commercial assay or kit	KOD Hot Start DNA polymerase	Merck	Cat# 710864	
Commercial assay or kit	Lipofectamine RNAiMAX	ThermoFisher Scientific	Cat# 13778150	
Commercial assay or kit	Muse Count and Viability kit	Merck	Cat# MCH600103	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay or kit	QIAmp DNA Blood Maxi Kit	Qiagen	Cat# 51194	
Commercial assay or kit	QIAquick PCR purification kit	Qiagen	Cat# 28104	
Commercial assay or kit	Qubit dsDNA HS assay kit	ThermoFisher Scientific	Cat# Q32851	
Commercial assay or kit	RIPA lysis buffer	ThermoFisher Scientific	Cat# 8990	
Commercial assay or kit	ZR GigaPrep Kit	Zymo Research	Cat# D4057	
Chemical compound, drug	Ampicillin	Merck	Cat# A5354	
Chemical compound, drug	Chloroquine diphosphate	Acros organics	Cat# 455240250	
Chemical compound, drug	cOmplete mini protease inhibitor	Roche	Cat# 11836153001	
Chemical compound, drug	DMSO	ThermoFisher Scientific	Cat# 20688	
Chemical compound, drug	Geneticin	Gibco	Cat# 10131035	
Chemical compound, drug	MK1775	Cambridge Bioscience	Cat# CAY21266	
Chemical compound, drug	PenStrep	ThermoFisher Scientific	Cat# 1507063	
Chemical compound, drug	PhenDC3	In-house synthesis	*De Cian et al., 2007a*	
Chemical compound, drug	Pimozide	Merck	Cat# P1793-500MG	
Chemical compound, drug	Pyridostatin (PDS)	In-house synthesis	*Rodriguez et al. (2008)*	
Chemical compound, drug	Sodium Butyrate	Merck	Cat# 303410	
Chemical compound, drug	TMB substrate	Merck	Cat# T4444	
Software, algorithm	Bowtie 2 v2.2.6	*Langmead and Salzberg, 2012*	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml	
Software, algorithm	ClueGO v3.5.1	*Bindea et al., 2009*; *Bindea et al., 2013*	http://www.ici.upmc.fr/cluego/cluegoDownload.shtml	
Software, algorithm	ColonyArea	*Guzmán et al., 2014*	Image J plugin	
Software, algorithm	Code used for shRNA screen data analysis	This paper	All scripts are available at: https://github.com/sblab-bioinformatics/GWscreen_G4sensitivity	
Software, algorithm	Combenefit	*Di Veroli et al., 2016*	https://sourceforge.net/projects/combenefit/	
Software, algorithm	Cytoscape v3.6.0	*Shannon et al., 2003*	http://www.cytoscape.org/	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Software, algorithm	edgeR v3.6	*Robinson et al., 2010*	http://bioconductor.org/packages/release/bioc/html/edgeR.html	
Software, algorithm	DAVID	*Huang et al., 2009a, Huang et al., 2009b*	https://david.ncifcrf.gov	
Software, algorithm	FastQC v0.11.3	*Andrews, 2010*	http://www.bioinformatics.babraham.ac.uk/projects/fastqc	
Software, algorithm	FASTX-Toolkit v0.0.14	*Gordon and Hannon, 2010*	http://hannonlab.cshl.edu/fastx_toolkit.html	
Software, algorithm	Graphpad Prism	GraphPad Prism (https://graphpad.com)	RRID:SCR_015807	Version 6
Software, algorithm	PolySearch2	*Liu et al., 2015a*	http://polysearch.cs.ualberta.ca/	
Software, algorithm	Python programming language v2.7.10	https://www.python.org		
Software, algorithm	R programming language v3.2.1	https://cran.r-project.org/		
Software, algorithm	Unix tools (cat, cut, awk, sort and uniq)	https://opengroup.org/unix		

Cell lines

HT1080 (RRID: CRL-1619) and A375 (RRID: CRL-121) were obtained from the American Type Culture Collection repository (ATCC) (LGC Standards, United Kingdom) and Plat-A (RRID: RV-102) was obtained from Cell Biolabs Incorporation. All cell lines were cultured in DMEM medium (Thermo-Fisher Scientific, cat #41966029) supplemented with 10% (v/v) heat inactivated FBS (ThermoFisher Scientific, cat #10500064) and grown at 37℃ in a 5% CO_2 humidified atmosphere. Cell lines were authenticated using small tandem repeat (STR) profiling and regularly checked to be mycoplasma-free by RNA-capture ELISA. All cell lines tested negative for *Mycoplasma* contamination. None of the cell lines used in our studies was mentioned in the list of commonly misidentified cell lines maintained by the International Cell Line Authentication Committee.

Quantification of live cell numbers

Live cell numbers (e.g. for plating cells for CellTitre-Glo assays, the screens and Incucyte experiments) were determined using the Muse Cell Analyzer (Merck), 'Count and Viability' assay according to manufacturer's instructions. Cells were diluted either 1:10 or 1:20 in 'Muse Count and Viability kit' solution (Merck, cat # MCH60013), to give a viable cell concentration of $1–2 \times 10^6$ cells/mL, with 'Events to Acquire' parameter set at 1000 events. Three cell counts were recorded.

Determination of G4 ligand concentration for shRNA screens

PDS and PhenDC3 (both synthesised in-house) (*De Cian et al., 2007b*; *Rodriguez et al., 2008*) were used as 100 mM stocks, dissolved in DMSO (Thermofisher Scientific, cat # 20688). GI_{20} values were calculated by treating A375 and HT1080 cells with serial dilutions of PDS and PhenDC3 for 96 hr and determining cell death via a CellTitre-Glo One Solution assay (Promega, cat # G8461) according to manufacturer's protocol. Each serial dilution was replicated four times for two-cell-seeding densities (1000/1500 cells per well). For both densities, curves were plotted averaging the four replicates in Prism (GraphPad v6) using a Non-Linear regression model, 'dose-response – inhibition' equation [log (inhibitor) vs. normalised response – variable slope] and GI_{20} values calculated. The GI_{20} concentrations used represent an average of three separate assays per cell line and yielded the following
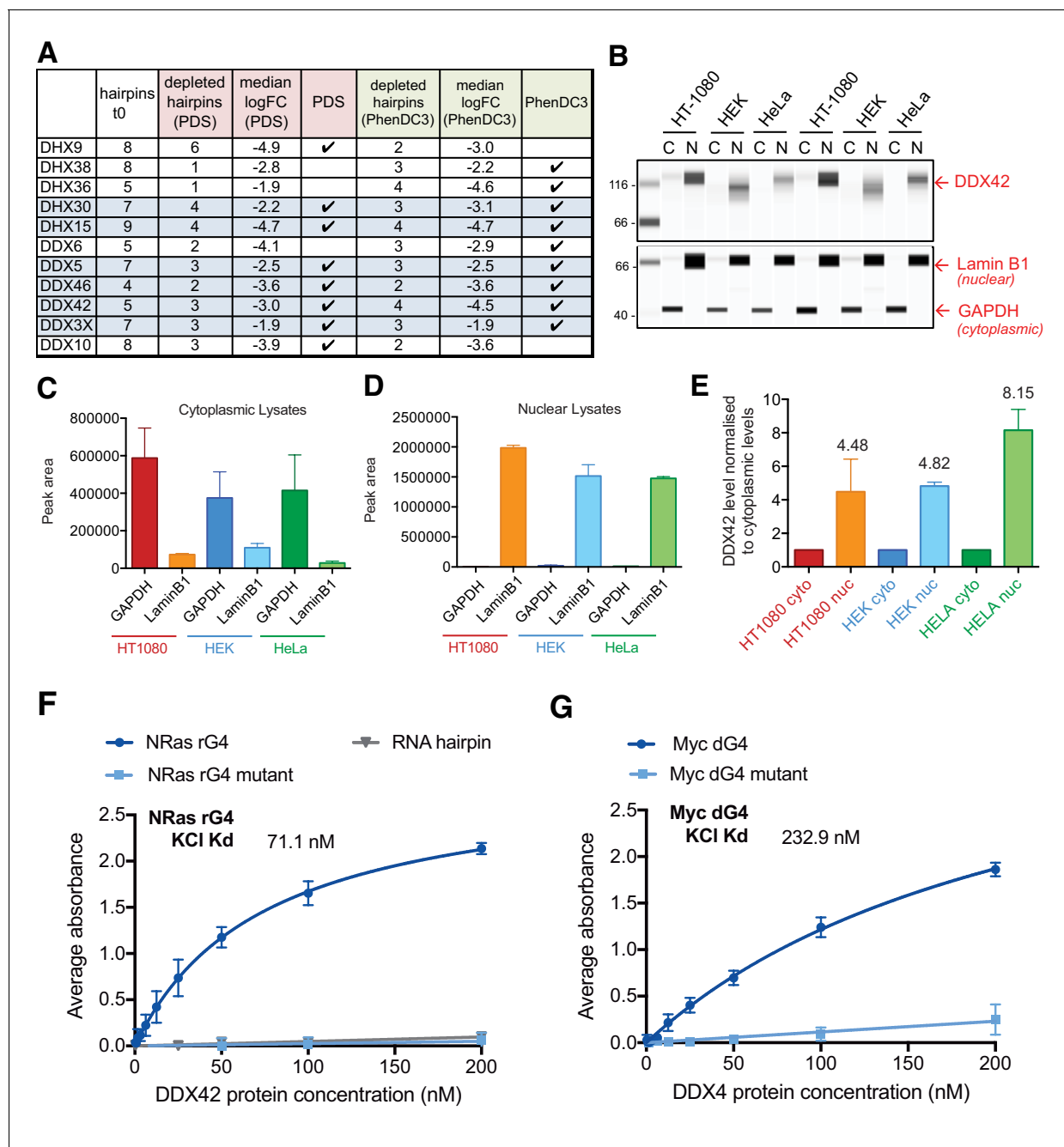
Figure 10. DDX42 is a predominantly nuclear G4-binding protein. (A) Number of depleted hairpins and median log$_2$FC values for DEAH/DEAD-box helicase genes within the 758 genes identified in the genome-wide screen. Those highlighted in blue caused sensitivity to both PDS and PhenDC3. (B) Representative immunoblots showing cytoplasmic (C) and nuclear (N) lysates for HT1080, human embryonic kidney (HEK) and HeLa cells probed for DDX42, laminB1 and GAPDH2. (C, D) GAPDH and laminB1 protein levels for (C) cytoplasmic and (D) nuclear lysates (mean for two biological replicates ± standard deviation). (E) DDX42 nuclear protein levels (normalised to cytoplasmic levels, mean for two biological replicates ± standard deviation). (F, G) DDX42 binding curves G4s by ELISA. (F) NRAS 5' UTR RNA G4 (rG4), mutated G4 sequence (rG4 mut) and RNA hairpin. (G) MYC DNA G4 (dG4) and mutated control (dG4 mut). Apparent K$_d$ is calculated from five replicates (values are indicative as the model assumes saturation kinetics).
DOI: https://doi.org/10.7554/eLife.46793.021

The following source data and figure supplement are available for figure 10:

Source data 1. Source files for western blots.
DOI: https://doi.org/10.7554/eLife.46793.023

Figure supplement 1. Circular dichroism spectroscopy for G4 oligonucleotides.
DOI: https://doi.org/10.7554/eLife.46793.022

