1   **Exosome component 1 cleaves single-stranded DNA and sensitizes kidney renal**

2   **clear cell carcinoma cells to poly(ADP-ribose) polymerase inhibitor**

3   Qi Xiao[1,†], Qiaoling Liu[1,†], Lina Wang[1], Na Wang[1], Kai Wang[1], Chengli Song[1,*] and

4   Qingkai Yang[1,*]

5

6   [1]Institute of Cancer Stem Cell, DaLian Medical University, 9 Western Lvshun South

7   Road, Dalian, Liaoning 116044, China.

8

9   †These authors contributed equally to this work.

10

11   *For correspondence: Yangqingkai@dmu.edu.cn; Chenglisong2015@hotmail.com.

12

13   **Abstract**

14   Targeting DNA repair pathway offers an important therapeutic strategy for cancers.

15   However, the failure of DNA repair inhibitors to markedly benefit patients

16   necessitates the development of new strategies. Here, we show that exosome

17   component 1 (EXOSC1) promotes DNA damages and sensitizes kidney renal clear

18   cell carcinoma (KIRC) cells to DNA repair inhibitor. Considering that endogenous

19   source of mutation (ESM) constantly assaults genomic DNA and likely sensitize

20   cancer cells to the inhibitor, we first analyzed the statistical relationship between the

21   expression of individual genes and the mutations for KIRC. Among the candidates,

22   EXOSC1 most notably promoted DNA damages and subsequent mutations via

23   preferentially cleaving C site(s) in single-stranded DNA. Consistently, EXOSC1 was

24   more significantly correlated with C>A transversions in coding strands than these in

25   template strands in KIRC. Notably, KIRC patients with high EXOSC1 showed a poor

26   prognosis, and EXOSC1 sensitized cancer cells to poly(ADP-ribose) polymerase

27   inhibitor. These results show that EXOSC1 acts as an ESM in KIRC, and targeting

28   EXOSC1 might be a potential therapeutic strategy.

29

30

31

32

33

34 **Introduction**

35   DNA damages and subsequent mutations are central to development, progression,

36   and treatment of nearly all cancers (Brown et al., 2017; Farmer et al., 2005; Jeggo et

37   al., 2016; Pearl et al., 2015; Roos et al., 2016). Cancer cells frequently decreases DNA

38   repair pathways and increases endogenous sources of mutation (ESM) to drive

39   mutations (Brown et al., 2017; Farmer et al., 2005; Jeggo et al., 2016; Pearl et al.,

40   2015; Roos et al., 2016). Hence, cancer cells are often more reliant on a subset of

41   DNA repair pathway(s) to survive DNA damages. While, targeting critical DNA

42   repair members, such as poly (ADP-ribose) polymerases (PARPs) (Brown et al., 2017;

43   Tubbs and Nussenzweig, 2017), offers a therapeutic strategy for cancers (Tutt et al.,

44   2010). Inhibition of PARPs by small-molecule compounds disrupts the ability of

45   cancer cells to survive ongoing DNA damage and results in cell cycle arrest and/or

46   cell death (Lord and Ashworth, 2012). However, failure of PARP inhibitors (PARPis)

47   to markedly benefit patients suggests the necessity for developing new strategies. Due

48   to the central role of ESM in ongoing DNA damages, there is a need for the

49   identification and understanding of ESM.

50   ESM constantly assaults genomic DNA and almost inevitably leads to mutations

51   (Figure 1A) (Jeggo et al., 2015; Roos et al., 2015). However, most of ESM studies

52   were focused on the deamination. The significance of deamination as an ESM is

53   supported mainly by two observations: (1) Transitions show higher frequency than

54   expected by chance, although there are twice as many possible transversions.

55   Nucleotide substitutions consist of two types: transition and transversion (Alexandrov

56   et al., 2013a; Burgess, 2019; Petljak et al., 2019). Transition is a substitution in which

57   one base is replaced by another of the same class (purine or pyrimidine), while

58   transversion is a substitution in which a purine is replaced with a pyrimidine or vice

59   versa (Figure 1B). (2) C>T transitions at methylated cytosine in CG base pairs display

60   a higher frequency than expected (Alexandrov et al., 2013a; Alexandrov et al., 2013b;

61   Burgess, 2019; Hutchinson, 2013; Petljak et al., 2019). Therefore, activation-induced

62   cytidine deaminase (AID) (Barlow et al., 2013; Greaves, 2018; Petersen-Mahrt et al.,

63   2002) and apolipoprotein B mRNA editing enzyme catalytic (APOBEC) family

64   (Alexandrov et al., 2013b; Buisson et al., 2019; Hutchinson, 2013; McGranahan et al.,

65   2017; Robertson et al., 2017) catalyzing the deamination of C were identified as

66   ESMs. Unfortunately, some cancers such as kidney renal clear cell carcinoma (KIRC)

67   show the low mutation proportion at CG base pairs and low APOBEC expression

68 (Burns et al., 2013b), raising the potential roles of unidentified ESMs.

69   Due to advances in sequencing technology and the great efforts of The Cancer

70 Genome Atlas (TCGA), it is now possible to explore the statistical relationship

71 between mutations and the expression of individual genes in multiple cancer types

72 (Weinstein et al., 2013). The majority of patients included in the TCGA database are

73 accompanied by data regarding both mutations and genome-wide expression of

74 individual genes (Weinstein et al., 2013). Because that DNA damages often comprise

75 a major source of mutation, the relativity between DNA damage and mutation allows

76 quantitative analyses of mutation to be taken as a proxy of DNA damage (Brown et al.,

77 2017; Jeggo et al., 2016; Pearl et al., 2015). Furthermore, the gene-specific correlation

78 between mRNA and protein levels allows quantitative analyses of individual gene

79 expression as an indicator for the corresponding protein (Peng et al., 2015; Uhlen et

80 al., 2017; Zhang et al., 2017). Hence, analyses of the cancer cohort may identify

81 candidate ESMs (Tubbs and Nussenzweig, 2017).

82   The exosome is an evolutionarily conserved multiprotein complex formed by

83 exosome components (EXOSCs) (Bousquet-Antonelli et al., 2000; Brown et al., 2000;

84 Tomecki et al., 2010). In eukaryotes, the exosome complex has a "ring complex"

85 (EXOSC4–EXOSC9) and a "cap" structure (EXOSC1–EXOSC3) (Figure 1C). The

86 human exosome complex may also contain two additional subunits, EXOSC10 and

87 EXOSC11(Jonathan et al., 2006; Lorentzen et al., 2008; Quansheng et al., 2006;

88 Wasmuth et al., 2014), which provide 3' to 5' exo- and/or endoribonuclease activities

89 (Januszyk and Lima, 2014; Kilchert et al., 2016). The exosome is well known to

90 degrade RNA (Januszyk and Lima, 2014; Kilchert et al., 2016; Ulmke et al., 2021).

