Published in final edited form as:

J Mol Biol. 2017 September 01; 429(18): 2780–2789. doi:10.1016/j.jmb.2017.07.020.

Trap^{Seq}: An RNA Sequencing-based pipeline for the identification of genetrapp insertions in mammalian cells

Cristina Mayor-Ruiz¹, Orlando Dominguez², and Oscar Fernandez-Capetillo^{1,3,*}

¹Genomic Instability Group, Spanish National Cancer Research Centre, 28029 Madrid, Spain

²Genomics Unit, Spanish National Cancer Research Centre, 28029 Madrid, Spain

³Science for Life Laboratory, Division of Genome Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institute, 171 21 Stockholm, Sweden

Abstract

The development of haploid mammalian cell lines, coupled to next generation sequencing, has recently facilitated forward genetic screenings in mammals. For mutagenesis, retrovirus- or transposon-based genetraps are frequently used. Current methods to map genetrapp insertions are based on inverse- or splinkerette-PCR, which despite their efficacy are prone to artifacts and do not provide information about expression of the targeted gene. Here, we describe a new RNA-sequencing based method (Trap^{Seq}) to map genetrapp insertions. By recognizing chimeric mRNAs containing genetrapp sequences spliced to an exon, our method identifies insertions that lead to productive trapping. When applied to individual mutant clones our method provides a fast and cost-effective way that not only identifies the insertion site but also reveals its impact on the expression of the trapped gene. As proof-of-principle, we conducted two independent screenings for resistance against 6-thioguanine and an ATR inhibitor, which identified mutations known to provide resistance to these reagents, and revealed ECT2 as a novel determinant for the sensitivity to ATR inhibition.

Keywords

Genetrapp; piggyBac transposon; RNAseq; haploid mammalian cells; genetic screenings

Introduction

Haploid organisms such as yeast or social insects have historically been excellent platforms for forward genetic screens [1, 2]. In contrast, their diploid genome has precluded similar

*To whom correspondence should be addressed. Tel: +34 917328000; Fax: +34 917328022; ofernandez@cniio.es.

Accession Numbers

RNAseq data associated to this study is available at the Gene Expression Omnibus database under the accession number GSE99483.

Author Contributions

C.M.-R. performed most of the experiments and wrote the manuscript; O.D. contributed to RNAseq and bioinformatics analyses; O.F. supervised the study and wrote the MS.

Competing Interests

The authors declare to have no conflict of interest related to this study.

approaches in mammalian cells. The isolation of a human haploid cell line derived from a leukemia patient and the implementation of methods to generate mammalian haploid embryonic stem cell lines has recently changed the situation [3–7]. Together with CRISPR/Cas9-mediated screens, genetrap mutagenesis in haploid mammalian cells has led to the emergence of successful genomewide screenings in mammals, with good concordance between both methods [8, 9]. Retroviral- or transposon-based genetrap are the most common method for the mutagenesis of mammalian genomes since, in addition to showing tropism to integrate near coding sequences, they tag the insertion facilitating the identification of the mutant locus [3, 6, 10, 11]. Such identification is currently done through a variety of PCR-based methods, including inverse- [12] and splinkerette-PCR [13]. In addition to the identification of a particular mutation, the availability of massive parallel sequencing pipelines has also enabled to perform screenings in populations of mutant cells, without the need to isolate individual clones [10]. However, and despite the efficacy of these methods, multi-template PCR-based strategies can suffer from artifacts and amplification biases and do not provide information about the expression of the targeted locus [14].

We report here on the development of an RNA sequencing (RNAseq)-based method named Trap^{Seq}, which provides a fast, direct, and cost-effective pipeline for the identification of genetrap insertions. By identifying chimeric mRNAs containing genetrap sequences spliced to an endogenous exon, our method reveals insertions that lead to actual trapping of the target gene. Moreover, when applied to individual mutant clones, our method reveals the impact of the genetrap on the expression of the mutated gene. As validation of our method, we show its efficacy in identifying mutations in two independent forward genetic screenings using the near-haploid human HAP1 cell line [4], together with an optimized piggyBac (pB) transposon system [15–19].