Table 1. Possible chemotherapeutic combinations for G4-stabilising ligands with clinically relevant pharmacological drugs

Gene	Oncogene/tumour suppressor	Combinatorial/ single agent	Available drug treatments	Cancer association summary	Reference
BRCA1/ 2	Tumour suppressor	Single agent	Olaparib CX-5461	Deficient in ovarian, breast and colorectal cancer.	*Lee et al., 2014; Xu et al., 2017; McLuckie et al., 2013; Zimmer et al., 2016*
CCDC6	Tumour suppressor	Single agent	Olaparib	Inactivated in thyroid and lung cancers. CCDC6-deficient tumours are cisplatin-resistant but olaparib sensitive.	*Puxeddu et al., 2005; Morra et al., 2015*
CDK12	Oncogene	Combinatorial	Dinaclib (SCH77965)	High-grade serous ovarian cancer, often exhibits gain-of-function CDK12.	*Parry et al., 2010; Bajrami et al., 2014*
KEAP1	Oncogene/Tumour suppressor	Combinatorial/ single agent	CDDO-Me CPUY192018	KEAP1 inactivated in multiple cancers including thoracic and endometrial; also has oncogenic role, CDDO-Me used for leukaemia and sold tumours.	*Sanchez-Vega et al., 2018; Abed et al., 2015; Lu et al., 2016; Wang et al., 2014*
PSMC2	Oncogene	Combinatorial	Proteosome inhibitors: Bortezomib CEP187710 Carfizomib	Ubiquitin is emerging as chemotherapeutic target, and general proteasome inhibitors clinically are used against multiple myeloma.	*Chen et al., 2011a; Mattern et al., 2012; Edelmann et al., 2011*
SMAD4	Tumour suppressor	Single agent	GSKi: NCT01632306 NCT01214603 NCT01287520	Inactivated in 50% of pancreatic adenocarcinomas. Negatively regulated by GSK, GSKis in clinical trials for metastatic pancreatic cancer and acute leukaemia.	*Schutte et al., 1996; Hahn et al., 1996; Demagny and De Robertis, 2016; McCubrey et al., 2014*
SRSF10	Oncogene	Combinatorial	E7107 1C8	Over-expressed in colon cancer. 1C8 inhibits SRSF10 and impairs HIV replication. FUS interacting protein. E7107 is a splicing inhibitor preventing spliceosome assembly.	*Zhou et al., 2014; Shkreta et al., 2017; Cheung et al., 2016; Kotake et al., 2007*
UBA3	Oncogene	Combinatorial	MLN4924	Upregulated in AML and multiple solid cancers. MLN4924 is in Phase I clinical trials.	*Soucy et al., 2009*
USP1	Oncogene/Tumour suppressor	Combinatorial/ single agent	Pimozide	Over-expressed in melanoma, gastric, cervical and NSCLC; under-expressed in leukaemia and lymphoma. Pimozide is a potent USP1-targeting drug.	*García-Santisteban et al., 2013; Chen et al., 2011b*
WEE1	Oncogene/Tumour suppressor	Combinatorial/ single agent	AZDMK1775	Over-expressed in several cancers, some NSCLC are deficient.	*Matheson et al., 2016; Richer et al., 2017; Backert et al., 1999; Yoshida et al., 2004*

Table 1 continued

Gene	Oncogene/tumour suppressor	Combinatorial/ single agent	Available drug treatments	Cancer association summary	Reference
WHSC1	Oncogene	Combinatorial	DA3003-1 PF-03882845 Chaetocin TC-LPA5-4 ABT-199	Over-expressed in prostate cancer, multiple myeloma and mantle cell lymphoma. five potent candidate inhibitors.	*Coussens et al., 2017; Bennett et al., 2017*

concentrations used for the screens - A375: 10 μM PhenDC3 and 1.5 μM PDS; HT1080: 1 μM PhenDC3 and 0.5 μM PDS.

Composition and recombinant DNA production of shRNA libraries

The genome-wide screen uses the transOMIC LMN shRNA library against the human protein coding genome, consisting of 113,002 total shRNAs, split between 12 pools for ease-of-handling (approximately 10,000 shRNAs per pool) with an average number of five optimised hairpins per gene. The G4 focused screen consists of a custom shRNA pool (transOMIC technologies) with the same LMN vector (8018 shRNAs); this includes 1247 genes (7436 shRNAs) uncovered in the genome-wide screen (751 sensitisers and 496 upregulated genes), 116 additional genes identified from the literature as potentially G4-associated (439 shRNAs) and shRNAs targeting 37 olfactory receptors as non-targeting controls (143 shRNAs). 496 upregulated genes (FDR $\leq$ 0.05, 50% or three hairpins; $\log_2 FC \geq 1$) were included to mimic the genome-wide screen on a smaller scale by maintaining the population ratio of sensitisation and resistance. In this custom pool, unlike the commercially available genome-wide library, we capped the number of shRNAs at seven per gene. The backbone of both libraries contains Neo^R and ZsGreen markers to allow monitoring of infected cell lines by Geneticin (Gibco, cat # 10131035) selection and fluorescence (MacsQUANT), respectively. Both libraries were provided as glycerol stocks. Bacterial density was determined by calculating the colony-forming units (CFU) from dilutions of the original glycerol stock after plating on agar plates (overnight, 37°C, 100 μg/mL ampicillin). Glycerol stocks were thawed completely with sufficient volume taken (based on CFU) to ensure a minimum of 1000-fold hairpin representation and inoculated into liquid culture (LB media + 100 μg/mL ampicillin). Plasmid DNA was isolated using ZR Gigaprep kit D4057 (Zymo research) according to manufacturer's protocol and DNA quantified by Nanodrop One^c (Thermo Fisher Scientific).

shRNA stable cell line creation

For the genome-wide screen, each pool was treated independently, necessitating the creation of 12 different polymorphic cell lines each containing an average of 10,000 shRNAs, for both HT1080 and A375, per replica (three replicas, 36 polymorphic cell lines). Virus was produced using the Platinum-A packaging cell line (4–6 × 15 cm plates per pool) and calcium phosphate transfection. 24 hr after plating Platimum-A cells (70–80% confluency), media was replaced with DMEM medium supplemented with 1% (v/v) PenStrep (Thermo Fisher Scientific, cat # 150763) and 10% (v/v) heat inactivated FBS, shRNA library plasmid (75 μg) was then mixed with pCMV-VSV-G plasmid (7.5 μg, Addgene cat # 8454), Pasha/DGCR8 siRNA (2.7 μM, Qiagen cat # 1027423) to increase viral titre and 0.25 M CaCl$_2$ in a total volume of 1.5 mL per 15 cm dish and bubbled with 1.5 mL 2 x HBS (50 mM HEPES, 10 mM KCl, 12 mM Dextrose, 280 mM NaCl, 1.5 mM Na$_2$PO$_4$ at pH 7.00) and added to the Platinum-A cells (containing 17mL media) in a dropwise fashion. Immediately before adding the DNA-Pasha-transfection mixture to the Platinum-A cells, chloroquine diphosphate (lysosomal inhibitor, Acros Organics cat # 455200250) was added to the plates at a final concentration 2.5 μM. 14–16 hr after transfection, fresh media was added with 1:1000 1 M sodium butyrate (Merck, cat # 303410) for enhanced mammalian expression of the shRNA LMN plasmid. Virus was then harvested 48 hr after transfection and filter sterilised (0.45 μM) and stored at 4°C for a maximum of 7 days. Viral titre was determined by performing mock infections and quantifying fluorescent cells, via flow cytometry (MacsQUANT, Miltenyi Biotec Ltd.) 48 hr after infection. For both the genome-wide and focused-

Table 2. Examples of cancer-associated genetic vulnerabilities to G4 ligands.

Gene category	Gene name	Function/pathway Summary	Cancer association summary	References
DNA damage repair	PALB2	Homologous recombination; binds BRCA2	Inactivating mutations predispose patients to myeloid leukaemia, Wilm's tumour and Fanconi anaemia.	*Harrigan et al., 2018; Nepomuceno et al., 2017*
	BAP1	Homologous recombination; binds BRCA1, deubiquitinase for Histone 2A and tumour suppressor HCFC-1	Inactivating mutations found in uveal melanoma and mesotheliomas.	*Harrigan et al., 2018; Carbone et al., 2013*
	USP1	Fanconi anaemia and translesion synthesis DDR; deubiquitinase required for FANCD2, FANCI and PCNA localisation to sites of DNA damage	USP1 mRNA over-expressed in melanoma, gastric, cervical and NSCLC; under-expressed in leukaemia and lymphoma.	*Harrigan et al., 2018; Nijman et al., 2005; Huang et al., 2006; García-Santisteban et al., 2013*
	TOP1	Relieves torsional stress during DNA replication; suppresses genomic instability at actively transcribed exogenous G4-forming sequences	Common cancer target, to induces DNA damage following pharmacological inhibition, lethal to cells.	*Yadav et al., 2014; Wang, 2002*
Helicase activity	RECQL4 RTEL1	Previously identified G4-helicases	RECQL4 (Rothmund-Thomsun syndrome) and RTEL1 (Hoyeraal-Hreidarsson Syndrome), deficiencies impart increased risks of cancer cancer, autoimmunity and premature ageing.	*Brosh, 2013*

Table 2 continued on next page

Table 2 continued

Gene category	Gene name	Function/pathway Summary	Cancer association summary	References
Chromatin remodellers	ANKRD11	Transcription factor; Recruits histone deacetylases	Tumour suppressor epigenetically silenced in breast cancers.	*Lim et al., 2012* *Neilsen et al., 2008* *Noll et al., 2012*
	MLL4	H3K4 lysine methyl transferase	Frequently inactivated in several cancers.	*Froimchuk et al., 2017* *Kadoch et al., 2013*; *Rao and Dou, 2015*
	SMARCA4 SMARCB1 SMARCE1	SWI/SNF ATP-dependent chromatin remodellers	Mutated in 20% of human cancers; doxorubicin resistant triple-negative breast cancer is associated with loss of SMARCB1, SMARCA4, or KEAP1 (a BRCA1 interactor).	*Kadoch et al., 2013*; *Shain and Pollack, 2013*
Ubiquitin	USP37	Deubiquitinating enzyme which stabilises MYC	Upregulated in lung cancer.	*Pan et al., 2015*
	NEDD4L	E3 ubiquitin ligase	Expression correlates with poor patient outcome in hepatocellular and gastric carcinomas.	*Zhao et al., 2018*; *Gao et al., 2012*
	RNF20	E3 ubiquitin ligase; chromatin remodelling and DDR	Tumour supressor down-regulated in several cancers. Deletion is main contributor to chromosomal instability in colorectal cancer.	*Moyal et al., 2011*; *Shema et al., 2008*; *Barber et al., 2008*
Splicing	FUS	Splicing component and known G4-interactor	Over-expressed in colon, breast and liposarcoma cancers, respectively.	*Crozat et al., 1993*; *Dvinge et al., 2016*; *Takahama et al., 2013*

screen, 3.6×10^6 target cells were infected with a viral volume predicted to cause 30% infection (MOI 0.3) to minimise multiple shRNA integrations per cell. This provides approximately 10×10^6 shRNA expressing cells (1000-fold shRNA representation). Virus was diluted in serum free media plus polybrene (8 µg/mL) with infections carried out in triplicate and treated as independent replicates hereafter. 48 hr after infection cells antibiotic selection was performed with 800 µg/mL (HT1080) and 1000 µg/mL (A375) geneticin for 7–9 days (antibiotic concentrations were determined from 7-day toxicity curves prior to transfection setup).

