Trends in Cancer

CellPress OPEN ACCESS

Review

Targeting synthetic lethal paralogs in cancer

Colm J. Ryan [1,2,*] Ishan Mehta,[3] Narod Kebabci,[1,4] and David J. Adams[3]

Synthetic lethal interactions, where mutation of one gene renders cells sensitive to inhibition of another gene, can be exploited for the development of targeted therapeutics in cancer. Pairs of duplicate genes (paralogs) often share common functionality and hence are a potentially rich source of synthetic lethal interactions. Because the majority of human genes have paralogs, exploiting such interactions could be a widely applicable approach for targeting gene loss in cancer. Moreover, existing small-molecule drugs may exploit synthetic lethal interactions by inhibiting multiple paralogs simultaneously. Consequently, the identification of synthetic lethal interactions between paralogs could be extremely informative for drug development. Here we review approaches to identify such interactions and discuss some of the challenges of exploiting them.

Highlights

Paralog pairs appear to be a rich source of synthetic lethality, and many of the synthetic lethal interactions in cancer reported to date involve paralog pairs.

Single- and double-gene CRISPR perturbation screens, combined with computational analyses, enable the discovery of new synthetic lethal paralog pairs at scale.

Synthetic lethal interactions between paralogs can potentially be exploited to target both tumour-suppressor mutations and passenger gene deletions. Identified interactions can serve as the starting point for the development of new targeted therapeutics.

Developing paralog-specific small-molecule inhibitors may present a challenge to exploiting synthetic lethal paralogs. However, existing small-molecule inhibitors may actually exploit synthetic lethality between paralogs because of a lack of specificity. Non-paralog-specific inhibitors may therefore be useful for targeting conditional synthetic lethal pairs.

Introduction – identifying synthetic lethal targets in cancer

Synthetic lethality has emerged as a useful paradigm for the identification of therapeutic targets in cancer [1]. The term was initially used to refer to an incompatibility between pairs of mutations in *Drosophila* where the presence of either mutation alone was well tolerated but their presence in combination resulted in lethality. However, in the context of cancer, the term is now used more broadly to refer to any instance where mutation of one gene renders cells especially sensitive to the perturbation of another gene or its protein product [1,2] (Figure 1B, Key figure). In principle, such relationships can be exploited for the selective killing of tumour cells that harbour a specific mutation, leaving normal cells unaffected (Figure 1C). In 2005, the first clinically useful example of a synthetic lethal interaction was identified – loss of either the *BRCA1* or *BRCA2* tumour-suppressor gene was shown to be associated with increased sensitivity to inhibition of PARP1 [3,4]. Based on this discovery and subsequent clinical trials, small-molecule inhibitors of PARP have now been approved for the treatment of select ovarian, breast, and prostate cancer patients whose tumours carry *BRCA1* or *BRCA2* mutations [5].

The clinical translation of the BRCA/PARP synthetic lethality, combined with the recent development of approaches based on **CRISPR** (see Glossary) to perform large-scale loss-of-function screens, has led to a significant increase in efforts to identify synthetic lethal interactions that may be exploited for the development of new targeted therapies [6]. In contrast to **oncogene addiction**, which can only be exploited to target alterations of druggable oncogenes, synthetic lethality can potentially be used to target a much wider variety of genetic and epigenetic alterations. These include not only alterations of both oncogenes and tumour suppressors but also **passenger gene mutations** [1,7]. Consequently, synthetic lethal approaches could potentially be used to target almost any alteration in cancer. However, a major challenge in the identification of new synthetic lethal interactions is that the search space for possible synthetic lethal pairs is enormous (~200 million gene pairs, assuming 20 000 protein coding genes) and synthetic lethal interactions appear to be extremely rare. Fortunately, pairs of duplicate genes (paralogs) appear to be an extremely rich source of synthetic lethality (Figure 1A,B). These appear to be significantly more likely than random gene pairs to be synthetic lethal, and many of the synthetic lethal interactions reported to date in cancer involve pairs of paralogs. Consequently, focusing

[1]Conway Institute and School of Computer Science, University College Dublin, Dublin, Ireland
[2]Systems Biology Ireland, University College Dublin, Dublin, Ireland
[3]Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK
[4]Science Foundation Ireland (SFI) Centre for Research Training in Genomics Data Science, University College Dublin, Dublin, Ireland

*Correspondence:
colm.ryan@ucd.ie (C.J. Ryan).

Key figure

Gene duplication and synthetic lethality

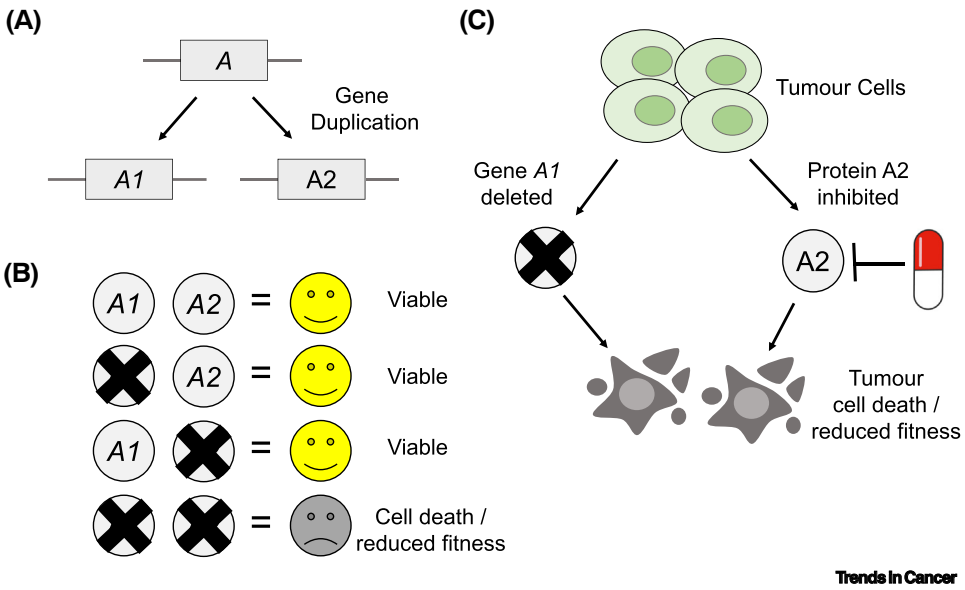

Figure 1. (A) Duplication of gene *A* results in the generation of a pair of paralogs, termed *A1* and *A2*. (B) Synthetic lethality between paralogs *A1* and *A2*. Either paralog can be inhibited/mutated individually with little consequence, but loss of both paralogs causes cell death or significantly reduced fitness. (C) Exploiting synthetic lethality between *A1* and *A2* to selectively kill tumour cells that have lost a functioning copy of gene *A1*.

discovery efforts on this relatively small subset of gene pairs is likely to result in the identification of new therapeutic targets (Figure 1C).

We discuss here the identification of synthetic lethal paralogs in cancer along with some of the potential challenges and opportunities of exploiting such synthetic lethal interactions for the development of targeted therapies.