91 Hence, the exosome was reported to protect cells from genomic instability via

92 degrading the DNA/RNA hybrids and restricting DNA strand mutational asymmetry

93 (Lim et al., 2017; Pefanis and Basu, 2015; Pefanis et al., 2015). Interestingly, the cap

94 unit EXOSC2 stably associated with the exosome complex, while EXOSC1 is not

95 stably associated (Dai et al., 2018; Malet et al., 2010), suggesting that EXOSC1 might

96 be involved in some functions independent of the complex.

97   In this study, we show that EXOSC1 acts as an ESM and sensitizes cancer cells to

98 PARPi in KIRC. Due to the role of exosome in maintaining genomic stability, these

99 results also indicate that a unit of multiprotein complex can play a role opposite to

100 that of the complex.

101

Results

Identification of candidate ESMs in KIRC

Because that ESMs constantly assaults genomic DNA, we hypothesized that ESMs likely sensitized cancer cells to the inhibitors of DNA repair pathways. Considering that substitution is the most abundant mutation in all cancers, we initiated this study to identify candidate ESMs responsible for substitution mutations. To identify the candidate ESMs other than deamination, we focused on KIRC for three reasons: (1) KIRC shows the lowest proportion of mutations at CG in major cancer types (Burns et al., 2013b), suggesting that the deamination contributes less to the mutations in KIRC. (2) only low expressions of AID and APOBECs were detected in KIRC (Burns et al., 2013b). (3) The kidney potentially suffers less from exogenous source of mutations (EOSMs) (Loeb, 2011; Roberts and Gordenin, 2014). RNA-seq and exomic mutation data corresponding to 532 KIRC patients and 30,254 somatic substitution mutations in the TCGA were retrieved from The cBio Cancer Genomics Portal (http://cbioportal.org) (Figure 1D). Because that ESMs likely show similar impact on the template and code DNA strands, the 12 types of substitution were groups into 6 types of complementary substitution (c-substitution) to simplify the analyses (Figure 1D).

Spearman's rank analysis was first performed to assess the correlation between each c-substitution type and the genome-wide expression of individual genes. Resultant p and r values were used for the further analyses. For example, GAPDH showed a $p = 0.0011$ and r = 0.16 correlation with C>A/G>T c-substitution, indicating that GAPDH expression displayed a positive correlation with C>A/G>T (Figure 1E). Similarly, CRB3 ($p = 0.0003$, r =0.17) was positively correlated with A>T/T>A (Figure 1E). Although the p values of multiple genes were lower than 0.05 (Figure 1F), only top ranked 200 genes (approximately 1% of the genome-wide genes) were taken as the candidates for each c-substitution type (Supplementary Table S1).

Student's t-test analysis was then used to determine whether the expression difference of individual gene between the high and low c-substitution groups is significant. The expression of individual genes in each patient was normalized by a house keeping gene, TATA-binding protein (TBP) as previously described (Burns et al., 2013a; Burns et al., 2013b). According to each c-substitution, 532 KIRC patients were groups into 3 groups (high, medium and low). The difference of individual gene between high and low c-substitution mutation groups was then analyzed by student's

136 *t*-test. Resultant *p* and fold of change (FC) values were used for the further analyses.

137 For example, ASNS with *p* = 0.0005 and FC = 1.39, indicating that ASNS was

138 increased in the high group (Figure 1G). Although the *p* values of multiple genes were

139 lower than 0.05 (Figure 1H), only top 200 genes with high FC and *p* < 0.05 were

140 taken as the candidates (Supplementary Table S2). Notably, none of APOBEC family

141 members were identified as candidate by correlation or student's *t*-test analyses,

142 supporting that deamination contributes less to the mutations in KIRC.

143 Next, we performed meta-analyses to determine which of the 6 c-substitution types

144 to be focused on. Mutation frequencies of c-substitution types were first analyzed in

145 five major cancers: breast adenocarcinoma (BRCA), glioblastoma multiforme (GBM),

146 bladder urothelial carcinoma (BLCA), acute myeloid leukemia (AML) and KIRC,

147 which potentially suffer less from the EOSMs. As shown in Figure 1I, KIRC

148 displayed higher frequencies of C>A/G>T, A>T/T>A, and A>C/T>G mutations.

149 Using tumor protein p53 (TP53) as control, we then assessed the frequencies of

150 c-substitution types in von Hippel-Lindau tumor suppressor (VHL), the most

151 frequently mutated gene in KIRC. Consistent with previous studies (Burns et al.,

152 2013a; Kandoth et al., 2013), the most frequent c-substitution type of TP53 mutations

153 in BRCA was C>T/G>A (Figure 1J), while the most frequent type of VHL mutations

154 in KIRC was C>A/G>T (*p* = 0.004, chi-squared test) (Figure 1K). Even after

155 normalization according to the base frequency, this phenomenon was still observed (*p*

156 = 0.021) (Figure 1—figure supplement 1A and B). Further Kaplan–Meier (KM)

157 analysis of overall survival (OS) indicated that patients with VHL C>A/G>T

158 mutations showed poor OS (Figure 1L). These observations raised the significance of

159 C>A/G>T c-substitutions in KIRC.

160 We then evaluated the expression difference of individual gene between the VHL

161 C>A/G>T mutation positive and negative patients. Student's *t*-tests analyses showed

162 that 66 genes displayed *p* < 0.05 (Supplementary Table S3). Further overlap analyses

163 demonstrated that cyclin B1 (CCNB1), exosome component 1 (EXOSC1), and RAB5

164 interacting factor (RAB5IF) were identified as candidate ESMs for C>A/G>T by all

165 of the above analyses (Figure 1M).

166

167 **EXOSC1 promotes mutations in *E. coli***

168 To evaluate the capability of the candidate gene to promote mutation,

169 rifampicin-resistant assay in *E. coli* was performed as previously described

170   (Petersen-Mahrt et al., 2002). Because that mutation of the rifampicin-targeted rpoB

171   gene to rifampicin resistance (Rif^R) occurs at a low frequency, the capability of a gene

172   to mutate rpoB to Rif^R can be evaluated by fluctuation analysis (Petersen-Mahrt et al.,

173   2002). AID, an known ESM (Petersen-Mahrt et al., 2002), was used as a positive

174   control. Four genes (CDK5, TARBP2, PSAT1 and NECAB3) were used as random

175   controls (Figure 2—figure supplement 1A and Supplementary Table S4). These genes

176   were expressed in *E. coli* under the regulation of a trp/lac (tac) hybrid promoter,

177   which could be activated by isopropyl β-D-1-thiogalactopyranoside (IPTG) (Figure

178   2A). Consistent with a previous study (Petersen-Mahrt et al., 2002), AID enhanced

179   mutation in *E. coli* (Figure 2B). Notably, EXOSC1 more significantly increased

180   mutations than AID did (p = 4.08 X 10^{-5}) (Figure 2B). We then evaluated the

181   capabilities of EXOSC1 homologs (EXOSC2–EXOSC9) to promote mutations

182   (Figure 2—figure supplement 1B). Among the members of exosome complex,

183   EXOSC1 most notably enhanced mutation in *E. coli* (p = 3.5 X 10^{-11}) (Figure 2C). To

184   determine whether the increase in mutation frequency stemmed from EXOSC1

185   protein itself, rifampicin-resistant assays were performed in the presence or absence

186   of IPTG, the transcriptional inducer. As shown in Figure 2E–F, IPTG absence notably

187   decreased the mutation frequency (p = 1.29 X 10^{-9}), indicating that the protein of

188   EXOSC1 promoted the mutation. Additionally, we evaluated the impact of EXOSC1

189   on the growth of *E. coli*. As shown in Figure 2—figure supplement 1C, EXOSC1

190   expression only slightly decreased cell growth, which might be due to the increase in

191   mutation burden (Schaaper and Dunn, 1987).