Cell culture for pilot, genome-wide and focused shRNAs pools

Following complete antibiotic selection, a reference time point was harvested (t0) and cells were split into 3×15 cm plates per replica: PDS, PhenDC3 and DMSO vehicle control, each containing $8–10 \times 10^6$ cells to maintain 1000-fold hairpin representation. Every 72 hr, cells were trypsinised, counted via Muse Cell Analyzer to determine the number of population doublings, and 10×10^6 (A375 genome-wide and focused) or 8×10^6 (HT1080) cells per replica re-plated in fresh drug/ DMSO and media (17 mL media per plate). At all times, sufficient cell numbers were used so that a

minimum of 1000 or 800 cells per shRNA was maintained (A375 and HT1080, respectively), to ensure maximal potential for uncovering phenotypic effects from each shRNA hairpin tested (*Knott et al., 2014*). The volume of DMSO used in the 'vehicle' condition is equal to the volume for 10 µM PhenDC3. The remaining drug treatments were supplemented with DMSO to match this volume to keep the same DMSO concentration between treatment cell lines and screens. For the pilot screen, two final timepoints were harvested after 7 and 15 population doublings (t7 and t15), pellets extracted and analysed as described below. Based on the pilot screen, discussed below, a final time point (tF) was harvested after 15 population doublings for subsequent genome-wide and focused screen. For each pool of the genome-wide screen, 12 samples were generated (t0, DMSO tF, PDS tF, PhenDC3 tF; three replicas each). Therefore, 144 samples of 10×10^6 cells were generated to cover the entire screen. For each cell line of the focused screen, 12 samples were generated (t0, DMSO tF, PDS tF, PhenDC3 tF, three replicas each).

Pilot screen technique to determine genome-wide parameters

To determine the most appropriate tF for the genome-wide screen, cells (from shRNA pool 8) were harvested after 7 and 15 population doublings (t7 and t15 respectively) and the average $\log_2$FC (tF/t0) counts for each hairpin were determined as described below. For t7, 13 and 115 shRNAs were significantly altered following PDS and PhenDC3 treatment respectively (FDR $\leq$ 0.05). At t15, more hairpins were significantly depleted following PDS and PhenDC3 treatment (746 and 93 shRNAs respectively, excluding those significantly changed in DMSO).

Barcode recovery, adapter ligation and sequencing

All PCR and sequencing oligonucleotides (Merck) are summarised in the table below. Cell pellets (10×10^6 cells) were resuspended in PBS and genomic DNA extracted using QIAmp DNA Blood Maxi Kit (Qiagen, cat # 51194) according to the manufacturer's spin protocol, eluted in a final volume of 1200 µL and quantified by Qubit DNA HS Assay Kit (Thermo Fisher Scientific, cat # Q32851). The shRNA inserts were PCR-amplified from all DNA in each sample, in multiple 50 µL reactions each using 1.5 µg gDNA, with KOD Hot Start DNA Polymerase (Merck, cat # 710864) and the following reagents (included within the kit): 5 µL 10 x buffer, 5 µL 2 mM each dNTPs, 4 µL MgSO$_4$ (25 mM), 1.5 µL polymerase, 4 µL DMSO. Forward (Mir-F) and reverse (PGKpro-R) primers flanking the loop and antisense sequence of the hairpin region were used at a final concentration of 300 nM. PCR was performed under the following conditions: 98°C for 5 min, then for 25 cycles of 98°C for 35 s, 58°C 35 s and 72°C for 35 s, followed by a final extension at 72°C for 5 min. 1.2 mL of pooled PCR reaction were cleaned-up using QIAquick PCR purification kit (Qiagen, cat # 28104) according to manufacturer's protocol. 2 µg purified PCR product were PCR amplified in a second step, using forward (P5-Seq-P-Mir-Loop) and reverse (P7-Index-n-TruSeq-PGKpro-R) primers containing the P5 and P7 flowcell adapters, respectively. PCR was performed in 8×50 µL reactions each with 500 ng template DNA. The reverse primer contains TruSeq adapter small RNA Indexes for multiplexing and a 6-nucleotide barcode, denoted 'nnnnnn' below. PCR reagents were as for the first PCR, with the exception of the primers, which were used at a final concentration 1.5 µM. The second PCR was performed under the following conditions: 98°C for 5 min, then for 25 cycles of 94°C for 35 s, 52°C 35 s and 72°C for 35 s, followed by a final extension at 72°C for 5 min. 200 µL of pooled secondary PCR product was cleaned up as previously and the desired product (~340 bp) was extracted using BluePippin (Sage Science) 2% Internal Standard Marker Kit (DF marker 100–600 bp; Sage Science, BDF2010), according to manufacturer's protocol using a broad range elution (300–400 bp). Individual samples were quantified with a KAPA library quantification kit (KAPA Biosystems, cat # 0796-6014-0001) using a BioRad CFX96 Real Time PCR instrument with no Rox according to manufacturer's protocols. Libraries were diluted to 4 nM in RNAse free water. For the genome-wide screen samples, 24 libraries (12 pools) and for the focused screen samples, 24 libraries (both cell lines) were combined to create a pooled 4 nM stock, with each sample having a unique TruSeq adapter. The genome-wide screen samples were sequenced in six batches; all focused screen samples were sequenced simultaneously. DNA-Seq libraries were prepared from these samples using the NextSeq Illumina Platform v2 High Output Kit 75 cycles, followed by 36 base pair single-read sequencing performed on an Illumina NextSeq instrument, using a custom sequencing primer.

Oligo name	Description	Sequence 5'−3'
Mir5-F	Primary PCR Forward Primer	5'-CAGAATCGTTGCCTGCACATCTTGGAAAC- 3'
PGKpro-R	Primary PCR Reverse Primer	5' -CTGCTAAA GCGCATGCTCCAGACTGC- 3'
P5-Seq-P-Mir-Loop	Secondary PCR forward Primer	5'-AATGATACGGCGACCACCGAGATCTACACT AGCCTGCGCACGTAGTGAAGCCACAGATGTA-3'
P7-Index-n-Truseq-PGKpro-R	Secondary PCR barcoded reverse primer	5'-CAAGCAGAAGACGGCATACGAGAT nnnnnnGTGACTGGAGTTCAGACGTGTGCTCTT CCGATCTCTGCTAAAGCGCATGCTCCAGACTGC – 3'
SeqPrimer MirLoop	Custom sequencing primer	5'-TAGCCTGCGCACGTAGTGAAGCCACAGATGTA-3

Sequencing, read processing, alignment and counting of shRNAs

Sequencing data have been deposited in ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) under accession number *E-MTAB-6367*. Reads were trimmed to 22 nucleotides, base qualities were evaluated with FastQC v0.11.3 (*Andrews, 2010*) and bases were filtered from the 3' end with a Phred quality threshold of 33 using the FASTX-Toolkit v0.0.14 (*Gordon and Hannon, 2010*). Trimmed reads were aligned to the 113,002 reference shRNA sequences provided by transOMIC technologies (*Knott et al., 2014*) using Bowtie 2 v2.2.6 with default parameters (*Langmead and Salzberg, 2012*), which resulted in overall alignment rates of 90–95% with an average of 98% of reference sequences detected. The generated SAM files were processed to obtain shRNA counts using Unix tools (https://opengroup.org/unix) and Python scripts (v2.7.10, https://www.python.org), and library purity and potential contaminations were investigated with stacked bar plots and multidimensional scaling (MDS) using the R programming language v3.2.1 (https://cran.r-project.org). The code and scripts developed during the development of the project are available in our group's GitHub website (*Martínez Cuesta, 2019*; copy archived at https://github.com/elifesciences-publications/GWscreen_G4sensitivity).

Filtering, normalisation, differential representation analysis and defining sensitisation

To discard shRNAs bearing low counts, each library was filtered based on a counts-per-million threshold of 0.5 for all initial time points (t0), for example in a library of 10M reads, shRNAs with at least five counts for all initial time points will pass this filter. Normalisation factors were calculated to scale the raw library sizes using the weighted trimmed mean of M-values (TMM) approach (*Robinson et al., 2010*). To compare groups of replicates (time points and chemical treatments) for each pool, differential representation analysis of shRNA counts was performed using edgeR (*Robinson et al., 2010*). Common and shRNA-specific dispersions were estimated to allow the fitting of a negative binomial generalised linear model to the treatment counts. Contrasts between the initial time point and the treatments were defined (PDS-t0, PhenDC3-t0, and DMSO-t0) and likelihood ratio tests were carried out accordingly (*Dai et al., 2014*). Fold changes (FC) were then computed for every shRNA, and false discovery rates (FDR) were estimated using the Benjamini-Hochberg method. A gene was defined as significantly differentially represented for a given treatment if at least 50% or a minimum of 3 shRNAs were significant (FDR $\leq$ 0.05); sensitisation was additionally determined by applying a $\log_2 FC \leq -1$.

Exploring genes associated to G4s in databases and biomedical literature

Three different approaches were developed to uncover genes linked to G4s in the literature and molecular biology databases. 18 high confidence G4-related genes were obtained by scanning for genes in which the corresponding UniprotKB (*The UniProt Consortium, 2017*) entry is annotated with the term 'quadruplex' or genes annotated with at least one of the following 11 GO terms with any evidence assertion method (*Ashburner et al., 2000*):

GO id	Name	Type	Link
GO:0051880	G-quadruplex DNA binding	Molecular function	https://www.ebi.ac.uk/QuickGO/term/GO:0051880
GO:0002151	G-quadruplex RNA binding	Molecular function	https://www.ebi.ac.uk/QuickGO/term/GO:0002151
GO:0061849	telomeric G-quadruplex DNA binding	Molecular function	https://www.ebi.ac.uk/QuickGO/term/GO:0061849
GO:0071919	G-quadruplex DNA formation	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:0071919
GO:0044806	G-quadruplex DNA unwinding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:0044806
GO:1905493	regulation of G-quadruplex DNA binding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:1905493
GO:1905494	negative regulation of G-quadruplex DNA binding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:1905494
GO:1905495	positive regulation of G-quadruplex DNA binding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:1905495
GO:1905465	regulation of G-quadruplex DNA unwinding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:1905465
GO:1905466	negative regulation of G-quadruplex DNA unwinding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:1905466
GO:1905467	positive regulation of G-quadruplex DNA unwinding	Biological process	https://www.ebi.ac.uk/QuickGO/term/GO:1905467

Furthermore, 55 confirmed human G4-interacting proteins as defined by the G4IPB database (*Mishra et al., 2016*) (http://bsbe.iiti.ac.in/bsbe/ipdb/index.php) were also used to determine predefined G4-interacting proteins from the genome-wide shRNA screen. For this, gene entries were removed where the only G4-relationship was a predicted G4-forming sequence in the mRNA or DNA (i.e. not a direct protein interaction) or where the protein was not human. To extend the list of G4-interacting proteins, text-mining using PolySearch2 (*Liu et al., 2015b*) was used. Human protein-coding gene names and G4-terms and synonyms are defined using the corresponding MeSH term id D054856 (https://www.ncbi.nlm.nih.gov/mesh/68054856) and the thesaurus of gene names obtained from the PolySearch2 website (http://polysearch.cs.ualberta.ca/). A relevancy score measures the strength of association between two text groups, and higher the score, the more likely terms from the two groups co-occur within the same abstract; the score also accounts for the distance between terms from the two groups. A total of 5477 pieces of text were identified in PubMed and PubMed Central where any of the G4 terms co-occur with more than 500 human gene names. Overall, this generated 526 G4-associated genes, with 54 (10%) uncovered as G4-sensitisers (https://github.com/sblab-bioinformatics/GWscreen_G4sensitivity), which were manually edited to 16 genes as discussed in the main text and figures.