Paralogs as a source of synthetic lethality

Gene duplication is the primary means by which new genes are created, and consequently the majority of human genes have an identifiable paralog. Paralog pairs can be broadly grouped into two categories – small-scale duplicates, which arise when a single gene or a limited number of genes is duplicated, and whole-genome duplicates, which arise when all chromosomes are duplicated simultaneously (reviewed in [8]). Small-scale duplicates typically arise from errors in DNA replication, resulting in the duplication of existing chromosomal segments, but can also arise from retrotransposition, where novel DNA sequences are introduced through reverse transcription of modified RNA transcripts [9,10]. The ancestor of all vertebrates, including humans, underwent two rounds of whole-genome duplication [11]. As a result of these two rounds of whole-genome duplication, along with additional small-scale duplication events, many human genes have multiple identifiable paralogs.

Paralog pairs are typically inferred computationally using some form of sequence comparison, and widely used resources such as EnsemblCompara incorporate information from many

Glossary

Cas enzymes: CRISPR-associated (Cas) proteins derived from a variety of bacterial species are used in combination with CRISPR guide RNAs (gRNAs) for genome editing.
CRISPR: clustered regularly interspaced short palindromic repeats – a bacterial defence system that has been repurposed as a gene-editing system.
CRISPR tiling screens: CRISPR screens performed using many sgRNAs that target the entire coding sequence of a given gene, facilitating the identification of important regions/domains.
Oncogene addiction: an increased dependency of tumour cells on the function of an oncogene, typically one that has been amplified or mutated.
Passenger gene mutations: mutations in genes whose alteration is not required for tumourigenesis and that do not provide a fitness advantage to tumour cells.
PROTACs: proteolysis-targeting chimera molecules that induce the degradation of target proteins. They include two ligands joined by a linker – one ligand binds to the protein of interest and the other recruits an E3 ligase.
Single guide RNA (sgRNA): a short piece of RNA that binds to a Cas enzyme and guides the Cas enzyme to a specific location in the genome which is complementary in sequence. If used in conjunction with Cas9 or Cas12a, these enzymes will induce a double-stranded DNA break and therefore disrupt the targeted genomic site. This term is used interchangeably with guide RNA or guide.
Synthetic sick interaction: a combination of genetic events that results in reduced fitness but not lethality. In high-throughput screens it can be difficult to distinguish between synthetic sick and synthetic lethal interactions, and the two are often treated in the same way.

species to identify paralogs [12,13]. A consequence of this computational inference is that there is no definitive list of paralogs in the human genome, and the paralog pairs identified vary across resources and also across different versions of the same resource. For simplicity in this review, when discussing paralog pairs, we use the defined set of paralog pairs from our previous publication – these were derived from Ensembl 93 with the additional requirement that a pair of genes required a minimum of 20% sequence identity in one direction to be considered paralogs [14]. By this definition, ~70% of human protein-coding genes have an identifiable paralog.

Over evolutionary time-periods after a duplication event, pairs of paralogs may diverge in terms of their functionality. In some cases one paralog may gain new functionality that the other does not possess, whereas in other cases existing functionality may be partitioned between the two paralogs (recently reviewed by Kuzmin *et al.* [8]). Nonetheless, in many cases pairs of paralogs retain sufficient common functionality to compensate for each other's loss in at least some circumstances, as revealed by synthetic lethal relationships (Figure 1B).

Before it was feasible to screen for synthetic lethal interactions in human cells, work in model organisms had established that paralog pairs are especially enriched for synthetic lethality [15–19]. For instance, in budding yeast 25–35% of paralog pairs display a **synthetic sick interaction** or a synthetic lethal interaction, whereas <5% of randomly selected gene pairs do [15,18]. These observations in model organisms have led many cancer researchers to prioritise paralog pairs in the search for novel synthetic lethal interactions.

Using paralog synthetic lethal relationships to target recurrent loss-of-function alterations in cancer

The most obvious application of synthetic lethality in cancer is to target genes that are subject to recurrent loss-of-function mutation or homozygous deletion in tumours (Figure 1C). Unlike oncogenes, these genes cannot be targeted directly with drugs because they no longer produce a protein product. The set of recurrently lost genes in cancer include both tumour suppressors and passenger genes. Tumour suppressors are especially promising targets because they offer tumour cells a selective advantage, and their alteration may therefore be subject to less heterogeneity within a tumour [20]. Of the 319 genes currently annotated as tumour suppressors in the Cancer Gene Census [21], 238 (74%) have an identifiable paralog and may therefore be potentially targeted using paralog synthetic lethal approaches. Tumour suppressors with reported paralog synthetic lethality include *SMARCA4*, *ARID1A*, *CREBBP*, and *STAG2* (Table 1).

Often when a tumour-suppressor gene is subject to homozygous deletion, multiple passenger genes adjacent to it on the chromosome are also subject to 'collateral loss'. Although not themselves drivers of cancer, these recurrently deleted passenger genes may expose vulnerabilities in cancer cells that are not present in normal cells [7,22]. The number of recurrently deleted passenger genes far outweighs the number of known tumour suppressors – a recent analysis identified >1700 passenger genes that are recurrently homozygous deleted in tumours [23]. The same paper identified that passenger genes with paralogs are more likely to be homozygously deleted than passenger genes without paralogs [23] suggesting that the set of potential genes to target using paralog synthetic lethality may be very large.

To our knowledge, the first reported paralog synthetic lethality in cancer involves a recurrently deleted passenger gene – the passenger gene *ENO1* is recurrently deleted in glioblastoma and its loss renders glioblastoma cells sensitive to inhibition of its paralog *ENO2* [24]. Subsequent work has identified multiple synthetic lethal interactions involving tumour suppressors, passenger genes, and genes of unknown status (Table 1).

Table 1. Selected synthetic lethal interactions involving paralogs that are recurrently mutated, deleted, or silenced in cancer