192   Next, the mutation spectra of Rif^R were analyzed by sequencing the rpoB gene

193   PCR products from rifampicin-resistant clones. Sequencing of rpoB gene in 25

194   randomly selected rifampicin-resistant clones indicated that most of mutations in

195   control clones were C>T/G>A transitions, while EXOSC1 frequently promoted

196   C>A/G>T transversion mutations (Figure 2G–I). Moreover, most of C>A/G>T

197   transversions in EXOSC1-transformed cells were clustered at C1576 (10/36 mutations)

198   and C1699 (6/36 mutations), whereas C>T/G>A transitions in control cells showed a

199   distinct distribution with major hot spots at C1565 (6/32 mutations) and C1721 (5/32

200   mutations) (Figure 2G). Hence, it was suggested that the mutations in control and

201   EXOSC1-transformed cells were promoted by distinct mechanism. Interestingly,

202   EXOSC1-transformed cells also showed a shift of C>A/G>T mutations from 6%

203   (2/32 mutations) to 69% (25/36 mutations) (p = 4.03 X 10^{-7}, chi-squared test) (Figure

204   2H–J). Even after normalization to the base frequency, this phenomenon was still

205   significant (p = 1.47 X 10^{-6}) (Figure 2—figure supplement 1D).

206

207   **EXOSC1 cleaves single-stranded DNA**

208   Considering that the exosome is well known to degrade RNA, we speculated that

209   EXOSC1 might promote mutation through cleaving DNA. EXOSC1 was expressed

210   and purified *in vitro* (Figure 3A). The resultant EXOSC1 protein was incubated with

211   generic single-stranded DNA (ssDNA), double-stranded DNA (dsDNA) or the hybrid

212   of DNA-RNA (Figure 3B). Polyacrylamide TBE-urea gel analyses of the resultant

213   mixtures indicated that EXOSC1 notably cleaved ssDNA, while it displayed no

214   detectable capability to cleave dsDNA or DNA-RNA hybrid (Figure 3C).

215   We then evaluated the capabilities of EXOSC1 homologs (EXOSC2–EXOSC9) to

216   cleave ssDNA. The EXOSC protein was separately incubated with ssDNA. Gel

217   analyses of the resultant mixtures indicated that, unlike EXOSC1, none of

218   EXOSC2–EXOSC9 detectably cleaved ssDNA (Figure 3—figure supplement 1A).

219   Considering that EXOSC1 is well known to form a complex with other exosome

220   members, we also assessed the impact of the exosome members on the cleavage

221   activity of EXOSC1. EXOSC1 was incubated with ssDNA in the presence of

222   individual EXOSC1 homolog. Interestingly, EXOSC6, EXOSC7 and EXOSC8

223   decreased EXOSC1 cleavage activity (Figure 3D), while EXOSC2, EXOSC3,

224   EXOSC4 and EXOSC9 showed no detectable impact (Figure 3—figure supplement

225   1B). We then evaluated the impact of the reaction components and pH on the cleavage

226   activity of EXOSC1. It was found that K^+ and Mg^{2+} enhanced the cleavage activity of

227   EXOSC1 (Figure 3E), and EXOSC1 showed the highest cleavage activity at pH 7.0

228   (Figure 3F). Further analyses indicated that the cleavage rate of EXOSC1 was

229   approximately 4 X 10^{-4}/min at 37 ℃ (Figure 3G and H).

230

231   **EXOSC1 prefers to cleave C sites in single-stranded DNA**

232   Considering that conserved exosome prefers to degrade the RNA with specific

233   sequence (Cvetkovic et al., 2017), we determined whether EXOSC1 preferred to

234   cleave some site(s) in ssDNA. EXOSC1 was incubated with DNAs containing

235   unbiased sequence and distinct 3' end. Consistent with the result of the generic DNA,

236   EXOSC1 cleaved unbiased ssDNA and displayed no detectable capability to cleave

237   dsDNA or DNA-RNA hybrid (Figure 4A and Figure 4—figure supplement 1A). And

238   only EXOSC1 displayed cleavage activity against unbiased ssDNA (Figure 4B).

239   Interestingly, the cleavage rate against unbiased ssDNA (approximately 1.2 X

240   10^{-3}/min) was higher than that against generic ssDNA (Figure 4—figure supplement

241   1B), suggesting that the DNA sequence might show some impact on the activity of

242   EXOSC1. Notably, mass spectrometry (MS) analyses demonstrated that the resultant

243   mixtures contained more free C than the other three base types (Figure 4E),

244   suggesting that EXOSC1 preferred to cleave C sites in ssDNA. Since EXOSC1 was

245   correlated with the C>A/G>T c-substitution type, it was likely that EXOSC1 cleaved

246   C sites in ssDNA and subsequently resulted in C>A mutations through "A" rule DNA

247   repair.

248   To evaluate the above hypothesis, we then determined whether EXOSC1-promoted

249   mutations displayed strand asymmetries. Considering that EXOSC1 cleaved C sites in

250   ssDNA but not DNA-RNA hybrid, we speculated that "transcribed" temple strands

251   likely bound by RNA were less cleaved by EXOSC1. As shown in Figure 4F, C>A

252   transversions in the "untranscribed" coding strand lead to C>A mutations in a gene,

253   while C>A transversions in the "transcribed" template strand result in G>T mutations.