KEGG pathway, gene ontology and protein domain enrichment analysis
The ClueGO v2.3.3 (*Bindea et al., 2009*; *Bindea et al., 2013*) plugin for Cytoscape (*Shannon et al., 2003*) (v3.5.1) was used to determine networks of enriched KEGG pathways and Gene Ontology terms (Biological Process and Molecular Function) for significantly depleted genes upon G4 ligand

treatment. Specifically, a right-sided (Enrichment) test based on the hyper-geometric distribution was performed on the corresponding Entrez gene IDs for each gene list and the Bonferroni adjustment ($p<0.05$) was applied to correct for multiple hypothesis testing. Only experimental evidence codes (EXP, IDA, IPI, IMP, IGI, IEP) were used. The Kappa-statistics score threshold was set to 0.4 and GO term fusion was used to diminish redundancy of terms shared by similar proteins. Other parameters include: GO level intervals (3–8 genes) and Group Merge (50%). Protein domains were investigated using DAVID (v6.8) to integrate GENE3D crystallographic data and PFAM sequence information and enrichment was considered significant if the EASE score $p<0.05$ (*Finn et al., 2016*; *Yeats et al., 2006*).

COSMIC analysis

Cancer mutation data (CosmicMutantExport.tsv) from the COSMIC database v82 (*Forbes et al., 2015*) was used to investigate the association between G4 sensitisers and cancer genes. ~150,000 were mutations available in COSMIC for 702 (93%) sensitiser genes, with some predicted to be pathogenic by the FATHMM algorithm embedded within the COSMIC database. The Cancer Gene Census (http://cancer.sanger.ac.uk/census) was used to investigate whether G4 sensitisers are enriched in genes containing mutations causally implicated in cancer. Fisher's exact tests as implemented in R were used to calculate fold enrichment significance of sensitisers that are cancer genes in COSMIC (compared to the percentage of protein-coding genes in COSMIC – 3.3%).

siRNA validation experiments – transfection, experimental outline, immunoblotting

ON-TARGETplus siRNAs (Dharmacon/GE healthcare) were used as summarised in the table below. Cells were transfected with either targeting or non-targeting control siRNAs using lipofectamine RNAiMAX (Thermo Fisher Scientific, cat # 13778150) and OptiMEM reduced serum medium (Thermo Fisher Scientific) according to the manufacturer's protocol (reagent protocol 2013) alongside a non-transfected control. 24 hr after transfection, cells were trypsinised, counted and re-plated in media supplemented with PDS, PhenDC3 or DMSO vehicle control (minimum two biological replicates per condition) in a 48-well plate format (seeding density - 8,000 cells per well A375; 4,000 cells per well HT1080). Cell growth was monitored for 144 hr using IncuCyteZOOM live cell analysis (Sartorius) and cell confluency calculated as a percentage of the well area covered. Scans were performed every 3 hr; nine scans per well. To monitor protein levels, cells transfected simultaneously with the same siRNA-reagent mixture were harvested 48 hr and 144 hr after transfection, by cell scraping and lysed on ice (30 min) with RIPA lysis buffer with protease inhibitor +EDTA (Thermo Fisher Scientific, cat # 8990). Lysates were analysed by capillary electrophoresis via the Protein Simple Wes platform according to manufacturer's protocol with antibodies summarised in the key resource table above. Lysates from non-transfected and siRNA-treated (targeting and non-targeting) samples were probed with antibodies against BRCA1 (Cell Signalling Technology, cat # 4970-CST), TOP1 (Abcam, cat # AB109374), GAR1 (NovusBio cat #NBP2-31742) or DDX42 (Abcam cat #AB80975), plus anti-beta actin antibody (mouse Merck cat # A5441; rabbit cat # 4970-CST) by multiplexing. For non-targeting and targeting lysates, the area of the desired band was normalised to beta-actin and then normalised to the protein level in the non-targeting sample, for three (48 hr after transfection lysates) or two independent Wes runs (144 hr after transfection). Protein depletion is expressed as an average of these normalised values. All lysates were used at a concentration of 0.8 mg/mL and antibody dilutions as follows: BRCA1 1:50; TOP1 1:250; GAR1 1:100; DDX42 1:250; rabbit-actin 1:500; mouse-actin 1:250.

siRNA	Catalogue number	Sequence 5'–3'
Non-targeting 2	D-001810-02-05	UGGUUUACAUGUUGUGUGA
BRCA1 (A375)	J-003461–09	CAACAUGCCCACAGAUCAA
BRCA1 (HT1080)	J-003461–12	GAAGGAGCUUUCAUCAUUUC
TOP1 (both cell lines)	J-005278–08	CGAAGAAGGUAGUAGAGUC
DDX42 (both cell lines)	J-012393–11	GGAGAUCGACUAACGGCAA
GAR1 (both cell lines)	J-013386–06	UCCAGAACGUGUAGUCUUA

G4 ligand and drug treatments

10 mM stocks in DMSO of PDS (in house synthesis), MK1775 (Cambridge Bioscience, cat# CAY21266) and pimozide (Merck, cat# P1793-500MG) were used as for synergy experiment. Cells were seeded in Corning, Tissue Culture-treated 96-well clear bottom plates (Thermofisher, cat#07-200-587) for HT1080 (1000 cells per well) and A375 (1500 cells per well) cell lines. 24 hr after plating, media was removed and cells were treated with different concentrations of PDS and MK1775 or pimozide in media in a final volume of 150 μL, alongside non-treated and solvent-treated controls. After 96 hr, cell death was determined via a CellTitre-Glo One Solution assay (Promega, cat # G8461) according to manufacturer's protocol using the PHERAstar FS (BMG labtech) to detect luminescence, using the recommended settings. Values were normalised to and expressed as a percentage of the untreated controls. This was performed for three biological replicas. Data were analysed via Combenefit software using the BLISS independence model since the molecule have independent targets (*Di Veroli et al., 2016*) to determine synergy. For the clonogenic cell survival assay, A375 (300 cells per well) and HT1080 (400 cells per well) were plated as single cells in 12-well plates. The next day, cells were treated with DMSO or the indicated doses of PDS, pimozide and/or MK1775 in media. After 8 days, colonies were fixed with 3% trichloroacetic acid (TCA) for 90 min at 4°C, washed with MiliQ, air dried and then stained with 0.057% (v/v) Sulforhodamine B solution (Merck, cat # 230162–5G) for 30 min at room temperature. Plates were then washed four times with 1% acetic acid, air dried and colonies visualised using GelCount (Oxford Optronix). Colony growth was determined using the 'colony intensity percentage' parameter in the ColonyArea Image J plugin (*Guzmán et al., 2014*), which considers both the intensity and percentage of area covered by the colonies. Values were normalised to and expressed as a percentage of the untreated controls and then further processed by Combenefit software, as described above, to determine synergy. A total of three independent biological replicates were performed.

Sub-cellular localisation of DDX42

HT1080, HEK and HeLa cells were harvested from a 70% confluent 15 cm plate by cell scraping in PBS on ice and pelleted by centrifugation (500 g, 5 min, 4°C). Pellets were resuspended in three volumes of low-salt buffer (20 mM HEPES pH7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT) plus protease inhibitor (cOmplete mini, Roche cat#11836153001), lysed on ice (15 min) and 0.5% Igepal added. Samples were vortexed (1 min) and centrifuged (900 g, 15 min, 4°C) and the supernatant collected for cytoplasmic extracts. Nuclear pellets were washed in low-salt buffer, supernatant discarded and then resuspended in high-salt buffer (20 mM HEPES pH7.4, 500 mM NaCl, 3 mM $MgCl_2$, 0.5% Igepal, 0.2 mM EDTA, 1 mM DTT) plus protease inhibitors, followed by lysis on ice with intermittent vortexing (30 min). Lysates were passed through a syringe needle to promote lysis and shear genomic DNA and followed by centrifugation (13,000 g, 10 min, 4°C). Lysis confirmed by trypan blue staining according to manufacturer's protocols (Thermofisher Scientific cat#15250061). The supernatant was then collected as nuclear extract. Cytoplasmic and nuclear lysates were quantified on a Direct Detect platform (Merck) and DDX42 expression analysed by immunoblotting using the Protein Simple Wes instrument as described above with a lysate concentration of 0.5 mg/mL. Samples were also immunoblotted with antibodies against nuclear laminB1 (CST 12586; 1:250) and cytoplasmic GAPDH to confirm subcellular fractionation efficiency (CST 5174, 1:50).

Oligonucleotide annealing

Biotinylated oligonucleotides for G4 and non-G4 forming sequences (IDT technologies; see Table below) were annealed in 10 mM TrisHCl pH 7.4, 100 mM KCl by heating at 95°C, 10 min followed by slow cooling to room temperature overnight at a controlled rate of 0.2°C/min. Annealed oligonucleotides were stored at 4°C for maximum 1 month.

Oligo	Rna/DNA	Sequence
NRAS G4	RNA	5' [Btn] UGU GGG AGG GGC GGG UCU GGG UGC 3'

Continued on next page

Continued

Oligo	Rna/DNA	Sequence
NRAS mut	RNA	5' [Btn] UGU AGA AAG AGC AGA UCU AGA UG 3'
Stem loop	RNA	5' [Btn] ACA GGG CUC CGC GAU GGC GGA GCC CAA 3'
Myc G4	DNA	5' [Btn] TGA GGG T GGG TA GGG T GGG TAA 3'
Myc mut	DNA	5' [Btn] TGA GAG T GAG TA GAG T GAG TAA 3'

Enzyme-Linked immunosorbent assay

Recombinant human DDX42 with an N-terminal GST tag was purchased from NovusBio (cat# H0001325-P01). Streptavidin-Coated High-Binding Capacity 96-well plates (ThermoScientific prod #15501) were hydrated with PBS (30 min) and coated with 50 nM biotinylated oligonucleotides (1 hr, shaking 450 rpm). Wells were washed three times with ELISA buffer (50 mM K_2HPO_4 pH 7.4 and 100 mM KCl/100 mM LiCl); 1 min shaking, 450 rpm. Wells were then blocked with 3% (w/v) BSA (Merck, cat# A7030) in ELISA buffer for 1 hr, at room temperature and then incubated with serial dilutions of DDX42 up to 200 nM for 1 hr. Wells were washed three times with 0.1% TWEEN-20 in ELISA buffer and then incubated for 1 hr with anti-GST HRP-conjugated antibody (Abcam AB3416) diluted 1:10,000 in blocking buffer. Wells were again washed three times with ELISA-Tween, and the bound anti-GST HRP detected with TMB substrate (Merck,cat#T4444) for 2 min. Reactions were stopped with 2 M HCl. Absorbance at 450 nm was measured with PheraSTAR plate reader (BMG labtech). Binding curves with standard error of the mean (SEM) were fitted using GraphPad Prism software, using a non-linear regression fit, one site, specific binding model with saturation kinetics. The following equation was used: $y=(Bmax*x)/(K_d +x)$, where x = concentration of DDX42 (nM) and Bmax is the maximum specific binding (i.e. saturation).