Altered gene	Vulnerability	Type	Altered gene details	Refs
ARID1A	ARID1B	Driver	*ARID1A* encodes a SWI/SNF subunit and is an established tumour suppressor in multiple cancer types	[32]
CREBBP	EP300	Driver	*CREBBP* encodes a histone acetylase and is an established tumour suppressor in multiple cancer types	[68]
SMARCA4	SMARCA2	Driver	*SMARCA4* encodes a SWI/SNF subunit and is an established tumour suppressor in multiple cancer types	[33,48,49]
STAG2	STAG1	Driver	*STAG2* encodes a cohesin subunit and is an established tumour suppressor in multiple cancer types	[69,70]
ENO1	ENO2	Passenger	*ENO1* encodes a glycolytic enzyme and is believed to be a passenger gene – it resides on 1p36, a locus that is subject to frequent homozygous deletion and that contains multiple candidate tumour suppressors	[24]
MAGOH	MAGOHB	Passenger	*MAGOH*, which encodes a subunit of the splicing-dependent exon junction complex, is located on chromosome 1p and is recurrently hemizygously lost in multiple cancer types	[28]
ME2	ME3	Passenger	*ME2*, which encodes a metabolic enzyme, is located near the tumour-suppressor *SMAD4* and hence is subject to frequent co-deletion	[71]
PTDSS2	PTDSS1	Passenger	*PTDSS2*, which encodes a phosphatidylserine synthetase, is located at 11p15.5 and is recurrently deleted in multiple cancer types	[72]
VPS4B	VPS4A	Passenger	*VPS4A* and *VPS4B* encode ESCRT subunits. *VPS4B* is adjacent to the tumour-suppressor *SMAD4* and is subject to frequent co-deletions as a result. Similarly, *VPS4A* is co-deleted with the tumour-suppressor *CDH1*. The reciprocal synthetic lethal interaction, where *VPS4A* is lost and *VPS4B* is a vulnerability, has also been reported	[40,73,74]
ASF1A	ASF1B	Unknown	*ASF1A* encodes a histone chaperone. It is unclear whether it is a tumour suppressor but it is recurrently deleted in multiple cancer types including prostate adenocarcinoma, diffuse large B cell lymphoma, and uveal melanoma	[14]
FAM50B	FAM50A	Unknown	The function of *FAM50B* is unclear, but it is frequently silenced in tumours [~4% of TCGA (The Cancer Genome Atlas) tumours] and is typically expressed in normal tissue	[35]
NXT2	NXT1	Unknown	*NXT2* encodes a nuclear export factor that is typically expressed in adult tissues but is frequently silenced in multiple paediatric tumours, including neuroblastoma. It has not yet been established whether the gene is also silenced in healthy paediatric tissue	[75]
UBB	UBC	Unknown	Silencing of the *UBB*, which encodes ubiquitin B, is common in ovarian and uterine tumours, although it is unclear whether this is a driver or passenger event.	[76]
VRK2	VRK1	Unknown	*VRK2*, which encodes a serine/threonine kinase, is silenced in adult and paediatric gliomas and in neuroblastoma	[77,78]

In Table 1 we highlight a set of paralog synthetic lethal interactions where there has been low-throughput experimental validation of the interaction and where there is a clear biomarker for the effect. Many additional putative synthetic lethal pairs have been identified using the screening approaches outlined below, but Table 1 focuses on those with low-throughput validation.

Experimental approaches to identify new synthetic lethal paralog pairs

Integrating genome-wide single-gene CRISPR screens in panels of cancer cell lines with molecular profiling

Genome-wide single-gene perturbation CRISPR screens report on the fitness effect of disrupting each protein-coding gene in a given cell line. Through large-scale systematic efforts, these screens have been performed across hundreds of cancer cell lines [25–27]. By combining the results of these loss-of-function screens with molecular profiling data of cancer cell lines, it is possible to determine genetic dependencies and the cellular contexts in which they operate. For a particular paralog pair, it is possible to identify cell lines which are characterised by loss or silencing of one member of the pair and to subsequently determine whether these cell lines exhibit an increased dependency on the other member (Figure 2A). Multiple analyses have systematically evaluated paralog pairs for synthetic lethality using such approaches [14,28–31]. Through these types of analyses, multiple synthetic lethal paralog pairs have been identified, including those which have been previously confirmed through low-throughput experimental validation such as *ARID1A*/*ARID1B* [32] as well as novel pairs such as *VPS26A*/*VPS26B* [14].

Before the development of CRISPR-Cas9 approaches, RNAi was the most widely used approach for identifying genetic dependencies in cancer cell lines. With robust on-target activity and lower off-target effects, CRISPR-Cas9 has now surpassed RNAi as the preferred approach for performing these loss-of-function genetic screens. However, analyses have been performed using RNAi-based single-gene perturbation screens to successfully identify a synthetic lethal interaction between the tumour-suppressor *SMARCA4* and its paralog *SMARCA2* [33]. In addition, results from RNAi-based and CRISPR-based single-gene perturbation screens have been successfully combined to identify a synthetic lethal interaction between *MAGOH* (whose deletion is frequently observed in human cancers owing to chromosome 1p loss) and its paralog *MAGOHB* [28].

Single-gene perturbation can also be performed to identify synthetic lethal interactions in isogenic cell lines; genome-wide single-gene perturbation CRISPR screens in *STAG2* wild-type cells and in purposely engineered isogenic *STAG2*-deficient cells demonstrated that *STAG1*, a paralog of *STAG2*, is selectively lethal in *STAG2*-deficient cells [34].

Performing combinatorial CRISPR screens focused on paralog pairs

An alternative approach to using single-gene perturbation screens for the identification of synthetic lethality is to use combinatorial or multiplexed CRISPR screens. Recent advances in CRISPR technology have facilitated multiplexed gene perturbation, whereby two genes can be disrupted simultaneously within the same cell, thus providing an opportunity to explore pairwise genetic interactions at a hitherto unprecedented scale and in diverse cellular contexts.

Pairwise genetic interactions can be thought of as deviations from the predicted phenotypic outcomes of simultaneous perturbation based on the effects of disrupting each gene individually; in the case of synthetic lethal interactions, simultaneous perturbation results in a lethal phenotype, whereas perturbation of each gene individually has a minimal effect on cellular viability. Multiplexed or combinatorial gene perturbation screens can test for such interactions across hundreds or thousands of gene pairs by disrupting both genes simultaneously as well as each of the pair individually in a pooled fashion. Synthetic lethal interactions can then be determined for those gene pairs where the observed fitness defect of simultaneous perturbation of a gene pair is significantly greater than that expected by disrupting each gene individually (Figure 2B). Several combinatorial gene perturbation systems have been developed in recent years (reviewed in Box 1).