254   Therefore, the capability of EXOSC1 to promote strand mutational asymmetry can be

255   evaluated by comparing C>A and G>T frequencies. We first analyzed the

256   distributions of C>A and G>T substitutions, instead of the distributions of C>A/G>T

257   c-substitution, in the EXOSC1-transformed *E. coli* cells described above. As shown in

258   Figure 4—figure supplement 1E and F, C>A substitution in rpoB gene were generated

259   at a much higher frequency than G>T. Compared with the control, EXOSC1 enhanced

260   C>A from 0% to 69% ($p = 2.67$ X 10^{-7}), while EXOSC1 only enhanced G>T from 6%

261   to 17% even without a significance ($p = 0.27$) (Figure 4—figure supplement 1E and

262   F). Next, we evaluated C>A strand asymmetry in KIRC using spearman's rank and

263   student's *t*-test analyses. Spearman's rank analyses indicated that EXOSC1 showed

264   the highest correlation (r) with C>A, and the correlation between EXOSC1 and G>T

265   was even lower than that between EXOSC1 and C>A/G>T (Figure 4G and Figure

266   4—figure supplement 1G). To evaluate the impact of group number on the further

267   student's *t*-test analyses, the KIRC patients were grouped into 2, 3, 4 or 5 groups

268   according to the mutation types C>A, G>T, C>A/G>T and total (12 substitution

269   types). As expected, the EXOSC1 differences between the low and high C>A/G>T

270   groups were more significant than those between the low and high groups of total (12

271   substitution types) mutations (Figure 4—figure supplement 1H and I). Importantly,

272   EXOSC1 showed more significant expression differences between the low and high

273   C>A groups than those between the low and high G>T groups (Figure 4H and Figure

274   4—figure supplement 1J), suggesting that EXOSC1 prefers to cleave C sites in

275   ssDNA *in vivo*.

276

277   **EXOSC1 enhances DNA damage and mutations in KIRC cells**

278   Considering that EXOSC1 cleaves DNA *in vitro*, we then evaluated the capability of

279   EXOSC1 to promote intracellular DNA damage using γ-H2AX staining and neutral

280   comet tail assays. The 769-P and TUHR14TKB KIRC cells were transfected with the

281   plasmid encoding EXOSC1 using empty vector as a control (Figure 5—figure

282   supplement 1A). γ-H2AX staining analyses of the resultant cells demonstrated that

283   EXOSC1 increased the γ-H2AX foci in the cells (Figure 5A). The number of

284   γ-H2AX-positive cells was increased approximately 7-fold by EXOSC1 (Figure 5B).

285   While knockdown of EXOSC1 reduced the γ-H2AX foci (Figure 5C, and Figure

286   5—figure supplement 1B and C). Consistent with the results of γ-H2AX staining,

287   comet tail analyses also indicated that EXOSC1 increased DNA damage (Figure 5D

288   and E).

289   Due to the central role of DNA damage in mutations, we performed differential

290   DNA denaturation PCR (3D-PCR) to determine whether EXOSC1 enhances mutation

291   in KIRC cells. Because that DNA sequences with more A/T content can be amplified

292   at lower denaturation temperatures than parental sequences, 3D-PCR enables

293   qualitative estimates of genomic C/G>A/T mutations in a population of cells. As

294   shown in Figure 5F, the enhanced expression of EXOSC1 (EXOSC1-OE) increased

295   lower temperature amplicons (LTAs) of VHL, suggesting that EXOSC1 increased the

296   mutations in VHL gene. Consistently, further sequencing analyses of the LTAs

297   indicated that EXOSC1-OE cells showed more C>A mutations in VHL gene (Figure

298   5F).

299   Considering that the "A" rule DNA repair is dependent on X-ray repair

300   cross-complementing 1 (XRCC1) (Sale et al., 2001), we knocked down XRCC1 to

301   evaluate the role of XRCC1 in EXOSC1-promoted mutations (Figure 5—figure

302   supplement 1D–F). 3D-PCR analyses of the resultant cells indicated that knockdown

303   of XRCC1 impaired the capability of EXOSC1 to increase the LTAs (Figure 5G).

304   Furthermore, both XRCC1 knockdown (XRCC1-KD) and EXOSC1 knockdown

305   (EXOSC1-KD) decreased the LTAs (Figure 5H and I). Additionally, a subcutaneous

306   xenograft tumor model was used to determine whether EXOSC1 enhances DNA

307   mutations *in vivo* (Figure 5J). Stable control (vector), EXOSC1-OE and EXOSC1-KD

308   769-P cells were subcutaneously implanted. After two weeks, 3D-PCR analyses of the

309   resultant tumors indicated that EXOSC1 increased the LTAs of VHL, whereas

310   knockdown of EXOSC1 reduced the LTAs (Figure 5K), suggesting that EXOSC1

311   enhanced mutations in KIRC.

312

313   **EXOSC1 sensitizes KIRC cells to PARP inhibitor**

314   Considering the central roles of mutation in the process of cancers, we evaluated the

315   potential clinical significance of EXOSC1 in KIRC using KM analyses. KM analyses

316   of disease-free survival (DFS) and OS were performed using the clinical data from

317   532 KIRC patients in TCGA. The fragments per kilo base per million mapped reads

318   (FPKMs) were used to evaluate the expression of EXOSC1 in KIRC (Figure

319   6—figure supplement 1A). The median-separation KM analyses indicated that high

320   EXOSC1 group showed poor DFS and OS (Figure 6A and B). The median DFS in the

321   low EXOSC1 group was 32.0-month longer than that in the high group (p = 9.78 X

322   10^{-8}, log-rank test) (Figure 6A). Consistently, the median OS in the low EXOSC1

323   group was 36.9-month longer than that in high group (p = 2.2 X 10^{-8}) (Figure 6B). As

324   expected, the best-separation KM analysis also indicated that high EXOSC1 group

325   significantly showed poor OS (p = 2.6 X 10^{-12}) (Figure 6—figure supplement 1B).

326   Due to the critical role of VHL mutation in KIRC, we then evaluated the potential

327   clinical significance of EXOSC1 in the presence and absence of VHL mutation. KM

328   analyses indicated that high EXOSC1 group showed poor OS in both presence

329   (median OS (high vs low group) = 65.2 vs 98.5 months, p = 0.015) and absence

330   (median OS (high vs low group) = 75.7 vs 104.5 months, p = 1.0 X 10^{-5}) of VHL

331   mutation (Figure 6C and Figure 6—figure supplement 1C).

332   Considering that EXOSC1 increases DNA damage, we speculated that EXOSC1

333   potentially sensitizes KIRC cells to the inhibitor of poly(ADP-ribose) polymerase

334   (PARP), which treats cancers via blocking DNA repair. As previously described (Li et

335   al., 2020), colony formation assays were performed to evaluate the role of EXOSC1

336   in the response to a PARP inhibitor, niraparib. Stable control (vector) and enhanced

337   EXOSC1 (EXOSC1-OE) KIRC cells were seeded and treated with serial dilutions of

338   niraparib until colonies were notably formed. As shown in Fig 6D–F, niraparib more

339   notably inhibited the 769-P and TUHR14TKB KIRC cells with enhanced EXOSC1,

340   suggesting that EXOSC1 sensitized the cells to the inhibitor. Consistent with the

341   results in 769-P and TUHR14TKB cells, EXOSC1 also sensitized SNU-1272 and

342   Caki-2 KIRC cells to the inhibitor (Figure 6—figure supplement 1D). Next, we

343   determined whether EXOSC1 could sensitize KIRC cells to niraparib in xenograft

344   mouse models. The control and EXOSC1-OE cells were subcutaneously injected.