Circular dichroism spectroscopy

200 µL of 10 µM oligonucleotide were prepared in assay buffer and annealed as described above. CD spectra were recorded on an Applied Photo-physics Chirascan CD spectropolarimeter using a 1 mm path length quartz cuvette. CD measurements were performed at 298 K over a range of 200–320 nm using a response time of 0.5 s, 1 nm pitch and 0.5 nm bandwidth. The recorded spectra represent a smoothed average of three scans, zero-corrected at 320 nm (Molar ellipticity θ is quoted in 10^5 deg cm^2 $dmol^{-1}$). The absorbance of the buffer was subtracted from the recorded spectra.

Acknowledgements

Additional information

Competing interests

Gregory J Hannon: Is associated with transOMIC Technologies, who have commercialised libraries constructed using the shERWOOD and ultramiR design strategies. Shankar Balasubramanian: Is founder, adviser and shareholder of Cambridge Epigenetix Ltd. The other authors declare that no competing interests exist.

Author contributions

Katherine G Zyner, Conceptualization, Data curation, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing, Designed and performed the genome-wide shRNA screen, library preparation and sequencing, Designed the focused RNAi screens, Performed the GO and Cytoscape analyses, Designed and performed the clonogenic assays and DDX42 binding experiments; Darcie S Mulhearn, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing, Designed and performed the genome-wide shRNA screen, library preparation and sequencing, Designed and performed the focused RNAi screens, Performed the GO and Cytoscape analyses, Designed and performed the drug synergy and DDX42 binding experiments; Santosh Adhikari, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing, Designed and performed the drug synergy experiments; Sergio Martínez Cuesta, Conceptualization, Data curation, Software, Formal analysis, Validation, Methodology, Writing—original draft, Writing—review and editing, Performed the computational analyses on all RNAi screens; Marco Di Antonio, Conceptualization, Resources, Data curation, Formal analysis, Methodology, Writing—original draft, Writing—review and editing, Synthesised all the G4 ligands; Nicolas Erard, Resources, Software, Formal analysis, Methodology, Writing—review and editing; Gregory J Hannon, Conceptualization, Resources, Supervision, Methodology, Writing—original draft, Writing—review and editing; David Tannahill, Conceptualization, Resources, Software, Formal analysis, Supervision, Methodology, Writing—original draft, Project administration, Writing—review and editing; Shankar Balasubramanian, Conceptualization, Formal analysis, Supervision, Funding acquisition, Methodology, Writing—original draft, Project administration, Writing—review and editing

Author ORCIDs

Katherine G Zyner (iD) https://orcid.org/0000-0002-4997-0150
Santosh Adhikari (iD) https://orcid.org/0000-0002-1501-2106
Sergio Martínez Cuesta (iD) https://orcid.org/0000-0001-9806-2805
Marco Di Antonio (iD) https://orcid.org/0000-0002-7321-1867
Gregory J Hannon (iD) https://orcid.org/0000-0003-4021-3898
David Tannahill (iD) https://orcid.org/0000-0002-3811-6864
Shankar Balasubramanian (iD) https://orcid.org/0000-0002-0281-5815

Decision letter and Author response
Decision letter https://doi.org/10.7554/eLife.46793.032
Author response https://doi.org/10.7554/eLife.46793.033

Additional files

Supplementary files

• Supplementary file 1. Supporting data for *Figure 3–7*. List of shRNAs/genes from venn diagrams and table statistics for KEGG, GO, DGIDb and Protein Domains analyses from *Figures 3–7*. Each data tab is labelled with its corresponding originating figure.
DOI: https://doi.org/10.7554/eLife.46793.026

• Supplementary file 2. Supporting data for *Figure 6—figure supplement 1* and *Figure 7—figure supplement 1*. List of shRNAs/genes from venn diagrams and table statistics for GO analyses from *Figure 6—figure supplement 1* and *Figure 7—figure supplement 1*. Each data tab is labelled with its corresponding originating figure.
DOI: https://doi.org/10.7554/eLife.46793.027

• Transparent reporting form
DOI: https://doi.org/10.7554/eLife.46793.028

Data availability

Sequencing data have been deposited in ArrayExpress under the accession number E-MTAB-6367.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Mulhearn DS, Zyner KG, Martinez Cuesta S, Balasubramanian S	2019	Systematic identification of G-quadruplex sensitive lethality by genome-wide genetic screening	https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-6367/	ArrayExpress, E-MTAB-6367

References

Abed DA, Goldstein M, Albanyan H, Jin H, Hu L. 2015. Discovery of direct inhibitors of Keap1-Nrf2 protein-protein interaction as potential therapeutic and preventive agents. *Acta Pharmaceutica Sinica B* **5**:285–299. DOI: https://doi.org/10.1016/j.apsb.2015.05.008, PMID: 26579458

Andrews S. 2010. FastQC. v0.11.3. Bioinformatics. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics* **25**:25–34. DOI: https://doi.org/10.1038/75556, PMID: 10802651

Azorsa DO, Gonzales IM, Basu GD, Choudhary A, Arora S, Bisanz KM, Kiefer JA, Henderson MC, Trent JM, Von Hoff DD, Mousses S. 2009. Synthetic lethal RNAi screening identifies sensitizing targets for gemcitabine therapy in pancreatic cancer. *Journal of Translational Medicine* **7**:43. DOI: https://doi.org/10.1186/1479-5876-7-43, PMID: 19519883

Backert S, Gelos M, Kobalz U, Hanski ML, Böhm C, Mann B, Lövin N, Gratchev A, Mansmann U, Moyer MP, Riecken EO, Hanski C. 1999. Differential gene expression in colon carcinoma cells and tissues detected with a cDNA array. *International Journal of Cancer* **82**:868–874. DOI: https://doi.org/10.1002/(SICI)1097-0215(19990909)82:6<868::AID-IJC16>3.0.CO;2-W, PMID: 10446455

Bajrami I, Frankum JR, Konde A, Miller RE, Rehman FL, Brough R, Campbell J, Sims D, Rafiq R, Hooper S, Chen L, Kozarewa I, Assiotis I, Fenwick K, Natrajan R, Lord CJ, Ashworth A. 2014. Genome-wide profiling of genetic synthetic lethality identifies CDK12 as a novel determinant of PARP1/2 inhibitor sensitivity. *Cancer Research* **74**:287–297. DOI: https://doi.org/10.1158/0008-5472.CAN-13-2541, PMID: 24240700

Balasubramanian S, Hurley LH, Neidle S. 2011. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nature Reviews Drug Discovery* **10**:261–275. DOI: https://doi.org/10.1038/nrd3428, PMID: 21455236

Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, Velculescu VE, Kinzler KW, Vogelstein B, Lengauer C, Hieter P. 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *PNAS* **105**:3443–3448. DOI: https://doi.org/10.1073/pnas.0712384105, PMID: 18299561

Bennett RL, Swaroop A, Troche C, Licht JD. 2017. The role of nuclear Receptor-Binding SET domain family histone lysine methyltransferases in cancer. *Cold Spring Harbor Perspectives in Medicine* **7**:a026708. DOI: https://doi.org/10.1101/cshperspect.a026708, PMID: 28193767

Besnard E, Babled A, Lapasset L, Milhavet O, Parrinello H, Dantec C, Marin JM, Lemaitre JM. 2012. Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nature Structural & Molecular Biology* **19**:837–844. DOI: https://doi.org/10.1038/nsmb.2339, PMID: 22751019

Biffi G, Tannahill D, McCafferty J, Balasubramanian S. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chemistry* **5**:182–186. DOI: https://doi.org/10.1038/nchem.1548, PMID: 23422559

Biffi G, Tannahill D, Miller J, Howat WJ, Balasubramanian S. 2014. Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues. *PLOS ONE* **9**:e102711. DOI: https://doi.org/10.1371/journal.pone.0102711, PMID: 25033211

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J. 2009. ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**:1091–1093. DOI: https://doi.org/10.1093/bioinformatics/btp101, PMID: 19237447

Bindea G, Galon J, Mlecnik B. 2013. CluePedia cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* **29**:661–663. DOI: https://doi.org/10.1093/bioinformatics/btt019, PMID: 23325622

Bochman ML, Paeschke K, Zakian VA. 2012. DNA secondary structures: stability and function of G-quadruplex structures. *Nature Reviews Genetics* **13**:770–780. DOI: https://doi.org/10.1038/nrg3296

Brázda V, Hároníková L, Liao JC, Fojta M. 2014. DNA and RNA quadruplex-binding proteins. *International Journal of Molecular Sciences* **15**:17493–17517. DOI: https://doi.org/10.3390/ijms151017493, PMID: 25268620

Brosh RM. 2013. DNA helicases involved in DNA repair and their roles in cancer. *Nature Reviews Cancer* **13**:542–558. DOI: https://doi.org/10.1038/nrc3560, PMID: 23842644

Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, Kyle S, Meuth M, Curtin NJ, Helleday T. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**:913–917. DOI: https://doi.org/10.1038/nature03443, PMID: 15829966

Bugaut A, Balasubramanian S. 2012. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Research* **40**:4727–4741. DOI: https://doi.org/10.1093/nar/gks068, PMID: 22351747

Cammas A, Millevoi S. 2016. RNA G-quadruplexes: emerging mechanisms in disease. *Nucleic Acids Research* **9**:gkw1280. DOI: https://doi.org/10.1093/nar/gkw1280

Carbone M, Yang H, Pass HI, Krausz T, Testa JR, Gaudino G. 2013. BAP1 and cancer. *Nature Reviews Cancer* **13**:153–159. DOI: https://doi.org/10.1038/nrc3459, PMID: 23550303

Chakraborty P, Grosse F. 2011. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair* **10**:654–665. DOI: https://doi.org/10.1016/j.dnarep.2011.04.013, PMID: 21561811

Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S. 2015. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nature Biotechnology* **33**:877–881. DOI: https://doi.org/10.1038/nbt.3295, PMID: 26192317

Chan DA, Giaccia AJ. 2011. Harnessing synthetic lethal interactions in anticancer drug discovery. *Nature Reviews Drug Discovery* **10**:351–364. DOI: https://doi.org/10.1038/nrd3374, PMID: 21532565

Chen D, Frezza M, Schmitt S, Kanwar J, Dou QP. 2011a. Bortezomib as the first proteasome inhibitor anticancer drug: current status and future perspectives. *Current Cancer Drug Targets* **11**:239–253. DOI: https://doi.org/10.2174/156800911794519752, PMID: 21247388

Chen J, Dexheimer TS, Ai Y, Liang Q, Villamil MA, Inglese J, Maloney DJ, Jadhav A, Simeonov A, Zhuang Z. 2011b. Selective and cell-active inhibitors of the USP1/ UAF1 deubiquitinase complex reverse cisplatin resistance in non-small cell lung cancer cells. *Chemistry & Biology* **18**:1390–1400. DOI: https://doi.org/10.1016/j.chembiol.2011.08.014, PMID: 22118673

