This is a repository copy of *Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/136933/

Version: Accepted Version

Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells

Harveer Dev[1,2], Ting-Wei Will Chiang[1Ψ], Chloe Lescale[3Ψ], Inge de Krijger[4Ψ], Alistair G. Martin[5], Domenic Pilger[1], Julia Coates[1], Matylda Sczaniecka-Clift[1], Wenming Wei[3], Matthias Ostermaier[6], Mareike Herzog[1], Jonathan Lam[1], Abigail Shea[5], Mukerrem Demir[1], Qian Wu[7], Fengtang Yang[8], Beiyuan Fu[8], Zhongwu Lai[9], Gabriel Balmus[1,8], Rimma Belotserkovskaya[1], Violeta Serra[10], Mark J. O'Connor[11], Alejandra Bruna[5], Petra Beli[6], Luca Pellegrini[7], Carlos Caldas[5], Ludovic Deriano[3ϕ], Jacqueline J.L. Jacobs[4ϕ], Yaron Galanty[1ϕ] and Stephen P. Jackson[1ϕ]

[1]The Wellcome Trust/Cancer Research UK Gurdon Institute and Department of Biochemistry, University of Cambridge, Cambridge CB2 1QN, UK. [2]Academic Urology Group, Department of Surgery, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK. [3]Genome Integrity, Immunity and Cancer Unit, Department of Immunology, Department of Genomes and Genetics, Institut Pasteur, 75015 Paris, France. [4]Division of Oncogenomics, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands. [5]Department of Oncology and Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge CB2 0RE, UK. [6]Institute of Molecular Biology (IMB), 55128 Mainz, Germany. [7]Department of Biochemistry, University of Cambridge, Cambridge, CB2 1GA, UK. [8]Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK. [9]AstraZeneca, Waltham, Massachusetts, 02451, USA. [10]Vall d'Hebron Institute of Oncology, Barcelona, Spain. [11]Astrazeneca, 1 Francis Crick Ave, Cambridge CB2 0RE, UK.

ΨEqual contribution
ϕ Co-corresponding authors: ludovic.deriano@pasteur.fr, j.jacobs@nki.nl, y.galanty@gurdon.cam.ac.uk, and s.jackson@gurdon.cam.ac.uk

Abstract

BRCA1 deficiencies cause breast, ovarian, prostate and other cancers, and render tumours hypersensitive to PARP inhibitors. To understand resistance mechanisms, we conducted whole-genome CRISPR-Cas9 synthetic-viability/resistance screens in BRCA1-deficient breast cancer cells treated with PARP inhibitors. We identified two previously uncharacterized proteins, C20orf196 and FAM35A, whose inactivation confers strong PARP-inhibitor resistance. Mechanistically, we show C20orf196 and FAM35A form a complex, "Shieldin" (SHLD1/2), with FAM35A interacting with single-stranded DNA via its C-terminal OB fold region. We establish that Shieldin acts as the downstream effector of 53BP1/RIF1/MAD2L2 to promote DNA double-strand break (DSB) end-joining through restricting DSB resection and counteract homologous recombination by antagonising BRCA2/RAD51 loading in BRCA1-deficient cells. Notably, Shieldin inactivation further sensitises BRCA1-deficient cells to cisplatin, suggesting how defining the SHLD1/2 status of

43 BRCA1-deficient tumours might aid patient stratification and yield new treatment
44 opportunities. Highlighting this potential, we document reduced SHLD1/2 expression in
45 human breast cancers displaying intrinsic or acquired PARP-inhibitor resistance.

46

47 **Introduction**
48 DNA double-strand breaks (DSBs) are highly cytotoxic cellular lesions that must be
49 effectively and accurately repaired to maintain genome stability and prevent premature aging,
50 neurodegeneration, immunodeficiency, cancer and other diseases[1-3]. In response to DSB
51 detection, the apical kinases ATM, ATR and PRKDC (DNA-PKcs) become activated and
52 phosphorylate numerous substrates to initiate the cellular DNA damage response (DDR)[4].
53 The ensuing cascade of molecular DDR events, which are promoted by various post-
54 translational modifications including protein phosphorylation, ubiquitylation, sumoylation
55 and poly (ADP-ribosyl)ation, impacts on a myriad of cellular components, amongst other
56 things leading to assembly of DDR factors at DNA-damage sites, arrest or slowing of cell-
57 cycle progression, and activation of DNA repair mechanisms[4, 5]. The two main types of
58 DSB-repair pathway are non-homologous end-joining (NHEJ) which is active throughout the
59 cell cycle, and homologous recombination (HR), which normally requires a sister chromatid
60 as a template and hence only operates in S and G2 phases of the cell cycle. DSB-repair
61 pathway choice is partly determined by functional antagonism between the HR-promoting
62 factor BRCA1 and NHEJ-promoting proteins such as TP53BP1 (53BP1), RIF1 and MAD2L2
63 (REV7)[6-13].

64

65 Inherited or acquired mutations in the *BRCA1* or *BRCA2* genes that result in protein loss or a
66 mutant BRCA1/2 protein cause breast, ovarian, prostate and other cancers, and render
67 tumours hypersensitive to PARP-inhibitor drugs such as olaparib[14-17]. Unfortunately, intrinsic
68 or acquired PARP-inhibitor resistance frequently leads to lack-of-response or to patient
69 relapse and tumour regrowth[15, 18]. In the clinic, the most common PARP-inhibitor resistance
70 mechanisms reported to date are restoration of BRCA1/2 expression or function. Notably,
71 53BP1 expression is lost in various triple-negative breast cancers[7], which may account for
72 certain clinically relevant examples of PARP-inhibitor resistance. Nevertheless, the
73 mechanisms driving PARP-inhibitor resistance in a large proportion of BRCA1/2-deficient
74 tumours remain unexplained[18, 19].

75

76   To systematically survey for genetic mechanisms of PARP-inhibitor resistance, we conducted
77   whole-genome CRISPR-Cas9 synthetic-viability/resistance screens in human BRCA1-
78   deficient breast cancer cells treated with PARP inhibitors. In addition to identifying known
79   resistance factors such as 53BP1, RIF1 and MAD2L2 loss[6-13], we identified two previously
80   uncharacterized proteins, C20orf196 and FAM35A, whose inactivation confers PARP-
81   inhibitor resistance to BRCA1-deficient cells. Our ensuing work lead us to define the
82   "Shieldin" ($SHLD1^{C20orf196}$/$SHLD2^{FAM35A}$) complex that promotes NHEJ by serving as the
83   downstream effector of 53BP1, RIF1 and MAD2L2, restricts DSB resection, and counteracts
84   HR in BRCA1-deficient cells by antagonising replacement of replication protein A (RPA)
85   with BRCA2 and RAD51 on resected single-stranded DNA (ssDNA). Finally, we report that
86     $SHLD1^{C20orf196}$/$SHLD2^{FAM35A}$ loss confers hypersensitivity to the DNA-crosslinking agent
87   cisplatin, and that reduced $SHLD1^{C20orf196}$ or $SHLD2^{FAM35A}$ expression is associated with
88   evolution of PARP-inhibitor resistance in a patient-derived BRCA1-deficient breast cancer
89   xenograft model and in BRCA1-mutant cancers displaying intrinsic PARP-inhibitor
90   resistance.

91

92   **Results**

93   **FAM35A or C20orf196 loss suppresses PARP-inhibitor sensitivity of *BRCA1*-mutant**
94   **cells**

95   To systematically explore genetic mechanisms imparting PARP-inhibitor resistance, we
96   carried out genome-wide CRISPR-Cas9 gene-inactivation screens with the GeCKO library[20]
97   in the *BRCA1*-mutant breast cancer cell line SUM149PT treated in parallel with the PARP
98   inhibitors olaparib, talazoparib (BMN673) or AZD2461 (Fig 1a, b, Supplementary Fig 1a-c).
99   In addition to identifying the known resistance genes *TP53BP1*, *RIF1* and *MAD2L2* whose
100   products form a complex[21], we identified several new suppressor candidates (Supplementary
101   Table 1, Supplementary Fig 1d-e). These included DYNLL1, a known 53BP1 interaction
102   partner[22], and TEN1, a component of the CST telomere-capping complex
103   (CTC1/STN1/TEN1) that also promotes telomere DNA replication[23]. In our ensuing studies,
104   however, we focused on the uncharacterised proteins FAM35A and C20orf196 that
105   collectively received the highest scores from our screens (Fig 1b and Supplementary Table
106   1). Thus, by carrying out short-interfering RNA (siRNA) mediated mRNA silencing in non-
107   transformed, hTERT immortalized human RPE1 cells (Supplementary Fig 1f), we established
108   that, as for 53BP1 loss[7], depletion of FAM35A or C20orf196 markedly suppressed PARP-
109   inhibitor sensitivity caused by BRCA1 inactivation while having no discernible effect in

110 BRCA1-proficient cells (Fig 1c, Supplementary Fig 1g). This conclusion was independently
111 confirmed by de novo CRISPR-Cas9 gene editing, with FAM35A or C20orf196 inactivation
112 alleviating the olaparib hypersensitivity of BRCA1-deficient cells in a manner counteracted
113 by re-introduction of wild-type FAM35A or C20orf196 (Fig 1d-e; Supplementary Fig 1h; as
114 shown in supplementary Fig 1i, these effects did not reflect altered cell-cycle profiles).

115

116 **The FAM35A/C20orf196 complex interacts with and acts downstream of**
117 **53BP1/RIF1/MAD2L2**
118 Sequence analyses indicated that FAM35A and C20orf196 are well conserved in vertebrates.
119 Moreover, structure prediction modelling (RaptorX; http://raptorx.uchicago.edu/) revealed
120 that FAM35A harbours a disordered N-terminus and an ordered C-terminal region containing
121 three OB folds, with the last C-terminal OB fold/FAM domain containing a CXXC-type zinc
122 finger motif (Fig 2a). Notably, this organization is highly similar to those of the RPA1
123 subunit of ssDNA binding protein RPA and the CTC1 subunit of the CST complex that also
124 binds ssDNA[23]. In this regard, we noted that while the C20orf196 N-terminus (residues 1-70)
125 is predicted to be intrinsically disordered, its C-terminal part is more structured and may
126 harbour one- or two-winged helix (WH) domains (Fig 2a) similar to those in the yeast CST
127 subunit Stn1[23, 24], suggesting that C20orf196 and Stn1 might play analogous or
128 complementary roles.

129

130 Through combining cellular co-localization and co-immunoprecipitation experiments, we
131 established that FAM35A and C20orf196 directly interact in a manner that is mainly, but not
132 exclusively, mediated by the FAM35A OB3/FAM domain (Fig 2b-c, Supplementary Fig 2a-
133 b). Because loss of FAM35A or C20orf196 had similar effects to loss of
134 53BP1/RIF1/MAD2L2 in BRCA1-deficient cells, we tested for possible interactions between
135 these factors. Thus, via co-immunoprecipitation and mass spectrometry (MS) studies, we
136 found that both C20orf196 and FAM35A interact with MAD2L2, the most distal factor of the
137 53BP1/RIF1/MAD2L2 axis mediating PARP-inhibitor sensitivity in BRCA1-deficient cells[6-
138 13] (Fig 2d, Supplementary Fig 2c).

139

140 Many DDR proteins accumulate at DSB sites within ionizing radiation (IR)-induced nuclear
141 foci (IRIF)[5]. We established that both FAM35A and C20orf196 formed IRIF, and by live-cell
142 imaging studies found that the proteins were also recruited to localised DNA-damage sites
143 induced by laser micro-irradiation (Supplementary Fig 2d). Furthermore, we determined by

144   confocal and super-resolution microscopy that FAM35A co-localised with the established

145   DSB markers phosphorylated histone H2AFX[25] (γH2AX) and 53BP1[26] (Fig 2e-f,

146   Supplementary Fig 2e). Notably, siRNA/shRNA-depletion experiments established that while

147   53BP1 IRIF and MAD2L2 levels and IRIF were not significantly impaired by FAM35A or

148   C20orf196 depletion (Supplementary Fig 2f-h), IRIF formation by FAM35A and C20orf196

149   required 53BP1, RIF1 and MAD2L2 but not PTIP (Fig 2e and Supplementary Fig 3a-c; note

150   that total levels of GFP-tagged FAM35A/C20orf196 were minimally affected by

151   53BP1/RIF1/MAD2L2 depletion). We also established that C20orf196 IRIF were almost

152   totally abrogated by FAM35A depletion, while C20orf196 depletion reduced but did not

153   abolish FAM35A IRIF (Fig 2e). In addition, FAM35A formed nuclear foci when cells were

154   treated with the DNA topoisomerase I inhibitor camptothecin (CPT; Supplementary Fig 3d).

155   Significantly, the FAM35A N-terminus was necessary and sufficient for its IRIF formation,

156   these IRIF depended on 53BP1, RIF1, MAD2L2 and C20orf196, and this region could be co-

157   immunoprecipitated with MAD2L2 (Fig 2f-g, Supplementary Fig 3e-f). Collectively, these

158   findings indicated that FAM35A and C20orf196 act as downstream components of the

159   53BP1/RIF1/MAD2L2 molecular assembly[21] at DSB sites.

160

161   **FAM35A and C20orf196 promote NHEJ**

162   Since 53BP1, RIF1 and MAD2L2 promote NHEJ[6-13], we tested whether FAM35A and

163   C20orf196 fulfil a similar role. Indeed, as for depletion of the NHEJ factor XRCC4, siRNA

164   depletion of 53BP1, FAM35A or C20orf196 impaired NHEJ as measured by random

165   integration of plasmid DNA into chromosomes[27] (Fig 3a). In addition, FAM35A or

166   C20orf196 depletion conferred IR hypersensitivity to both human and mouse cells (Fig 3b

167   and Supplementary Fig 4a). 53BP1 and associated factors promote NHEJ-mediated class-

168   switch recombination (CSR) at the immunoglobulin heavy-chain locus, a process that allows

169   B-lymphocytes to change antibody production from one type to another[28]. By CRISPR-Cas9

170   gene editing in mouse CH12F3 (CH12) B-lymphocytes[29] we established that, as for

171   53BP1/RIF1/MAD2L2 inactivation[6, 8, 10-12, 30, 31], loss of FAM35A or C20orf196 markedly

172   reduced CSR (Fig 3c-d and Supplementary Fig 4b-d). Furthermore, analysis of metaphase

173   chromosomal spreads of such cells revealed that FAM35A or C20orf196 inactivation led to

174   chromosomal breaks and translocations symptomatic of aberrant CSR[32] (Fig 3e-f; note from

175   Supplementary Fig 4e-f that CSR effects were not associated with defects in cell

176   proliferation, or in Aid or germ-line Sα switch region transcription).

177

FAM35A and C20orf196 antagonise DNA-end resection

178

To explore FAM35A and C20orf196 function further, we carried out assays in mouse cells

179

180   harbouring a temperature-sensitive allele of the telomere-associated factor TRF2 (TRF2ts).

181   TRF2ts inactivation at higher temperatures results in de-protected chromosome ends and

182   causes NHEJ-mediated telomere fusions (Fig 4a)[6, 33]. Strikingly, through use of short-hairpin

183   RNA (shRNA) mediated mRNA silencing, we found that such chromosome fusions were

184   diminished by FAM35A or C20orf196 depletion like that elicited by MAD2L2 depletion (Fig

185   4b, Supplementary Fig 5a-b).

186

187   The impacts of 53BP1, RIF1 or MAD2L2 depletion in the TRF2ts system are connected to

188   these factors counteracting DSB resection[6-13, 34]. We thus explored whether FAM35A and

189   C20orf196 might also have this function. Indeed, as for 53BP1/RIF1/MAD2L2 inactivation[6-

190   13, 35], loss of FAM35A or C20orf196 in human cells enhanced DSB-resection as measured by

191   RPA and ssDNA staining intensity in pre-extracted nuclei after treatment with camptothecin

192   (Fig 4c-e; RPA1 kinetics at DNA damage sites induced by laser micro-irradiation were not

193   altered by FAM35A or C20orf196 depletion, Supplementary Fig 5c). In line with this

194   resection being mediated by canonical pathways, it was diminished by depletion of the

195   resection promoting factors RBBP8 (CtIP) or BLM (Fig 4f). Furthermore, we established that

196   recruitment of BLM to sites of laser micro-irradiation was enhanced by FAM35A or

197   C20orf196 depletion (Fig 4g). Similarly, as shown for 53BP1 depletion in mouse cells[6],

198   FAM35A or C20orf196 depletion in such cells led to higher levels of the DNA-end resection

199   marker, Ser4/8 phosphorylated RPA2, after IR treatment (Fig 4d). Together with our other

200   findings, these data established FAM35A and C20orf196 as crucial components of

201   53BP1/RIF1/MAD2L2-mediated chromosomal NHEJ, and suggested that their pro-NHEJ

202   function is connected to limiting DSB resection.

203

The FAM35A OB fold region interacts with ssDNA and promotes IR survival

204

Consistent with our prediction of structural similarity between FAM35A and RPA1, the

205

206   FAM35A C-terminus could be retrieved from cell extracts via interaction with a ssDNA

207   oligonucleotide (Fig 5a). Sequence alignment to RPA1 and structural modelling of FAM35A

208   identified two Trp (W) residues predicted to be at the protein-ssDNA interface, based on

209   analogous residues critical for RPA binding to ssDNA (Fig 5b, Supplementary Fig 5d). In

210   accord with this prediction, we found via electrophoretic gel-mobility shift assays (EMSAs)

211   that the bacterially-expressed, purified FAM35A C-terminal region bound preferentially to

212   ssDNA rather than double-stranded DNA (Fig 5c, Supplementary Fig 5e), and ssDNA

213   binding was reduced when the two Trp residues were mutated to Ala (W489/W640A; Fig

214   5c). Furthermore, while full-length FAM35A bearing these mutations (FAM35A$^{W489/W640A}$)

215   still interacted with C20orf196 (Supplementary Fig 5f) and formed IRIF in cells, these IRIF

216   were consistently less pronounced than those of the wild-type FAM35A protein (Fig 5d).

217   This suggested that following IRIF recruitment via its N-terminal region, the FAM35A C-

218   terminal ssDNA binding region may allow further FAM35A recruitment, retention and/or

219   stabilization. In addition, unlike the wild-type protein, FAM35A$^{W489/W640A}$ did not confer

220   significant IR resistance when reintroduced into FAM35A null cells (Fig 5e). In parallel

221   studies, expression of the FAM35A C-terminus did not complement the IR hypersensitivity

222   of FAM35A null cells. Moreover, expression of the FAM35A N-terminus rendered cells IR

223   hypersensitive irrespective of whether they expressed endogenous FAM35A, implying that

224   the N-terminal IRIF-forming domain of FAM35A may have a dominant-negative effect on

225   NHEJ (Fig 5f; overexpression of these FAM35A derivatives did not affect olaparib

226   sensitivity in a wild-type background, Supplementary Fig 5g).

227

228   **FAM35A and C20orf196 antagonise HR in BRCA1-deficient cells**

229   PARP inhibitors generate replication-associated DNA lesions that require BRCA1-mediated

230   HR for their effective repair[19], and loss of 53BP1/RIF1/MAD2L2 partly restores the ability of

231   BRCA1-deficient cells to repair such lesions[6-13]. This has led to a model in which BRCA1

232   and 53BP1/RIF1/MAD2L2 play antagonistic roles in channelling DSBs towards HR or

233   NHEJ, respectively. We thus speculated that BRCA1 might antagonise FAM35A/C20orf196

234   action. Accordingly, both FAM35A and C20orf196 IRIF, but not 53BP1 IRIF, were

235   significantly enhanced in number and intensity upon BRCA1 but not BRCA2 depletion (Fig

236   6a-b, Supplementary Fig 6a-b).