345   Resultant tumor-bearing mice were grouped and treated by vehicle or niraparib.

346   Consistent with the *ex vivo* results, niraparib more notably inhibited the tumor with

347   enhanced EXOSC1 (Figure 6G and I), indicating that EXOSC1 sensitized KIRC

348   xenografts to the inhibitor. No significant weight loss was observed throughout the

349   study, suggesting that the niraparib treatment was well tolerated (Figure 6H and J).

350

351   **Discussion**

352   The genomic integrity of human cells is constantly assaulted by ESMs. Although

353   human cells possess multi DNA repair mechanisms to counteract these constant

354   assaults, not all lesions are correctly repaired and almost inevitably result in mutations.

355   The central roles of these acquired mutations in nearly all cancers (Jeggo et al., 2015;

356   Roos et al., 2015) emphasize the identification and understanding of the ESMs. Here,

357   we show that EXOSC1 cleaves ssDNA and acts as an ESM in KIRC. Consistent with

358   the capability of EXOSC1 to promote DNA damage and mutations, KIRC patients

359   with high EXOSC1 showed a poor prognosis, and EXOSC1 also sensitized cancer

360   cells to the PARP inhibitor.

361   Our results show that a unit of multiprotein complex can play a role distinct from

362   the function(s) of the complex. Biological processes frequently require fine control

363   over the formation of a multiprotein complex in a particular region of the cell. The

364   exosome complex is well known for its roles in RNA degradation (Januszyk and Lima,

365   2014; Kilchert et al., 2016). However, the role of exosome complex members other

366   than RNA binding and degradation remains elusive. Interestingly, EXOSC1 can

367   disassociate from the exosome complex (Dai et al., 2018; Malet et al., 2010),

368   suggesting that EXOSC1 might be involved in some functions independent of the

369   exosome. Our study demonstrated that EXOSC1 acts as an ESM to promote

370   mutagenesis. Conversely, previous studies have described that the exosome, as a

371   multiprotein complex, protects cells from genomic instability by preventing the

372   formation of DNA/RNA hybrids and restricting DNA strand mutational asymmetry

373   (Lim et al., 2017; Pefanis and Basu, 2015; Pefanis et al., 2015). This phenomenon can

374   at least partially be explained by the finding that some exosome members (EXOSC7

375   and EXOSC8) can block the activity of EXOSC1 to cleave DNA. Therefore, a single

376   EXOSC1 protein can show different responses depending on the presence/absence of

377   its interacting partners. Further studies are needed to better understand the roles of the

378   individual exosome member.

379   The potential pathological significance of EXOSC1 is supported by its association

380   with poor DFS and OS in KIRC. Due to the capability of EXOSC1 to cleave DNA

381   and promote mutations, EXOSC1 might enhance mutations and consequently provide

382   genetic fuel for cancer development, metastasis, and even therapy resistance.

383   Therefore, EXOSC1 might represent not only a KIRC marker but also a target to

384   decrease the rate of KIRC evolution and stabilize the targets of existing therapeutics.

385   Furthermore, targeting DNA repair in cancers by inhibiting PARPs offers an

386   important therapeutic strategy (Cleary et al., 2020). Unfortunately, the failure of

387   PARP inhibitors to markedly benefit patients enforces the necessity for developing

388   new strategies to improve their efficacy (Cleary et al., 2020; Dizon, 2017; Lord and

389   Ashworth, 2017). Our study demonstrated that EXOSC1 sensitized KIRC cells to

390   PARP inhibitor, suggesting inhibition of PARPs might be a penitential strategy to treat

391   KIRC patients with high EXOSC1. We also noticed that KIRC patients with high

392   EXOSC1 and VHL mutations showed the poorest OS. Considering the DNA cleavage

393   activity of EXOSC1 and the role of VHL in stabilizing the genome (Thoma et al.,

394   2009; Zhang et al., 2018), we speculate that patients with the VHL mutation and high

395   EXOSC1might show higher possibility to benefit from PARP inhibitor(s).

396   However, several limitations of this study should be noted. First, we observed a

397   notable variation in terms of the correlation with a different c-substitution type for a

398   given gene, implying the need for further studies. Second, although we showed that

399   EXOSC1 could cleave ssDNA and act as an ESM, we did not directly identify the

400   mechanism responsible for turning the DNA cleavages into mutations. The role of

401   XRCC1 in EXOSC1-promoted mutations was only briefly evaluated. Hence, we

402   cannot exclude the possibility that other proteins might contribute to this process.

403   Third, because that DNA cleavage by EXOSC1 should be independent of the cancer

404   type, we propose that EXOSC1 likely contributes to the mutations in the cancers other

405   than KIRC. And more work is still needed. Despite these limitations, our results still

406   indicate that EXOSC1 acts as an ESM in KIRC.

407

408 **Materials and methods**

409 **Sample preparation**

410 Samples of 532 KIRC patients from TCGA used for expression and mutation analyses

411 were collected through The cBio Cancer Genomics Portal (http://cbioportal.org) as

412 described in our previous studies (Li et al., 2020; Zhou et al., 2019).

413

414 **Cell culture**

415 All cell lines were obtained from the American Type Tissue Collection. The 769-P,

416 SNU-1272 and Caki-2 cells were maintained in RPMI 1640 medium containing 10%

417 heat-inactivated FBS, 100 U/ml penicillin and streptomycin at 37 ℃ under a

418 humidified atmosphere of 5% CO_2. TUHR14TKB cells were maintained in DMEM

419 containing 10% heat-inactivated FBS, 100 U/ml penicillin and streptomycin at 37 ℃

420 under a humidified atmosphere of 5% CO_2. Transfections were performed using

421 lipofectamine 2000 (Thermo Fisher Scientific Inc., Waltham, MA, USA).

422

423 **Antibodies, reagents and plasmids**

424 Anti-Flag (Cat. number: F7425) and anti-His (Cat. number: SAB1306085) antibodies

425 were from Sigma-Aldrich (St. Louis, MS, USA). Anti-phospho-γ-H2AX (Ser139)

426 (Cat. number: 05-636) and anti-XRCC1 (Cat. number: SAB1306085) were from

427 Millipore (Billerica, MA, USA). Anti-EXOSC1 (Cat. number: EPR13526) was from

428 Abcam (Cambridge, MA, USA). Niraparib (Cat. number: HY-10619) and rifampicin

429 (Cat. number: R3501) were from MedChem Express and Sigma-Aldrich, respectively.