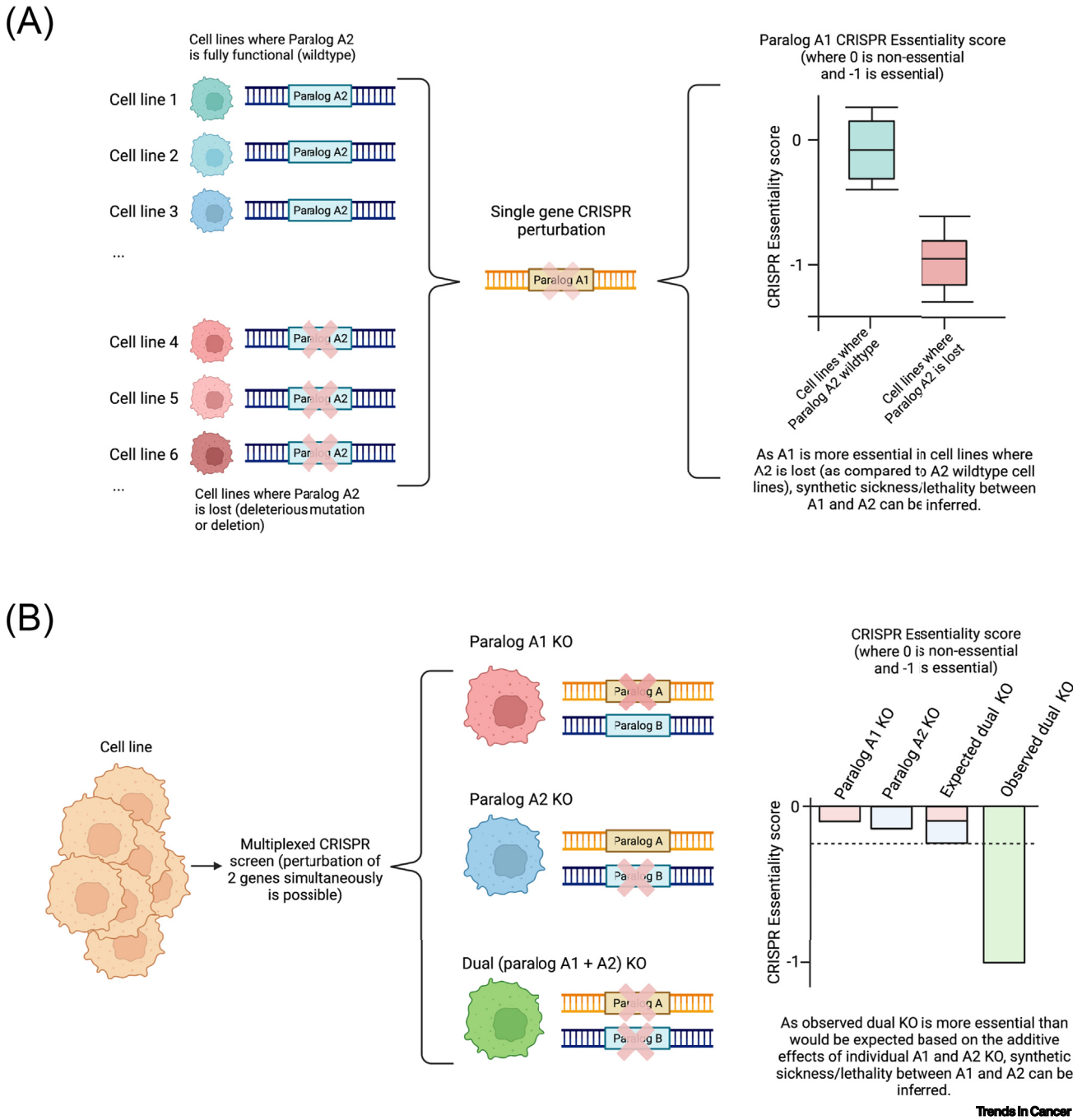

Figure 2. Experimental approaches to identify new synthetic lethal paralog pairs. (A) By combining the results of single-gene perturbation CRISPR screens across hundreds of cell lines with the molecular profiles of these cell lines, it is possible to determine the context in which genetic dependencies operate. For a particular paralog pair, *A1* and *A2*, it is possible to identify cell lines which are characterised by loss or silencing of *A2* and to subsequently determine the fitness effect of disrupting *A1* in these lines and compare this to the effect in cell lines with preserved *A2* function (wild-type). If the cell lines which have lost *A2* exhibit an increased dependency on *A1* compared to cell lines with wild-type *A2*, synthetic lethality/sickness can be inferred. (B) Multiplexed or combinatorial CRISPR screening makes dual perturbation possible in a given cell line. By designing the screen to include individual *A1* and *A2* perturbation as well as dual perturbation, the observed fitness effect of dual perturbation can be compared to the expected effect of disrupting *A1* and *A2* individually. If the observed fitness defect of dual perturbation is significantly greater than expected, synthetic lethality/sickness can be inferred. Abbreviation: KO, knockout. Figure created with BioRender.com.

Multiple independent combinatorial screening efforts focusing on pairs of paralogs have now been reported in diverse cell lines [30,35–38]. These screens have identified previously validated synthetic lethal paralog interactions, such as the relationship between *MAP2K1* (MEK1) and *MAP2K2* (MEK2) [36,38], as well as novel interactions such as between *FAM50A* and *FAM50B* [35] and between *CCNL1* and *CCNL2* [36].

Advantages of the two approaches

These two approaches to identify synthetic lethal interactions have different strengths. A significant advantage of identifying synthetic lethal interactions through the analysis of single-gene CRISPR screens in panels of molecularly characterised cell lines is that any interactions identified using this approach should be somewhat robust to the genetic heterogeneity observed in the cell line panel [39,40]. A disadvantage is that the approach can only be used to identify synthetic lethal interactions for genes that are recurrently lost/silenced in the cell line panel. This makes it challenging to identify synthetic lethal interactions for genes that are altered infrequently or that are only lost in specific cancer types. Furthermore, to gain the statistical power required for this approach, typically the results of single-gene CRISPR screens in cell lines from diverse cancer types are aggregated together, negating the ability to identify context-specific relationships. By contrast, combinatorial gene perturbation systems permit the interrogation of any gene pair in any cellular context. This makes it possible to identify synthetic lethal interactions involving rarely mutated genes and also to identify context-specific effects. An emerging challenge of this approach is in determining those pairs that are synthetic lethal in a particular cellular context rather than being unique to a single cell line.

Lessons learned from screening for paralog synthetic lethality

Some paralog pairs are more likely to be synthetic lethal than others

Computational analyses of paralogs identified as synthetic lethal either in single- or double-perturbation CRISPR screens have identified that some types of paralogs are more likely to be synthetic lethal than others. In general, pairs with higher protein sequence identity are more likely to be synthetic lethal [14,29,30,35], as are pairs from smaller gene families and pairs that arose through whole-genome duplication [14,29]. These mirror findings from budding yeast [17,18,41,42] and are general trends rather than rules – some synthetic lethal paralog pairs do have relatively low sequence identity and some come from larger families.

Although combinatorial screens have been used to test thousands of paralog pairs in a single experiment, they have not yet been scaled up to assess all possible paralog pairs in the human genome. Indeed, to date only 10–15% of paralog pairs have been experimentally tested for synthetic lethality in any cell line. Consequently, there is still a need to be selective in choosing pairs for experimental testing. That some paralog pairs are more likely to be synthetic lethal than others suggests that it may be possible to prioritise those gene pairs with the highest probability of being synthetic lethal for inclusion in screens. Recently, De Kegel *et al.* developed a machine learning approach to rank all paralog pairs according to their likelihood of being synthetic lethal, and demonstrated that this can be used to accurately predict the results of combinatorial CRISPR screens [14].