237

238   Collectively, our results suggested that FAM35A/C20orf196 act at the interface between the

239   opposing functions of BRCA1 and 53BP1/RIF1/MAD2L2 to regulate DSB-repair pathway

240   choice. While this action could operate at least in part through control of DSB resection,

241   which is misregulated and of slower kinetics in BRCA1-deficient cells[11, 36], we reasoned that

242   FAM35A/C20orf196 might also contribute to the severe defect in BRCA2-mediated RAD51

243   loading at DNA-damage sites in BRCA1-deficient cells[11, 35, 37]. Indeed, as for 53BP1

244   inactivation, loss of FAM35A or C20orf196 restored RAD51 IRIF formation in *BRCA1*-null

245 cells (Fig 6c). While exploring the mechanism for this effect, we found elevated resection
246 levels in FAM35A and C20orf196 knockout cells, as measured by RPA recruitment at DNA-
247 damage sites, were still maintained in the absence of BRCA1 (Supplementary Fig 6c-d).
248 Furthermore, FAM35A/C20orf196 depletion also alleviated the impaired recruitment of
249 BRCA2 to DNA-damage sites in BRCA1-deficient cells (Fig 6d, Supplementary Fig 6e).
250 Accordingly, studies with a cell-based chromosomal traffic light reporter (TLR) HR system
251 [38, 39] established that FAM35A or C20orf196 depletion in BRCA1-deficient cells restored HR
252 to levels similar to those acquired upon 53BP1 depletion in this setting (Fig 6e,
253 Supplementary 6f). In addition, removing FAM35A or C20orf196 rescued the spontaneous
254 genomic instability phenotype of BRCA1 knockout cells (Fig 6f). Building on our findings
255 that the FAM35A N-terminal region largely mediates its localisation to IRIF (Fig 2f,
256 Supplementary Fig 3e-f), introducing the FAM35A N-terminus, but not the C-terminus,
257 enhanced PARP-inhibitor sensitivity of BRCA1/FAM35A null cells (Fig 6g, Supplementary
258 Fig 6g). Furthermore, FAM35A inactivation was epistatic to 53BP1 inactivation in relation to
259 conferring PARP-inhibitor resistance in BRCA1 knockout cells (Supplementary Fig 6h).
260 Considering our findings together, we propose that C20orf196 and FAM35A be named
261 SHLD1 and SHLD2, respectively, or collectively as the "Shieldin complex" because it
262 shields DSBs from inappropriate activities and promotes appropriate modes of DSB repair.
263

264 **FAM35A/C20orf196 loss correlates with PARP inhibitor resistance in cancers**
265 Having identified SHLD1^{C20orf196} and SHLD2^{FAM35A} as mediating the PARP-inhibitor
266 sensitivity of a BRCA1-deficient breast cancer cell line, we speculated that this might also
267 apply in more physiological settings. Consequently, we employed a patient-derived xenograft
268 (PDX) model of BRCA1-deficient breast cancer propagated in mice in the presence (cohort
269 2) or absence (cohort 1) of olaparib (Fig 7a). The resistant tumour was further serially
270 passaged into new hosts that were treated in the presence (cohort 4) or absence (cohort 3) of
271 olaparib to confirm and sustain drug resistance (Fig 7a; see also Supplementary Fig 7a). The
272 tumours were then harvested and subjected to whole-genome RNA-sequencing. Notably, our
273 analyses revealed that in contrast to the other cohorts, nearly all resistant tumours after
274 chronic olaparib treatment (cohort 4) correlated with reduced mRNA expression of
275 SHLD1^{C20orf196}, SHLD2^{FAM35A}, 53BP1 and/or PARP1 (Fig 7a; each heatmap column
276 represents one tumour/mouse sample). Because this tumour model is polyclonal[40], our data
277 suggested that olaparib resistance mechanisms might arise through parallel evolutionary
278 trajectories converging on loss of Shieldin activity. Furthermore, when we stratified a cohort

279   of BRCA1-deficient PDX tumours by SHLD1/2 expression, ensuing analyses indicated that

280   low SHLD1^{C20orf196} transcript levels correlated with intrinsic olaparib resistance (Fig 7b). One

281   of the olaparib resistant models (PDX127) demonstrated concomitant loss of both

282   SHLD1^{C20orf196} and SHLD2^{FAM35A}, while two other resistant models exhibiting normal

283   SHLD1/2 transcript levels harboured deleterious 53BP1 mutations. Notably, several of these

284   resistant PDX models also display BRCA1 nuclear foci[41], suggesting the presence of multiple

285   mechanisms of resistance, due to tumour heterogeneity and/or mechanistic cooperation.

286

287   Finally, we found that in contrast to 53BP1 deficiency[42], SHLD1^{C20orf196} or SHLD2^{FAM35A}

288   loss increased the sensitivity of BRCA1-proficient and BRCA1-null cells to IR, and even

289   more markedly enhanced their sensitivity to the DNA crosslinking agent cisplatin (Fig 7c-d,

290   Supplementary Fig 7b-c). Furthermore, enhanced cisplatin sensitivity upon SHLD1^{C20orf196} or

291   SHLD2^{FAM35A} inactivation was associated with increased DNA-damage focus formation by

292   the FANCD2 protein that is involved in detection and repair of DNA crosslinks (Fig 7e,

293   Supplementary Fig 7d). These findings therefore suggested that, if loss/reduced expression of

294   SHLD1/2 occurs in patients, it may provide collateral therapeutic vulnerabilities that could be

295   exploited clinically.

296

297   **Discussion**

298   Over the past two decades, it has become evident that eukaryotic cells have evolved multiple

299   mechanisms of DNA DSB repair that are regulated in complex and sophisticated ways to

300   optimise genome stability. In particular, much attention has focused on how cells

301   strategically employ the two prime modes of DSB repair – NHEJ and HR – which antagonise

302   one another, operate optimally in different contexts and whose relative usage is regulated by

303   factors such as chromatin structure and cell-cycle stage. In addition to being of academic

304   interest, work on such subjects is also of clinical relevance, particularly in cancer therapy

305   where DSB-inducing chemotherapeutic agents are frequently used, and molecularly-targeted

306   drugs such as PARP inhibitors are being increasingly employed in specific settings. Intrinsic

307   or arising PARP-inhibitor resistance in patients with *BRCA1/2* mutations is nevertheless an

308   increasing clinical problem. Using whole-genome CRISPR-Cas9 synthetic-

309   viability/resistance screens, we have uncovered two, previously uncharacterised proteins –

310   SHLD1^{C20orf196} and SHLD2^{FAM35A} – whose loss mediates PARP-inhibitor resistance and

311   which we have shown act as the most distal factors of the 53BP1/RIF1/MAD2L2 molecular

312   axis to promote NHEJ and restrict HR in BRCA1-deficient settings. Our screens have also

313   identified additional candidate PARP-inhibitor resistance factors that await validation in

314   future studies.

315

316   Mechanistically, we have shown that SHLD1^{C20orf196} and SHLD2^{FAM35A} form a complex,

317   termed Shieldin, with SHLD1^{C20orf196} recruitment to DNA-damage sites via its interactions

318   with SHLD2^{FAM35A} and other factors, and SHLD2^{FAM35A} interacting with single-stranded

319   DNA via its C-terminal OB fold region. Moreover, we have established that SHLD1^{C20orf196}

320   and SHLD2^{FAM35A} promote NHEJ in a manner that may be mediated via their effects on

321   restricting DNA-end resection, and serve as a barrier to HR by antagonising the replacement

322   of RPA with BRCA2/RAD51 on resected ssDNA in a manner counteracted by BRCA1. Our

323   work is in line with a recent report[43], that independently identified SHLD1^{C20orf196} and

324   SHLD2^{FAM35A} as NHEJ-promoting factors and antagonists of HR in BRCA1-defective cells.

325   This study also identified a third component, RINN1/SHLD3$^{CTC-534A2.2}$ that is proposed to

326   serve as a molecular bridge from RIF1 to MAD2L2 and SHLD1/2.

327

328   While it seems possible that Shieldin loss contributes to HR restoration in BRCA1-deficient

329   cells through effects on both resection and on BRCA2/RAD51 loading, the relative

330   importance of these mechanisms needs further investigation. We note that more extensive

331   and possibly faster resection in the absence of Shieldin might enhance BRCA2/RAD51

332   loading. Alternatively, or in addition, Shieldin might serve as a physical barrier to

333   BRCA2/RAD51 loading at dsDNA/ssDNA junctions in BRCA1-deficient cells – perhaps

334   through it being tethered to DSB-flanking chromatin via the 53BP1/RIF1/MAD2L2 complex

335   allowing interactions between the C-terminus of distal SHLD2^{FAM35A} and ssDNA (see Fig 7f

336   for a model, and Supplementary Fig 7e for SHLD2^{FAM35A} domain function summary).

337   Nevertheless, we found that overexpression of the SHLD2^{FAM35A} N- but not the C-terminus

338   confers olaparib hypersensitivity to *BRCA1/FAM35A* double knockout cells, suggesting that

339   at least in this context, chromatin binding by SHLD2^{FAM35A} plays a dominant role in

340   restricting HR. By contrast, we found that both the SHLD2^{FAM35A} N- and C-termini are

341   important for IR resistance (in BRCA1-proficient cells). As IR sensitivity in Shieldin-

342   deficient cells likely reflects impaired NHEJ, we speculate that Shieldin potentiates NHEJ by

343   restricting DSB resection as well as by assembling with other NHEJ-promoting factors to

344   tether DSB ends together to facilitate their juxtaposition and repair.

345

346　Notably, we have found that SHLD1^{C20orf196} or SHLD2^{FAM35A} inactivation confers enhanced

347　cisplatin sensitivity to BRCA1-null or BRCA1-proficient cells. Such sensitivity probably

348　does not reflect the roles for Shieldin in promoting NHEJ, restricting DNA-end resection or

349　antagonising BRCA1-mediated BRCA2/RAD51 loading because, in our hands, 53BP1 loss

350　does not have pronounced effects on cisplatin sensitivity. Intra-strand DNA crosslinks (ICLs)

351　generated by cisplatin and other compounds are detected and repaired by the Fanconi anemia

352　(FA) pathway, with a key FA protein being FANCD2, which forms foci at sites of these

353　lesions[44]. We have observed that following cisplatin treatment, FANCD2 foci were more

354　pronounced in cells in which SHLD1^{C20orf196} or SHLD2^{FAM35A} was inactivated. It will thus be

355　of interest to determine if Shieldin − like MAD2L2, which (with REV3L) functions as a

356　regulatory subunit of the trans-lesion DNA synthesis (TLS) polymerase Pol zeta, and whose

357　biallelic inactivation has been associated with FA[45-47] − might also promote ICL repair via

358　TLS mechanisms.

359

360　Finally, based on our findings, it will be interesting to evaluate SHLD1/2 expression in

361　tumour biopsies from patients, establish whether this information can be used in patient

362　stratification for PARP-inhibitor therapies, and determine whether SHLD1/2 expression

363　changes arise in patients whose BRCA1-deficient cancers develop resistance after PARP-

364　inhibitor treatment. In this regard, we note that if down-regulation of Shieldin components in

365　BRCA1-deficient cancers does confer clinical resistance, this might allow alternative

366　treatments, such as ones based on platinum compounds.

367

368　Acknowledgements

Author contributions

407 TWC and SPJ conceived the project and TWC initiated the project by performing the
408 CRISPR-Cas9 screens, with MH doing the bioinformatic analyses. LD supervised, and CL
409 and WW performed class switch recombination and *Igh* locus instability experiments. JJ
410 supervised, and IK performed recombinant MAD2L2 co-IPs, MAD2L2 IRIF/IB, pS4/8-
411 RPA2 IB and telomere fusion experiments. MO, PB performed mass-spectrometry. JC
412 performed and analysed clonogenic survival experiments and random plasmid integration
413 assay; MD and MS generated human knockout cell lines and MS performed *in vitro*
414 pulldown experiments; JL carried out oligonucleotide interaction studies; DP performed end-
415 resection assays; TWC and RB generated RPE1 p53ko, RPE1 p53ko/BRCAko and
416 p53ko/BRCA1ko/53BP1ko cell lines; YG and SPJ supervised the above. GB performed IR
417 survivals in mouse cells, mouse sgRNA cloning, and FISH with FY and BF. LP performed
418 structural analysis and FAM35A modelling. QW performed purification of bacterially
419 expressed recombinant FAM35A proteins and EMSA. AM, AS, AB and CC performed
420 patient-derived xenograft experiments on PARPi-induced resistance. VS, MO'C and ZL
421 established, performed, analysed and characterised PDXs in the experiments on intrinsic
422 PARPi resistance. HD assisted with many of the above, and devised and performed all other

423 experiments. HD, YG and SPJ wrote the manuscript with input from all others. LD, JJ, YG

424 and SPJ supervised the project.

425

Competing financial interests

431

References

1. Goldstein, M. & Kastan, M.B. The DNA damage response: implications for tumor responses to radiation and chemotherapy. *Annu Rev Med* **66**, 129-143 (2015).
2. Jackson, S.P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071-1078 (2009).
3. Schwertman, P., Bekker-Jensen, S. & Mailand, N. Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nat Rev Mol Cell Biol* **17**, 379-394 (2016).
4. Blackford, A.N. & Jackson, S.P. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell* **66**, 801-817 (2017).
5. Bekker-Jensen, S. & Mailand, N. Assembly and function of DNA double-strand break repair foci in mammalian cells. *DNA Repair (Amst)* **9**, 1219-1228 (2010).
6. Boersma, V. *et al.* MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5' end resection. *Nature* **521**, 537-540 (2015).
7. Bunting, S.F. *et al.* 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243-254 (2010).
8. Chapman, J.R. *et al.* RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol Cell* **49**, 858-871 (2013).
9. Daley, J.M. & Sung, P. RIF1 in DNA break repair pathway choice. *Mol Cell* **49**, 840-841 (2013).
10. Di Virgilio, M. *et al.* Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. *Science* **339**, 711-715 (2013).
11. Escribano-Diaz, C. *et al.* A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol Cell* **49**, 872-883 (2013).
12. Xu, G. *et al.* REV7 counteracts DNA double-strand break resection and affects PARP inhibition. *Nature* **521**, 541-544 (2015).
13. Zimmermann, M., Lottersberger, F., Buonomo, S.B., Sfeir, A. & de Lange, T. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science* **339**, 700-704 (2013).

463 14. Capoluongo, E. *et al.* Guidance Statement On BRCA1/2 Tumor Testing in Ovarian
464 Cancer Patients. *Semin Oncol* **44**, 187-197 (2017).

465 15. Lord, C.J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science*
466 **355**, 1152-1158 (2017).

467 16. Bryant, H.E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of
468 poly(ADP-ribose) polymerase. *Nature* **434**, 913-917 (2005).

469 17. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a
470 therapeutic strategy. *Nature* **434**, 917-921 (2005).

471 18. Lord, C.J. & Ashworth, A. Mechanisms of resistance to therapies targeting BRCA-
472 mutant cancers. *Nat Med* **19**, 1381-1388 (2013).

473 19. Basu, B., Sandhu, S.K. & de Bono, J.S. PARP inhibitors: mechanism of action and their
474 potential role in the prevention and treatment of cancer. *Drugs* **72**, 1579-1590
475 (2012).

476 20. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells.
477 *Science* **343**, 84-87 (2014).

478 21. Simonetta, M. *et al.* H4K20me2 distinguishes pre-replicative from post-replicative
479 chromatin to appropriately direct DNA repair pathway choice by 53BP1-RIF1-
480 MAD2L2. *Cell Cycle*, 1-13 (2018).

481 22. Lo, K.W. *et al.* The 8-kDa dynein light chain binds to p53-binding protein 1 and
482 mediates DNA damage-induced p53 nuclear accumulation. *J Biol Chem* **280**, 8172-
483 8179 (2005).

484 23. Giraud-Panis, M.J., Teixeira, M.T., Geli, V. & Gilson, E. CST meets shelterin to keep
485 telomeres in check. *Mol Cell* **39**, 665-676 (2010).

486 24. Sun, J. *et al.* Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev* **23**,
487 2900-2914 (2009).

488 25. Scully, R. & Xie, A. Double strand break repair functions of histone H2AX. *Mutat Res*
489 **750**, 5-14 (2013).

490 26. Kakarougkas, A. & Jeggo, P.A. DNA DSB repair pathway choice: an orchestrated
491 handover mechanism. *Br J Radiol* **87**, 20130685 (2014).

492 27. Galanty, Y. *et al.* Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses
493 to DNA double-strand breaks. *Nature* **462**, 935-939 (2009).

494 28. Feldman, S. *et al.* 53BP1 Contributes to Igh Locus Chromatin Topology during Class
495 Switch Recombination. *J Immunol* **198**, 2434-2444 (2017).

496 29. Nakamura, M. *et al.* High frequency class switching of an IgM+ B lymphoma clone
497 CH12F3 to IgA+ cells. *Int Immunol* **8**, 193-201 (1996).

498 30. Manis, J.P. *et al.* 53BP1 links DNA damage-response pathways to immunoglobulin
499 heavy chain class-switch recombination. *Nat Immunol* **5**, 481-487 (2004).

500 31. Ward, I.M. *et al.* 53BP1 is required for class switch recombination. *J Cell Biol* **165**,
501 459-464 (2004).

502 32. Lieber, M.R. Mechanisms of human lymphoid chromosomal translocations. *Nat Rev*
503 *Cancer* **16**, 387-398 (2016).

504 33. Konishi, A. & de Lange, T. Cell cycle control of telomere protection and NHEJ
505 revealed by a ts mutation in the DNA-binding domain of TRF2. *Genes Dev* **22**, 1221-
506 1230 (2008).

507 34. Dimitrova, N., Chen, Y.C., Spector, D.L. & de Lange, T. 53BP1 promotes non-
508 homologous end joining of telomeres by increasing chromatin mobility. *Nature* **456**,
509 524-528 (2008).

510 35. Ochs, F. *et al.* 53BP1 fosters fidelity of homology-directed DNA repair. *Nat Struct Mol*
511   *Biol* **23**, 714-721 (2016).
512 36. Cruz-Garcia, A., Lopez-Saavedra, A. & Huertas, P. BRCA1 accelerates CtIP-mediated
513   DNA-end resection. *Cell Rep* **9**, 451-459 (2014).
514 37. Gudmundsdottir, K. & Ashworth, A. The roles of BRCA1 and BRCA2 and associated
515   proteins in the maintenance of genomic stability. *Oncogene* **25**, 5864-5874 (2006).
516 38. Certo, M.T. *et al.* Tracking genome engineering outcome at individual DNA
517   breakpoints. *Nat Methods* **8**, 671-676 (2011).
518 39. Schmidt, C.K. *et al.* Systematic E2 screening reveals a UBE2D-RNF138-CtIP axis
519   promoting DNA repair. *Nat Cell Biol* **17**, 1458-1470 (2015).
520 40. Bruna, A. *et al.* A Biobank of Breast Cancer Explants with Preserved Intra-tumor
521   Heterogeneity to Screen Anticancer Compounds. *Cell* **167**, 260-274 e222 (2016).
522 41. Cruz, C. *et al.* RAD51 foci as a functional biomarker of homologous recombination
523   repair and PARP inhibitor resistance in germline BRCA mutated breast cancer. *Ann*
524   *Oncol* (2018).
525 42. Bunting, S.F. *et al.* BRCA1 functions independently of homologous recombination in
526   DNA interstrand crosslink repair. *Mol Cell* **46**, 125-135 (2012).
527 43. Gupta, R. *et al.* DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of
528   NHEJ and PARP Inhibitor Sensitivity. *Cell* **173**, 972-988 e923 (2018).
529 44. Ceccaldi, R., Sarangi, P. & D'Andrea, A.D. The Fanconi anaemia pathway: new players
530   and new functions. *Nat Rev Mol Cell Biol* **17**, 337-349 (2016).
531 45. Makarova, A.V. & Burgers, P.M. Eukaryotic DNA polymerase zeta. *DNA Repair (Amst)*
532   **29**, 47-55 (2015).
533 46. Tomida, J. *et al.* REV7 is essential for DNA damage tolerance via two REV3L binding
534   sites in mammalian DNA polymerase zeta. *Nucleic Acids Res* **43**, 1000-1011 (2015).
535 47. Bluteau, D. *et al.* Biallelic inactivation of REV7 is associated with Fanconi anemia. *J*
536   *Clin Invest* **126**, 3580-3584 (2016).
537

Figure legends

Figure 1. CRISPR-Cas9 screens identify suppressors of PARP-inhibitor sensitivity in *BRCA1*-mutant cells. a, Schematic of screen procedure. **b,** MAGeCK analysis of guide enrichments following specified drug treatments; false discovery rate (FDR) of 0.1 indicated by dotted line; n=3 technical replicates per drug treatment. **c,** siRNA mediated verification of hits in clonogenic survival assays; lower panels show area under the curve (AUC); n=3 independent experiments **d,** De novo Cas9 mediated knockout (ko) verification and complementation for FAM35A in clonogenic survival assays (multiple ko clones are shown in AUC); n=4 independent experiments except FAM35Ako(#14) (n=2), FAM35Ako(#40) (n=3), BRCA1ko/FAM35Ako(#34) (n=2), and BRCA1ko/FAM35Ako(#2) +FAM35A (n=3). **e,** As (d) but for C20orf196; n=3 independent experiments except BRCA1ko/C20orf196ko + C20orf196 (n=2). **c-e** Bars represent mean ± SEM, one-way Anova; *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns=not significant (p≥0.05). Individual data points are plotted

551   over bars, and statistical source data including the precise p values can be found in

552   Supplementary Table 5.

553

554   **Figure 2. FAM35A and C20orf196 domains, interactions and IRIF formation. a,**

555   FAM35A and C20orf196 predicted domains and variants used, OB fold (OB), FAM domain

556   (OB3/FD). **b,** Recruitment of FAM35A/derivatives GFP-fusions to a chromosomal Lac-

557   operator array via mCherry-LacR-C20orf196. Data shown represent 3 experiments with

558   quantifications shown in Supplementary Fig 2a. Scale bar 10μm. **c,** (left and middle panel)

559   Purified recombinant GST-FAM35A directly interacts with recombinant His-C20orf196. c,

560   (right panel) Cell extracts expressing GFP-FAM35A/derivatives and HA-C20orf196 analysed

561   by co-immunoprecipitation and immunoblotting. **d,** V5-FAM35A co-immunoprecipitates

562   with GFP-MAD2L2; interaction with C20orf196 shown in Supplementary Fig 2c. **e,**

563   Quantification of inducible GFP-FAM35A (left panel) and GFP-C20orf196 (right panel)

564   IRIF in γH2AX positive cells 5 h after IR (5Gy) treated with indicated siRNAs. N=4

565   independent experiments except (left panel) si53BP1 (n=3), siRIF1 and siMAD2L2 (n=2);

566   and (right panel) siCTRL(n=5), siRIF1(n=3), siFAM35A(n=3). **f,** As in (e) but for inducible

567   GFP-FAM35A N-terminus; n=4 independent experiments except siRIF1 (n=3). **g,**

568   Endogenous MAD2L2 co-immunoprecipitates with GFP-FAM35A N-terminus. **e-f,** Bars

569   represent mean ± SEM, one-way Anova; *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$,

570   ns=not significant ($p\geq0.05$); individual data points plotted over bars. Statistical source data

571   including the precise p values are shown in Supplementary Table 5. All immunoblots are

572   representative of two independent experiments; unprocessed scans of immunoblots are shown

573   in Supplementary Fig 8.