Results

To generate a library of mutants, 100 million human haploid HAP1 cells were cotransfected with a pB gene-trapping transposon vector [15] and a plasmid expressing an optimized version of the pB transposase [20] (Fig. 1a). We selected pB for mutagenesis since the mutation can be reverted by transposase without leaving footprints [16, 21], and because it presents a broader pattern of integrations than retroviral vectors [18, 19]. Given that the trapping leads to the splicing of an endogenous exon to a puromycin resistance cassette, mutant cells carrying productive trapping events were selected with puromycin (Fig. 1b). To determine the efficiency of our mutagenesis process, we first performed a screening against 6-thio-guanine (6TG). 6TG is not toxic *per se*, but is converted into a cytotoxic antimetabolite within cells by the action of the hypoxanthine phosphoribosyltransferase 1 (HPRT1). The mutant library of HAP1 cells was exposed to 20μM 6TG for two weeks, a dose and time at which no wild type cell remained alive. At this point, several individual clones were picked and the rest of the resistant cells were pooled together and saved for further analyses.

In pB-based vectors, mutagenesis is achieved either through the direct insertion of the genetrap into a coding exon or more frequently through its integration in an intron, followed by the splicing of the splice acceptor En2 present in the genetrap to an exon of the targeted

gene. This leads to the expression of a truncated gene product, which is fused to a transcriptional terminating sequence together with a reporter, in our case puromycin (Fig. 1c). We reasoned that by performing RNAseq we could identify chimeric mRNAs carrying the En2 sequence together with sequences from the endogenous exon that was spliced to the genetrap. In addition, we used paired-end RNAseq since this method provides two independent ~50bp reads separated by distances ranging from 200bp to 500 bp, hence increasing the chances of detecting chimeric products (Fig. 1c). For the analysis, we multiplexed 12 samples from 6TG resistant cells per one lane of paired-end RNAseq. On an Illumina HiSeq2000, this yielded between 8.5 and 12.5 million filtered short read pairs per sample, which were then used for the identification of the chimeric mRNAs through an ad-hoc bioinformatic pipeline that was developed for these analyses (Fig. 1d). The key step of the strategy is based on the incorporation of pB sequences to Bowtie indexes along the hg19 reference sequence of the human genome (Fig. 1d). The analysis provides two outputs: the location of the exon that was spliced to the genetrap, and the expression of the trapped gene (Fig. 1d). To facilitate the interpretation of the outputs, tracks for the genome browser at University of California Santa Cruz (UCSC) are also generated which provide a visual and simple overview of the results.

When performing Trap^{Seq} on the pooled population of 6TG-resistant cells, 73% of the identified chimeric reads mapped to *HPRT1* (Fig. 2a; also Additional file 1: Table S1), with trapping events being detected along the whole gene and overall HPRT1 mRNA levels being significantly reduced in the resistant pool (Fig. 2b; also Additional file 1: Table S2). The method also worked efficiently on individual 6TG-resistant clones, revealing both the intron where the insertion occurred as well as its effect on the expression of the trapped gene and on the rest of the transcriptome (Fig. 2c; also Additional file 1: Table S3). To strengthen the validation of the method, we generated a new pB-mutagenized library in HAP1 cells and repeated the 6TG screen. In this case, 86% of the chimeric reads from the 6TG-resistant pool mapped to *HPRT1* (Additional file 1: Figure S1A, Table S1). To confirm the insertions detected by Trap^{Seq}, we designed primers to specifically amplify the predicted chimeric mRNAs. RT-PCR with these primers only amplified the expected chimeric products on RNA from the 6TG resistant pools (Fig. 2d). As further proof, equal amounts of RNA from three individual clones isolated from the mutant library were mixed and analyzed by our method. 24% of the chimeric reads in this analysis mapped to *PSMD1*, 32% to *MSH6* and 44% to *PTGSE3* (Additional file 1: Figure S2A, Table S4). To compare our method with the one typically used, the insertion of each clone was also identified by the analysis of their genomic DNA through splinkerette-PCR and Sanger sequencing, which identified the insertions in the same three genes (Additional file 1: Figure S2B). Note, however, that as mentioned DNA-based methods do not provide information on whether the insertion affects the expression of the trapped gene. Interestingly, our analyses also revealed the frequent usage of a previously described cryptic splice acceptor sequence 74-bp downstream of the pB 5' terminal repeat, which contributes to the trapping in addition to the canonical splice acceptor En2 [18] (Additional file 1: Figure S3).