Many paralog pairs are synthetic lethal in a context-specific fashion

Previous work has demonstrated that many synthetic lethal interactions appear to be highly context-specific – a gene pair that is synthetic lethal in one cell line may not be synthetic lethal in another [39,40,43–45]. The extent of this context specificity varies across studies and is likely due to a combination of both biological and technical factors [39,44]. Technical factors include the quality and reproducibility of individual screens, which may result in both false positive and

false negative hits, and biological factors include both cancer type and genetic background [39,44]. Preliminary analyses suggest that synthetic lethal interactions between paralog pairs may be somewhat less context-specific than other synthetic lethal interactions [35,40]. For instance, Thompson *et al.* performed a combinatorial CRISPR screen in three cell lines using a library that contained both paralog and non-paralog pairs, but all of the hits common to all three cell lines were paralog pairs [35]. Nonetheless, although they may be less context-specific than other gene pairs, the majority of paralog synthetic lethal pairs reported in combinatorial screens still appear to be synthetic lethal only in specific cell lines [14,35–38]. In at least some cases, it seems that gene pairs may be synthetic lethal only in the presence of an additional genetic alteration. Such higher-order genetic interactions, also termed complex genetic interactions, have been studied extensively for paralog pairs in yeast [46]. A recent paralog-focused combinatorial screen performed in 11 cancer cell lines identified the paralog pair *DUSP4*/*DUSP6* as specifically synthetic lethal in the context of *NRAS* mutant melanoma [38]. Similarly, a combinatorial CRISPR screen performed in two isogenic cell lines identified *RALA*/*RALB* as synthetic lethal in the context of *KRAS* mutant lung cancer [47]. When there is a clear biomarker (e.g., *KRAS* mutation) for these conditional synthetic lethal interactions, they may also be exploited for the development of targeted therapeutics.

From screen hits to new drugs

The ultimate goal of many synthetic lethal screening efforts is the identification of new targets that can form the basis for drug development efforts. However, even in the case of targets that have been well validated using genetic approaches, several challenges must be overcome to develop these into clinically viable therapeutics. Some of these challenges are common to all drug development efforts, but targeting synthetic lethal paralog pairs also presents some relatively unique challenges owing to their structural similarity.

From target genes to domains

Typically, CRISPR- and RNAi-based screens identify genes as targets. However, many small-molecule inhibitors function by disrupting the activity of specific functional sites within a protein (e.g., kinase domains) rather than of whole proteins. Consequently, a first step for drug

Box 1. Combinatorial CRISPR screens permit dual gene perturbation

Several combinatorial gene perturbation systems have been developed in recent years that exploit CRISPR gene perturbation approaches using a variety of **Cas enzymes**. A commonly used system involves the action of *Streptococcus pyogenes* Cas9 (SpCas9) combined with coexpression of two structurally similar SpCas9 **single guide RNAs (sgRNAs)** driven by two different promoters (Figure IA) [35,36]. Although this approach has been used to successfully identify novel genetic interactions, it may be limited by issues arising from (i) recombination between homologous promoter regions and guide scaffold regions, or (ii) unequal guide expression as a result of each being driven by different promoters (promoter bias). Recent advances have made it possible to derive multiple Cas9 guides from a single transcript. An array of Cas9 guides separated by a tRNA spacer can be processed into multiple Cas9 guides by the endogenous tRNA-processing system of the cell (Figure IB) [65]. Therefore, multiple guides can be generated from a single transcript, thus avoiding the use of a second promoter and the risk of promoter bias. Another approach which seeks to minimise the issue of recombination involves orthologous Cas9 endonucleases, namely SpCas9 and *Staphylococcus aureus* Cas9 (SaCas9), and the coexpression of two structurally distinct guides: one SpCas9 guide and one SaCas9 guide expressed from two promoters (Figure IC) [66]. The Cas12a family of CRISPR nucleases have also been used in multiplexed screening systems and are particularly attractive because, unlike Cas9 nuclease, Cas12a can intrinsically process multiple guides from a single, compact transcript. One such approach uses an engineered Cas12a variant derived from *Acidaminococcus*, (en)AsCas12a, in conjunction with the expression of a multiplexed array of Cas12a guides from a single transcript (Figure ID) [67]. The authors of this approach demonstrated that a single construct containing three guides could effectively target three distinct genes, suggesting that (en)AsCas12a approaches could be employed to explore higher-order genetic interactions (e.g., between three members of a paralog family). Another approach uses both SpCas9 and *Lachnospiraceae* bacterium Cas12a (LbCas12a), again with the expression of a multiplexed array of guides encoded by a single transcript, including a single Cas9 guide and single or multiple Cas12a guides (Figure IE) [37]. The approaches illustrated in Figure IB–E rely on a single transcript and therefore only require one promoter, thus ensuring equal expression of the multiple component guides.

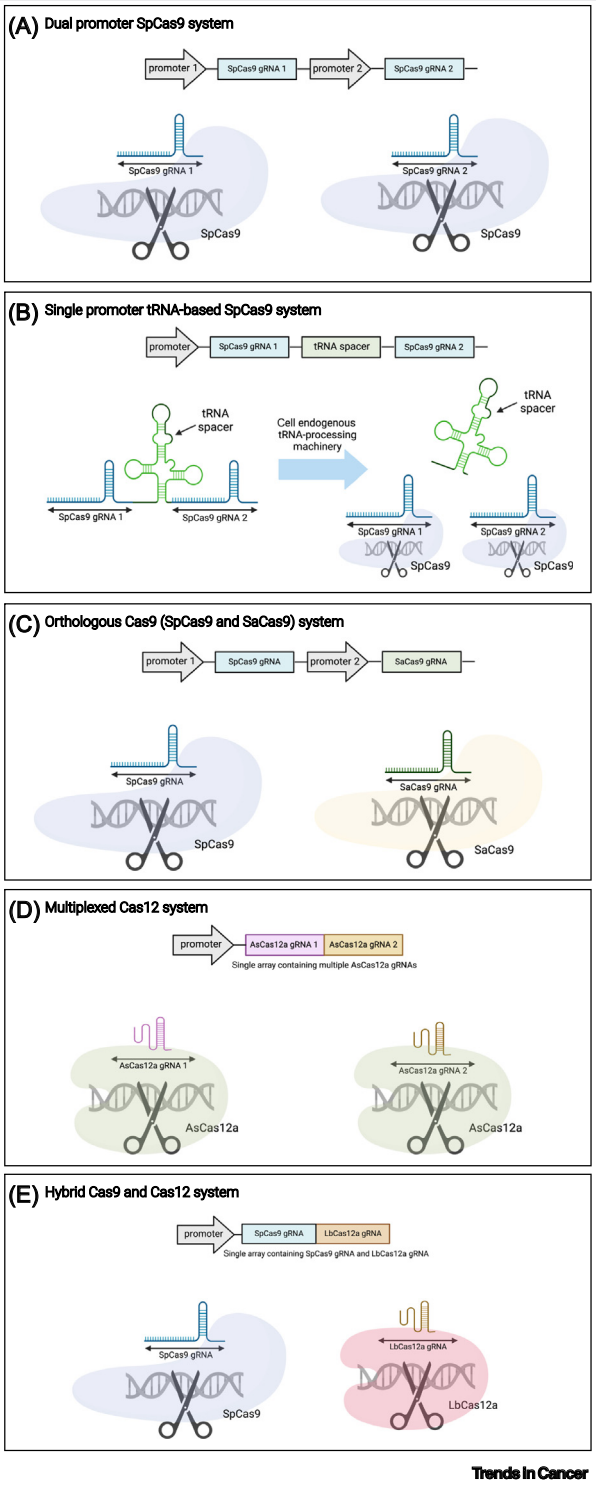

Figure I. Combinatorial CRISPR perturbation approaches to identify genetic interactions. Figure created with BioRender.com.

development may be the identification of the domains or regions that are responsible for a given synthetic lethal relationship. For example, multiple genetic screens identified *SMARCA4* and *SMARCA2* as having a synthetic lethal relationship [33,48,49]. However, both proteins contain a bromodomain and an ATPase domain, and consequently subsequent work was necessary to identify the domain responsible for the synthetic lethal effect [50]. This demonstrated that the SMARCA2 bromodomain could be inhibited in *SMARCA4* mutant cells without causing a significant fitness defect, but that a functioning ATPase domain was required for growth [50]. Consequently, the authors suggest that the development of small-molecule inhibitors to exploit this synthetic lethal interaction should focus on the ATPase domain rather than on the more readily druggable bromodomain. More recently, **CRISPR tiling screens** have been developed as a means to rapidly dissect the domains responsible for a given genetic defect [51–54].