574

575   **Figure 3. FAM35A and C20orf196 promote NHEJ and immunoglobulin class-switch**

576   **recombination. a,** Random plasmid integration assay. **b,** FAM35Ako and C20orf196ko cells

577   were treated with IR and analysed for clonogenic survival, right panel shows AUC. **a-b,** Bars

578   represent mean ± SEM, one-way Anova; n=3 independent experiments, except C20orf196ko

579   in b (n=4), with individual data points plotted over bars; statistical source data can be found

580   in Supplementary Table 5. **c,** Schematic representation of class-switch recombination and

581   chromosomal instability in murine IgM$^+$ B cells (germline configuration with C$_\mu$

582   transcription) induced to express AID and undergo CSR to IgA (switch configuration with C$_\alpha$

583   transcription) upon addition of anti-CD40, IL4 and TGF-β. CSR levels are measured as the %

584 of IgA positive cells after 72 h cytokine stimulation, and DNA fluorescence in situ

585 hybridization (FISH) is performed using a chromosome 12-specific paint (grey chromosome)

586 and *Igh* locus specific probes (red and green spots) for the measurement of chromosomal

587 instability at the *Igh* locus upon induction of CSR. **d,** CSR levels in Fam35Ako and

588 C20orf196ko CH12-Cas9 cells are reduced compared with wild-type (WT) CH12-Cas9 cells.

589 Bars represent mean ± SEM, one-way Anova. N=4 independent experiments of 3 clones

590 except 53BP1ko +cytokine where n=3 of 2 clones, and 53BP1ko -cytokine where n=2 of 2

591 clones; with individual data points plotted over bars. **e,** Representative images of *Igh*

592 translocation and breaks in aberrant metaphases, quantified in f. **f,** Quantification of *Igh*

593 breaks and translocations in metaphases of the indicated CH12-Cas9 cells. Horizontal bars

594 represent means, Fisher's Exact test; n=2 independent experiments except Fam35ako and

595 C20orf196ko where n=3. For a, b, d and f, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001,

596 ns=not significant (p≥0.05); statistical source data including the precise p values for these

597 panels can be found in Supplementary Table 5.

598

599 **Figure 4. FAM35A and C20orf196 promote telomere-mediated fusions and limit DNA-**

600 **end resection. a,** Schematic of TRF2ts experimental setup. **b,** shRNA depletion of FAM35A

601 (left panel) or C20orf196 (right panel) reduces un-capped telomere-mediated chromosome

602 fusions. Bars represent means. The experiments were performed twice with ≥1300

603 chromosomes counted per condition, and individual data points plotted over bars; source data

604 can be found in Supplementary Table 5. **c,** FAM35Ako and C20orf196ko RPE1 cells labelled

605 with BrdU (10μM) for 48 h then treated with 1μM camptothecin (CPT) for 1 h, pre-extracted,

606 fixed and stained for BrdU under non-denaturing conditions to visualise ssDNA. Box and

607 whisker plot with centre line at median, box limits at $25^{th}/75^{th}$ centiles and whiskers

608 ±1.5xIQR; one-way Anova; n=3 independent experiments. **d,** IR-induced pRPA(S4/8) is

609 enhanced in MEFs due to Fam35a or C20orf196 silencing. Bars represent means. The

610 experiments were performed twice with individual data points plotted over bars; source data

611 can be found in Supplementary Table 5. **e,** RPE1-FAM35Ako or -C20orf196ko cells display

612 hyper DNA-end resection (cells treated with 1μM camptothecin for 1h). Representative

613 images from 3 independent experiments. Scale bar 10μm. **f,** RPE1-FAM35Ako or -

614 C20orf196ko cells display BLM and CtIP dependent markers of excessive DNA-end

615 resection. Box and whisker plot with centre line at median, box limits at $25^{th}/75^{th}$ centiles and

616 whiskers ±1.5xIQR; one-way Anova; n=3 independent experiments. **g,** Enhanced BLM

617 accrual in FAM35Ako and C20orf196ko compared with wild-type (WT) RPE1 cells fixed
618 and stained 2 h after laser micro-irradiation. Representative images shown in left panel and
619 quantification in right panel. Scale bar 10μm. Box and whisker plot with centre line at
620 median, box limits at $25^{th}/75^{th}$ centiles and whiskers ±1.5xIQR; one-way Anova; n=3
621 independent experiments. For c, f and g, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001,
622 ns=not significant (p≥0.05); statistical source data including the precise p values can be found
623 in Supplementary Table 5.

624

625 **Figure 5. FAM35A OB folds mediate ssDNA interaction and is required for IR**
626 **resistance. a,** Schematic of FAM35A with residues W489/W640 mutated to A (top panel).
627 Predicted 3D structure of wild-type FAM35A with W489 and W640 positions (lower left
628 panel). FAM35A W489/W640 promote efficient ssDNA binding in cellular extracts (lower
629 right panel). **b,** Alignment of yRPA1 with FAM35A C-terminus; amino acids critical for
630 yRPA1 ssDNA binding and the corresponding amino acid residues in FAM35A are boxed. **c,**
631 EMSAs on native (non-denaturing) gels with 10nM ssDNA or dsDNA, and the indicated
632 purified, bacterially expressed FAM35A C-terminus or W489/W640A mutant in μM. **d,**
633 Inducible GFP-FAM35A W489/W640A fails to efficiently form IRIF (12 h after 5Gy of IR).
634 Scale bar 10μm. Representative images from 3 independent experiments. **e,** FAM35Ako
635 RPE1 cells complemented with FAM35A derivatives in clonogenic survival assays; right
636 panel shows AUC. **f,** Overexpression of FAM35A N-terminus but not C-terminus or full-
637 length FAM35A sensitises wild-type cells to IR in clonogenic assays; right panel shows
638 AUC. **e-f,** Bars represent mean ± SEM, one-way Anova; *p<0.05, **p<0.01, ***p<0.001,
639 ****p<0.0001, ns=not significant (p≥0.05); n=3 independent experiments except group 2 and
640 4 in e (n=2), with individual data points plotted over bars; statistical source data including the
641 precise p values can be found in Supplementary Table 5. All immunoblots are representative
642 of two independent experiments; unprocessed scans of immunoblots are shown in
643 Supplementary Fig 8.

644

645 **Figure 6. FAM35A or C20orf196 loss restores HR in BRCA1-deficient cells. a,**
646 Quantification of GFP-FAM35A (left panel) and GFP-C20orf196 (right panel) IRIF in U2OS
647 cells after BRCA1 or BRCA2 depletion (5 h after 5Gy). Bars represent mean ± SEM, one-
648 way Anova; n=3 independent experiments, except FAM35A siCTRL (n=4), FAM35A
649 siBRCA2 (n=2), and C20orf196 siCTRL (n=5); with individual data points plotted over bars.

650   **b,** Quantification of 53BP1 and inducible GFP-FAM35A IRIF in U2OS cells with or without

651   BRCA1 depletion (5Gy, indicated time points). Bars represent mean ± SEM, one-way

652   Anova; n=4 independent experiments, except 53BP1 1.5h siCTRL (n=2), 53BP1 1.5h

653   siBRCA1 and 53BP1 16h siCTRL (n=3), FAM35A 1.5h siCTRL (n=5); with individual data

654   points plotted over bars. **c,** Representative images (left panel) and quantification (right panel)

655   of RAD51 IRIF (5.5 h after 5Gy) in Cyclin A (CycA) positive RPE1ko cell lines as indicated.

656   Bars represent mean ± SEM, one-way Anova; n=3 independent experiments, with individual

657   data points plotted over bars. Scale bar 10μm. **d,** FAM35A/C20orf196 loss restore BRCA2

658   recruitment 2 h after laser-induced DNA-damage sites in *BRCA1*-null cells (for quantification

659   see Supplementary Fig 6e). Scale bar 10μm. **e**, HR assay in U2OS-TLR cells treated with

660   indicated siRNAs (for gating strategy see Supplementary Fig 6f). Bars represent mean ±

661   SEM, one-way Anova; n=4 independent experiments, with individual data points plotted over

662   bars. **f,** Formation of spontaneous chromosomal aberrations in BRCA1ko cells is alleviated

663   by FAM35A/C20orf196 inactivation. Representative images of metaphase spreads shown,

664   and quantified in graph; bars represent means, n=2 independent experiments except

665   FAM35Ako and C20orf196ko (n=1), with individual data points plotted over bars. **g**,

666   Olaparib clonogenic survival assay with indicated RPE1ko and complemented cell lines. Bars

667   represent mean ± SEM, one-way Anova; n=4 independent experiments, except group 4 and 5

668   (n=3) and group 3 (n=2); AUC is shown in Supplementary Fig 6g. For a-c and e, *p<0.05,

669   **p<0.01, ***p<0.001, ****p<0.0001, ns=not significant (p≥0.05); statistical source data

670   including the precise p values can be found in Supplementary Table 5.

671

672   **Figure 7. FAM35A or C20orf196 loss correlates with PARP inhibitor resistance in**

673   **cancers. a,** Schematic of in vivo PDX study (top panel). Heat map generated from mRNA-

674   sequencing showing scaled expression levels of indicted genes from corresponding PDX

675   samples (lower panel); n=6, 5, 7, 8 mice for cohorts 1-4 respectively. **b**, Expression of

676   C20orf196/FAM35A in breast and ovarian cancer PDXs derived from BRCA1-deficient

677   tumours. y-axis: log2 transcript per million. Lines represent mean ± SEM; n=12, 4, 15, 1 for

678   SHLD1-high, SHLD1-low, SHLD2-high, SHLD2-low groups respectively; two-tailed

679   unpaired student t-test; ***p=0.0003. Statistical source data for PDXs can be found in

680   Supplementary Table 5 and methods. **c-d,** Clonogenic survival assay after IR (c) or cisplatin

681   treatment (d) in the indicated RPE1ko cell lines with AUC shown in Supplementary Fig 7b

682   and 7c, respectively. Data shown represent mean ± SEM (n = 3 independent experiments

683   except for group 7 in c and group 7 in d where n = 2) **e**, Loss of FAM35A/C20orf196 leads to

684   increased cisplatin-induced FANCD2 foci. Bars represent mean ± SEM, one-way Anova;

685   *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$, ns=not significant ($p\geq0.05$); n=4

686   independent experiments, with individual data points plotted over bars; statistical source data

687   can be found in Supplementary Table 5. Scale bar 10μm. **f,** Proposed model for the action of

688   SHLD1/2 in DSB repair in the presence or absence of functional BRCA1.

689

690

Figure 1

Figure 2

Figure 3

a

b

c

d

e

f

Figure 4

Figure 5

a

FAM35A (W489/W640A)

1 ... 489 ... 640 ... 904

OB1 OB2 OB3/FD

Zinc-binding (OB3/FD)