Next, we conducted an additional screening in HAP1 cells to identify mutations that can confer resistance to inhibitors of the ATR kinase, the main responder to a type of genomic damage known as replication stress (RS) [22, 23]. Inhibitors of the ATR kinase have shown

antitumoral properties *in vitro* and *in vivo* and are currently on clinical trials [24–31]. Using a genomewide CRISPR/Cas9-based screening, we recently reported that low levels of the CDC25A phosphatase render mouse embryonic stem cells (mES) resistant to ATR inhibitors [32]. However, this approach was conducted in isolated ATR inhibitor-resistant clones, and to what extent a global analysis would also identify CDC25A as a determinant of the sensitivity to these agents in human cells was unclear. To this end, and using an ATR inhibitor previously described by our group (ATRi) [31], we conducted a new screening. A library of pB-mutagenized HAP1 cells was exposed to ATRi for 9 days at 0.3 μ M, a dose and time at which no wild type cells survive. RNA was isolated from the pool of ATRi-resistant cells and subjected to Trap^{Seq}. The analysis revealed 24 different genetrapp insertions in ATRi-resistant cells, one of them in *CDC25A* (Fig. 3a, b; Additional file 1: Table S5). Noteworthy, *CDC25A* was the only common hit between our previous CRISPR/Cas9-based ATRi-resistant screen in mES [32] and the present one in HAP1 cells. As in the case of mES, no insertions were found on *CDC25B* or *CDC25C* phosphatases, underscoring a very particular role of CDC25A in the ATR response.

To confirm whether CDC25A deficiency in HAP1 cells increases their resistance against ATRi, we sought to mutate the gene coding for this phosphatase through CRISPR/Cas9. To this end, and to increase the efficiency of CRISPR/Cas9-induced deletions, we first generated a doxycycline (dox)-inducible Cas9 line in HAP1 cells (HAP1^{Cas9}) by co-infecting with lentiviruses expressing Cas9, rtTA and GFP (Additional file 1: Figure S4A). To determine the efficiency of CRISPR/Cas9 editing in this cell line, we infected HAP1^{Cas9} cells with lentiviruses expressing a *GFP*-targeting single guide RNA (sgRNA). 7 days of treatment with dox led to a significant reduction in GFP levels, validating the efficiency of this approach (Additional file 1: Figure S4B). Next, we infected HAP1^{Cas9} cells with lentiviruses expressing sgRNAs against *CDC25A*. Whereas, in contrast to mES cells [32], nullzygous mutants of *CDC25A* could not be obtained in HAP1 cells, we were able to isolate several *CDC25A* heterozygous clones that had spontaneously become diploid during the gene editing process (Fig. 3c). Similar to our observation in mES cells, reduced levels of CDC25A in HAP1 cells led to increased resistance to ATR inhibitors (Fig. 3d), which was associated with a lower amount of ATRi-induced chromosomal breakage as measured by evaluating the phosphorylation of RPA and histone H2AX by Western Blotting (Fig. 3e).

Finally, we evaluated the response to ATR inhibition in all the hits that were identified in the ATRi screening. To this end, U2OS cells were transfected with endoribonuclease-prepared short interfering RNAs (esiRNAs) targeting each of the 24 genes, and the sensitivity to ATRi was evaluated with a viability assay performed 36 hrs after exposure to the drug (Fig. 4a). Of note, we chose U2OS since this cell line has been reported as specially sensitive to ATR inhibition [28]. Once again *CDC25A* depletion increased the resistance to ATRi in this experiment. From all genes tested, downregulation of Epithelial Cell Transforming 2 (*ECT2*) provided the highest resistance in this assay (Fig. 4a,b). The genetrapp insertion in *ECT2* present in the screening led to a truncated protein product (Fig. 4c). Finally, to confirm the role of ECT2 in the sensitivity to ATRi U2OS cells were transfected with *ECT2*-targeting siRNAs, which efficiently depleted its expression (Fig. 4d). Western blot analysis of RPA and H2AX phosphorylation revealed that *ECT2* knockdown reduced the levels of DNA

damage induced by ATRi (Fig. 4e). Furthermore, ECT2 depletion significantly increased the resistance of U2OS cells to ATRi (Fig. 4f).

Discussion

We here present Trap^{Seq}, a new method to map genetrap insertions by paired-end RNA sequencing that should facilitate forward genetic screenings in mammalian cells. As proof of principle, we show the capacity of Trap^{Seq} to identify mutations that render human cells resistant to 6TG or ATR inhibitors. These screenings found mutations already known to provide resistance to the compounds, and also led to the identification of ECT2 as a new determinant of the sensitivity to ATR inhibition in mammalian cells. While ECT2 has been mostly studied due to its role in cytokinesis as a Guanine nucleotide exchange factor (GEF) [33], a recent report revealed that ECT2 formed DNA damage induced nuclear foci, and that reduced levels of ECT2 increased the resistance to Topoisomerase II inhibition [34]. In fact, and besides its GEF domain, ECT2 presents several BRCT and XRCC1 domains, which are linked to the DNA damage response (Fig. 4b). The present study supports a role for ECT2 in DNA repair and checkpoint responses, specifically in the response to replication stress. To what extent variations in ECT2 levels affect the response to DNA damaging chemotherapies, including ATR inhibitors, should be further investigated.