A promising alternative to the use of small-molecule inhibitors to disrupt specific protein domains or functional sites is the use of **PROTACs** for target protein degradation [55]. In this case, because the entire protein is degraded, the druggability of a specific domain is of less concern. Indeed, PROTACs have been developed to target SMARCA2 and SMARCA4 with a view to exploiting their synthetic lethal relationship [56–58].

Paralog binding specificity as a challenge and an opportunity

It is typically assumed that pairs of paralog genes exhibit synthetic lethality because their encoded proteins share sufficient structural similarity to carry out each other's functions within the cell. However, this structural similarity may pose problems for drug development when the goal is to inhibit the function of one specific paralog.

The synthetic lethal interactions highlighted in Table 1 all demonstrate that loss-of-function alterations in one paralog gene (e.g., *ARID1A*) are associated with increased sensitivity to perturbation of a second gene (e.g., *ARID1B*). To exploit these findings for the development of new therapies, it would seem preferable to use drugs that specifically inhibit the function of the target paralog, for example, by using an ARID1B inhibitor in *ARID1A* mutant cancers (Figure 3A). In this way cell death is specific to cancerous cells because normal cells have a functioning copy of *ARID1A*. However, the structural similarity of these two synthetic lethal paralogs may mean that most inhibitors developed against one protein will have off-target effects against the other. If a drug inhibits the function of both paralogs equally, for example, by inhibiting the function of both *ARID1A* and *ARID1B*, it is unclear whether any tumour selectivity will be retained because the dual inhibitor may also be toxic for normal cells (Figure 3B). In some cases the loss of one paralog may still make tumour cells more sensitive to inhibitors that target both members of a paralog pair, thus continuing to provide a therapeutic window. However, it is unclear how often this will be the case.

Many existing small-molecule drugs inhibit the function of multiple paralogs from the same family [59,60]. For example, palbociclib, a small-molecule inhibitor used to treat select breast cancer patients, targets both CDK4 and its paralog CDK6 [61]. Similarly, lapatinib, a small molecule also used in breast cancer, inhibits both ERBB2 *(HER2)* and its paralog EGFR (*ERBB1)* [62]. In some cases more than two paralogs from the same family may be targeted by an inhibitor – for example, the AKT inhibitor ipatasertib inhibits AKT1, AKT2, and AKT3 [63]. This targeting of multiple paralogs by inhibitors can be by design or may simply reflect challenges in the development of specific inhibitors. Whether by design or limitation, in some cases multi-targeting drugs likely function by exploiting synthetic lethal relationships between paralogs. For instance, the MEK inhibitor trametinib is currently used as a therapy for *BRAF* mutant melanomas [64]. MEK inhibitors function by inhibiting the function of the paralogs MEK1 and MEK2. However, in CRISPR screens, disrupting either MEK1 (*MAP2K1*) or MEK2 (*MAP2K2*) alone is rarely lethal, even in *BRAF* mutant

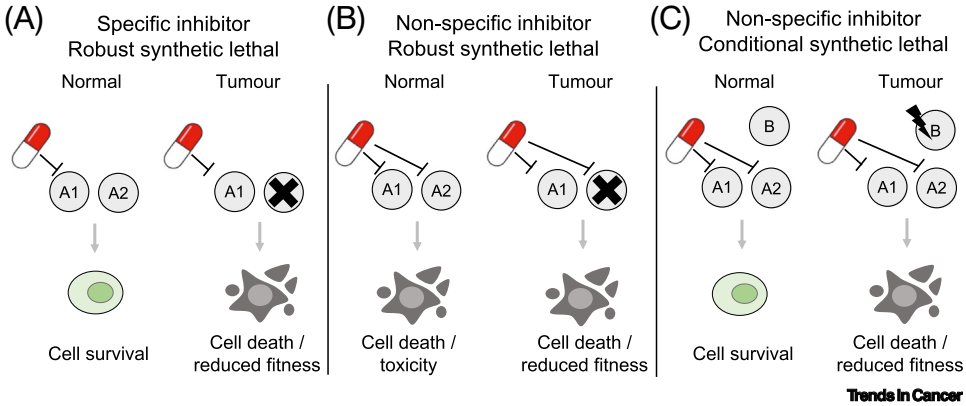

Figure 3. Exploiting synthetic lethality between paralogs using inhibitors. (A) Genes *A1* and *A2* represent a robust synthetic lethal pair – a pair of paralogs that are synthetic lethal across many contexts. A selective inhibitor of protein A1 could be used as a targeted therapy for patients whose tumours lack a functional copy of *A2*, resulting in the selective killing of tumour cells. (B) *A1* and *A2* again represent a robust synthetic lethal pair. A non-selective inhibitor, that inhibits the function of both protein A1 and protein A2, may cause cell death regardless of the status of A2, and will therefore induce toxicity in normal cells. (C) *A1* and *A2* again represent a pair of paralogs that are only synthetic lethal in the presence of a mutation in gene *B*. A non-selective inhibitor, that inhibits the function of both A1 and A2, could be used as a targeted therapy for patients whose tumours have a mutation in gene *B*, resulting in selective killing of tumour cells.

Outstanding questions

How context-specific are most paralog synthetic lethalities? Can we identify/predict in which contexts a paralog synthetic lethal might be useful (e.g., MEK1 and MEK2 in *BRAF* mutant cancers)?

Many paralogs belong to larger gene families, including genes of interest in cancer (e.g., *CDH1* and *AKT*). How do we deal with these higher-order effects?

Can the synthetic lethal paralog pairs identified to date be translated into new drugs?

Can we develop specific inhibitors for most paralog pairs or only for a minority?

melanoma models [25]. However, combinatorial CRISPR screens have identified that combined disruption of *MAP2K1* and *MAP2K2* is synthetic lethal [36,38]. Consequently, the use of MEK inhibitors in melanoma can be considered as an example of conditional synthetic lethality – in the context of *BRAF* mutant melanoma, the combined inhibition of MEK1 and MEK2 is lethal. From this perspective, the challenges in developing specific inhibitors for individual paralogs may also be considered as an opportunity for exploiting conditional synthetic lethalities (Figure 3C). Currently, relatively few conditional synthetic lethalities associated with specific biomarkers (such as the association between *RALA*/*RALB* and *KRAS*) have been identified. However, many more seem likely to be identified as combinatorial screens are performed in larger panels of cell lines or in isogenic models of different mutations [38,47].