W640

W489

OB2

OB1

FAM35A C-terminus

b

c

d

e

f

Figure 6

Figure 7

Genotypes of knockouts confirmed by topo-clone sequencing
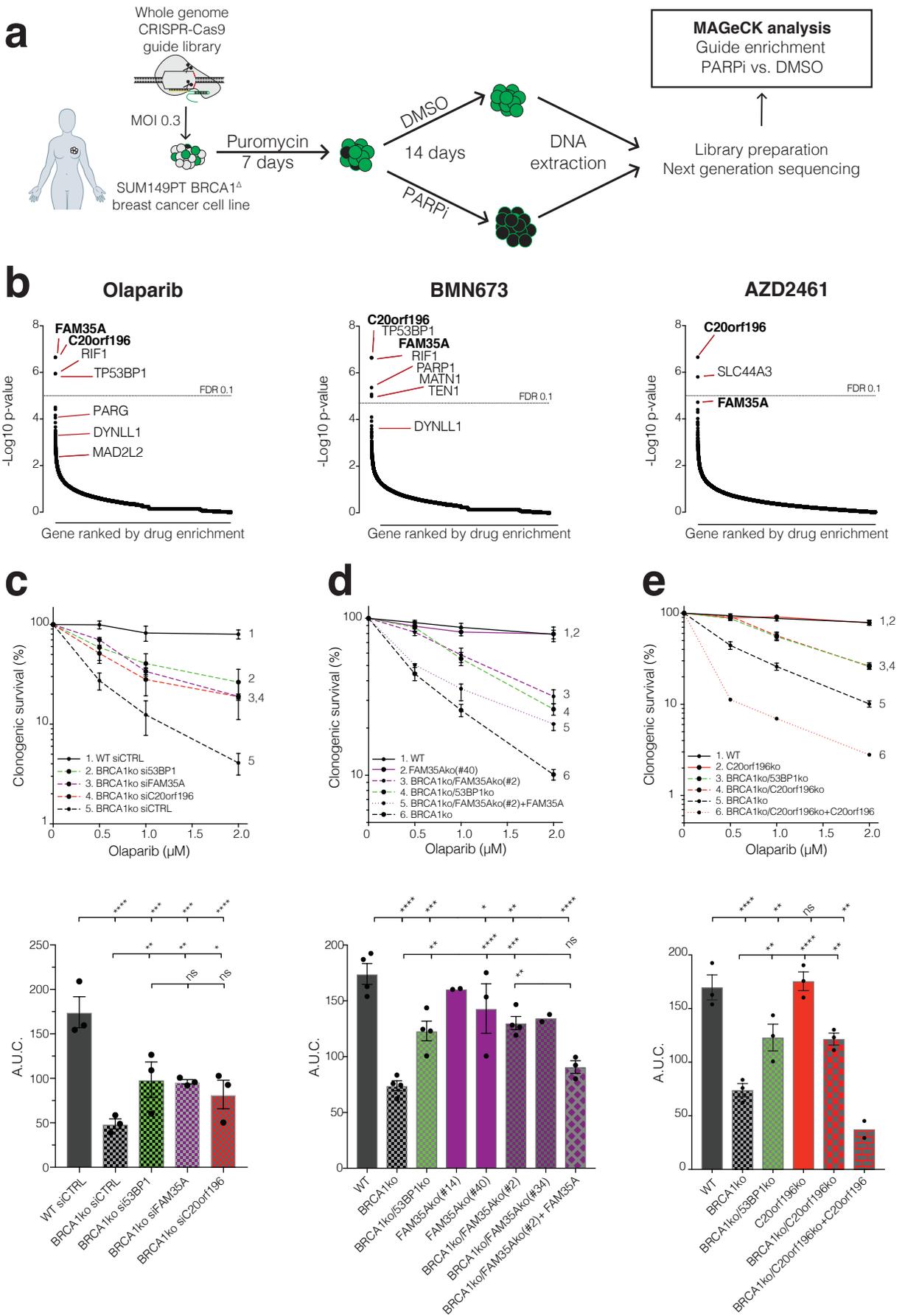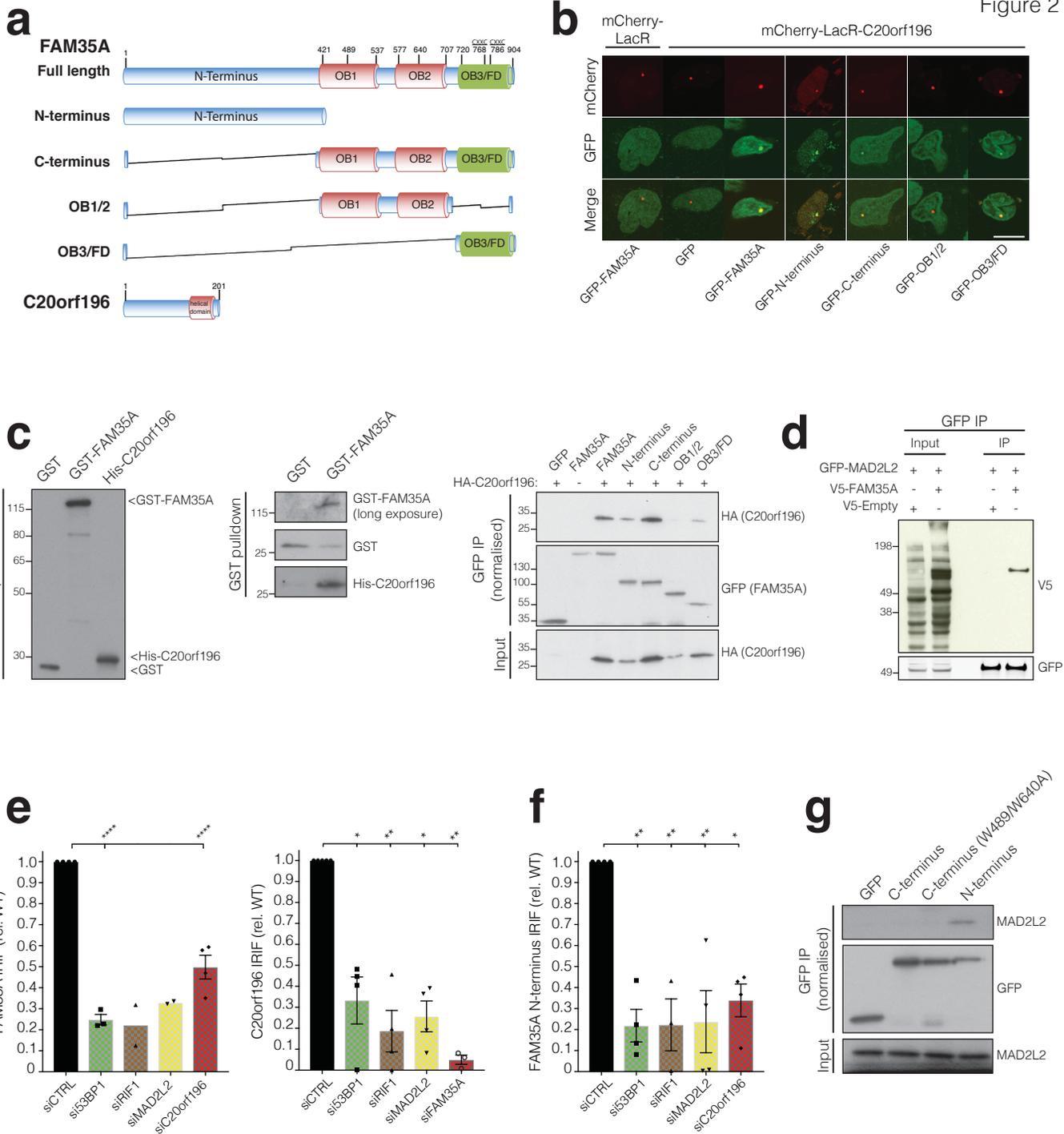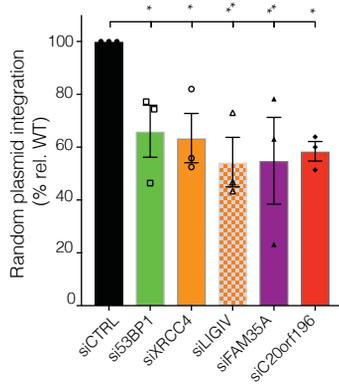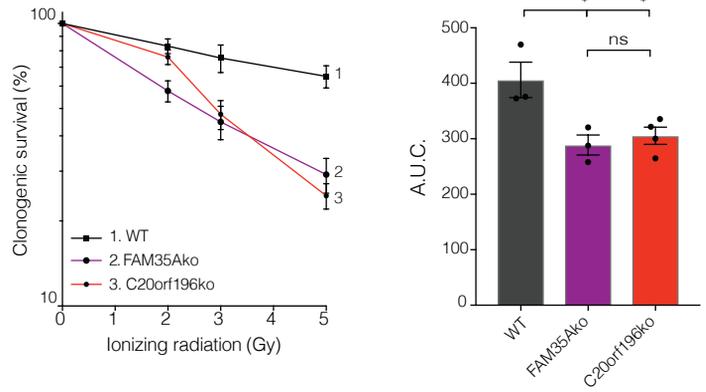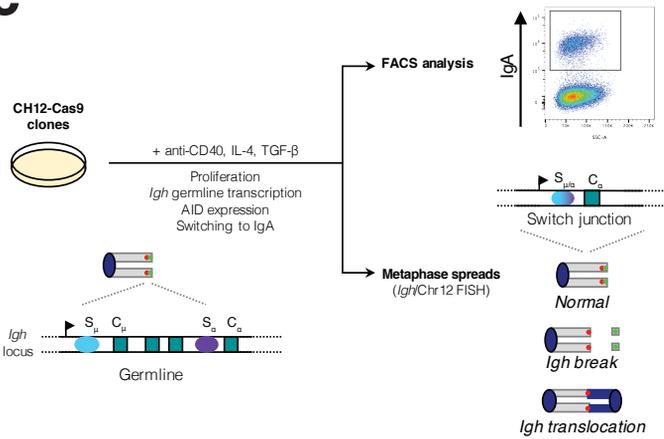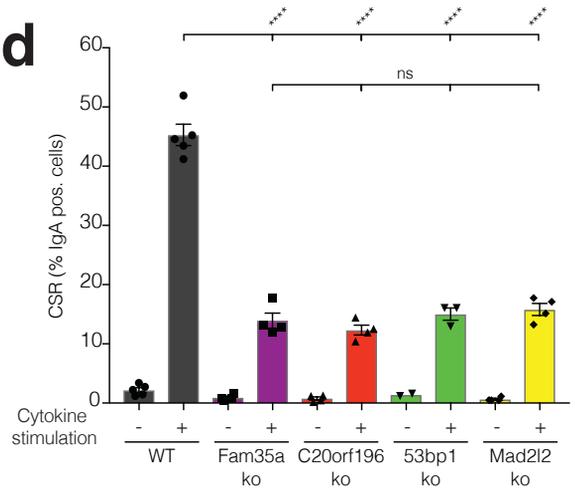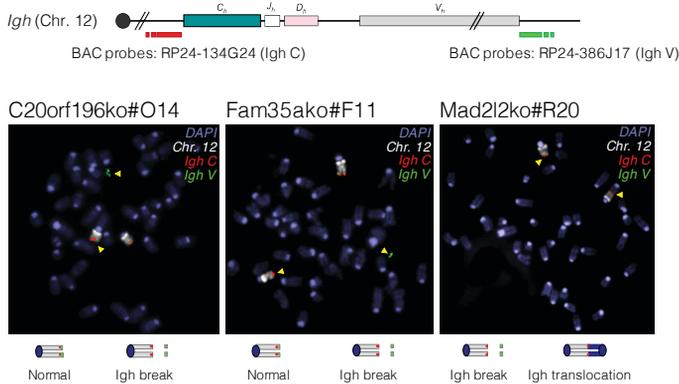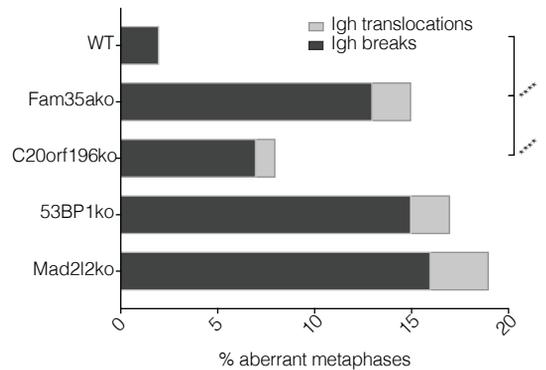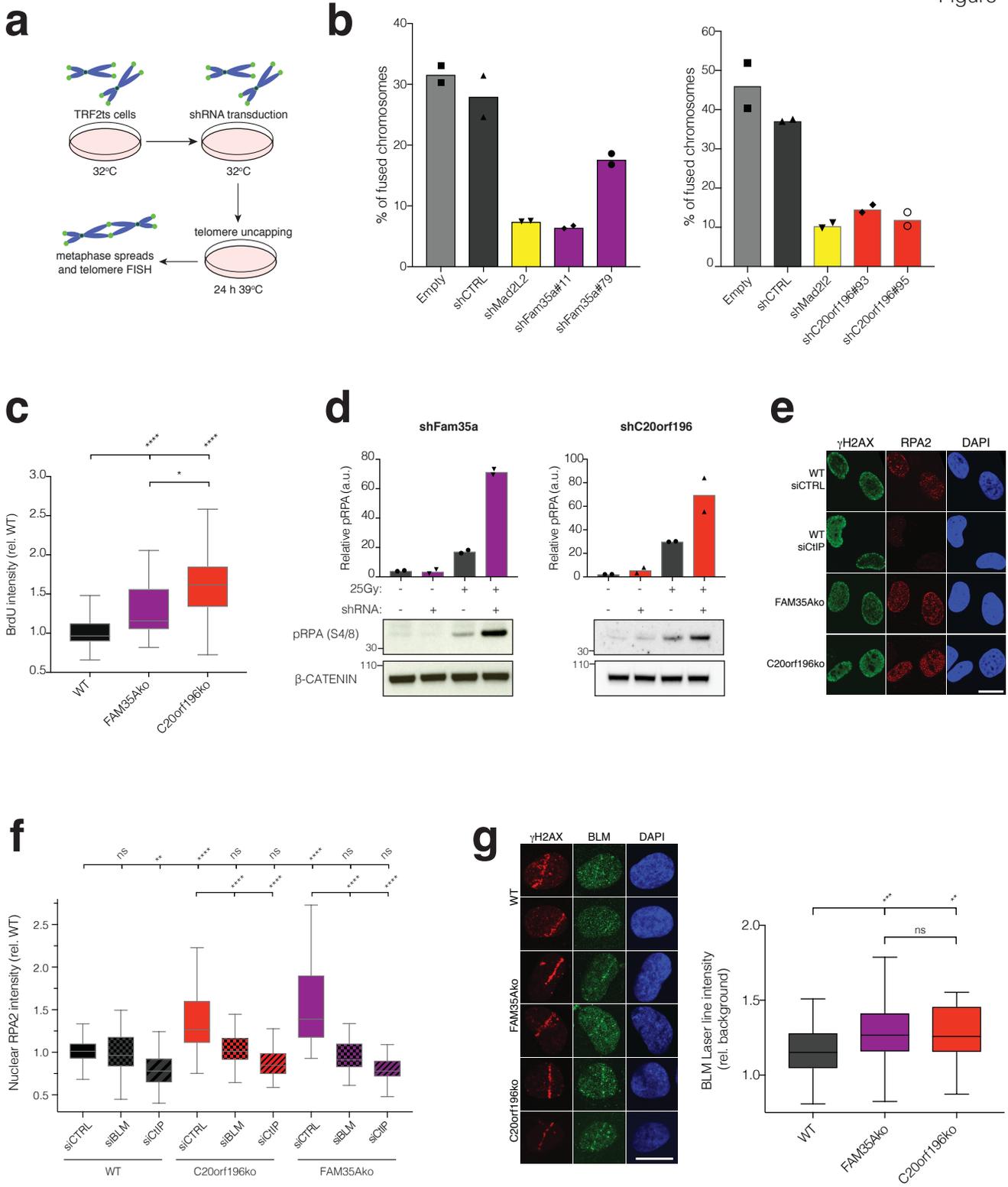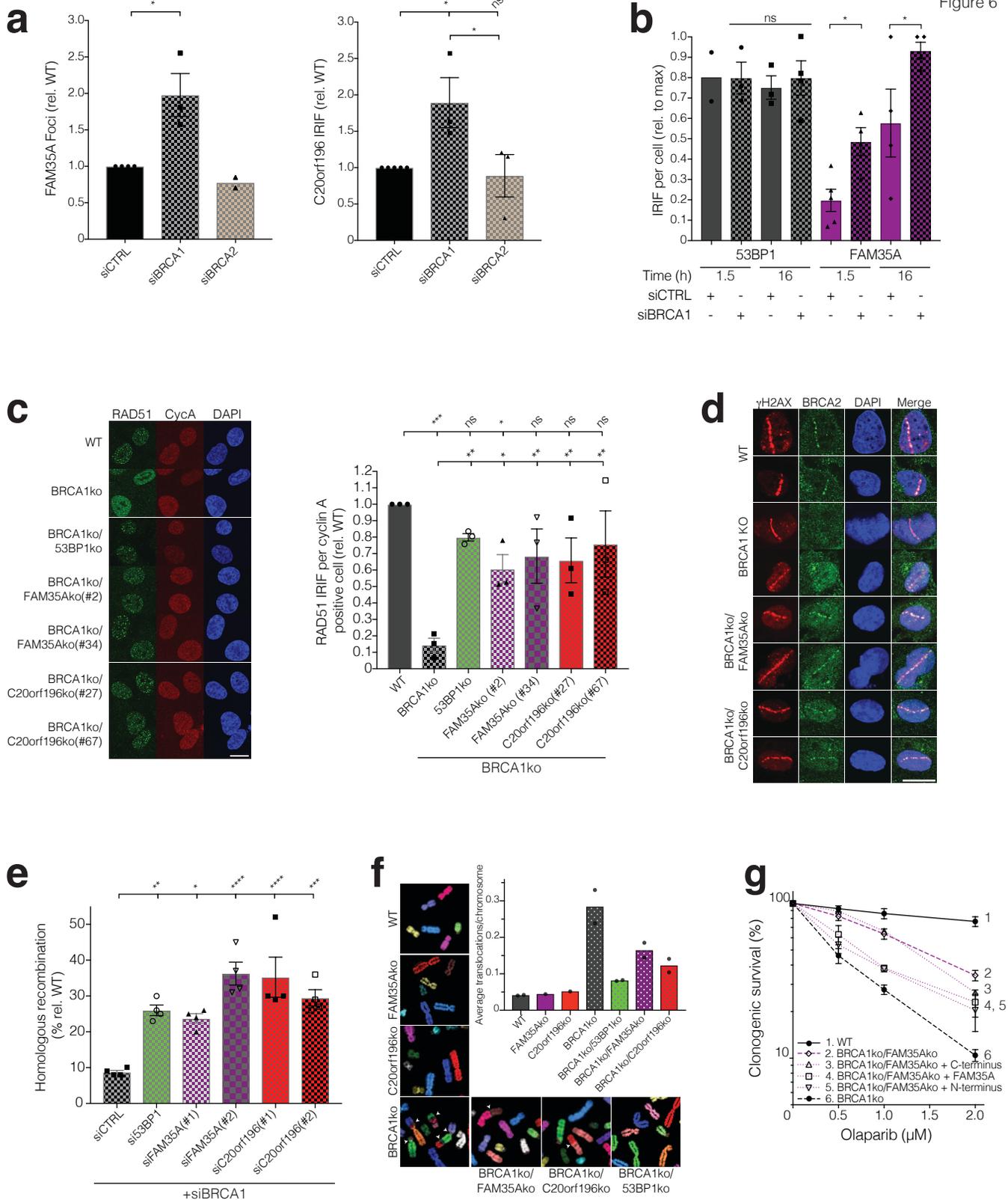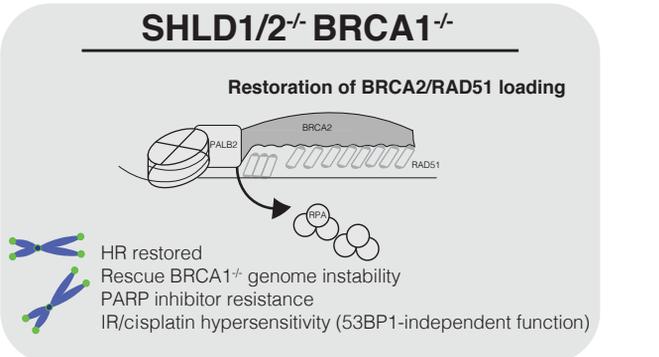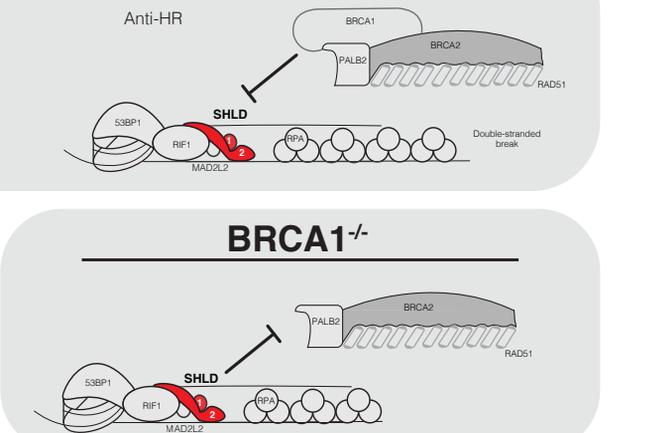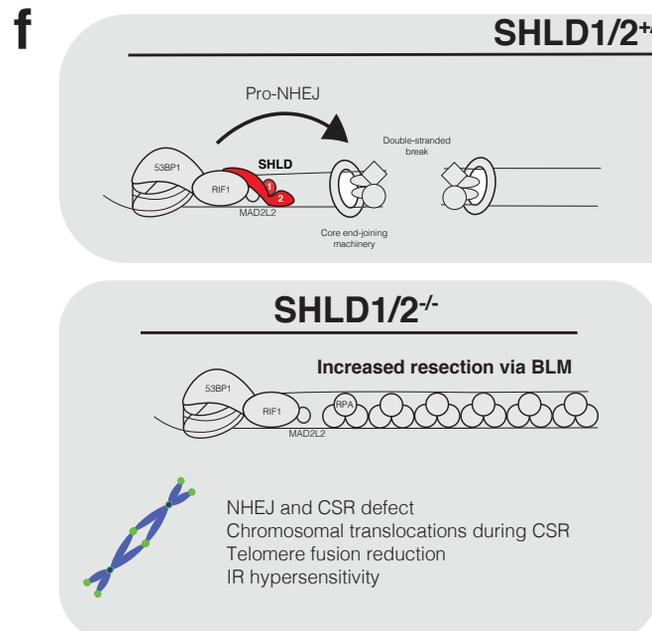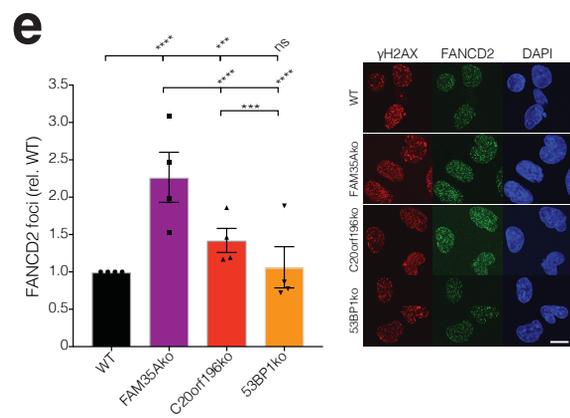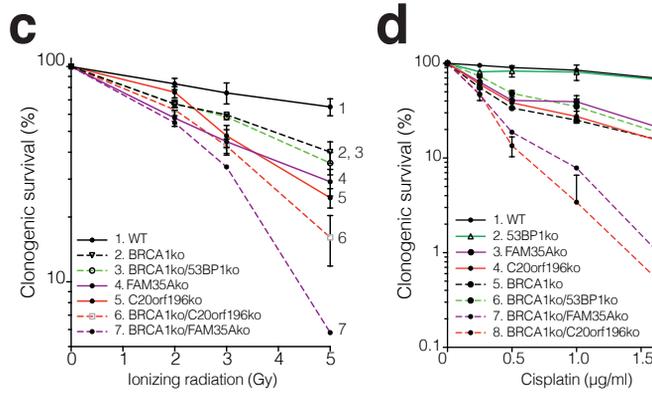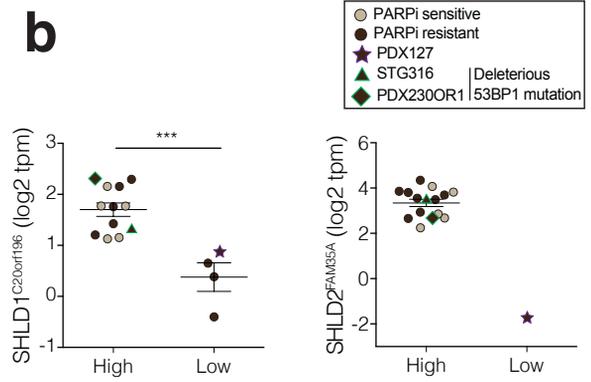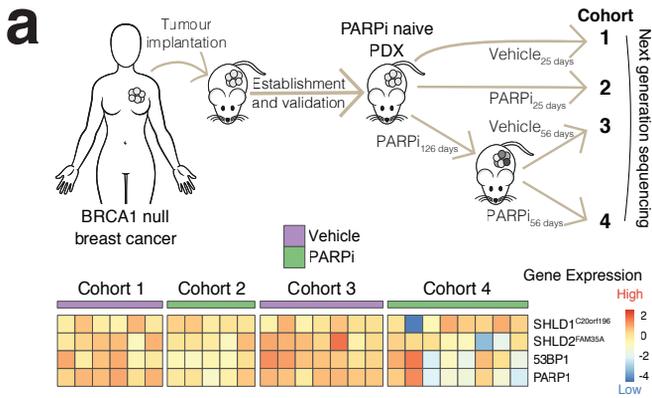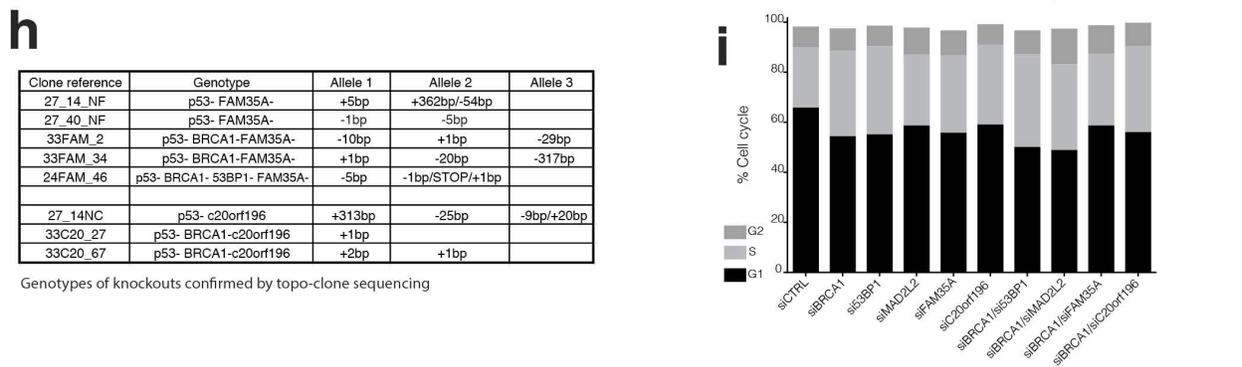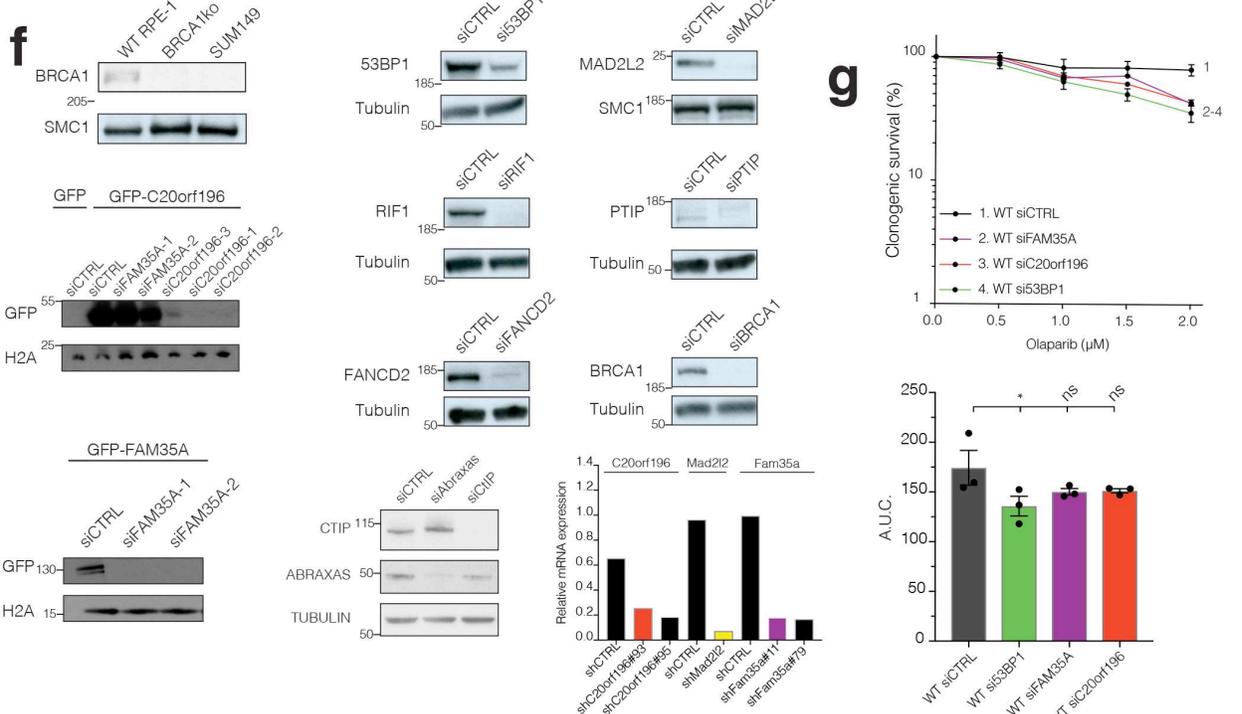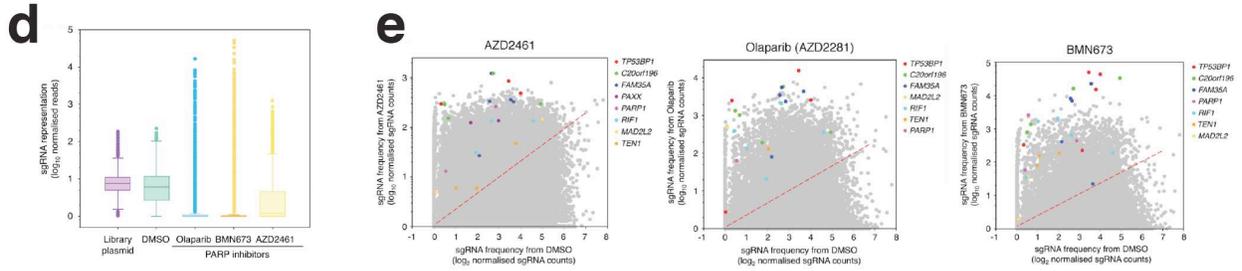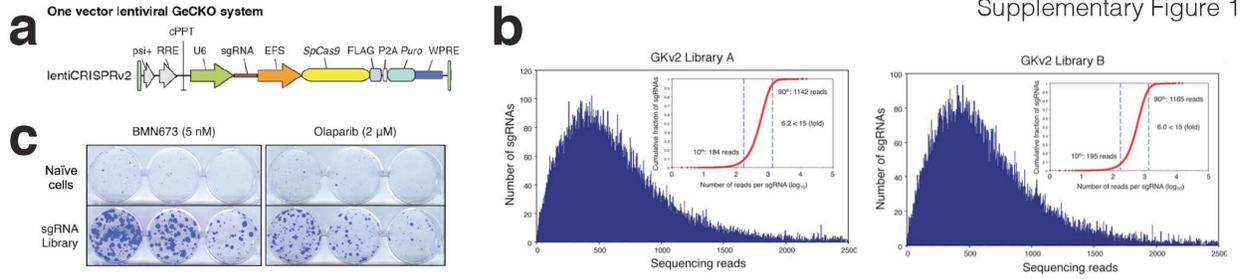
Supplementary Figure 1