In what regards to the methodology, one limitation of Trap^{Seq} when compared to DNA-based mapping methods is that the insertions are not mapped to the nucleotide level but rather to the intron where they landed. As a consequence, TrapSeq could miss insertions where the transposition perturbs the trapped gene by other events than splicing (e.g. direct insertion into an exon or by altering a promoter). However, and in contrast to DNA sequencing methods where each mutational event is considered individually regardless of whether the mutation has an impact on the trapped gene, one of the strengths of Trap^{Seq} is that the analysis is already filtered on mutations that (a) are expressed on the studied cells and (b) led to a productive trapping event. Hence, while Trap^{Seq} might fail to provide a panoramic view of all insertions, it provides a fast and efficient alternative to identify bona-fide hits, as exemplified in the 3 proof-of-principle screenings provided in this study. Moreover, when applied to individual clones, Trap^{Seq} also provides information on the effect of the genetrap on the expression of the trapped gene as well as how the mutation affects the overall transcriptome. Of note, the recent improvements in single-cell transcriptomics [35] have led to the implementation of RNAseq based methods such as Perturb-seq [36, 37] or CRISPR-seq [38] for genomewidescreenings based on CRISPR-Cas9. In principle, since TrapSeq is also based on RNAseq, it could also be coupled to such technologies in order to provide a global view of how each genetraping event impacts on overall transcription. Finally, we want to note that our method is also markedly cost-effective. So far we have multiplexed up to 12 screening samples per one paired-end RNAseq lane of 150M reads, and it is probable that this can be taken beyond this level. Collectively, all of the above validate our Trap^{Seq} method as a fast and cost-effective alternative for the mapping of genetrap insertions in mammalian cells.

Materials and Methods

Generation of pB-mutagenized HAP1 libraries

HAP1 cells (kind gift of T. Brummelkamp) were maintained in IMDM (Invitrogen) supplemented with 15% FBS and 1% penicillin-streptomycin. In order to ensure haploidy, cell sorting for DNA content was performed after staining with $10 \mu\text{g ml}^{-1}$ Hoechst 33342 (Invitrogen) on an Influx flow sorter (BD Biosciences). The haploid G1 peak (1n) was purified and expanded. 100 millions of recently sorted HAP1 cells were co-transfected with piggyBac transposase and gene-trap vectors [15, 20] (kind gift of J. Cardianos) using Amaxa® Nucleofector (reactive L, X-001 program). Two days after electroporation, piggyBac insertions leading to productive trapping were selected with $0.5 \mu\text{g/ml}$ puromycin for 8 days. After selection, cells were expanded and frozen down for future screens.

6-TG and ATRi resistance screens

For 6TG, 6 million cells of two independent pB-mutagenized libraries were plated at low density to avoid cross killing by 6-TG-sensitive cells. A concentration of $20 \mu\text{M}$ 6-TG was used during the complete incubation period (15 days). A pool of 6TG-resistant cells as well as individual clones were isolated. In the case of ATRi, 6 million cells of a pB library were exposed for 9 days to $0.3 \mu\text{M}$ of ATRi [31].

Splinkerette PCR and Sanger sequencing

DNA from three clones from a pB-mutagenized library was extracted and used to map genetrap integration sites. Host sequences flanking piggyBac insertions were amplified using the Splinkerette protocol as previously described [17] and subjected to Sanger sequencing.

Paired-end RNAseq

RNA from HAP1 cells was isolated using Agilent Technologies' Absolutely RNA Miniprep Kit following manufacturers instructions. RNA integrity was subsequently verified on an Agilent 2100 Bioanalyzer. For the generation of sequencing libraries, the poly(A)-positive fraction was purified, randomly fragmented, converted to double stranded cDNA and processed through subsequent enzymatic treatments of end-repair, dA-tailing, and ligation to adapters as described in Illumina's "TruSeq Stranded mRNA Sample Preparation Part # 15031047 Rev. D". Adapter-ligated library was completed by PCR with Illumina paired-end primers (8 cycles). The resulting purified cDNA library was applied to an Illumina flow cell for cluster generation and sequenced in a PairEnd 2x50 bases format on an Illumina HiSeq2000 instrument by following manufacturer's protocols.