Concluding remarks

Synthetic lethality represents a promising paradigm by which we might target a variety of genetic lesions in cancer. Pairs of paralogous genes are especially enriched in synthetic lethality, and focusing on these gene pairs is therefore likely to result in the identification of new therapeutic targets. We have reviewed experimental and computational approaches for the identification of synthetic lethality between paralog pairs and have highlighted some of the promising synthetic lethal pairs that have been identified to date. Although some of these interactions have been validated in preclinical models by multiple groups, none have yet been developed into new therapeutics. A major challenge in translating these discoveries into new therapeutics will be the identification of inhibitors that can inhibit specific paralogs (for robust synthetic lethal interactions) or pairs of paralogs (for context-specific synthetic lethal interactions) (see Outstanding questions).

Multiple combinatorial CRISPR screens for paralog pairs have now been published, but the approach is still relatively new and the vast majority of paralog pairs have not yet been tested for synthetic lethality in any context. Consequently, a large space of paralog pairs remains to be explored. Furthermore, although there are hundreds of cancer cell lines in common use, only a small number of cell lines have been screened in any combinatorial CRISPR screen to date. A

major challenge will be to develop scalable approaches for screening paralog pairs in panels of cancer cell lines or to use computational approaches to identify the most promising pairs to screen in a given context.

Acknowledgments

We thank Barbara De Kegel for useful conversations and Anjan Venkatesh for careful reading of the manuscript. C.J.R. is funded by an Irish Research Council 2017/2018 Laureate Award and by SFI under grant number 20/FFP-P/8641. N.K. was funded by SFI through the SFI Centre for Research Training in Genomics Data Science under grant 18/CRT/6214 and in part by the EU Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant H2020-MSCA-COFUND-2019-945385. D.J.A. was funded by OpenTargets and I.M by a Wellcome Trust Clinical PhD studentship.

Declaration of interests

The authors declare no conflicts of interest.

References

1. Huang, A. *et al.* (2020) Synthetic lethality as an engine for cancer drug target discovery. *Nat. Rev. Drug Discov.* 19, 23–38
2. O'Neil, N.J. *et al.* (2017) Synthetic lethality and cancer. *Nat. Rev. Genet.* 18, 613–623
3. Farmer, H. *et al.* (2005) Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* 434, 917–921
4. Bryant, H.E. *et al.* (2005) Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* 434, 913–917
5. Lord, C.J. and Ashworth, A. (2017) PARP inhibitors: synthetic lethality in the clinic. *Science* 355, 1152–1158
6. Mullard, A. (2017) Synthetic lethality screens point the way to new cancer drug targets. *Nat. Rev. Drug Discov.* 16, 589–591
7. Muller, F.L. *et al.* (2015) Collateral lethality: a new therapeutic strategy in oncology. *Trends Cancer Res.* 1, 161–173
8. Kuzmin, E. *et al.* (2022) Retention of duplicated genes in evolution. *Trends Genet.* 38, 59–72
9. Kaessmann, H. *et al.* (2009) RNA-based gene duplication: mechanistic and evolutionary insights. *Nat. Rev. Genet.* 10, 19–31
10. Carelli, F.N. *et al.* (2016) The life history of retrocopies illuminates the evolution of new mammalian genes. *Genome Res.* 26, 301–314
11. Dehal, P. and Boore, J.L. (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3, e314
12. Vilella, A.J. *et al.* (2009) EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res.* 19, 327–335
13. Yates, A.D. *et al.* (2020) Ensembl 2020. *Nucleic Acids Res.* 48, D682–D688
14. De Kegel, B. *et al.* (2021) Comprehensive prediction of robust synthetic lethality between paralog pairs in cancer cell lines. *Cell Syst.* 12, 1144–1159
15. Dean, E.J. *et al.* (2008) Pervasive and persistent redundancy among duplicated genes in yeast. *PLoS Genet.* 4, e1000113
16. Costanzo, M. *et al.* (2016) A global genetic interaction network maps a wiring diagram of cellular function. *Science* 353, aaf1420
17. Musso, G. *et al.* (2008) The extensive and condition-dependent nature of epistasis among whole-genome duplicates in yeast. *Genome Res.* 18, 1092–1099
18. DeLuna, A. *et al.* (2008) Exposing the fitness contribution of duplicated genes. *Nat. Genet.* 40, 676–681
19. Tischler, J. *et al.* (2006) Combinatorial RNA interference in *Caenorhabditis elegans* reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. *Genome Biol.* 7, R69
20. Fisher, R. *et al.* (2013) Cancer heterogeneity: implications for targeted therapeutics. *Br. J. Cancer* 108, 479–485
21. Sondka, Z. *et al.* (2018) The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* 18, 696–705
22. Frei, E., 3rd (1993) Gene deletion: a new target for cancer chemotherapy. *Lancet* 342, 662–664
23. De Kegel, B. and Ryan, C.J. (2022) Paralog dispensability shapes homozygous deletion patterns in tumor genomes. *bioRxiv* Published online October 31, 2022. https://doi.org/10.1101/2022.06.20.496722
24. Muller, F.L. *et al.* (2012) Passenger deletions generate therapeutic vulnerabilities in cancer. *Nature* 488, 337–342
25. Pacini, C. *et al.* (2021) Integrated cross-study datasets of genetic dependencies in cancer. *Nat. Commun.* 12, 1661
26. Meyers, R.M. *et al.* (2017) Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* 49, 1779–1784
27. Behan, F.M. *et al.* (2019) Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* 568, 511–516
28. Viswanathan, S.R. *et al.* (2018) Genome-scale analysis identifies paralog lethality as a vulnerability of chromosome 1p loss in cancer. *Nat. Genet.* 50, 937–943
29. De Kegel, B. and Ryan, C.J. (2019) Paralog buffering contributes to the variable essentiality of genes in cancer cell lines. *PLoS Genet.* 15, e1008466
30. Dede, M. *et al.* (2020) Multiplex enCas12a screens detect functional buffering among paralogs otherwise masked in monogenic Cas9 knockout screens. *Genome Biol.* 21, 262
31. Köferle, A. *et al.* (2022) Interrogation of cancer gene dependencies reveals paralog interactions of autosome and sex chromosome-encoded genes. *Cell Rep.* 39, 110636
32. Helming, K.C. *et al.* (2014) ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat. Med.* 20, 251–254
33. Hoffman, G.R. *et al.* (2014) Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proc. Natl. Acad. Sci. U. S. A.* 111, 3128–3133
34. van der Lelij, P. *et al.* (2020) STAG1 vulnerabilities for exploiting cohesin synthetic lethality in STAG2-deficient cancers. *Life Sci. Alliance* 3, e202000725
35. Thompson, N.A. *et al.* (2021) Combinatorial CRISPR screen identifies fitness effects of gene paralogues. *Nat. Commun.* 12, 1302
36. Parrish, P.C.R. *et al.* (2021) Discovery of synthetic lethal and tumor suppressor paralog pairs in the human genome. *Cell Rep.* 36, 109597
37. Gonatopoulos-Pournatzis, T. *et al.* (2020) Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9-Cas12a platform. *Nat. Biotechnol.* 38, 638–648
38. Ito, T. *et al.* (2021) Paralog knockout profiling identifies DUSP4 and DUSP6 as a digenic dependence in MAPK pathway-driven cancers. *Nat. Genet.* 53, 1664–1672
39. Ryan, C.J. *et al.* (2018) Synthetic lethality and cancer – penetrance as the major barrier. *Trends Cancer Res.* 4, 671–683
40. Lord, C.J. *et al.* (2020) Integrative analysis of large-scale loss-of-function screens identifies robust cancer-associated genetic interactions. *eLife* 9, e58925
41. Guan, Y. *et al.* (2007) Functional analysis of gene duplications in *Saccharomyces cerevisiae*. *Genetics* 175, 933–943