Whole Genome CRISPR screen data and validation studies

a, Schematic of the one vector lentiviral GeCKOv2 system. **b,** Histograms of sgRNA representation of GeCKOv2 (GKv2) library A (left panel) and B (right panel). Inset: cumulative distribution of sequencing reads. The number of sequencing reads for the 10th and 90th sgRNA percentiles are indicated by the dashed vertical blue lines and text labels. The representation of sgRNAs is indicated by the fold-difference between the 10th and 90th percentile. **c,** Representative surviving clones after treatment with PARP inhibitors, representative of 2 independent experiments. Top panel: naïve, un-transduced SUM149PT cells; bottom panel: GeCKOv2 library-transduced cells. **d,** Distributions of sgRNA frequencies arising in different conditions; Box and whisker plot with centre line at median, box limits at 25th/75th centiles and whiskers ±1.5xIQR; n=3 technical replicates. **e,** sgRNA enrichments after treatments with the indicated drugs; each dot represents one sgRNA, with coloured dots representing the indicated target genes. **f,** Verification of BRCA1 mutant SUM149PT, BRCA1ko RPE1 and siRNAs and shRNAs used in this paper, by immunoblot or RT-qPCR (bars represent means; one experiment performed in triplicates). **g,** Clonogenic survival assay using the indicated siRNAs in BRCA1-proficient cells (WT); lower panel shows AUC. Bars represent mean ± SEM, one-way Anova; *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns=not significant (p≥0.05); n=3 independent experiments, with individual data points plotted over bars; statistical source data including the precise p values can be found in Supplementary Table 5. **h,** Genotypes of human knockout clones used in this work confirmed by Topo-cloning and Sanger sequencing. **i,** Cell cycle profiles of cells transfected with the indicated siRNAs used in this work (bars represent means derived from two independent experiments). All immunoblots are representative of two independent experiments; unprocessed scans of immunoblots are shown in Supplementary Fig 8.
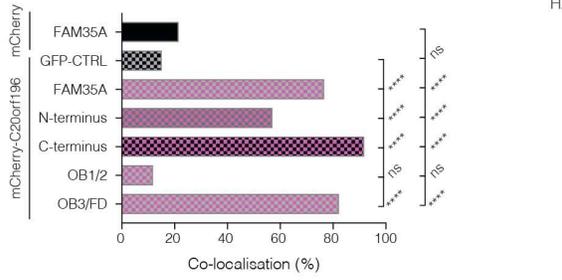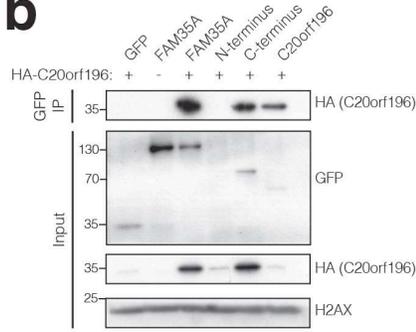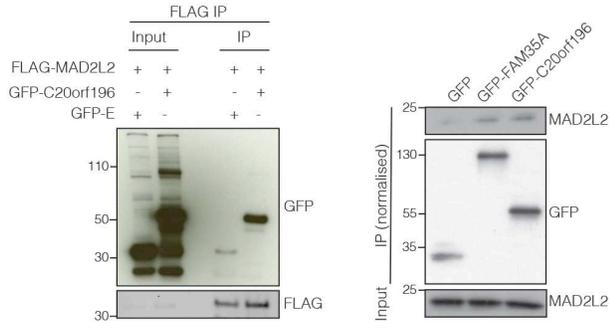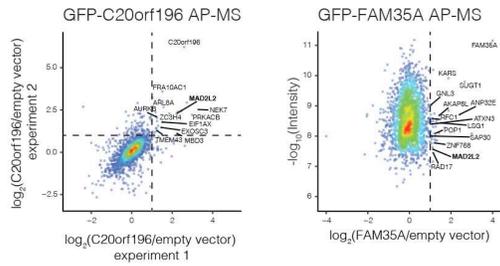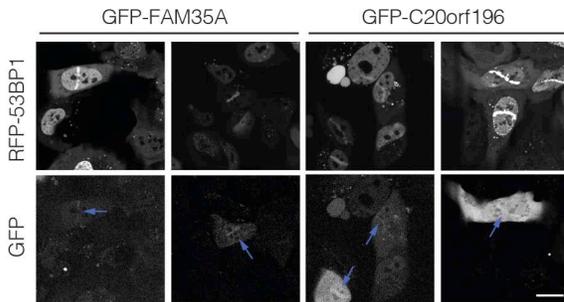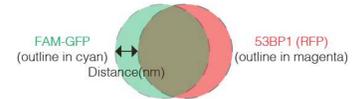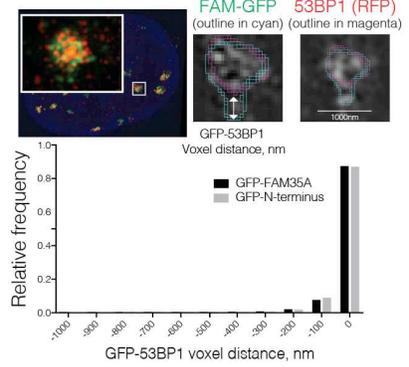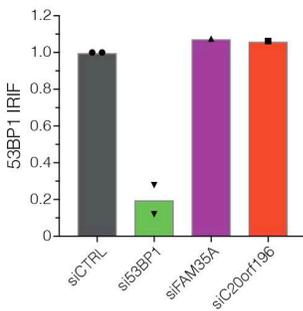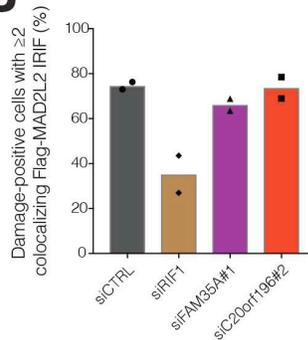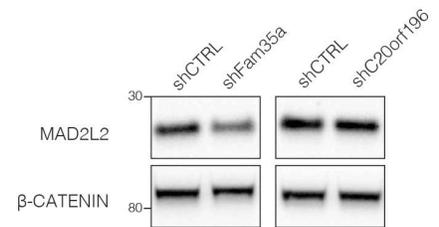
Supplementary Figure 2

C20orf196/FAM35A interactions and localisation to DNA damage sites

a, Co-localisation quantification of FAM35A/derivatives GFP-fusions with mCherry-LacR-C20orf196. Horizontal bars represent means, one-way Anova; *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$, ns=not significant ($p\geq0.05$); n=3 independent experiments; statistical source data including the precise p values can be found in Supplementary Table 5. **b,** C-terminus of FAM35A interacts with C20orf196 in cells (without normalisation). **c,** Immunoprecipitation-mass spectrometric analysis of protein interactors of GFP-C20orf196 or GFP-FAM35A (MAD2L2 is detected in both; far left and middle left panels). GFP-C20orf196 co-immunoprecipitates with Flag-MAD2L2 in HEK293 cells (middle right IB panel). Endogenous MAD2L2 co-immunoprecipitates with GFP-FAM35A and GFP-C20orf196 (far right IB panel) in HEK293 cells. **d,** Live-cell imaging of GFP-FAM35A or GFP-C20orf196 transiently expressed in U2OS cells stably expressing RFP-53BP1. Recruitment of GFP-FAM35A and GFP-C20orf196 to laser tracks was visible 30 min after laser micro-irradiation; representative image from 3 independent experiments. **e,** GFP-FAM35A and GFP-FAM35A N-terminus co-localise with 53BP1 in IRIF by super-resolution microscopy; histogram of n=11 cells per condition. **f,** Depletion of FAM35A or C20orf196 does not affect 53BP1 IRIF (U2OS cells). **g,** as in f but for MAD2L2 IRIF. Bars represent means derived from 2 independent experiments, with individual data points plotted over bars. **h,** Depletion of FAM35A or C20orf196 does not affect MAD2L2 protein levels. All immunoblots are representative of two independent experiments; unprocessed scans of immunoblots are shown in Supplementary Fig 8.

Supplementary Figure 3
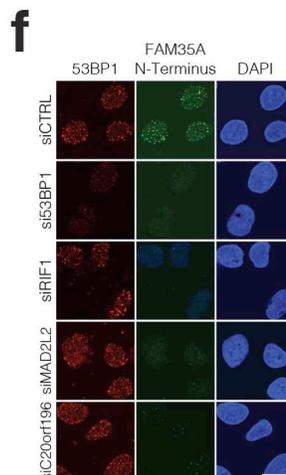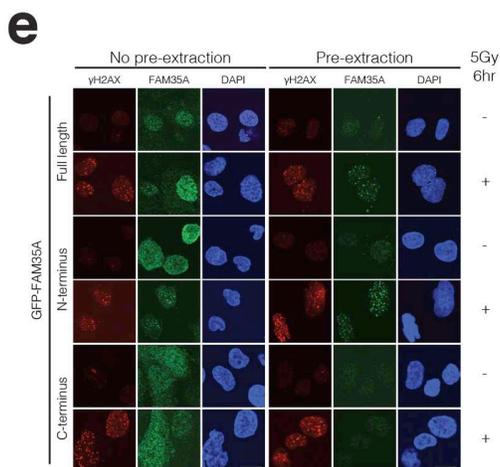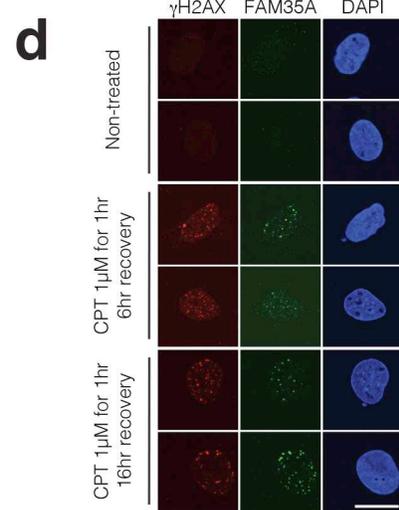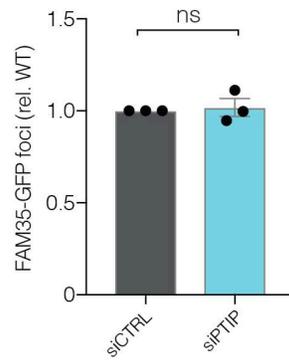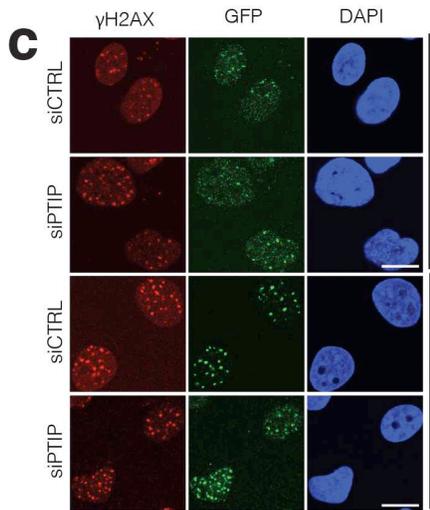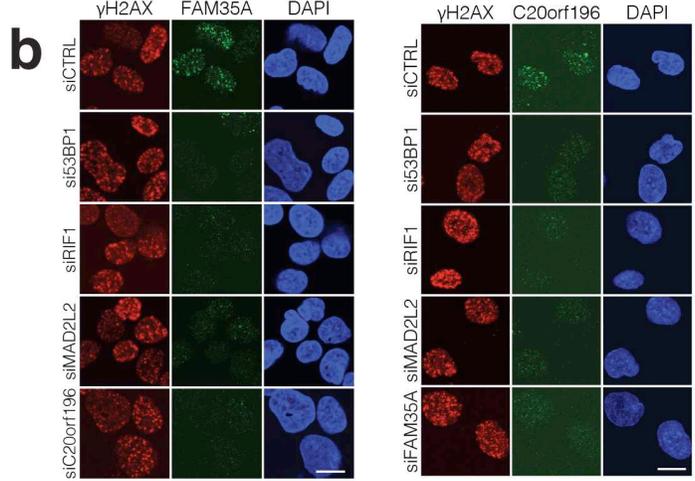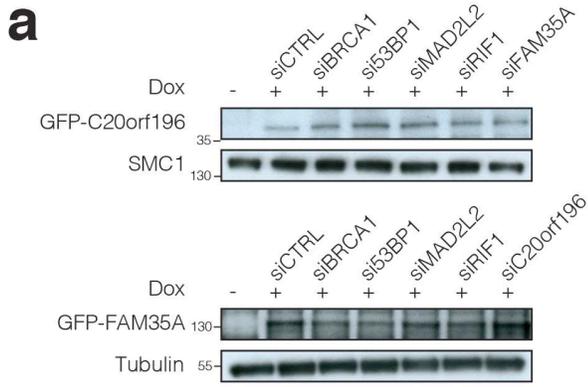
DNA damage response and IRIF factor dependencies of FAM35A and C20orf196

a, Minimal variation of doxycycline induced GFP-FAM35A (U2OS) and GFP-C20orf196 (RPE1) in cells treated with the indicated siRNAs. Immunoblots shown are representative of two independent experiments with unprocessed scans of immunoblots in Supplementary Fig 8. **b,** Representative images of GFP-FAM35A (left panel) and GFP-C20orf196 (right panel) IRIF in γH2AX positive cells quantified in Fig 2e. Scale bar 10μm. **c,** Depletion of PTIP does not affect GFP-C20orf196 or GFP-FAM35A IRIF. Bars represent mean ± SEM, one-way Anova; ns=not significant (p≥0.05); n=3 independent experiments, with individual data points plotted over bars; statistical source data can be found in Supplementary Table 5. Scale bar 10μm. **d,** Camptothecin induced GFP-FAM35A foci. Scale bar 10μm. **e,** Representative images of GFP-FAM35A derivatives with/without pre-extraction ±IR; d-e representative of 2 independent experiments. Scale bar 10μm. **f,** Representative images of GFP-FAM35A N-terminus IRIF dependencies quantified in Fig 2f. Scale bar 10μm.

b

CH12 cell clone reference	Genotype	Allele 1	Allele 2	Allele 3	Allele 4
F10	*Fam35a-*	-5bp	-1bp	-171bp	-166bp
F11	*Fam35a-*	-1bp	-18bp		
F12	*Fam35a-*	-33bp/+8b	-1bp		
O1	*c20orf196-*	-58bp			
O6	*c20orf196-*	-17bp			
O12	*c20orf196-*	-25bp	-186bp		
R13	*Mad2l2-*	-7bp/+4bp			
R14	*Mad2l2-*	-1bp	-4bp		
R20	*Mad2l2-*	-1bp	-10bp		

c

CH12-Cas9 clones

d

e

f
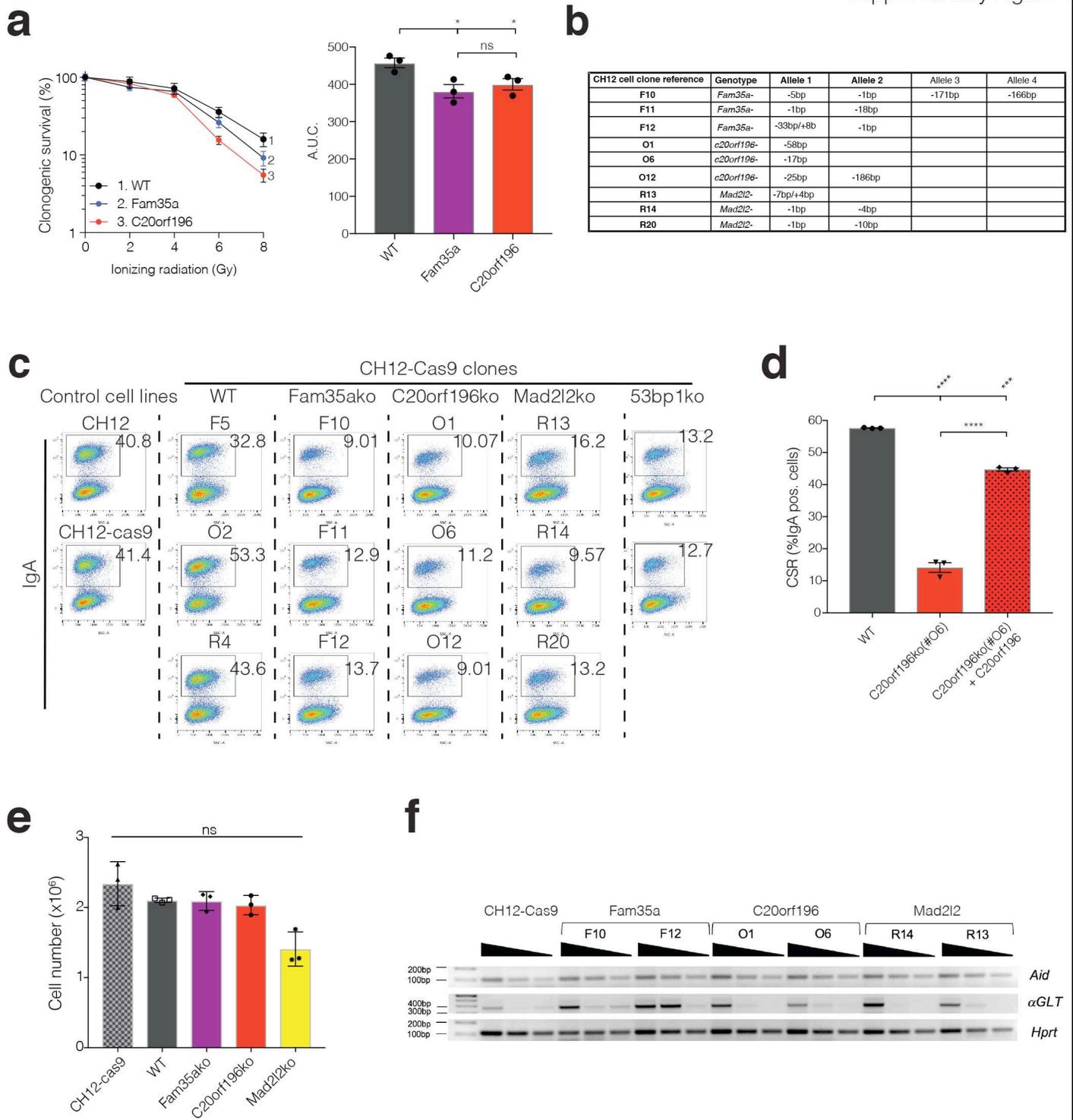
Supplementary Figure 4

FAM35A and C20orf196 directly affect class switch recombination

a, Clonogenic survival assay following IR treatment using wild-type, Fam35ako or C20orf196ko mouse