Bioinformatic analysis

For the mapping of pB insertions, a custom Bowtie index [39] was prepared to attach selected pB fragments to the human hg19 reference sequence. Specifically, the 3'TR and up to 1,300bp from the start of 5'TR to the start of Puro [15] were added as fasta fragments to the multifasta UCSC hg19 file and indexed with bowtie-build. The size of the pB sequence could have been shorter, to include just TR boundaries and official SA location plus an

allowance for library fragment size, but some cryptic SA sites suggested the use of an extended version.

Pair-end RNAseq reads were aligned to the pB-hg19 database with TopHat (v. 2.0.10) with –fusion-search option [40]. Other options were --bowtie1, --no-coverage-search, -r 30, --mate-std-dev 80, --max-intron-length 10000, --fusion-min-dist 200000000, --fusion-anchor-length 13, --fusion-ignore-chromosomes chrM.” In order to extract the pB insertions, Tophat accepted_hits.bam file outputs were converted to sam format with samtools view. Lines containing “XF:Z:” or “XP:Z:” as markers for fusion alignments were extracted from accepted_hits.sam files with egrep. Read names from extracted lines were used to fetch corresponding read1 and read2 sequences with egrep from the initial fastq files. These putative pB-hg19 fusions were further inspected to discard homopolymeric repeats and to verify individual insertion positions. For the visualization of insertion sites in the UCSC browser, sam formatted paired reads containing “XF:Z” and “XP:Z” annotations were converted to bed12 files which were loaded into UCSC browser as custom tracks. Thus, Trap^{Seq} involves a successive processing through SAMtools view and BEDtools bamToBed [41, 42].

For the evaluation of gene expression, RNAseq reads were also aligned to human reference transcriptome build hg19 from UCSC using TopHat [40]. Mapped sequences were subsequently assembled into transcripts and analyzed for differential expression by Cuffdiff (Cufflinks-2.1.1) [43], which yielded gene_exp.diff output files. Mapped reads were also uploaded into UCSC genome browser for visualization. For this purpose the corresponding accepted_hits.bam result files were pretreated with SAMtools sort and index scripts, BEDtools genomeCoverageBed and finally bedGraphToBigWig.

Reverse transcription PCR (RT-PCR)

RNA was converted to complementary DNA using SuperScript® III according to manufacturer’s protocol (Invitrogen) and PCR was conducted with GoTaq® (Promega). The sequences of the oligos to confirm the presence of specific HPRT1-pB chimeric transcripts are shown as a Table in the Supplementary Material.

Plasmids

To generate the HAP1^{Cas9} line, a Cas9 expression vector from Addgene (42230) was used to amplify the coding sequences of Cas9 by PCR. The PCR product was digested with EcoRI and subcloned into a FU-tet-o plasmid from Addgene (19778). The plasmid FU-delta-GW-rtTA (Addgene, 19780) was also used. eGFP coding sequence was amplified, digested with EcoRI and subcloned into the FU-tet-o plasmid. The lentiviral plasmid pKLV-U6gRNA-PGKpuro2ABFP (Addgene, 50946) was used to express sgRNAs in HAP1^{Cas9} cells. The sequences of the sgRNAs against *CDC25A* and GFP were designed using the MIT CRISPR design tool (<http://www.genome-engineering.org/crispr/>). All oligo sequences used in this study are available as a Table in the Supplementary Material.

Cell viability assays

Cells were seeded at 4,000 cells per well in 96-well tissue culture plates and treated with the indicated concentrations of ATRi or left untreated. 36 hours after the treatment, cell viability was measured using a luminescent system (CellTiter-Glo, Promega), according to the manufacturer's protocols.

Western blot

Cell pellets were lysed in 50 mM Tris/150 mM NaCl/1% TritonX-100 or in 50 mM Tris pH 7.9/ 8M Urea/1%Chaps and incubated with shaking at 4°C for 30 min. 20 µg of each extract was run and transferred for detection by using the following antibodies: CDC25A (1:1000; Cell Signaling 3652), ACTIN (1:10000, Sigma A5441), γH2AX (1:1000, Millipore 05-636), RPA (1:1000, Cell Signaling 2208), pRPA (1:750, Bethyl A300-245), TUBULIN (1:50000, Sigma T9026) and ECT2 (1:500, Santa Cruz 514750).

Knockdown experiments

Exponentially growing U2OS were transfected in 96-well plates with 7.5nM of each esiRNA from a custom library targeting the hits of the ATRi screen (MISSION esiRNA, Sigma-Aldrich), by using Lipofectamine RNAiMAX reagent (Thermo Fisher Scientific). U2OS cells were transfected in suspension with 20nM of human siRNAs targeting ECT2 (SMART pools, Dharmacon) by using Lipofectamine RNAiMAX reagent (Thermo Fisher Scientific). For each siRNA independent experiment, the efficiency of the knockdown on transfected cells was evaluated by WB.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank the Flow Cytometry Unit of the CNIO for their help in work with human HAP1 cells, Sergio Ruiz for his help with the plasmids to generate the HAP1^{Cas9} line and José María Fernández from the Spanish National Bioinformatics Institute for his advice.