Chen MC, Tippana R, Demeshkina NA, Murat P, Balasubramanian S, Myong S, Ferré-D'Amaré AR. 2018. Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* **558**:465–469. DOI: https://doi.org/10.1038/s41586-018-0209-9, PMID: 29899445

Cheung PK, Horhant D, Bandy LE, Zamiri M, Rabea SM, Karagiosov SK, Matloobi M, McArthur S, Harrigan PR, Chabot B, Grierson DS. 2016. A parallel synthesis approach to the identification of novel Diheteroarylamide-Based compounds blocking HIV replication: potential inhibitors of HIV-1 Pre-mRNA alternative splicing. *Journal of Medicinal Chemistry* **59**:1869–1879. DOI: https://doi.org/10.1021/acs.jmedchem.5b01357, PMID: 26878150

Coussens NP, Kales SC, Henderson MJ, Lee OW, Horiuchi KY, Wang Y, Hall MD. 2017. Small molecule inhibitors of the human histone lysine methyltransferase NSD2 / WHSC1 / MMSET identified from a quantitative High-Throughput screen with nucleosome substrate. *bioRxiv*. DOI: https://doi.org/10.1101/208439

Creacy SD, Routh ED, Iwamoto F, Nagamine Y, Akman SA, Vaughn JP. 2008. G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *Journal of Biological Chemistry* **283**:34626–34634. DOI: https://doi.org/10.1074/jbc.M806277200, PMID: 18842585

Crozat A, Aman P, Mandahl N, Ron D. 1993. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* **363**:640–644. DOI: https://doi.org/10.1038/363640a0, PMID: 8510758

Dai Z, Sheridan JM, Gearing LJ, Moore DL, Su S, Wormald S, Wilcox S, O'Connor L, Dickins RA, Blewitt ME, Ritchie ME. 2014. edgeR: a versatile tool for the analysis of shRNA-seq and CRISPR-Cas9 genetic screens. *F1000Research* **3**:95. DOI: https://doi.org/10.12688/f1000research.3928.2, PMID: 24860646

Damerla RR, Knickelbein KE, Strutt S, Liu FJ, Wang H, Opresko PL. 2012. Werner syndrome protein suppresses the formation of large deletions during the replication of human telomeric sequences. *Cell Cycle* **11**:3036–3044. DOI: https://doi.org/10.4161/cc.21399, PMID: 22871734

De Cian A, Cristofari G, Reichenbach P, De Lemos E, Monchaud D, Teulade-Fichou MP, Shin-Ya K, Lacroix L, Lingner J, Mergny JL. 2007a. Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. *PNAS* **104**:17347–17352. DOI: https://doi.org/10.1073/pnas.0707365104, PMID: 17954 919

De Cian A, Delemos E, Mergny JL, Teulade-Fichou MP, Monchaud D. 2007b. Highly efficient G-quadruplex recognition by bisquinolinium compounds. *Journal of the American Chemical Society* **129**:1856–1857. DOI: https://doi.org/10.1021/ja067352b, PMID: 17260991

Demagny H, De Robertis EM. 2016. Point mutations in the tumor suppressor Smad4/DPC4 enhance its phosphorylation by GSK3 and reversibly inactivate TGF-β signaling. *Molecular & Cellular Oncology* **3**:e1025181. DOI: https://doi.org/10.1080/23723556.2015.1025181, PMID: 27308538

Di Veroli GY, Fornari C, Wang D, Mollard S, Bramhall JL, Richards FM, Jodrell DI. 2016. Combenefit: an interactive platform for the analysis and visualization of drug combinations. *Bioinformatics* **32**:2866–2868. DOI: https://doi.org/10.1093/bioinformatics/btw230, PMID: 27153664

Dos Santos PC, Panero J, Stanganelli C, Palau Nagore V, Stella F, Bezares R, Slavutsky I. 2017. Dysregulation of H/ACA ribonucleoprotein components in chronic lymphocytic leukemia. *PLOS ONE* **12**:e0179883. DOI: https://doi.org/10.1371/journal.pone.0179883, PMID: 28666010

Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. 2016. RNA splicing factors as oncoproteins and tumour suppressors. *Nature Reviews Cancer* **16**:413–430. DOI: https://doi.org/10.1038/nrc.2016.51, PMID: 27282250

Edelmann MJ, Nicholson B, Kessler BM. 2011. Pharmacological targets in the ubiquitin system offer new ways of treating cancer, neurodegenerative disorders and infectious diseases. *Expert Reviews in Molecular Medicine* **13**:e35. DOI: https://doi.org/10.1017/S1462399411002031, PMID: 22088887

Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, Santarosa M, Dillon KJ, Hickson I, Knights C, Martin NM, Jackson SP, Smith GC, Ashworth A. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**:917–921. DOI: https://doi.org/10.1038/nature03445, PMID: 1582 9967

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* **44**:D279–D285. DOI: https://doi.org/10.1093/nar/gkv1344, PMID: 26673716

Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ. 2015. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Research* **43**:D805–D811. DOI: https://doi.org/10.1093/nar/gku1075, PMID: 25355519

Froimchuk E, Jang Y, Ge K. 2017. Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **627**:337–342. DOI: https://doi.org/10.1016/j.gene.2017.06.056, PMID: 28669924

Gao C, Pang L, Ren C, Ma T. 2012. Decreased expression of Nedd4L correlates with poor prognosis in gastric cancer patient. *Medical Oncology* **29**:1733–1738. DOI: https://doi.org/10.1007/s12032-011-0061-3, PMID: 21 909941

García-Santisteban I, Peters GJ, Giovannetti E, Rodríguez JA. 2013. USP1 deubiquitinase: cellular functions, regulatory mechanisms and emerging potential as target in cancer therapy. *Molecular Cancer* **12**:91. DOI: https://doi.org/10.1186/1476-4598-12-91, PMID: 23937906

Giri B, Smaldino PJ, Thys RG, Creacy SD, Routh ED, Hantgan RR, Lattmann S, Nagamine Y, Akman SA, Vaughn JP. 2011. G4 resolvase 1 tightly binds and unwinds unimolecular G4-DNA. *Nucleic Acids Research* **39**:7161–7178. DOI: https://doi.org/10.1093/nar/gkr234, PMID: 21586581

González V, Hurley LH. 2010. The c-MYC NHE III(1): function and regulation. *Annual Review of Pharmacology and Toxicology* **50**:111–129. DOI: https://doi.org/10.1146/annurev.pharmtox.48.113006.094649, PMID: 1 9922264

Gordon A, Hannon G. 2010. *FASTQ/A Short-Reads Pre-Processing Tools.* http://hannonlab.cshl.edu/fastx_toolkit/download.html

Griffith M, Griffith OL, Coffman AC, Weible JV, McMichael JF, Spies NC, Koval J, Das I, Callaway MB, Eldred JM, Miller CA, Subramanian J, Govindan R, Kumar RD, Bose R, Ding L, Walker JR, Larson DE, Dooling DJ, Smith SM, et al. 2013. DGIdb: mining the druggable genome. *Nature Methods* **10**:1209–1210. DOI: https://doi.org/10.1038/nmeth.2689, PMID: 24122041

Guzmán C, Bagga M, Kaur A, Westermarck J, Abankwa D. 2014. ColonyArea: an ImageJ plugin to automatically quantify colony formation in clonogenic assays. *PLOS ONE* **9**:e92444. DOI: https://doi.org/10.1371/journal.pone.0092444, PMID: 24647355

Hahn SA, Schutte M, Hoque AT, Moskaluk CA, da Costa LT, Rozenblum E, Weinstein CL, Fischer A, Yeo CJ, Hruban RH, Kern SE. 1996. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* **271**:350–353. DOI: https://doi.org/10.1126/science.271.5247.350, PMID: 8553070

Halder R, Riou JF, Teulade-Fichou MP, Frickey T, Hartig JS. 2012. Bisquinolinium compounds induce quadruplex-specific transcriptome changes in HeLa S3 cell lines. *BMC Research Notes* **5**:138. DOI: https://doi.org/10.1186/1756-0500-5-138, PMID: 22414013

Hänsel-Hertsch R, Beraldi D, Lensing SV, Marsico G, Zyner K, Parry A, Di Antonio M, Pike J, Kimura H, Narita M, Tannahill D, Balasubramanian S. 2016. G-quadruplex structures mark human regulatory chromatin. *Nature Genetics* **48**:1267–1272. DOI: https://doi.org/10.1038/ng.3662, PMID: 27618450

Harrigan JA, Jacq X, Martin NM, Jackson SP. 2018. Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nature Reviews Drug Discovery* **17**:57–78. DOI: https://doi.org/10.1038/nrd.2017.152, PMID: 28959952

Henderson A, Wu Y, Huang YC, Chavez EA, Platt J, Johnson FB, Brosh RM, Sen D, Lansdorp PM. 2014. Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Research* **42**:860–869. DOI: https://doi.org/10.1093/nar/gkt957, PMID: 24163102

Herdy B, Mayer C, Varshney D, Marsico G, Murat P, Taylor C, D'Santos C, Tannahill D, Balasubramanian S. 2018. Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. *Nucleic Acids Research* **46**:11592–11604. DOI: https://doi.org/10.1093/nar/gky861, PMID: 30256975

Huang TT, Nijman SMB, Mirchandani KD, Galardy PJ, Cohn MA, Haas W, Gygi SP, Ploegh HL, Bernards R, D'Andrea AD. 2006. Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nature Cell Biology* **8**:341–347. DOI: https://doi.org/10.1038/ncb1378

Huang daW, Sherman BT, Lempicki RA. 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**:44–57. DOI: https://doi.org/10.1038/nprot.2008.211, PMID: 19131956

Huang daW, Sherman BT, Lempicki RA. 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* **37**:1–13. DOI: https://doi.org/10.1093/nar/gkn923, PMID: 19033363

Huang X, Dixit VM. 2016. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Research* **26**:484–498. DOI: https://doi.org/10.1038/cr.2016.31, PMID: 27002218

Huppert JL. 2008. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chemical Society Reviews* **37**:1375 . DOI: https://doi.org/10.1039/b702491f

Huppert JL, Balasubramanian S. 2007. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Research* **35**:406–413 . DOI: https://doi.org/10.1093/nar/gkl1057

Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, Crabtree GR. 2013. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature Genetics* **45**:592–601. DOI: https://doi.org/10.1038/ng.2628

Kikuma T, Ohtsu M, Utsugi T, Koga S, Okuhara K, Eki T, Fujimori F, Murakami Y. 2004. Dbp9p, a member of the DEAD box protein family, exhibits DNA helicase activity. *Journal of Biological Chemistry* **279**:20692–20698. DOI: https://doi.org/10.1074/jbc.M400231200, PMID: 15028736

Knott SRV, Maceli A, Erard N, Chang K, Marran K, Zhou X, Gordon A, Demerdash OE, Wagenblast E, Kim S, Fellmann C, Hannon GJ. 2014. A computational algorithm to predict shRNA potency. *Molecular Cell* **56**:796–807. DOI: https://doi.org/10.1016/j.molcel.2014.10.025, PMID: 25435137