ES cells (right panel shows AUC). Bars represent means ± SEM, one-way Anova; n=3 independent experiments, with individual data points plotted over bars. **b,** Genotypes of CH12-Cas9 cell knockout clones used CSR assays confirmed by Topo-cloning and Sanger sequencing. **c,** Flow cytometry profiles showing the percentage of IgA+ cells for indicated CH12-Cas9 cell clones (genotypes) after 3 days stimulation with anti-CD40, IL-4 and TGF-β. Cell clone numbers are indicated; representative of 3 independent experiments. **d,** CSR assay in C20orf196ko cells complemented with C20orf196. Bars represent means ± SEM, one-way Anova; n=3 independent experiments, with individual data points plotted over bars. **e,** CH12-Cas9 clones were plated at 50,000 cells/ml and counted after 3 days stimulation with anti-CD40, IL4, and TGF-β. Bars represent means ± SEM, one-way Anova; n=3 independent experiments, with individual data points plotted over bars. For a, d and e, *p<0.05, ***p<0.001, ****p<0.0001, ns=not significant (p≥0.05); statistical source data including the precise p values can be found in Supplementary Table 5. **f,** *Igh*, α germ-line transcripts (αGLT) and *Aid* mRNA were quantified by semi-quantitative RT–PCR using 2.5-fold serial dilutions of cDNA made from CH12-Cas9 cells and indicated CH12-Cas9 knockout cell clones after 2 days stimulation with anti-CD40, IL4, and TGF-β. *Hprt* was used as a control for transcript expression. Immunoblots are representative of two independent experiments with unprocessed scans of immunoblots in Supplementary Fig 8.
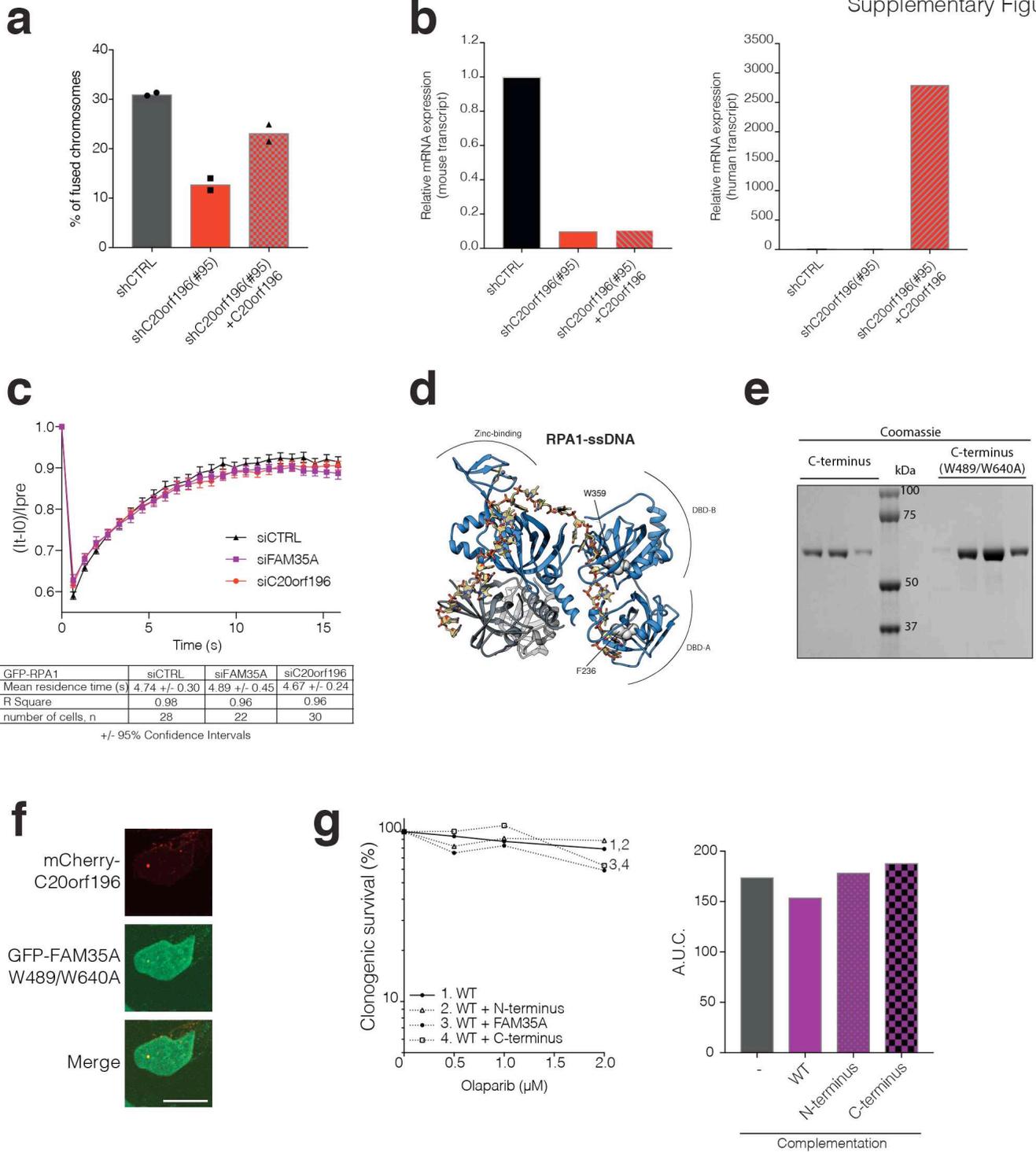
Supplementary Figure 5

Effects of FAM35A and C20orf196 on telomere fusions, DNA binding and DNA-end resection

a, Telomere fusion assay as shown in Fig 4b but complemented with shRNA resistant human C20orf196.

Bars represent means derived from 2 independent experiments with ≥1300 chromosomes counted per condition, and individual data points plotted over bars; source data can be found in Supplementary Table 5. **b,** qRT-PCR of mouse (left) and human (right) transcripts in MEFs. Bars represent means from one experiment performed in triplicates. **c,** FRAP of GFP-RPA1 in stably expressing U2OS cells, depleted of FAM35A or C20orf196. Points represent mean ± 95% confidence intervals; residence time calculated as previously described[50]; n=28 independent experiments (siCTRL), n=22 (siFAM35A) and n=30 (siC20orf196). **d,** Structure of yeast RPA1 (yRPA1) with ssDNA. **e,** Coomassie stained SDS-PAGE gel showing the bacterial purified FAM35A variants used in EMSAs. Immunoblots are representative of two independent experiments with unprocessed scans of immunoblots in Supplementary Fig 8. **f,** GFP-FAM35A W489/W640A is able to interact with mCherry-LacR-C20orf196 in cells; representative of two independent experiments, scale bar 10µm. **g,** Overexpression of FAM35A or derivatives does not sensitise wild-type cells to olaparib, adjacent panel shows AUC. Bars represent means from one experiment performed in triplicates.

Supplementary Figure 6
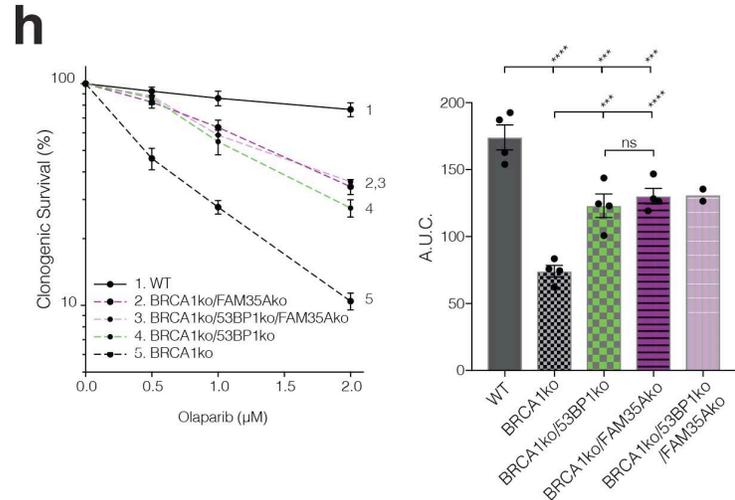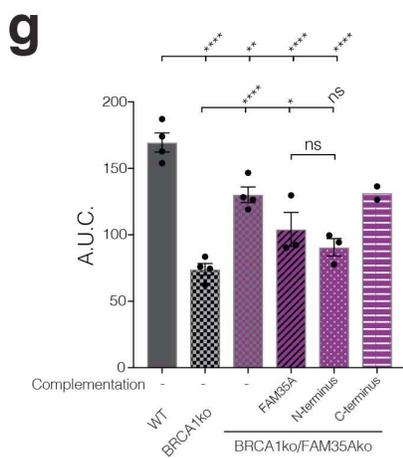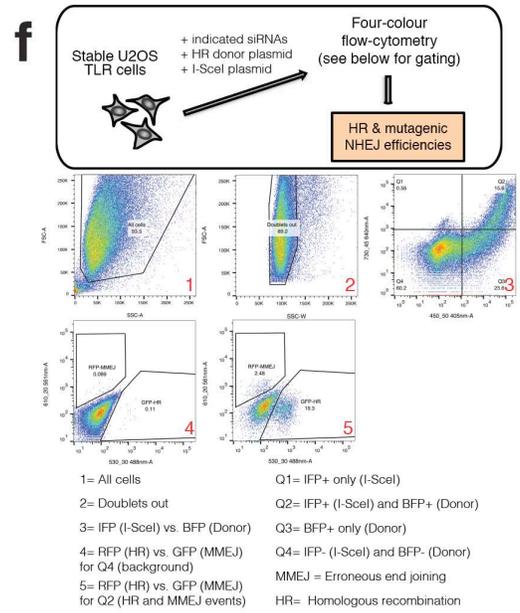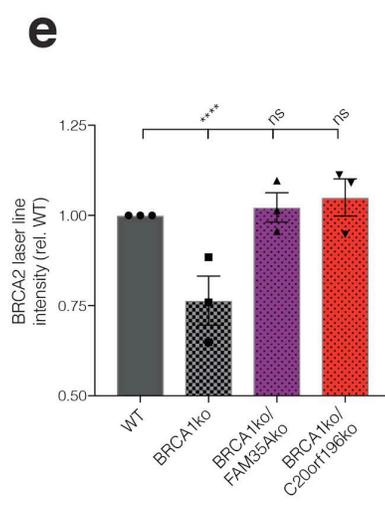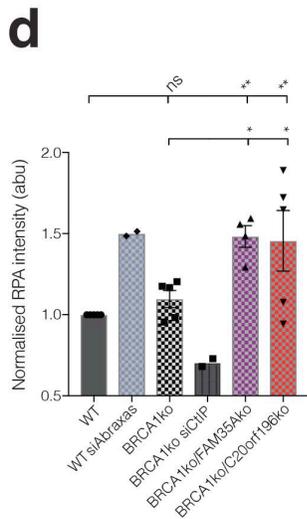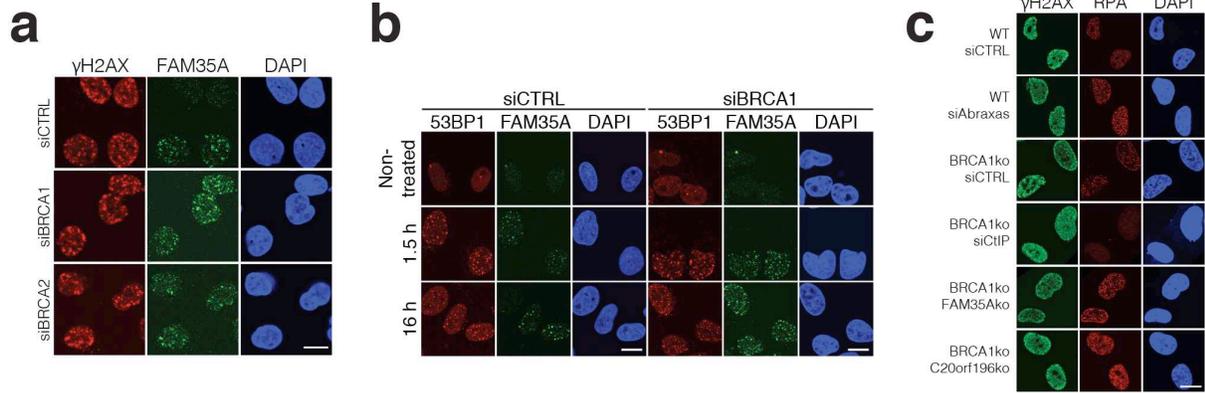
FAM35A and C20orf196 functions relating to homologous recombination

a, Representative images for quantifications of GFP-FAM35A presented in Fig 6a; scale bar 10μm. **b,** Representative images for quantifications presented in Fig 6b; scale bar 10μm. **c,** Representative images of FAM35A and C20orf196 effects on DNA-end resection in wild-type and BRCA1ko cells as measured by RPA nuclear intensity (after pre-extraction) following camptothecin treatment in the indicated genotypes, quantified in Supplementary Fig 6d; scale bar 10μm. **d,** Quantification of nuclear RPA intensity; n=5 independent experiments, except WT siAbraxas and BRCA1ko siCtIP (n=2) and BRCA1ko/FAM35Ako (n=4), with individual data points plotted over bars. **e,** Quantification of BRCA2 accrual at laser micro-irradiated RPE1 cells with the indicated genotypes for the representative images presented in Fig 6d. n=3 independent experiments, with individual data points plotted over bars. **f,** Gating strategy employed for TLR assay. **g,** AUC for clonogenic survival assay presented in Fig 6g. N=4 independent experiments, except BRCA1ko/FAM35Ako +FAM35A and +N-terminus where n=3, and +C-terminus where n=2; with individual data points plotted over bars. **h,** FAM35A and 53BP1 effects on olaparib resistance in BRCA1ko cells are not additive as measured by clonogenic survival assay (left panel), AUC (right panel). N=4 independent experiments, except BRCA1ko/53BP1ko/FAM35Ako where n=2; with individual data points plotted over bars. In d, e, g and h, bars represent mean ± SEM, one-way Anova; *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns=not significant (p≥0.05); statistical source data including the precise p values can be found in Supplementary Table 5.

a

b

c

d

5hrs after 5Gy

e

SHLD2
(FAM35A)

| | | | | | | CXXC | CXXC | |
| | 421 | 489 | 537 | 577 | 640 | 707 720 768 | 786 | 904 |

N-Terminus | OB1 | OB2 | OB3/FD

Mediate ssDNA binding
via W489/W640

IRIF forming region;
binding chromatinised
53BP1 via MAD2L2

Contributes to
IRIF formation

Sensitise WT cells to IR (dominant negative)

Necessary for IR resistance

Necessary and sufficient for PARPi
resensitisation in
FAM35Ako/BRCA1ko cells

Dispensable for PARPi
resensitisation in
FAM35Ako/BRCA1ko cells
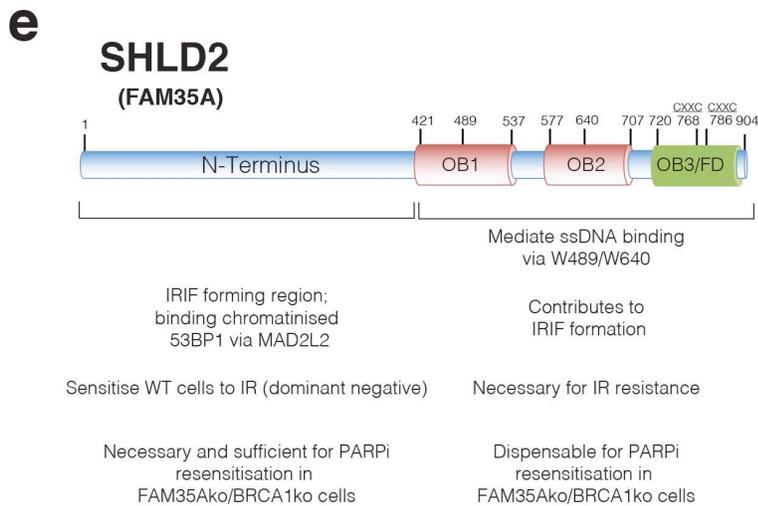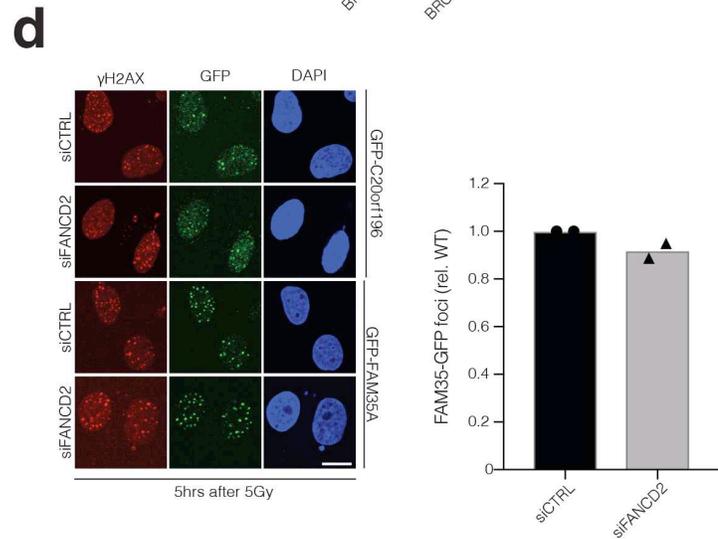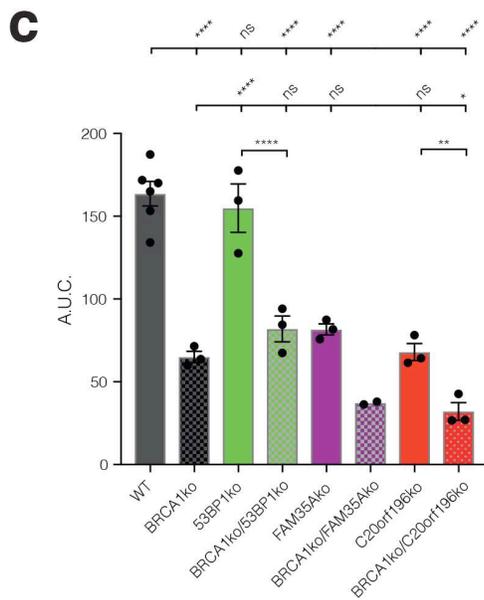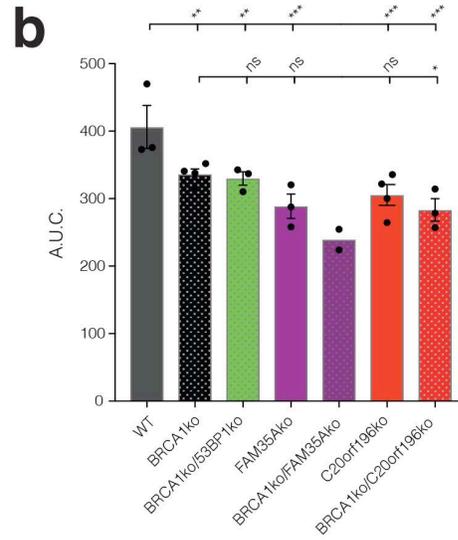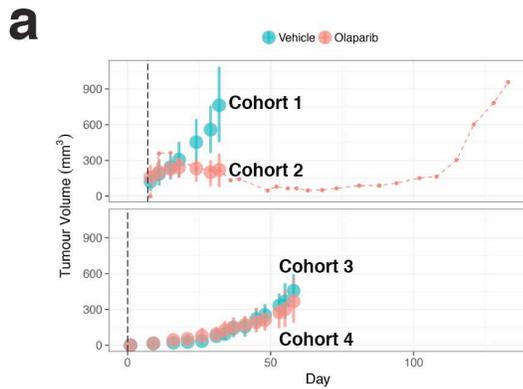
Supplementary Figure 7

Tumour growth curves in mice and cell sensitivities of SHLD mutant cells to DNA damaging agents

a, Tumour growth curves of PDX mice cohorts treated with vehicle or olaparib in Fig 7a; points are means, with lines representing s.d. for each of cohorts 1-4. **b,** AUC for clonogenic survival assay presented in Fig 7c. N=3 independent experiments except BRCA1ko and C20orf196ko where n=4, and BRCA1ko/FAM35Ako where n=2. **c,** AUC for clonogenic survival assay presented in Fig 7d. N=3 independent experiments except WT where n=6 and BRCA1ko/FAM35Ako where n=2. **b-c** Bars represent mean ± SEM, one-way Anova; *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns=not significant (p≥0.05). Individual data points plotted over bars; statistical source data including the precise p values can be found in Supplementary Table 5. **d,** GFP-FAM35A foci are not affected by depletion of FANCD2; representative images (left panel) and quantification (right panel). Bars represent means from 2 independent experiments, with individual data points plotted over bars. Scale bar 10μm. **e,** Graphical summary of SHLD2^{FAM35A} domains and their function.