Funding

Research was funded by Fundación Botín, Banco Santander, through its Santander Universities Global Division and by grants from the Spanish Ministry of Economy and Competitiveness (MINECO) (SAF2014-59498-R; SAF2014-57791-REDC), Fundació La Marató de TV3, Howard Hughes Medical Institute and the European Research Council (ERC-617840) to OF; and by a PhD fellowship from La Caixa Foundation to CM.

Abbreviations

RNAseq	RNA sequencing
pB	piggyBac
6TG	6-thio-guanine
mES	mouse embryonic stem cells

ATRI	ATR inhibitor
dox	doxycycline
esiRNAs	endoribonuclease-prepared short interfering RNAs.

References

- [1]. Hartwell LH, Culotti J, Pringle JR, Reid BJ. Genetic control of the cell division cycle in yeast. *Science* (New York, NY). 1974; 183:46–51.
- [2]. Otto SP, Jarne P. Evolution. Haploids--hapless or happening? *Science* (New York, NY). 2001; 292:2441–3.
- [3]. Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuethrich I, Godarova A, et al. Haploid genetic screens in human cells identify host factors used by pathogens. *Science* (New York, NY). 2009; 326:1231–5.
- [4]. Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature*. 2011; 477:340–3. [PubMed: 21866103]
- [5]. Elling U, Taubenschmid J, Wirnsberger G, O'Malley R, Demers SP, Vanhaelen Q, et al. Forward and reverse genetics through derivation of haploid mouse embryonic stem cells. *Cell stem cell*. 2011; 9:563–74. [PubMed: 22136931]
- [6]. Leeb M, Wutz A. Derivation of haploid embryonic stem cells from mouse embryos. *Nature*. 2011; 479:131–4. [PubMed: 21900896]
- [7]. Sagi I, Chia G, Golan-Lev T, Peretz M, Weissbein U, Sui L, et al. Derivation and differentiation of haploid human embryonic stem cells. *Nature*. 2016; 532:107–11. [PubMed: 26982723]
- [8]. Marceau CD, Puschnik AS, Majzoub K, Ooi YS, Brewer SM, Fuchs G, et al. Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens. *Nature*. 2016; 535:159–63. [PubMed: 27383987]
- [9]. Timms RT, Menzies SA, Tchasovnikarova IA, Christensen LC, Williamson JC, Antrobus R, et al. Genetic dissection of mammalian ERAD through comparative haploid and CRISPR forward genetic screens. *Nat Commun*. 2016; 7:11786. [PubMed: 27283361]
- [10]. Carette JE, Guimaraes CP, Wuethrich I, Blomen VA, Varadarajan M, Sun C, et al. Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. *Nat Biotechnol*. 2011; 29:542–6. [PubMed: 21623355]
- [11]. Pettitt SJ, Rehman FL, Bajrami I, Brough R, Wallberg F, Kozarewa I, et al. A genetic screen using the PiggyBac transposon in haploid cells identifies Parp1 as a mediator of olaparib toxicity. *PloS one*. 2013; 8:e61520. [PubMed: 23634208]
- [12]. Ochman H, Gerber AS, Hartl DL. Genetic applications of an inverse polymerase chain reaction. *Genetics*. 1988; 120:621–3. [PubMed: 2852134]
- [13]. Devon RS, Porteous DJ, Brookes AJ. Splinkerettes--improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res*. 1995; 23:1644–5. [PubMed: 7784225]
- [14]. Kanagawa T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J Biosci Bioeng*. 2003; 96:317–23. [PubMed: 16233530]
- [15]. Cadinanos J, Bradley A. Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res*. 2007; 35:e87. [PubMed: 17576687]
- [16]. Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*. 2005; 122:473–83. [PubMed: 16096065]
- [17]. Li MA, Pettitt SJ, Yusa K, Bradley A. Genome-wide forward genetic screens in mouse ES cells. *Methods in enzymology*. 2010; 477:217–42. [PubMed: 20699144]
- [18]. Wang W, Lin C, Lu D, Ning Z, Cox T, Melvin D, et al. Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*. 2008; 105:9290–5. [PubMed: 18579772]