Kotake Y, Sagane K, Owa T, Mimori-Kiyosue Y, Shimizu H, Uesugi M, Ishihama Y, Iwata M, Mizui Y. 2007. Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nature Chemical Biology* **3**:570–575. DOI: https://doi.org/10.1038/nchembio.2007.16, PMID: 17643112

Kumari S, Bugaut A, Huppert JL, Balasubramanian S. 2007. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nature Chemical Biology* **3**:218–221. DOI: https://doi.org/10.1038/nchembio864, PMID: 17322877

Kwok CK, Marsico G, Sahakyan AB, Chambers VS, Balasubramanian S. 2016. rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nature Methods* **13**:841–844. DOI: https://doi.org/10.1038/nmeth.3965, PMID: 27571552

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. *Nature Methods* **9**:357–359. DOI: https://doi.org/10.1038/nmeth.1923, PMID: 22388286

Le DD, Di Antonio M, Chan LKM, Balasubramanian S. 2015. G-quadruplex ligands exhibit differential G-tetrad selectivity. *Chemical Communications* **51**:8048–8050. DOI: https://doi.org/10.1039/c5cc02252e

Lee JM, Ledermann JA, Kohn EC. 2014. PARP inhibitors for BRCA1/2 mutation-associated and BRCA-like malignancies. *Annals of Oncology* **25**:32–40. DOI: https://doi.org/10.1093/annonc/mdt384, PMID: 24225019

Lim SP, Wong NC, Suetani RJ, Ho K, Ng JL, Neilsen PM, Gill PG, Kumar R, Callen DF. 2012. Specific-site methylation of tumour suppressor ANKRD11 in breast Cancer. *European Journal of Cancer* **48**:3300–3309. DOI: https://doi.org/10.1016/j.ejca.2012.03.023, PMID: 22538187

Liu J, Shaik S, Dai X, Wu Q, Zhou X, Wang Z, Wei W. 2015a. Targeting the ubiquitin pathway for cancer treatment. *Biochimica Et Biophysica Acta (BBA) - Reviews on Cancer* **1855**:50–60. DOI: https://doi.org/10.1016/j.bbcan.2014.11.005

Liu Y, Liang Y, Wishart D. 2015b. PolySearch2: a significantly improved text-mining system for discovering associations between human diseases, genes, drugs, metabolites, toxins and more. *Nucleic Acids Research* **43**:W535–W542. DOI: https://doi.org/10.1093/nar/gkv383, PMID: 25925572

Liu HY, Zhao Q, Zhang TP, Wu Y, Xiong YX, Wang SK, Ge YL, He JH, Lv P, Ou TM, Tan JH, Li D, Gu LQ, Ren J, Zhao Y, Huang ZS. 2016. Conformation selective antibody enables genome profiling and leads to discovery of parallel G-Quadruplex in human telomeres. *Cell Chemical Biology* **23**:1261–1270. DOI: https://doi.org/10.1016/j.chembiol.2016.08.013, PMID: 27693060

Lu MC, Ji JA, Jiang YL, Chen ZY, Yuan ZW, You QD, Jiang ZY. 2016. An inhibitor of the Keap1-Nrf2 protein-protein interaction protects NCM460 colonic cells and alleviates experimental colitis. *Scientific Reports* **6**: 26585. DOI: https://doi.org/10.1038/srep26585, PMID: 27215610

Martens-de Kemp SR, Brink A, van der Meulen IH, de Menezes RX, Te Beest DE, Leemans CR, van Beusechem VW, Braakhuis BJ, Brakenhoff RH. 2017. The FA/BRCA pathway identified as the major predictor of cisplatin response in head and neck cancer by functional genomics. *Molecular Cancer Therapeutics* **16**:540–550. DOI: https://doi.org/10.1158/1535-7163.MCT-16-0457, PMID: 27980104

Martínez Cuesta S. 2019. GWscreen_G4sensitivity. *GitHub*. 9264d3b. https://github.com/sblab-bioinformatics/GWscreen_G4sensitivity

Matheson CJ, Backos DS, Reigan P. 2016. Targeting WEE1 kinase in cancer. *Trends in Pharmacological Sciences* **37**:872–881. DOI: https://doi.org/10.1016/j.tips.2016.06.006, PMID: 27427153

Mattern MR, Wu J, Nicholson B. 2012. Ubiquitin-based anticancer therapy: carpet bombing with proteasome inhibitors vs surgical strikes with E1, E2, E3, or DUB inhibitors. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research* **1823**:2014–2021. DOI: https://doi.org/10.1016/j.bbamcr.2012.05.005

McCubrey JA, Steelman LS, Bertrand FE, Davis NM, Sokolosky M, Abrams SL, Montalto G, D'Assoro AB, Libra M, Nicoletti F, Maestro R, Basecke J, Rakus D, Gizak A, Demidenko ZN, Cocco L, Martelli AM, Cervello M. 2014. GSK-3 as potential target for therapeutic intervention in cancer. *Oncotarget* **5**:2881–2911. DOI: https://doi.org/10.18632/oncotarget.2037, PMID: 24931005

McLuckie KI, Di Antonio M, Zecchini H, Xian J, Caldas C, Krippendorff BF, Tannahill D, Lowe C, Balasubramanian S. 2013. G-quadruplex DNA as a molecular target for induced synthetic lethality in cancer cells. *Journal of the American Chemical Society* **135**:9640–9643. DOI: https://doi.org/10.1021/ja404868t, PMID: 23782415

Mendoza O, Bourdoncle A, Boulé JB, Brosh RM, Mergny JL. 2016. G-quadruplexes and helicases. *Nucleic Acids Research* **44**:1989–2006. DOI: https://doi.org/10.1093/nar/gkw079, PMID: 26883636

Mishra SK, Tawani A, Mishra A, Kumar A. 2016. G4IPDB: a database for G-quadruplex structure forming nucleic acid interacting proteins. *Scientific Reports* **6**:38144. DOI: https://doi.org/10.1038/srep38144, PMID: 27905517

Morra F, Luise C, Visconti R, Staibano S, Merolla F, Ilardi G, Guggino G, Paladino S, Sarnataro D, Franco R, Monaco R, Zitomarino F, Pacelli R, Monaco G, Rocco G, Cerrato A, Linardopoulos S, Muller MT, Celetti A. 2015. New therapeutic perspectives in CCDC6 deficient lung cancer cells. *International Journal of Cancer* **136**: 2146–2157. DOI: https://doi.org/10.1002/ijc.29263, PMID: 25302833

Moyal L, Lerenthal Y, Gana-Weisz M, Mass G, So S, Wang SY, Eppink B, Chung YM, Shalev G, Shema E, Shkedy D, Smorodinsky NI, van Vliet N, Kuster B, Mann M, Ciechanover A, Dahm-Daphi J, Kanaar R, Hu MC, Chen DJ, et al. 2011. Requirement of ATM-dependent monoubiquitylation of histone H2B for timely repair of DNA double-strand breaks. *Molecular Cell* **41**:529–542. DOI: https://doi.org/10.1016/j.molcel.2011.02.015, PMID: 21362549

Neidle S. 2010. Human telomeric G-quadruplex: the current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS Journal* **277**:1118–1125. DOI: https://doi.org/10.1111/j.1742-4658.2009.07463.x, PMID: 19951354

Neidle S. 2017. Quadruplex nucleic acids as targets for anticancer therapeutics. *Nature Reviews Chemistry* **1**: 0041. DOI: https://doi.org/10.1038/s41570-017-0041

Neilsen PM, Cheney KM, Li CW, Chen JD, Cawrse JE, Schulz RB, Powell JA, Kumar R, Callen DF. 2008. Identification of ANKRD11 as a p53 coactivator. *Journal of Cell Science* **121**:3541–3552. DOI: https://doi.org/10.1242/jcs.026351, PMID: 18840648

Nepomuceno T, De Gregoriis G, de Oliveira FMB, Suarez-Kurtz G, Monteiro A, Carvalho M. 2017. The role of PALB2 in the DNA damage response and Cancer predisposition. *International Journal of Molecular Sciences* **18**:1886. DOI: https://doi.org/10.3390/ijms18091886

Nijman SM, Huang TT, Dirac AM, Brummelkamp TR, Kerkhoven RM, D'Andrea AD, Bernards R. 2005. The deubiquitinating enzyme USP1 regulates the fanconi anemia pathway. *Molecular Cell* **17**:331–339. DOI: https://doi.org/10.1016/j.molcel.2005.01.008, PMID: 15694335

Nijman SM. 2011. Synthetic lethality: general principles, utility and detection using genetic screens in human cells. *FEBS Letters* **585**:1–6. DOI: https://doi.org/10.1016/j.febslet.2010.11.024, PMID: 21094158

Noll JE, Jeffery J, Al-Ejeh F, Kumar R, Khanna KK, Callen DF, Neilsen PM. 2012. Mutant p53 drives multinucleation and invasion through a process that is suppressed by ANKRD11. *Oncogene* **31**:2836–2848. DOI: https://doi.org/10.1038/onc.2011.456, PMID: 21986947

Pan J, Deng Q, Jiang C, Wang X, Niu T, Li H, Chen T, Jin J, Pan W, Cai X, Yang X, Lu M, Xiao J, Wang P. 2015. USP37 directly deubiquitinates and stabilizes c-Myc in lung cancer. *Oncogene* **34**:3957–3967. DOI: https://doi.org/10.1038/onc.2014.327, PMID: 25284584

Parry D, Guzi T, Shanahan F, Davis N, Prabhavalkar D, Wiswell D, Seghezzi W, Paruch K, Dwyer MP, Doll R, Nomeir A, Windsor W, Fischmann T, Wang Y, Oft M, Chen T, Kirschmeier P, Lees EM. 2010. Dinaciclib (SCH 727965), a novel and potent cyclin-dependent kinase inhibitor. *Molecular Cancer Therapeutics* **9**:2344–2353. DOI: https://doi.org/10.1158/1535-7163.MCT-10-0324, PMID: 20663931

Puxeddu E, Zhao G, Stringer JR, Medvedovic M, Moretti S, Fagin JA. 2005. Characterization of novel non-clonal intrachromosomal rearrangements between the H4 and PTEN genes (H4/PTEN) in human thyroid cell lines and papillary thyroid cancer specimens. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **570**:17–32. DOI: https://doi.org/10.1016/j.mrfmmm.2004.09.010

Rao RC, Dou Y. 2015. Hijacked in Cancer: the KMT2 (MLL) family of methyltransferases. *Nature Reviews Cancer* **15**:334–346. DOI: https://doi.org/10.1038/nrc3929, PMID: 25998713

Rhodes D, Lipps HJ. 2015. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Research* **43**: 8627–8637. DOI: https://doi.org/10.1093/nar/gkv862, PMID: 26350216

Richer AL, Cala JM, O'Brien K, Carson VM, Inge LJ, Whitsett TG. 2017. WEE1 kinase inhibitor AZD1775 has preclinical efficacy in LKB1-Deficient Non-Small cell lung cancer. *Cancer Research* **77**:4663–4672. DOI: https://doi.org/10.1158/0008-5472.CAN-16-3565, PMID: 28652249

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–140. DOI: https://doi.org/10.1093/bioinformatics/btp616, PMID: 19910308