Fig 2c (left panel)
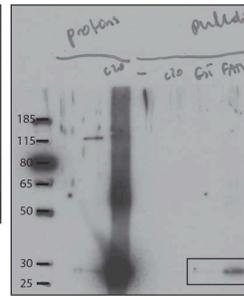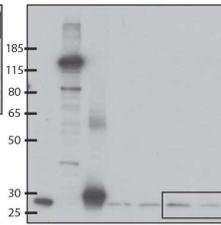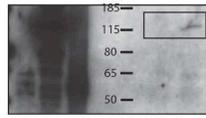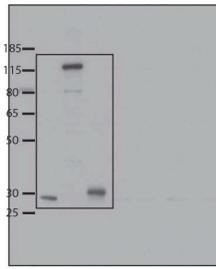
Fig 2c (right panel)
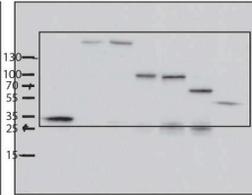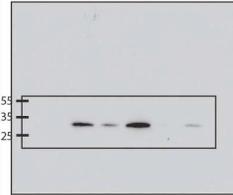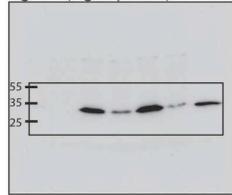
Fig 2d
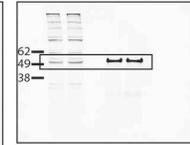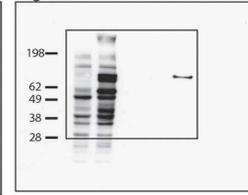
Fig 2g
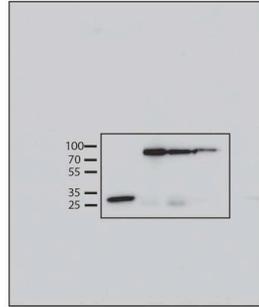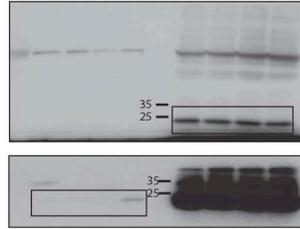
Fig 5a
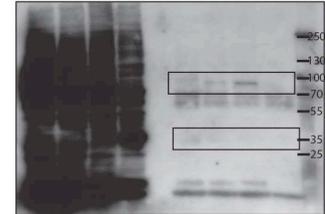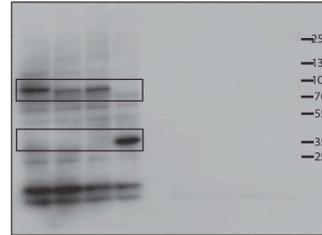
Fig 5c (native polyacrylamide gel)
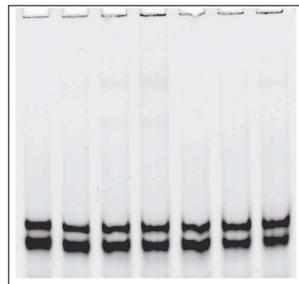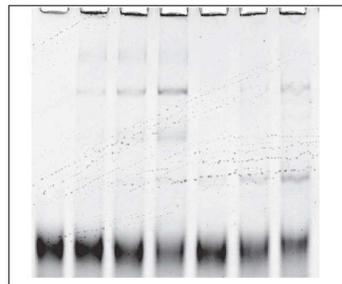
Supplementary Fig 1f (BRCA1)
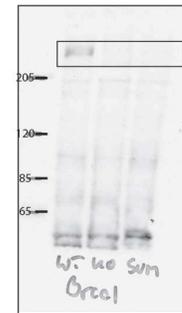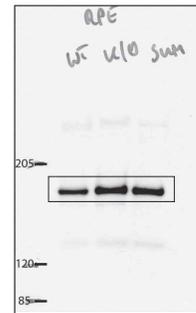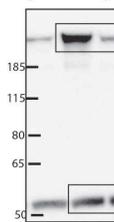
(si53BP1)   (siMAD2L2)   (siPTIP)   (siBRCA1)   (siRIF1)   (siFANCD2)   (siCTIP)   (siABRAXAS)
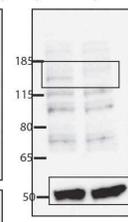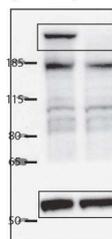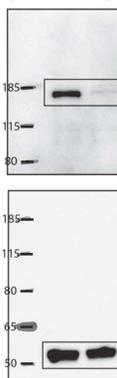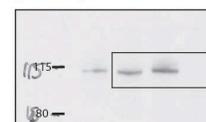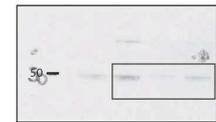
Supplementary Figure 8

Uncropped blots

Supplementary Fig1f (GFP-C20orf196)   (GFP-FAM35A)
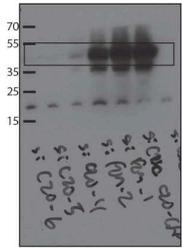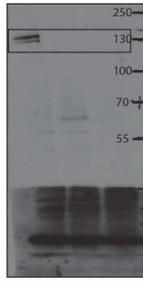
Supplementary Fig 2b
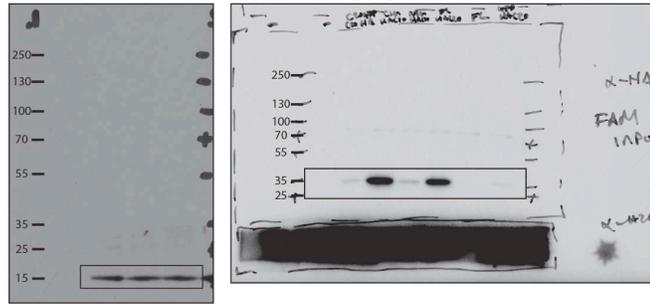
Supplementary Fig 2b
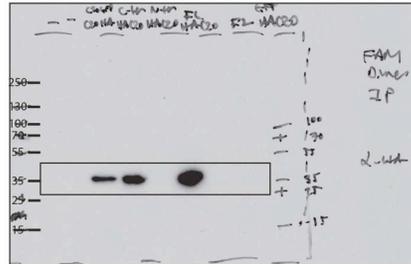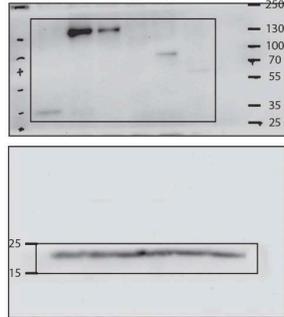
Supplementary Fig 2c
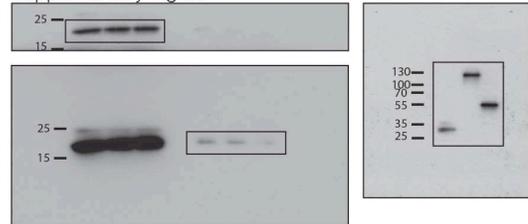
Supplementary Fig 2c
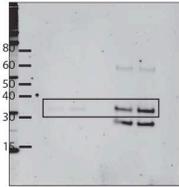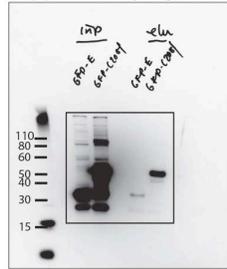
Supplementary Fig 3a (GFP-C20orf196)   (GFP-FAM35A)
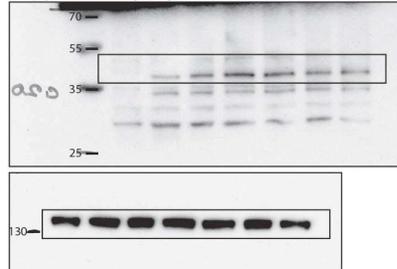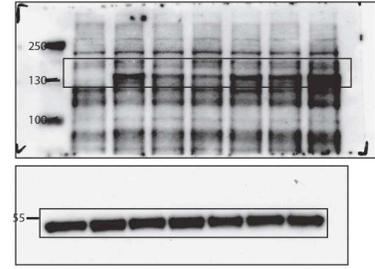
Supplementary Fig 2h
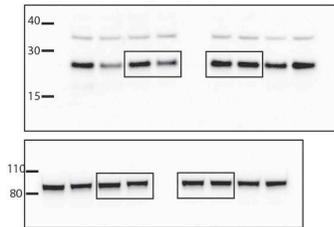
Supplementary Fig 4f
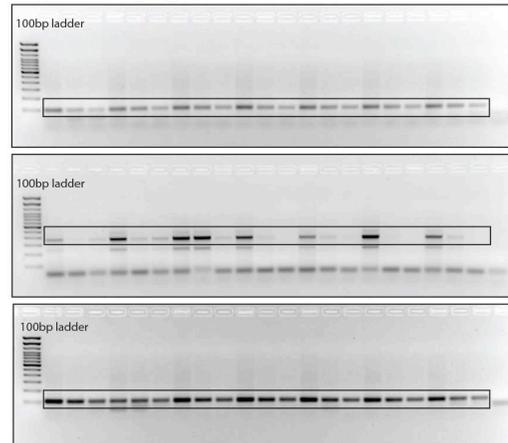
Fig 4d
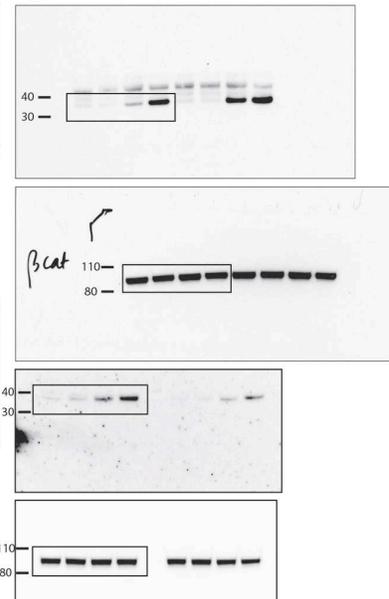
Supplementary Fig 5e
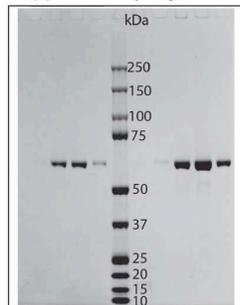
Supplementary Figure 8 continued

Uncropped blots

Supplementary Table 1. CRISPR-Cas9 screen results
Supplementary Table 2. Antibodies
Supplementary Table 3. Plasmids
Supplementary Table 4. Oligos, siRNA/shRNA, CRISPR-cas9 sgRNAs sequences
Supplementary Table 5. Statistics source data

Methods

CRISPR-Cas9 screen. Performed using genome-scale (GeCKO) v2.0[20]. SUM149PT cells were transduced at multiplicity of infection (MOI) of 0.3 and 250-fold coverage of the library. Cells were then selected with puromycin for 7 days prior to treatment with 3 different PARPi for a further 14 days. IC's used were; Olaparib IC95-2 µM, BMN673 IC95-5 nM, AZD2461 IC70-4 µM. Surviving clones from each condition were collected, genomic DNA (gDNA) isolated (Blood & Cell Culture DNA Midi Kit, Qiagen) and subjected to PCR with Illumina-compatible primers, followed by Illumina sequencing. Genes enriched or depleted in the inhibitor-treated samples were determined by the software package MAGeCK version 0.5.5 (see commands in the section of "Code availability").

Cell culture. U2OS, U2OS-derived, HEK293, HEK293T-LentiX cells were cultured as in[39]. RPE1 p53 null FRT[48] and RPE1 p53 null FRT-derived cells were cultured in F-12 (Ham's F-12; Sigma-Aldrich) supplemented with 17 ml $NaHCO_3$ 7.5% per 500ml (Sigma-Aldrich). All media was supplemented with 10% (v/v) foetal bovine serum (FBS; BioSera), 100 U/ml penicillin, 100 µg/ml streptomycin (Sigma-Aldrich) and 2 mM L-glutamine. SUM149PT cells were cultured in Ham's F12 Nutrient Mixture (ThermoFisher) supplemented with 5% (v/v) FBS (BioSera), 10 mM HEPES, 1 µg/ml hydrocortisone, 5 µg/ml insulin, antibiotic as described above. For maintenance and selection of RPE1 FRT p53 null or U20S Trex cells stably expressing GFP or GFP-tagged constructs, 2 µg/ml blasticidin (Sigma-Aldrich) and 0.5 mg/ml G418 (Invitrogen) were used. U2OS-TLR were cultured as in[39]. In addition to RPE1 p53 null FRT-derived cells, U20S Trex cells stably expressing inducible constructs were cultured with 1 µg/ml doxycycline (Sigma-Aldrich) for 24-48h to induce expression of GFP constructs. All cells were originally obtained from the ATCC cell repository and routinely tested to be mycoplasma free. The U2OS and RPE1 cell lines were recently authenticated using Affymetrix SNP6 copy number analysis. *Trf2-/-;p53-/-;*TRF2(Ile468Ala) MEFs (TRF2ts MEFs) as described previously[33, 49]. CH12F3 (CH12)[29] and CH12-Cas9 cell lines were cultured in RPMI 1640 supplemented with 10% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 50µM 2-mercaptoethanol, 1xMEM non-essential amino acids, 1mM sodium pyruvate and 10mM HEPES.

Generation of human stable cell lines and knockouts. U2OS Trex or RPE1 p53 null FRT-derived cells stably expressing inducible GFP-tagged constructs, were generated by

34 transfection of pcDNA5/FRT/TO-neo containing the GFP-tagged construct and pOG44 (1:4

35 respectively). Selection began at 48 h using 0.5 mg/ml G418 (Invitrogen). Knockouts were

36 generated in RPE1 p53 null cells by transfecting an 'All-in-one' plasmid[48]. Single-cell sorting

37 by GFP expression was done using MoFlo (Beckman Coulter). Single clones were expanded,

38 genomic DNA extracted and screened by PCR, TOPO-cloning and sequencing. Validated

39 mouse Embryonic Stem Cell (mESC) knockouts of Fam35a and C20orf196 were obtained

40 from Haplobank ([www.haplobank.at)](www.haplobank.at).

41

42 **Plasmids and cloning.** See Supplementary Table 3.

43

44 **siRNA and plasmid transfection.** siRNAs were obtained from MWG or IDT and transfected

45 using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's protocol.

46 Plasmid transfections were carried out using TransIT-LT1 (Mirus Bio) according to the

47 manufacturer's protocol. For siRNA and DNA co-transfections, plasmids were transfected 8h

48 after siRNA treatment. See Supplementary Table 4.

49

50 **Random plasmid integration assay.** Performed as previously described[27].

51

52 **DNA-damage induction using chemical agents, ionizing radiation and laser micro-**

53 **irradiation.** Performed as previously described[39].

54

55 **FRAP and association kinetics.** Performed as previously described[50].

56

57 **TLR assays.** The Traffic Light Reporter (TLR) assay and the constructs used herein have

58 been described in detail previously[38, 39].

59

60 **Cell-cycle profiling.** Performed as previously described[39].

61

62 **Clonogenic survival assays.** Performed as previously described[27,39].

63

64 **Whole cell extracts and immunoblotting.** Were performed as previously described[39]. For

65 detection of phospho-RPA (pS4/S8-RPA2), lysates were prepared by scraping cells in 2xSDS

66 buffer followed by SDS-PAGE using 4-12% Bis-Tris gel (Invitrogen), and immunoblotting

67 was done using SuperSignal West Pico PLUS (Thermo Scientific). IRDye800CW- and

68 IRDye680-labelled secondary antibody was used for detection on the Odyssey Infrared

69 imager (LI-COR). Quantification of blots was performed using ImageJ. All protein

70 concentrations were determined using a BCA assay (Pierce). All antibodies are listed in

71 Supplementary Table 2.

72

73 **Immunoprecipitation.** All immunoprecipitation procedures performed twice as previously

74 described[39]. For co-immunoprecipitation shown in Fig 2d (FAM35A), 293T cells were co-

75 transfected with pMSCV-blas-eGFP-MAD2L2 and either pLX304-blast-V5-Empty or

76 pLX304-V5-FAM35A. 72h post-transfection cells were exposed to 25Gy IR followed by 3h

77 recovery. GFP-Trap_MA beads (ChromoTek) were used, and immunoprecipitation was

78 performed according to the manufacturer's protocol. For C20orf196 (Supplementary Fig 2c),

79 293T cells were co-transfected with pMSCV-blas-3xFlag-hMAD2L2 and either pcDNA5.1-

80 GFP or pcDNA5.1-GFP-C20orf196. 72h post-transfection, cells were exposed to 25Gy IR

81 followed by 3h recovery. After washing with cold PBS, cells were lysed in 1 ml lysis buffer

82 (50mM Tris HCl pH7.4; 150mM NaCl; 1mM EDTA; 1% Triton X-100) supplemented with

83 the same inhibitors as above. After 30min incubation on ice followed by centrifugation

84 (16,000g), anti-Flag M2 Magnetic Beads (M8823, Sigma-Aldrich) pre-washed with TBS

85 (50mM Tris HCl, 150 mM NaCl pH7.4), were added to the lysate and rotated over-night at

86 4°C. Immune-complexes were eluted by 5min boiling.

87

88 **DNA pulldown experiments.** Procedures were described in detail in[39] using oligos with the

89 sequence:

90 5'BiosG/ATCGCATTGGCATTGGCAATGCGATACGACTGATCGAGGGTACTCAGCT

91 AGCTGATTCCGATCGGCTTATTCCGTGTACATACATCGGAT-3' (IDT)

92

93 **In vitro GST pull-down.** Gluathione sepharose beads (GE Healthcare) were washed with

94 ice-cold PBS and blocked for 30min with PBS supplemented with 10% bacterial lysate (non-

95 induced BL21 cells, lysed using PBS/lysozyme) then resuspended in binding buffer (10mM

96 Tris pH7.5, 150 mM NaCl, 0.5% NP40, 0.5 mM EDTA, 0.5 % BSA). Purified GST (bacterial

97 expression), GST-FAM35A (Novus Biologicals), and His-C20orf196 (Creative BioMart)

98 were added to the beads at 2 pmol and incubated for 30min at 4°C. Beads were washed 5x

99 with 10 mM Tris, pH 7.5, 250 mM NaCl, 0.5 % NP40, 0.5 mM EDTA and eluted with 100

100 mM Tris pH 8, 20 mM reduced glutathione, 120 mM NaCl for 15min rotating at 4°C. The

101 eluates were boiled for 5min, loaded on 4-12 % Bis-Tris gel (Invitrogen) and subjected to

102 western blotting. The blots were probed with the indicated antibodies.

103

104 **Recombinant protein purifications and Electrophoretic Mobility Shift Assays (EMSAs).**

105 Wild-type and mutant FAM35A C-terminal domains were purified using the same method.

106 Harvested cells were lysed by sonication in 50 mM Tris pH 8.0, 5% glycerol, 150 mM NaCl,

107 2 mM β-mercaptoethanol, 10mM imidazole, protease inhibitor (Roche) and 40 g/ml

108 deoxyribonuclease I (Sigma). After centrifugation at 30,000g for 30min, supernatant was

109 loaded onto a gravity column containing Ni-NTA affinity resin (Qiagen) pre-equilibrated

110 with 50 mM Tris pH 8.0, 5% glycerol, 150 mM NaCl, 2 mM β-mercaptoethanol and 10 mM

111 imidazole. After washing beads with the same buffer for 10x column volume, protein was

112 eluted using 50 mM Tris pH 8.0, 5% glycerol, 150 mM NaCl, 2 mM β-mercaptoethanol and

113 100 mM imidazole. The eluate was dialysed with Q column (GE healthcare) buffer A (20

114 mM Tris pH 8.0, 50 mM NaCl, 5% glycerol and 2 mM β-mercaptoethanol) and loaded onto a

115 5 ml Q column. Protein was eluted in a gradient against buffer B (20 mM Tris pH 8.0, 1 M

116 NaCl, 5% glycerol and 2mM β-mercaptoethanol). Fractions containing FAM35A protein

117 were collected and further purified by running through Superdex 200 10/300 column (GE

118 Healthcare) equilibrated in buffer GF (20 mM Tris pH 8, 150 mM NaCl, 5% glycerol and 5

119 mM DTT). Protein samples during each step of purification were analysed on 4-12% Bis-Tris

120 gels (Invitrogen). Final purified samples were concentrated and stored at -80°C. Both

121 forward and reverse 90-bases DNA oligos (IDT) (F: 6-FAM (6-carboxyfluorescein)-

122 ATCGCATTGGCATTGGCAATGCGATACGACTGATCGAGGGTACTCAGCTAGCTG

123 ATTCCGATCGGCTTATTCCGTGTACATACATCGGAT;   R:6-FAM-

124 ATCCGATGTATGTACACGGAATAAGCCGATCGGAATCAGCTAGCTGAGTACCCT

125 CGATCAGTCGTATCGCATTGCCAATGCCAATGCGAT) were dissolved in annealing

126 buffer (10 mM Tris pH 8.0, 50 mM NaCl and 1 mM EDTA) to a final concentration of 100

127 μM. DNA oligo F was used as ssDNA for EMSA. Equal volumes of DNA oligo F and R

128 were mixed and annealed (heated to 95°C for 2min and cooled to 25°C over 45min) to

129 generate dsDNA. Each 20 μl of EMSA reaction contained 10 nM of ssDNA/dsDNA

130 incubated with different concentrations of proteins in 20 mM Tris-HCl pH 7.5, 50 mM KCl,

131 5% (v/v) glycerol, 100 μM DTT, 10 μg/ml BSA. Samples were incubated at 37°C for 15min

132 and applied onto a 5% polyacrylamide native gel in 0.5xTBE buffer for electrophoresis at

133 4°C. DNA was visualized by Typhoon 9000 (GE Healthcare).