- [19]. Wang W, Bradley A, Huang Y. A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. *Genome Res.* 2009; 19:667–73. [PubMed: 19233961]
- [20]. Yusa K, Zhou L, Li MA, Bradley A, Craig NL. A hyperactive piggyBac transposase for mammalian applications. *Proc Natl Acad Sci U S A.* 2011; 108:1531–6. [PubMed: 21205896]
- [21]. Yusa K, Rad R, Takeda J, Bradley A. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nat Methods.* 2009; 6:363–9. [PubMed: 19337237]
- [22]. Lopez-Contreras AJ, Fernandez-Capetillo O. The ATR barrier to replication-born DNA damage. DNA repair. 2010; 9:1249–55. [PubMed: 21036674]
- [23]. Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. *Nature reviews Molecular cell biology.* 2008; 9:616–27. [PubMed: 18594563]
- [24]. Nieto-Soler M, Morgado-Palacin I, Lafarga V, Lecona E, Murga M, Callen E, et al. Efficacy of ATR inhibitors as single agents in Ewing sarcoma. *Oncotarget.* 2016
- [25]. Morgado-Palacin I, Day A, Murga M, Lafarga V, Anton ME, Tubbs A, et al. Targeting the kinase activities of ATR and ATM exhibits antitumoral activity in mouse models of MLL-rearranged AML. *Sci Signal.* 2016; 9:ra91. [PubMed: 27625305]
- [26]. Kwok M, Davies N, Agathangelou A, Smith E, Oldreive C, Petermann E, et al. ATR inhibition induces synthetic lethality and overcomes chemoresistance in TP53- or ATM-defective chronic lymphocytic leukemia cells. *Blood.* 2016; 127:582–95. [PubMed: 26563132]
- [27]. Krajewska M, Fehrmann RS, Schoonen PM, Labib S, de Vries EG, Franke L, et al. ATR inhibition preferentially targets homologous recombination-deficient tumor cells. *Oncogene.* 2015; 34:3474–81. [PubMed: 25174396]
- [28]. Flynn RL, Cox KE, Jeitany M, Wakimoto H, Bryll AR, Ganem NJ, et al. Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science.* 2015; 347:273–7. [PubMed: 25593184]
- [29]. Cottini F, Hideshima T, Suzuki R, Tai YT, Bianchini G, Richardson PG, et al. Synthetic lethal approaches exploiting DNA damage in aggressive myeloma. *Cancer Discov.* 2015
- [30]. Santos MA, Faryabi RB, Ergen AV, Day AM, Malhowski A, Canela A, et al. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. *Nature.* 2014; 514:107–11. [PubMed: 25079327]
- [31]. Toledo LI, Murga M, Zur R, Soria R, Rodriguez A, Martinez S, et al. A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. *Nature structural & molecular biology.* 2011; 18:721–7.
- [32]. Ruiz S, Mayor-Ruiz C, Lafarga V, Murga M, Vega-Sendino M, Ortega S, et al. A Genome-wide CRISPR Screen Identifies CDC25A as a Determinant of Sensitivity to ATR Inhibitors. *Mol Cell.* 2016; 62:307–13. [PubMed: 27067599]
- [33]. Tatsumoto T, Xie X, Blumenthal R, Okamoto I, Miki T. Human ECT2 is an exchange factor for Rho GTPases, phosphorylated in G2/M phases, and involved in cytokinesis. *J Cell Biol.* 1999; 147:921–8. [PubMed: 10579713]
- [34]. He D, Xiang J, Li B, Liu H. The dynamic behavior of Ect2 in response to DNA damage. *Sci Rep.* 2016; 6:24504. [PubMed: 27074761]
- [35]. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun.* 2017; 8:14049. [PubMed: 28091601]
- [36]. Adamson B, Norman TM, Jost M, Cho MY, Nunez JK, Chen Y, et al. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell.* 2016; 167:1867–82 e21. [PubMed: 27984733]
- [37]. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell.* 2016; 167:1853–66 e17. [PubMed: 27984732]
- [38]. Jaitin DA, Weiner A, Yofe I, Lara-Astiaso D, Keren-Shaul H, David E, et al. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell.* 2016; 167:1883–96 e15. [PubMed: 27984734]
- [39]. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009; 10:R25. [PubMed: 19261174]

- [40]. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009; 25:1105–11. [PubMed: 19289445]
- [41]. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25:2078–9. [PubMed: 19505943]
- [42]. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010; 26:841–2. [PubMed: 20110278]
- [43]. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010; 28:511–5. [PubMed: 20436464]

Fig. 1. Pipeline for genetrap-based genetic screenings with Trap^{Seq}.