Rodriguez R, Müller S, Yeoman JA, Trentesaux C, Riou JF, Balasubramanian S. 2008. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *Journal of the American Chemical Society* **130**:15758–15759. DOI: https://doi.org/10.1021/ja805615w, PMID: 18975896

Rodriguez R, Miller KM, Forment JV, Bradshaw CR, Nikan M, Britton S, Oelschlaegel T, Xhemalce B, Balasubramanian S, Jackson SP. 2012. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nature Chemical Biology* **8**:301–310. DOI: https://doi.org/10.1038/nchembio.780, PMID: 22306580

Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, Dimitriadoy S, Liu DL, Kantheti HS, Saghafinia S, Chakravarty D, Daian F, Gao Q, Bailey MH, Liang WW, Foltz SM, Shmulevich I, Ding L, Heins Z, Ochoa A, et al. 2018. Oncogenic signaling pathways in the cancer genome atlas. *Cell* **173**:321–337. DOI: https://doi.org/10.1016/j.cell.2018.03.035, PMID: 29625050

Schaffitzel C, Berger I, Postberg J, Hanes J, Lipps HJ, Plückthun A. 2001. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with stylonychia lemnae macronuclei. *PNAS* **98**:8572–8577. DOI: https://doi.org/10.1073/pnas.141229498, PMID: 11438689

Schutte M, Hruban RH, Hedrick L, Cho KR, Nadasdy GM, Weinstein CL, Bova GS, Isaacs WB, Cairns P, Nawroz H, Sidransky D, Casero RA, Meltzer PS, Hahn SA, Kern SE. 1996. DPC4 gene in various tumor types. *Cancer Research* **56**:2527–2530. PMID: 8653691

Senft D, Qi J, Ronai ZA, , . 2018. Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nature Reviews Cancer* **18**:69–88. DOI: https://doi.org/10.1038/nrc.2017.105, PMID: 29242641

Shain AH, Pollack JR. 2013. The spectrum of SWI/SNF mutations, ubiquitous in human cancers. *PLOS ONE* **8**: e55119. DOI: https://doi.org/10.1371/journal.pone.0055119, PMID: 23355908

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**:2498–2504. DOI: https://doi.org/10.1101/gr.1239303, PMID: 14597658

Shema E, Tirosh I, Aylon Y, Huang J, Ye C, Moskovits N, Raver-Shapira N, Minsky N, Pirngruber J, Tarcic G, Hublarova P, Moyal L, Gana-Weisz M, Shiloh Y, Yarden Y, Johnsen SA, Vojtesek B, Berger SL, Oren M. 2008. The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes & Development* **22**:2664–2676. DOI: https://doi.org/10.1101/gad.1703008, PMID: 18832071

Shkreta L, Blanchette M, Toutant J, Wilhelm E, Bell B, Story BA, Balachandran A, Cochrane A, Cheung PK, Harrigan PR, Grierson DS, Chabot B. 2017. Modulation of the splicing regulatory function of SRSF10 by a novel compound that impairs HIV-1 replication. *Nucleic Acids Research* **45**:4051–4067. DOI: https://doi.org/10.1093/nar/gkw1223, PMID: 27928057

Sims D, Mendes-Pereira AM, Frankum J, Burgess D, Cerone M-A, Lombardelli C, Mitsopoulos C, Hakas J, Murugaesu N, Isacke CM, Fenwick K, Assiotis I, Kozarewa I, Zvelebil M, Ashworth A, Lord CJ. 2011. High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. *Genome Biology* **12**:R104. DOI: https://doi.org/10.1186/gb-2011-12-10-r104

Soucy TA, Smith PG, Milhollen MA, Berger AJ, Gavin JM, Adhikari S, Brownell JE, Burke KE, Cardin DP, Critchley S, Cullis CA, Doucette A, Garnsey JJ, Gaulin JL, Gershman RE, Lublinsky AR, McDonald A, Mizutani H, Narayanan U, Olhava EJ, et al. 2009. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**:732–736. DOI: https://doi.org/10.1038/nature07884, PMID: 19360080

Spillare EA, Wang XW, von Kobbe C, Bohr VA, Hickson ID, Harris CC. 2006. Redundancy of DNA helicases in p53-mediated apoptosis. *Oncogene* **25**:2119–2123. DOI: https://doi.org/10.1038/sj.onc.1209242, PMID: 16288211

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. 2017. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research* **45**:D362–D368. DOI: https://doi.org/10.1093/nar/gkw937

Takahama K, Takada A, Tada S, Shimizu M, Sayama K, Kurokawa R, Oyoshi T. 2013. Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS. *Chemistry & Biology* **20**:341–350. DOI: https://doi.org/10.1016/j.chembiol.2013.02.013, PMID: 23521792

The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **45**:D158–D169. DOI: https://doi.org/10.1093/nar/gkw1099, PMID: 27899622

Uhlmann-Schiffler H, Jalal C, Stahl H. 2006. Ddx42p–a human DEAD box protein with RNA chaperone activities. *Nucleic Acids Research* **34**:10–22. DOI: https://doi.org/10.1093/nar/gkj403, PMID: 16397294

Valton AL, Prioleau MN. 2016. G-Quadruplexes in DNA replication: a problem or a necessity? *Trends in Genetics* **32**:697–706. DOI: https://doi.org/10.1016/j.tig.2016.09.004, PMID: 27663528

Vaughn JP, Creacy SD, Routh ED, Joyner-Butt C, Jenkins GS, Pauli S, Nagamine Y, Akman SA. 2005. The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving

activity in HeLa cell lysates. *Journal of Biological Chemistry* **280**:38117–38120. DOI: https://doi.org/10.1074/jbc.C500348200, PMID: 16150737

Wang JC. 2002. Cellular roles of DNA topoisomerases: a molecular perspective. *Nature Reviews Molecular Cell Biology* **3**:430–440. DOI: https://doi.org/10.1038/nrm831, PMID: 12042765

Wang YY, Yang YX, Zhe H, He ZX, Zhou SF. 2014. Bardoxolone methyl (CDDO-Me) as a therapeutic agent: an update on its pharmacokinetic and pharmacodynamic properties. *Drug Design, Development and Therapy* **8**: 2075–2088. DOI: https://doi.org/10.2147/DDDT.S68872, PMID: 25364233

Wang Y, Yang J, Wild AT, Wu WH, Shah R, Danussi C, Riggins GJ, Kannan K, Sulman EP, Chan TA, Huse JT. 2019. G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nature Communications* **10**:943. DOI: https://doi.org/10.1038/s41467-019-08905-8, PMID: 30808951

Watson LA, Solomon LA, Li JR, Jiang Y, Edwards M, Shin-ya K, Beier F, Bérubé NG. 2013. Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. *Journal of Clinical Investigation* **123**: 2049–2063. DOI: https://doi.org/10.1172/JCI65634, PMID: 23563309

Wei W, Lin HK, . 2012. The key role of ubiquitination and sumoylation in signaling and cancer: a research topic. *Frontiers in Oncology* **2**:187. DOI: https://doi.org/10.3389/fonc.2012.00187, PMID: 23227454

Will CL, Urlaub H, Achsel T, Gentzel M, Wilm M, Lührmann R. 2002. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *The EMBO Journal* **21**: 4978–4988. DOI: https://doi.org/10.1093/emboj/cdf480, PMID: 12234937

Wolfe AL, Singh K, Zhong Y, Drewe P, Rajasekhar VK, Sanghvi VR, Mavrakis KJ, Jiang M, Roderick JE, Van der Meulen J, Schatz JH, Rodrigo CM, Zhao C, Rondou P, de Stanchina E, Teruya-Feldstein J, Kelliher MA, Speleman F, Porco JA, Pelletier J, et al. 2014. RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513**:65–70. DOI: https://doi.org/10.1038/nature13485, PMID: 25079319

Wu W, Rokutanda N, Takeuchi J, Lai Y, Maruyama R, Togashi Y, Nishikawa H, Arai N, Miyoshi Y, Suzuki N, Saeki Y, Tanaka K, Ohta T. 2018. HERC2 facilitates BLM and WRN helicase complex interaction with RPA to suppress G-Quadruplex DNA. *Cancer Research* **78**:6371–6385. DOI: https://doi.org/10.1158/0008-5472.CAN-18-1877, PMID: 30279242

Wu Y, Brosh RM. 2010. G-quadruplex nucleic acids and human disease. *FEBS Journal* **277**:3470–3488. DOI: https://doi.org/10.1111/j.1742-4658.2010.07760.x, PMID: 20670277

Xu H, Di Antonio M, McKinney S, Mathew V, Ho B, O'Neil NJ, Santos ND, Silvester J, Wei V, Garcia J, Kabeer F, Lai D, Soriano P, Banáth J, Chiu DS, Yap D, Le DD, Ye FB, Zhang A, Thu K, et al. 2017. CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nature Communications* **8**:14432. DOI: https://doi.org/10.1038/ncomms14432, PMID: 28211448

Yadav P, Harcy V, Argueso JL, Dominska M, Jinks-Robertson S, Kim N. 2014. Topoisomerase I plays a critical role in suppressing genome instability at a highly transcribed G-quadruplex-forming sequence. *PLOS Genetics* **10**: e1004839. DOI: https://doi.org/10.1371/journal.pgen.1004839, PMID: 25473964

Yang D, Hurley LH. 2006. Structure of the biologically relevant G-Quadruplex in the c-MYC promoter. *Nucleosides, Nucleotides and Nucleic Acids* **25**:951–968. DOI: https://doi.org/10.1080/15257770600809913

Yeats C, Maibaum M, Marsden R, Dibley M, Lee D, Addou S, Orengo CA. 2006. Gene3D: modelling protein structure, function and evolution. *Nucleic Acids Research* **34**:D281–D284. DOI: https://doi.org/10.1093/nar/gkj057, PMID: 16381865

Yoshida T, Tanaka S, Mogi A, Shitara Y, Kuwano H. 2004. The clinical significance of cyclin B1 and Wee1 expression in non-small-cell lung cancer. *Annals of Oncology* **15**:252–256. DOI: https://doi.org/10.1093/annonc/mdh073, PMID: 14760118

Zhao F, Gong X, Liu A, Lv X, Hu B, Zhang H. 2018. Downregulation of Nedd4L predicts poor prognosis, promotes tumor growth and inhibits MAPK/ERK signal pathway in hepatocellular carcinoma. *Biochemical and Biophysical Research Communications* **495**:1136–1143. DOI: https://doi.org/10.1016/j.bbrc.2017.11.139, PMID: 29175326

Zhou X, Li X, Cheng Y, Wu W, Xie Z, Xi Q, Han J, Wu G, Fang J, Feng Y. 2014. BCLAF1 and its splicing regulator SRSF10 regulate the tumorigenic potential of colon cancer cells. *Nature Communications* **5**:4581. DOI: https://doi.org/10.1038/ncomms5581, PMID: 25091051

Zimmer J, Tacconi EMC, Folio C, Badie S, Porru M, Klare K, Tumiati M, Markkanen E, Halder S, Ryan A, Jackson SP, Ramadan K, Kuznetsov SG, Biroccio A, Sale JE, Tarsounas M. 2016. Targeting BRCA1 and BRCA2 deficiencies with G-Quadruplex-Interacting compounds. *Molecular Cell* **61**:449–460. DOI: https://doi.org/10.1016/j.molcel.2015.12.004, PMID: 26748828