430 Full-length EXOSC1 was cloned into the X*ba* I and N*he* I sites of the lentivirus

431 vector pCDH-CMV-MCS-EF1-Puro (System Biosciences, Mountain View, CA, USA)

432 to construct pCDH-Flag EXOSC1. CCGAGTTCCTACAGACCTAAG and

433 CGAGGAACTATCCGCAAAGAA sequences were cloned into pLKO.1 to construct

434 the pLKO shEXOSC1-1 and pLKO shEXOSC1-2 plasmids, respectively. Similarly,

435 CCAGTGCTCCAGGAAGATATA and CGATACGTCACAGCCTTCAAT sequences

436 were cloned into pMKO.1 to construct the pMKO shXRCC1-1 and pMKO

437 shXRCC1-2 plasmids, respectively. According to the knockdown efficiency (Figure

438 5—figure supplement 1), pLKO shEXOSC1-1 and pMKO.1 shXRCC1-1 with higher

439 knockdown efficiencies were used to generate EXOSC1-KD and XRCC1-KD cells.

440 AID (NM_020661.4), CDK5 (NM_004935.4), TARBP2 (NM_134323.1), EXOSC1

441 (NM_016046.5), RAB5IF (NM_018840.5), CCNB1 (NM_031966.4), PSAT1

442   (NM_058179.4), NECAB3 (NM_031232.3), EXOSC2 (NM_014285.7), EXOSC3

443   (NM_016042.4), EXOSC4 (NM_019037.3), EXOSC5 (NM_020158.4), EXOSC6

444   (NM_058219.3), EXOSC7 (NM_015004.4), EXOSC8 (NM_181503.3) and EXOSC9

445   (NM_005033) were amplified and cloned into the pET-28a(+) vector (Novogen

446   Limited, Hornsby Westfield, NSW, Australia) to construct the pET-28a-Gene-6XHis *E.*

447   *coli* expression plasmids. PCR primers for the amplification of above genes are

448   described in Supplementary Table S5.

449

450   **Immunoblotting and immunofluorescence**

451   Immunoblotting and immunofluorescence were carried out as described in our

452   previous study (Song et al., 2018; Wang et al., 2020).

453

454   **Rifampicin-resistant assay in *E. coli***

455   Rifampicin-resistant assays were carried out as described previously (Petersen-Mahrt

456   et al., 2002). Briefly, rifampicin-resistant assays for each gene were performed using

457   30 independent cultures grown overnight to saturation in rich medium supplemented

458   with 50 mg/L kanamycin and 1 mM IPTG. Rif^R mutants were selected on medium

459   containing 50 mg/L rifampicin. Mutation frequencies were assessed by determining

460   the median number of rifampicin-resistant clones per 10^9 viable plated cells. The

461   mutation spectra of Rif^R were analyzed by sequencing the amplified rpoB 627-bp

462   PCR products using 5'-TTGGCGAAATGGCGGAAAACC-3' and

463   5'-CACCGACGGATACCACCTGCTG-3' primers.

464

465   **Expression and purification of EXOSC proteins**

466   Expression and purification of EXOSC proteins were carried out as described in our

467   previous study (Wang et al., 2020). pET-28a-EXOSCs-6XHises encoding His-tagged

468   EXOSCs were introduced into BL21 (DE3)-pLysS, which were grown in nutrient-rich

469   medium with 32Y (containing 3.2% (w/v) yeast extract, 0.8% (w/v) peptone and

470   0.58% (w/v) NaCl) in 10 mM Tris–HCl at 30°C and pH 7.4. Protein expression was

471   induced with 0.4 mM IPTG at 20°C for 20 h after the cells reached an OD_{600} of

472   0.4~0.5. Induced BL21 (DE3)-pLysS host cells without any plasmid were used as a

473   negative control. The resultant cells were harvested by centrifugation at 5000 g for 10

474   min and washed twice with ice-cold PBS. The collected cells were resuspended in

475   PBS (1 g of wet weight cells per 10 mL of PBS) containing 1 mM $MgCl_2$, 20 mM

476   imidazole, 1 tablet/50 mL protease inhibitor cocktail, and 100 U/mL DNase.

477   Resuspended cells were broken by an ultrasonic wave. Cell lysates were centrifuged

478   at 20,000 g at 4°C for 30 min to remove unbroken cells and debris.

479   After pre-equilibration with 10 column volumes (CV) of binding buffer (PBS

480   containing 10% (v/v) glycerol and 20 mM imidazole, pH 7.6), Ni sepharose 6 Fast

481   flow (GE Healthcare, New York, NY, USA) was applied for the purification of

482   EXOSCs. The resins were washed 5 times and eluted using elution buffer (binding

483   buffer containing 300 mM imidazole). EXOSCs were concentrated using an Amicon

484   Ultrafree centrifugal filter (Millipore Corporation, Billerica, MA, USA) and

485   pre-equilibrated with 10 mM HEPES buffer (pH 7.4) containing 150 mM NaCl and

486   10% glycerol. Size-exclusion chromatography (SEC) with a Superdex-200 HiLoad

487   10/600 column (GE Healthcare, New York, NY, USA) was used to further purify

488   EXOSCs. The purity of the fractions was analyzed by coomassie blue staining. The

489   protein concentration was determined using a BCA assay according to the

490   manufacturer's instructions (Pierce, Rockland, IL, USA).

491

492   **Cleavage activity assay *in vitro***

493   The cleavage assays of EXOSCs were carried out in reaction buffer modified from a

494   previous study (Quansheng et al., 2006). Briefly, 50 μL of reaction mixture containing

495   1 μM oligonucleotides, 1 μM EXOSC protein, 70 mM KCl, 700 μM $MgCl_2$, 1 mM

496   DTT, and 20 mM Tris-HCl pH 7.0 was incubated at 37°C for 4 h. The reaction was

497   stopped by addition of 10 μM proteinase K at 58°C for 10 min and heating at 90°C for

498   30 s. The resultant samples were then analyzed using 15% polyacrylamide TBE-urea

499   gels.

500

501   **LC-MS/MS analysis**

502   LC-MS/MS analyses of deoxyadenine (A), deoxythymidine (T), deoxyguanine (G)

503   and deoxycytocine (C) were carried out as described in our previous studies (Song et

504   al., 2018; Wang et al., 2020).

505

506   **Generation of stable cell lines**

507   Stable cell lines were generated as described in our previous study (Song et al., 2018;

508 Wang et al., 2020). Briefly, the TUHR14TKB, SNU-1272, 769-P and Caki-2 cells

509 were infected with pCDH-CMV-MCS-EF1-Puro (empty vector used as control),

510 pCDH-Flag EXOSC1 (EXOSC1-OE), pLKO.1-scramble shRNA (empty vector used

511 as control), pLKO shEXOSC1-1 (EXOSC1-KD) or pLKO shEXOSC1-2 lentiviral

512 particles, which were generated following the manufacturer's protocol (System

513 Biosciences, Mountain View, CA, USA). The resultant cells were selected with

514 puromycin for 2 weeks. These stable cells were then infected with virus encoding

515 pMKO.1, pMKO.1 shXRCC1-1 (XRCC1-KD) or pMKO.1 shXRCC1-2. The

516 resultant cells were selected with hygromycin B for 2 weeks to generate stable

517 XRCC1 knockdown cells. According to the knockdown efficiency (Figure 5—figure

518 supplement 1), shEXOSC1-1 and shXRCC1-1 with higher knockdown efficiencies

519 were used as EXOSC1-KD and XRCC1-KD in this study.