42. VanderSluis, B. *et al.* (2010) Genetic interactions reveal the evolutionary trajectories of duplicate genes. *Mol. Syst. Biol.* 6, 429
43. Downward, J. (2015) RAS synthetic lethal screens revisited: still seeking the elusive prize? *Clin. Cancer Res.* 21, 1802–1809
44. Henkel, L. *et al.* (2019) Context-dependent genetic interactions in cancer. *Curr. Opin. Genet. Dev.* 54, 73–82
45. Ku, A.A. *et al.* (2020) Integration of multiple biological contexts reveals principles of synthetic lethality that affect reproducibility. *Nat. Commun.* 11, 2375
46. Kuzmin, E. *et al.* (2020) Exploring whole-genome duplicate gene retention with complex genetic interaction analysis. *Science* 368, eaaz5667
47. Kelly, M.R. *et al.* (2020) Combined proteomic and genetic interaction mapping reveals new RAS effector pathways and susceptibilities. *Cancer Discov.* 10, 1950–1967
48. Wilson, B.G. *et al.* (2014) Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. *Mol. Cell. Biol.* 34, 1136–1144
49. Oike, T. *et al.* (2013) A synthetic lethality-based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1. *Cancer Res.* 73, 5508–5518
50. Vangamudi, B. *et al.* (2015) The SMARCA2/4 ATPase domain surpasses the bromodomain as a drug target in SWI/SNF-mutant cancers: insights from cDNA rescue and PFI-3 inhibitor studies. *Cancer Res.* 75, 3865–3878
51. Munoz, D.M. *et al.* (2016) CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. *Cancer Discov.* 6, 900–913
52. Shi, J. *et al.* (2015) Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* 33, 661–667
53. He, W. *et al.* (2019) De novo identification of essential protein domains from CRISPR-Cas9 tiling-sgRNA knockout screens. *Nat. Commun.* 10, 4541
54. Bhogal, B. *et al.* (2022) The methyltransferase domain of DNMT1 is an essential domain in acute myeloid leukemia independent of DNMT3A mutation. *Commun. Biol.* 5, 1174
55. Lai, A.C. and Crews, C.M. (2017) Induced protein degradation: an emerging drug discovery paradigm. *Nat. Rev. Drug Discov.* 16, 101–114
56. Xiao, L. *et al.* (2021) Targeting SWI/SNF ATPases in enhancer-addicted prostate cancer. *Nature* 601, 434–439
57. Farnaby, W. *et al.* (2019) BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. *Nat. Chem. Biol.* 15, 672–680
58. Kofink, C. *et al.* (2022) A selective and orally bioavailable VHL-recruiting PROTAC achieves SMARCA2 degradation in vivo. *Nat. Commun.* 13, 5969
59. Kruger, F.A. and Overington, J.P. (2012) Global analysis of small molecule binding to related protein targets. *PLoS Comput. Biol.* 8, e1002333
60. Ochoa, D. *et al.* (2021) Open Targets Platform: supporting systematic drug-target identification and prioritisation. *Nucleic Acids Res.* 49, D1302–D1310
61. Sherr, C.J. *et al.* (2016) Targeting CDK4 and CDK6: from discovery to therapy. *Cancer Discov.* 6, 353–367
62. Medina, P.J. and Goodin, S. (2008) Lapatinib: a dual inhibitor of human epidermal growth factor receptor tyrosine kinases. *Clin. Ther.* 30, 1426–1447
63. Lin, J. *et al.* (2013) Targeting activated Akt with GDC-0068, a novel selective Akt inhibitor that is efficacious in multiple tumor models. *Clin. Cancer Res.* 19, 1760–1772
64. Flaherty, K.T. *et al.* (2012) Improved survival with MEK inhibition in BRAF-mutated melanoma. *N. Engl. J. Med.* 367, 107–114
65. Zhao, Y. *et al.* (2019) A one-step tRNA-CRISPR system for genome-wide genetic interaction mapping in mammalian cells. *Sci. Rep.* 9, 14499
66. Najm, F.J. *et al.* (2017) Orthologous CRISPR–Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.* 36, 179
67. DeWeirdt, P.C. *et al.* (2021) Optimization of AsCas12a for combinatorial genetic screens in human cells. *Nat. Biotechnol.* 39, 94–104
68. Ogiwara, H. *et al.* (2016) Targeting p300 addiction in CBP-deficient cancers causes synthetic lethality by apoptotic cell death due to abrogation of MYC expression. *Cancer Discov.* 6, 430–445
69. van der Lelij, P. *et al.* (2017) Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. *eLife* 6, e26980
70. Benedetti, L. *et al.* (2017) Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1. *Oncotarget* 8, 37619–37632
71. Dey, P. *et al.* (2017) Genomic deletion of malic enzyme 2 confers collateral lethality in pancreatic cancer. *Nature* 542, 119–123
72. Yoshihama, Y. *et al.* (2022) Potent and selective PTDSS1 inhibitors induce collateral lethality in cancers with PTDSS2 deletion. *Cancer Res.* 82, 4031–4043
73. Szymańska, E. *et al.* (2020) Synthetic lethality between VPS4A and VPS4B triggers an inflammatory response in colorectal cancer. *EMBO Mol. Med.* 12, e10812
74. Neggers, J.E. *et al.* (2020) Synthetic lethal interaction between the ESCRT paralog enzymes VPS4A and VPS4B in cancers harboring loss of chromosome 18q or 16q. *Cell Rep.* 33, 108493
75. Malone, C.F. *et al.* (2021) Selective modulation of a pan-essential protein as a therapeutic strategy in cancer. *Cancer Discov.* 11, 2282–2299
76. Tsherniak, A. *et al.* (2017) Defining a cancer dependency map. *Cell* 170, 564–576
77. So, J. *et al.* (2022) VRK1 as a synthetic lethal target in VRK2 promoter-methylated cancers of the nervous system. *JCI Insight* 7, e158755
78. Shields, J.A. *et al.* (2022) VRK1 is a synthetic-lethal target in VRK2-deficient glioblastoma. *Cancer Res.* 82, 4044–4057