134

GFP-Trap pull down for mass spectrometry. HEK293T cells were cultured in SILAC media containing either L-arginine and L-lysine, or L-arginine [$^{13}C_6$, $^{15}N4$] and L-lysine [$^{13}C_6$, $^{15}N_2$] (Cambridge Isotope Laboratories) as described previously[51]. Cells were lysed 48h post-transfection in modified RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% sodium deoxycholate) supplemented with protease, phosphatase inhibitors and N-ethylmaleimide. Lysates were cleared by centrifugation at 16,000×g for 15min at 4°C and protein concentrations were estimated using QuickStart Bradford Protein assay (BioRad). Per SILAC condition, 20 µl of pre-equilibrated GFP-Trap-A beads (ChromoTek) were added to 2 mg of lysate and incubated 1h at 4°C rotating, followed by 3x washes with modified RIPA buffer. Bound proteins were eluted in NuPAGE® LDS Sample Buffer (Life Technologies) supplemented with 1 mM dithiothreitol, heated at 70 ºC for 10min and alkylated with 5.5 mM chloroacetamide at RT. Samples were loaded onto 4-12% gradient SDS-PAGE gels, proteins were stained using the Colloidal Blue Staining Kit (Life Technologies) and digested in-gel using trypsin. Peptides were extracted from the gel and desalted on reversed phase C18 StageTips[52].

Mass spectrometry analysis. Peptide fractions were analyzed on a quadrupole Orbitrap mass spectrometer (Q Exactive Plus, Thermo Scientific) equipped with an EASY-nLC 1000 (Thermo Scientific) as described[53]. Peptide samples were loaded onto C18 reversed phase columns and eluted with a linear gradient from 8 to 40% acetonitrile containing 0.1% formic acid for 2h. The mass spectrometer was operated in data dependent mode, automatically switching between MS and MS2 acquisition. Survey full-scan MS spectra (m/z 300–1650) were acquired in the Orbitrap. The ten most intense ions were sequentially isolated and fragmented by HCD[54]. Fragment spectra were acquired in the Orbitrap mass analyzer. Raw data files were analyzed using MaxQuant (development version 1.5.2.8)[55]. Parent ion and MS2 spectra were searched against a database containing 92,578 human protein sequences obtained from the UniProtKB released in December 2016 using Andromeda search engine[56]. Spectra were searched with a mass tolerance of 6 ppm in MS mode, 20 ppm in HCD MS2 mode, strict trypsin specificity and allowing up to three miscleavages. Cysteine carbamidomethylation was searched as a fixed modification, whereas protein N-terminal acetylation, methionine oxidation, n-ethylmaleimide modification of cysteines were searched as variable modifications. The dataset was filtered based on posterior error probability to arrive at a false discovery rate below 1% estimated using a target-decoy approach[57].

169   **Immunofluorescence and microscopy imaging.** Confocal imaging for γH2AX, RAD51,
170   RPA, ssDNA (BrdU), BLM, BRCA2, FANCD2, Cyclin A and GFP (FAM35A and
171   C20orf196) was performed as described in[39], for RAD51 and Cyclin A the pre-extraction step
172   was omitted and cells were permeabilised for 15min in 0.2% Triton X-100 (Sigma) in PBS
173   after fixation. Super-resolution images were acquired using a Deltavision OMX 3D-SIM
174   System V3 BLAZE (Applied Precision, a GE Healthcare company) equipped with 3 sCMOS
175   cameras, 405, 488, 592.5nm diode laser illumination, an Olympus Plan Apo N 60x 1.42NA
176   oil objective, and standard excitation and emission filter sets. Imaging of each channel was
177   done sequentially using three angles and five phase shifts of the illumination pattern as
178   described[58]. Sections were acquired at 0.125 μm z steps. Raw OMX data was reconstructed
179   and channel registered in SoftWoRx software version 6.5.2 (Applied Precision, a GE
180   Healthcare company). Voxelwise nearest-neighbour distances were measured for GFP-
181   FAM35A signal relative to 53BP1 signal using a custom script (Butler R) for Fiji
182   (https://github.com/gurdon-institute/OMX-Spatial-Analysis). The script maps signal volumes
183   using Kapur's maximum entropy thresholding method[59] and measures distances using the
184   exact signed 3D Euclidean distance transform with internal distances set to zero for display
185   on the histogram. For all images, scale bars = 10μm.
186
187   **Multiplex fluorescence in situ hybridisation (M-FISH).** Human 24-colour multiplex FISH
188   (M-FISH) probe preparation and slides treatments followed[60]. For each human cell sample,
189   10-30 metaphases were karyotyped based on the M-FISH classification and DAPI-banding
190   pattern. FISH on metaphases spreads using BAC probes was performed as previously
191   described[61] and counted manually. For class-switch recombination (CSR) assays, DNA FISH
192   on metaphases spreads was performed as previously described[61] and counted manually. At
193   least 470 metaphases were evaluated per genotype, using at least 2 independent clones for
194   each condition. For telomere uncapping, cell harvesting, preparation of metaphase spreads
195   and telomere FISH with an Alexa488-(CCCTAA repeat) peptide nucleic acid custom probe
196   (PN-TC060-005, Panagene/Eurogentec), metaphase chromosome analysis was done as
197   described previously[6]. These data represent 2 independent experiments, ≥1300 chromosomes
198   for each condition, counted manually after blinding the genotypes.
199
200   **Telomere fusion assays: MEFs viral transduction.** Cells were transduced as before[49] with
201   pLKO-puro shRNA lentiviruses obtained from the MISSION shRNA library (Sigma), against

202   mouse genes as described or left untransduced ('empty'). **Assessment of telomere NHEJ**.

203   TRF2ts MEFs were grown for 24h at the non-permissive temperature of 39°C to inactivate

204   TRF2 and induce NHEJ-dependent chromosome end-to-end fusions because of telomere

205   uncapping.

206

207   **Class-switch recombination assays: CH12 and CH12-Cas9 cell lines.** The CH12-Cas9 cell

208   line was generated by transducing CH12-Cas9 cells made using spin-infect with lentivirus

209   particles packaged in HEK293T. Plasmids: pKLV2-EF1aBsd2Acas9-W, pxPAX2 (Addgene

210   #12260), VSV-G and pMD2.G (Addgene #12259). Blasticidin selection (10 μg/ml) started

211   48h after transduction for one week. **sgRNA expression plasmids for class-switch**

212   **recombination assays. s**gRNAs were used to target Fam35a, C20orf196 and Mad2l2/Rev7

213   mouse genes (2 sgRNAs per target gene, sequences listed in Supplementary Table 4. sgRNAs

214   were cloned into pKLV-flipped U6gRNA_CCDB_PB_BbsI_PGKpuro2ABFP vector[62].

215   **Generation of wild-type and knockout CH12-Cas9 cell clones.** 53bp1 null CH12 cell

216   clones (gift from Fred Alt) were as previously described[63]. 12 million CH12-Cas9 cells were

217   nucleofected with 2.2 μg of each sgRNA-1 and sgRNA-2 and 0.6 μg of piggyBac transposase

218   expression vector[64], using an Amaxa Nucleofector, Nucleofector® Kit V solution (Lonza)

219   and program X-001. Two days later, BFP-positive/puromycin-resistant CH12-Cas9 cells

220   were selected with 3 μg/ml puromycin for one week. Cells were then single cell diluted into

221   96-well plates, further cultured and screened by PCR and Sanger sequenced using PCR

222   primers listed in Supplementary Table 4. **Class-switch recombination and cell**

223   **proliferation assays.** CH12 cells were plated at 50,000 cells/ml in complete RPMI

224   supplemented with anti-CD40 antibody (1 μg/ml, Miltenyi), IL-4 (20 ng/ml, Miltenyi) and

225   TGF-β (1 ng/ml, R&D Biotech) to induce IgM to IgA switching. After 3 days, cells were

226   assayed for class-switching by flow cytometry using an IgA-PE antibody (eBiosciences) and

227   a Canto II analyser (BD Biosciences). Viable cells were counted using a Casy cell counter

228   (Roche). CSR and proliferation assays were done on: 3xwild-type (WT), 3xFam35a knockout

229   (Fam35a), 3xC20orf196 knockout (C20orf196), 2x53bp1 knockout (53bp1) and 3xMad2l2

230   knockout (Mad2l2) in three independent experiments. **RT-PCR analysis.** *Igh*, *α* germ-line

231   transcripts (*αGLT*) and *Aid* mRNA were quantified as previously described[12]. Primers are

232   listed in Supplementary Table 4.

233

234   **Patient-derived tumour xenografts.** PDXs were generated and established from consented
235   breast or ovarian cancer patients' samples as previously described[40]. The research was done
236   with the appropriate approval by the National Research Ethics Service, Cambridgeshire 2
237   REC (REC reference number: 08/H0308/178), and by the Vall d'Hebron Hospital Clinical
238   Investigation Ethical Committee (PR(AG)183/2012). STG201, the PDX model used in this
239   study, is a BRCA-null model featuring BRCA1 promoter methylation, loss of BRCA1
240   mRNA and protein expression. We have previously shown its sensitivity in vivo and in PDX
241   derived cells to PARP inhibitors, including olaparib. STG201 is also linked to deep molecular
242   and drug sensitivity annotation[40] and http://caldaslab.cruk.cam.ac.uk/bcape/. All other PDXs
243   were derived from breast or ovarian tumours from *BRCA1*-mutation carriers or *BRCA1*
244   epigenetic silencing due to promoter hypermethylation[41]. PDX127 did not show any co-
245   expression of BRCA1 but it was low in both FAM35A and C20orf196 expression. None of
246   the 5 PARPi-sensitive PDXs exhibited low levels of C20orf196, FAM35A or 53BP1 loss nor
247   BRCA1 hypomorphs. The study was compliant with all relevant ethical regulations regarding
248   research involving animal use and human participants.
249
250   **Generation of acquired drug resistance in vivo.** AZD2281 (Olaparib/Lynparza) as a PARP
251   inhibitor (50mg/kg, 5IW) was administered to immunocompromised tumour bearing mice
252   upon randomization as previously described (50mg/kg, 5 days/week)[40]. To classify the
253   response of the subcutaneous implants we modified the RECIST criteria to be based on the %
254   tumour volume change following continuous olaparib treatment: complete response (CR),
255   best response≤-95%; partial response (PR), -95%<best response≤-30%; stable disease (SD), -
256   30%<best response≤+20%; progressive disease (PD), % tumour volume change at day 21 of
257   treatment >+20%. PARPi-resistant PDXs exhibited PD while PARPi-sensitive models
258   exhibited SD, PR or CR. For STG201, time matched vehicle and olaparib treated samples
259   were collected 25 days after treatment (PARPi naïve PDX) and processed for RNA
260   extractions and sequencing. A couple of mice in the study were left with continued exposure
261   to olaparib until tumour regrowth. One of these resistant tumours was serially passaged 126
262   days after treatment into new host mice (PARPi resistant PDX) and treated with further
263   vehicle or olaparib. 58 days after treatment the resistance phenotype was confirmed and
264   samples were collected and processed for RNA-sequencing as described below. Growth
265   curves show average and standard deviation of at least 5 independent tumour volumes per
266   trial arm. All experimental procedures were approved by the University of Cambridge
267   Animal Welfare and Ethical Review Committee and by the Vall d'Hebron Hospital Clinical

268   Investigation Ethical Committee and Animal Use Committee. **RNA-sequencing.** RNA was

269   extracted from all samples using the Qiagen miRNeasy or RNeasy Mini kit (Cat ID, 217004

270   or 74104) as per manufacturer's instructions. Libraries for Illumina sequencing were

271   prepared using TruSeq Stranded mRNA HT kit or Total RNA Library Prep kit with Ribo-

272   Zero Gold (Cat ID, RS-122-2103 or RS-122-2301, Illumina). 500ng of total RNA with RNA

273   Integrity Numbers (RINs) above 8 was used for library preparation. Samples were processed

274   following manufacturer's HS (High-Sample) instructions (part# 15031048 Rev. E, Illumina)

275   with 12 or 15 cycles of PCR used at the Enrichment of DNA Fragments step. All libraries

276   were quantified using KAPA Library Quantification Kit Illumina ROX Low (Cat ID,

277   KK4873, KAPA Biosystems) and normalised. Libraries were pooled in equal volumes and

278   pools were used for clustering on HiSeq4000 sequencing flow cell following manufacturer's

279   instructions. Sequencing was performed using 150bp or 100bp paired-end run type for dual-

280   indexed libraries. Prior to alignment, sequencing quality of the reads was enforced using

281   Trim Galore! (v0.4.2) http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

282   Then, as described in Callari et al. or Ahdesmaki et al[65] reads were aligned to a combined

283   human (hg19) and mouse (mm10) reference genome using STAR (v2.5.2b)[66, 67]. Counts were

284   assigned to genome features using featureCounts (v1.5.2), whereby the alignment score is

285   used to discern accurately between reads sourced from human and mouse[68]. Counts from

286   multiple sequencing runs were merged and then normalised using the edgeR package[69, 70].

287

288   **Code availability** A custom FIJI script used in OMX analysis can be found at

289   https://github.com/gurdon-institute/OMX-Spatial-Analysis.

290

291   The MAGeCK commands used for CRISPR-Cas9 screens were:

292   mageck test -k counts.csv -c DMSO -t WC_2461 -n WC_2461

293   mageck test -k counts.csv -c DMSO -t WC-673 -n WC-673

294   mageck test -k counts.csv -c DMSO -t WC-2281 -n WC-2281

295

296

297   **Statistics and Reproducibility**

298   Unless stated otherwise Prism v7.0b (GraphPad Software) was used to generate graphs,

299   perform statistical tests and calculate p values. Error bars, statistical tests and number of

300   independent repeats (n) are indicated in figure legends with statistical source data including

301   the precise p values provided in Supplementary Table 5. Statistical tests included two-tailed

302   Student t-tests, Fisher's Exact test, and one-way analysis of variance (ANOVA), the latter all

303   being corrected as recommended for multiple comparisons. Microscopy image analyses were

304   performed using ImageJ/FIJI or Volocity 6.3 (Perkin-Elmer). CRISPR screens were

305   performed with three clones per drug treatment. Mass spectrometry of GFP-FAM35A and

306   GFP-C20orf196 was performed in two independent experiments. RNA-sequencing was

307   performed as three replicates for each trial arm, due to sequencing across multiple lanes

308   (which were merged prior to any further analysis). This was performed for the following

309   number of independent biological samples: six PDXs in cohort 1, five PDXs in cohort 2,

310   seven PDXs in cohort 3 and eight PDXs in cohort 4. For the SHLD1 high and low expression

311   cohorts 12 and 4 independent PDXs were evaluated respectively. All immunofluorescence

312   assay quantification data represent means ±SEM's of 3 independent biological repeats and

313   n≥30 cells per condition unless otherwise specified. All immunoblots are representative of

314   two independent experiments with unprocessed scans of immunoblots shown in

315   Supplementary Fig 8.

316

317   **Data availability.**

318   The raw data files for the whole genome CRISPR-Cas9 screen in SUM149 cells is available

319   on NIH Sequence Read Archive (SRA) via the accession number PRJNA471892. Raw data

320   files for the PDX RNA sequencing is available on NIH Sequence Read Archive (SRA) via

321   the accession number PRJNA473981. Raw data files for the mass spectrometry are available

322   via the ProteomeXchange Consortium on the PRIDE partner repository with the dataset

323   identifier PXD009830. Source data for figures can be found in Supplementary Table 5. All

324   other data supporting the findings of this study are available from the corresponding authors

325   on reasonable request.

326

References

328
329   48.   Chiang, T.W., le Sage, C., Larrieu, D., Demir, M. & Jackson, S.P. CRISPR-Cas9(D10A)
330   nickase-based genotypic and phenotypic screening to enhance genome editing. *Sci*
331   *Rep* **6**, 24356 (2016).
332   49.   Peuscher, M.H. & Jacobs, J.J. DNA-damage response and repair activities at
333   uncapped telomeres depend on RNF8. *Nat Cell Biol* **13**, 1139-1145 (2011).
334   50.   Galanty, Y., Belotserkovskaya, R., Coates, J. & Jackson, S.P. RNF4, a SUMO-targeted
335   ubiquitin E3 ligase, promotes DNA double-strand break repair. *Genes Dev* **26**, 1179-
336   1195 (2012).

337 51. Ong, S.E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a
338   simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-
339   386 (2002).
340 52. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment,
341   pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc*
342   **2**, 1896-1906 (2007).
343 53. Michalski, A. *et al.* Mass spectrometry-based proteomics using Q Exactive, a high-
344   performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics*
345   **10**, M111 011015 (2011).
346 54. Olsen, J.V. *et al.* Higher-energy C-trap dissociation for peptide modification analysis.
347   *Nat Methods* **4**, 709-712 (2007).
348 55. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates,
349   individualized p.p.b.-range mass accuracies and proteome-wide protein
350   quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).
351 56. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant
352   environment. *J Proteome Res* **10**, 1794-1805 (2011).
353 57. Elias, J.E. & Gygi, S.P. Target-decoy search strategy for increased confidence in large-
354   scale protein identifications by mass spectrometry. *Nat Methods* **4**, 207-214 (2007).
355 58. Gustafsson, M.G. *et al.* Three-dimensional resolution doubling in wide-field
356   fluorescence microscopy by structured illumination. *Biophys J* **94**, 4957-4970 (2008).
357 59. Kapur, J.N., Sahoo, P.K. & Wong, A.K.C. A New Method for Gray-Level Picture
358   Thresholding Using the Entropy of the Histogram. *Comput Vision Graph* **29**, 273-285
359   (1985).
360 60. Agu, C.A. *et al.* Successful Generation of Human Induced Pluripotent Stem Cell Lines
361   from Blood Samples Held at Room Temperature for up to 48 hr. *Stem Cell Reports* **5**,
362   660-671 (2015).
363 61. Lescale, C. *et al.* RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG
364   complex in DNA repair. *Nat Commun* **7**, 10529 (2016).
365 62. Metzakopian, E. *et al.* Enhancing the genome editing toolbox: genome wide CRISPR
366   arrayed libraries. *Sci Rep* **7**, 2244 (2017).
367 63. Panchakshari, R.A. *et al.* DNA double-strand break response factors influence end-
368   joining features of IgH class switch and general translocation junctions. *Proc Natl*
369   *Acad Sci U S A* **115**, 762-767 (2018).
370 64. Yusa, K., Zhou, L., Li, M.A., Bradley, A. & Craig, N.L. A hyperactive piggyBac
371   transposase for mammalian applications. *Proc Natl Acad Sci U S A* **108**, 1531-1536
372   (2011).
373 65. Ahdesmaki, M.J., Gray, S.R., Johnson, J.H. & Lai, Z. Disambiguate: An open-source
374   application for disambiguating two species in next generation sequencing data from
375   grafted samples. *F1000Res* **5**, 2741 (2016).
376 66. Callari, M. *et al.* Computational approach to discriminate human and mouse
377   sequences in patient-derived tumour xenografts. *BMC Genomics* **19**, 19 (2018).
378 67. Casper, J. *et al.* The UCSC Genome Browser database: 2018 update. *Nucleic Acids Res*
379   **46**, D762-D769 (2018).
380 68. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program
381   for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930
382   (2014).

383   69.   McCarthy, D.J., Chen, Y. & Smyth, G.K. Differential expression analysis of multifactor
384   RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40**,
385   4288-4297 (2012).
386   70.   Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for
387   differential expression analysis of digital gene expression data. *Bioinformatics* **26**,
388   139-140 (2010).
389
390