520

521 **γ-H2AX staining and neutral comet tail assays**

522 γ-H2AX staining and neutral comet tail assays were performed as described

523 previously (Li et al., 2020).

524

525 **3D-PCR and sequencing**

526 3D-PCRs of VHL mutations were carried out as described previously (Burns et al.,

527 2013a) using first (5'-GAGTACGGCCCTGAAGAAGA-3' and

528 5'-TCAATCTCCCATCCGTTGAT-3') and nested

529 (5'-TGCGCTAGGTGAACTCGC-3' and 5'-GCGGCAGCGTTGGGTAGG-3') PCR

530 primers. PCR products were then analyzed by gel electrophoresis, cloned into

531 pMD20-T vector, and sequenced.

532

533 **Colony-forming assay**

534 The colony-forming assays were performed as described in our previous study (Li et

535 al., 2020).

536

537 **Subcutaneous xenograft tumor growth *in vivo***

538 The following animal-handling procedures were approved by the Animal Care and

539 Use Committee of Dalian Medical University. Xenograft models were carried out as

540 described in our previous studies (Li et al., 2020; Song et al., 2018; Yang et al., 2010).

541 Briefly, 2×10^6 stable control/EXOSC1-OE 769-P and Caki-2 cells were suspended

542 and injected subcutaneously into the flank of 6-week-old nude mice. After 7 days,

543 these tumor-bearing mice were randomized into 4 groups (6 mice per group) and

544 treated by oral gavage twice a day with vehicle or niraparib (4 mg/kg). The mice were

545 observed daily and weighed once per week. Tumor size was measured using a caliper,

546 and the tumor volume was calculated using the following formula: $0.52 \times L \times W^2$,

547 where L is the longest diameter and W is the shortest diameter. Mice were euthanized

548 when the tumors reached 1500 mm^3 or showed necrosis.

549

550 **Statistical analyses**

551 *P* values were calculated by the two-tailed student's *t*-test, log-rank test, fisher's

552 exact-test, chi-squared test, or spearman correlation analyses as noted. *P* values < 0.05

553 were considered statistically significant.

554

555 **Availability**

556 All data associated with this study are available in the main text or the supplementary

557 materials.

558

559 **Experimental replicates and reproducibility**

560 All data presented in this paper are representative of 2–4 independent experiments

561 with comparable results.

562

563 **Acknowledgements**

571

572 **Authors' contributions**

573 Q.Y. and C.S. designed the experiments and wrote the paper. Q.X., Q.L. and L.W. performed

574 most experiments. Q.X. and Q.L. carried out EXOSC protein extraction. C.S. and W.L.

575 performed bioinformatics analyses. N.W. and K.W. maintained the cells and performed some

576 of the western blot analyses. Q.L. and L.W. carried out the γ-H2AX staining and comet

577 assays.

578

579 **Conflict of interest**

580 The authors declare no competing interests.

581

582 **References**

583 Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin,

584 A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L.*, et al.* (2013a).

585 Signatures of mutational processes in human cancer. Nature *500*, 415-421.

586 Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin,

587 A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L.*, et al.* (2013b).

588 Signatures of mutational processes in human cancer. Nature *500*, 415-421.

589 Barlow, J.H., Faryabi, R.B., Callen, E., Wong, N., Malhowski, A., Chen, H.T.,

590 Gutierrez-Cruz, G., Sun, H.W., McKinnon, P., Wright, G.*, et al.* (2013). Identification

591 of early replicating fragile sites that contribute to genome instability. Cell *152*,

592 620-632.

593 Bousquet-Antonelli, C., ., Presutti, C., ., and Tollervey, D., . (2000). Identification of a

594 regulated pathway for nuclear pre-mRNA turnover. Cell *102*, 765-775.

595 Brown, J.S., O'Carrigan, B., Jackson, S.P., and Yap, T.A. (2017). Targeting DNA

596 Repair in Cancer: Beyond PARP Inhibitors. Cancer Discov *7*, 20-37.

597 Brown, J.T., Bai, X., and Johnson, A.W. (2000). The yeast antiviral proteins Ski2p,

598 Ski3p, and Ski8p exist as a complex in vivo. Rna-a Publication of the Rna Society *6*,

599 449.

600 Buisson, R., Langenbucher, A., Bowen, D., Kwan, E.E., Benes, C.H., Zou, L., and

601 Lawrence, M.S. (2019). Passenger hotspot mutations in cancer driven by APOBEC3A

602 and mesoscale genomic features. Science *364*.

603 Burgess, D.J. (2019). Switching APOBEC mutation signatures. Nat Rev Genet *20*,

604 253.

605 Burns, M.B., Lackey, L., Carpenter, M.A., Rathore, A., Land, A.M., Leonard, B.,

606 Refsland, E.W., Kotandeniya, D., Tretyakova, N., Nikas, J.B.*, et al.* (2013a).

607 APOBEC3B is an enzymatic source of mutation in breast cancer. Nature *494*,

608 366-370.

609   Burns, M.B., Temiz, N.A., and Harris, R.S. (2013b). Evidence for APOBEC3B
610   mutagenesis in multiple human cancers. Nat Genet *45*, 977-983.

611   Cleary, J.M., Aguirre, A.J., Shapiro, G.I., and D'Andrea, A.D. (2020).
612   Biomarker-Guided Development of DNA Repair Inhibitors. Mol Cell *78*, 1070-1085.

613   Cvetkovic, M.A., Wurm, J.P., Audin, M.J., Schutz, S., and Sprangers, R. (2017). The
614   Rrp4-exosome complex recruits and channels substrate RNA by a unique mechanism.
615   Nat Chem Biol *13*, 522-528.

616   Dai, L., Zhao, T., Bisteau, X., Sun, W., Prabhu, N., Lim, Y.T., Sobota, R.M., Kaldis, P.,
617   and Nordlund, P. (2018). Modulation of Protein-Interaction States through the Cell
618   Cycle. Cell *173*, 1481-1494 e1413.

619   Dizon, D.S. (2017). PARP inhibitors for targeted treatment in ovarian cancer. Lancet
620   *390*, 1929-1930.

621   Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N., Johnson, D.A., Richardson, T.B.,
622   Santarosa, M., Dillon, K.J., Hickson, I., Knights, C.*, et al.* (2005). Targeting the DNA
623   repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917-921.

624   Greaves, M. (2018). A causal mechanism for childhood acute lymphoblastic
625   leukaemia. Nat Rev Cancer *18*, 471-484.

626   Hutchinson, L. (2013). Genetics: Signatures of mutational processes in cancer-a big
627   step closer. Nat Rev Clin Oncol *10*, 545.