- Scheme followed for pB-based mutagenesis and genetic screenings on the near-haploid human cell line HAP1.
- Example of a typical library of mutant alleles generated by pB transposons. The pB genetrap vector was electroporated with or without a pB transposase (pBase) expressing plasmid into HAP1 cells, which were subsequently cultured in the presence of puromycin.

- c.** Illustration of the trapping induced by pB. Paired-end RNAseq is used to sequence fragments from both ends, increasing the chance of identifying chimeric transcripts containing sequences from the genetrap together with the trapped exon. (*TR*: terminal repeat; *En2SA*: splice acceptor En2; *IRES*: internal ribosomal entry site; *puro tk*: positive-negative selection marker, puromycin can be used to select for integrations and FIAU to select for loss of the transposon; *pA*: polyA).
- d.** Trap^{Seq} bioinformatics pipeline (see Experimental Procedures for a full description).

Fig. 2. Output from a 6TG-resistance screen with Trap^{Seq}.

- a. Percentage of RNAseq reads of chimeric transcripts obtained with Trap^{Seq} from a pool of 6TG-resistant HAP1 cells.
- b. Trap^{Seq}-based mapping of insertion sites and differential expression in the 6TG-resistant pool. UCSC insertion tracks for genetrap insertions (pink, above) illustrate the exons joined to the splice acceptors of the pB genetrap vectors into the gene *HPRT1*. Each red vertical line indicates an exon that was found spliced to the pB (horizontal lines correspond to reads covering both exons). UCSC expression tracks (below) show the reduction in *HPRT1* mRNA levels found on the 6TG-resistant pool compared to the unselected library.

- c. Trap^{Seq}-based mapping of the genetrapped insertion and its effect on the expression on the target gene (*HRPT1*) in a 6TG-resistant single clone.
- d. PCR confirming the presence of specific HPRT1-pB chimeric transcripts in 6TG resistant cells (6TG-R pool). cDNA from wild type HAP1 cells (WT) and the unselected libraries (pB-library) was used as control.

See also Additional File 1 (Fig. S1-3 and Tables S1,2).

Fig. 3. ATR inhibitor resistance screen.

- UCSC insertion track showing the insertion in *CDC25A* detected with Trap^{Seq} from the ATRi-resistant pool.
- Sequence of the *CDC25A*-pB chimeric transcript identified above.
- WB illustrating the reduction in *CDC25A* levels observed in 2 independent clones of *CDC25A*^{+/-} cells that were generated in HAP1^{Cas9} cells with *CDC25A*-targeting sgRNAs. β ACTIN levels are shown as a loading control.
- CellTiter-Glo viability assay in WT and *CDC25A*^{+/-} HAP1^{Cas9} cells exposed to ATRi for 36h. Data are representative of 3 independent experiments. Error bars indicate s.d.
- WB illustrating the levels of γ H2AX and RPA (S4/8) phosphorylation in WT and *CDC25A*^{+/-} HAP1^{Cas9} cells exposed to ATRi for 8h (0.3, 0.9 and 2 μ M). TUBULIN and total RPA levels are shown as loading controls.

See also Additional File 1 (Fig. S4 and Table S3).

Fig. 4. Identification of ECT2 as a determinant to the sensitivity to ATRi

- CellTiter-Glo viability assay in U2OS cells transfected with esiControl or esiRNAs targeting the hits from the ATRi screen and exposed to ATRi for 36h at 3 μ M. Error bars indicate s.d.
- Schematic illustration of the domain structure of the ECT2 protein. (N, Amino-terminal region; XRCC1, X-ray repair complementing defective repair in Chinese hamster cells 1 domain; Cyclin B6, cyclin B6-like domain; BRCT, BRAC1 C-terminal domain; S, small central region; DH, Dbl homology domain; PH, pleckstrin homology domain; C, Carboxyl-terminal region. NLS, nuclear localization sequence.)
- UCSC insertion track showing the insertion in *ECT2* detected with Trap^{Seq} from the ATRi-resistant pool.
- WB illustrating the reduction of ECT2 in U2OS cells transfected with control or ECT2-targeting siRNAs 72h after siRNA transfection.

- e. WB illustrating the levels of γ H2AX and RPA (S4/8) phosphorylation in U2OS transfected with control or ECT2-targeting siRNAs and exposed to ATRi for 4h (0.5, 1 and 2 μ M). TUBULIN and total RPA levels are shown as loading controls.
- f. CellTiter-Glo viability assay in U2OS cells transfected with control or ECT2-targeting siRNAs and exposed to ATRi for 36h at the indicated doses. Data are representative of 3 independent experiments. Error bars indicate s.d.