628   Januszyk, K., and Lima, C.D. (2014). The eukaryotic RNA exosome. Curr Opin Struct
629   Biol *24*, 132-140.

630   Jeggo, P.A., Pearl, L.H., and Carr, A.M. (2015). DNA repair, genome stability and
631   cancer: a historical perspective. Nature Reviews Cancer *16*, 35.

632   Jeggo, P.A., Pearl, L.H., and Carr, A.M. (2016). DNA repair, genome stability and
633   cancer: a historical perspective. Nat Rev Cancer *16*, 35-42.

634   Jonathan, H., John, L.C., and David, T. (2006). RNA-quality control by the exosome.
635   Nat Rev Mol Cell Biol *7*, 529-539.

636   Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q.,
637   McMichael, J.F., Wyczalkowski, M.A.*, et al.* (2013). Mutational landscape and
638   significance across 12 major cancer types. Nature *502*, 333-339.

639   Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of
640   the nuclear RNA exosome complex. Nat Rev Mol Cell Biol *17*, 227-239.

641   Li, S., Zhang, Y., Wang, N., Guo, R., Liu, Q., Lv, C., Wang, J., Wang, L., and Yang,

642   Q.K. (2020). Pan-cancer analysis reveals synergistic effects of CDK4/6i and PARPi
643   combination treatment in RB-proficient and RB-deficient breast cancer cells. Cell
644   Death Dis *11*, 219.

645   Lim, J., Giri, P.K., Kazadi, D., Laffleur, B., Zhang, W., Grinstein, V., Pefanis, E.,
646   Brown, L.M., Ladewig, E., Martin, O.*, et al.* (2017). Nuclear Proximity of Mtr4 to
647   RNA Exosome Restricts DNA Mutational Asymmetry. Cell *169*, 523-537 e515.

648   Loeb, L.A. (2011). Human cancers express mutator phenotypes: origin, consequences
649   and targeting. Nat Rev Cancer *11*, 450-457.

650   Lord, C.J., and Ashworth, A. (2012). The DNA damage response and cancer therapy.
651   Nature *481*, 287-294.

652   Lord, C.J., and Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic.
653   Science *355*, 1152-1158.

654   Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A., and Conti, E. (2008).
655   Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of
656   substrate recruitment in the RNase II nuclease family. Molecular Cell *29*, 717-728.

657   Malet, H., Topf, M., Clare, D.K., Ebert, J., Bonneau, F., Basquin, J., Drazkowska, K.,
658   Tomecki, R., Dziembowski, A., Conti, E.*, et al.* (2010). RNA channelling by the
659   eukaryotic exosome. EMBO Rep *11*, 936-942.

660   McGranahan, N., Rosenthal, R., Hiley, C.T., Rowan, A.J., Watkins, T.B.K., Wilson,
661   G.A., Birkbak, N.J., Veeriah, S., Van Loo, P., Herrero, J.*, et al.* (2017). Allele-Specific
662   HLA Loss and Immune Escape in Lung Cancer Evolution. Cell *171*, 1259-1271
663   e1211.

664   Pearl, L.H., Schierz, A.C., Ward, S.E., Al-Lazikani, B., and Pearl, F.M. (2015).
665   Therapeutic opportunities within the DNA damage response. Nat Rev Cancer *15*,
666   166-180.

667   Pefanis, E., and Basu, U. (2015). RNA Exosome Regulates AID DNA Mutator
668   Activity in the B Cell Genome. Adv Immunol *127*, 257-308.

669   Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao,
670   J., Elliott, O., Liu, Z.P.*, et al.* (2015). RNA exosome-regulated long non-coding RNA
671   transcription controls super-enhancer activity. Cell *161*, 774-789.

672   Peng, L., Bian, X.W., Li, D.K., Xu, C., Wang, G.M., Xia, Q.Y., and Xiong, Q. (2015).
673   Large-scale RNA-Seq Transcriptome Analysis of 4043 Cancers and 548 Normal
674   Tissue Controls across 12 TCGA Cancer Types.   *5*, 13413.

675 Petersen-Mahrt, S.K., Harris, R.S., and Neuberger, M.S. (2002). AID mutates E. coli
676 suggesting a DNA deamination mechanism for antibody diversification. Nature *418*,
677 99.

678 Petljak, M., Alexandrov, L.B., Brammeld, J.S., Price, S., Wedge, D.C., Grossmann, S.,
679 Dawson, K.J., Ju, Y.S., Iorio, F., Tubio, J.M.C.*, et al.* (2019). Characterizing
680 Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC
681 Mutagenesis. Cell *176*, 1282-1294 e1220.

682 Quansheng, L., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and
683 structure of the eukaryotic RNA exosome. Cell *127*, 1223-1237.

684 Roberts, S.A., and Gordenin, D.A. (2014). Hypermutation in human cancer genomes:
685 footprints and mechanisms. Nat Rev Cancer *14*, 786-800.

686 Robertson, A.G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A.D.,
687 Hinoue, T., Laird, P.W., Hoadley, K.A., Akbani, R.*, et al.* (2017). Comprehensive
688 Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell *171*, 540-556
689 e525.

690 Roos, W.P., Thomas, A.D., and Kaina, B. (2015). DNA damage and the balance
691 between survival and death in cancer biology. Nature Reviews Cancer *16*, 20.

692 Roos, W.P., Thomas, A.D., and Kaina, B. (2016). DNA damage and the balance
693 between survival and death in cancer biology. Nat Rev Cancer *16*, 20-33.

694 Sale, J.E., Calandrini, D.M., Takata, M., ., Takeda, S., ., and Neuberger, M.S. (2001).
695 Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic
696 hypermutation. Nature *412*, 921-926.

697 Schaaper, R.M., and Dunn, R.L. (1987). Spectra of spontaneous mutations in
698 Escherichia coli strains defective in mismatch correction: the nature of in vivo DNA
699 replication errors. Proceedings of the National Academy of Sciences of the United
700 States of America *84*, 6220-6224.

701 Song, C., Wang, L., Wu, X., Wang, K., Xie, D., Xiao, Q., Li, S., Jiang, K., Liao, L.,
702 Yates, J.R., 3rd*, et al.* (2018). PML Recruits TET2 to Regulate DNA Modification and
703 Cell Proliferation in Response to Chemotherapeutic Agent. Cancer Res *78*,
704 2475-2489.

705 Thoma, C.R., Toso, A., Gutbrodt, K.L., Reggi, S.P., Frew, I.J., Schraml, P., Hergovich,
706 A., Moch, H., Meraldi, P., and Krek, W. (2009). VHL loss causes spindle
707 misorientation and chromosome instability. Nat Cell Biol *11*, 994-1001.

708 Tomecki, R., Drazkowska, K., and Dziembowski, A. (2010). Mechanisms of RNA
709 degradation by the eukaryotic exosome. Chembiochem *11*, 938-945.

710 Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of
711 Genomic Instability in Cancer. Cell *168*, 644-656.

712 Tutt, A., Robson, M., Garber, J.E., Domchek, S.M., Audeh, M.W., Weitzel, J.N.,
713 Friedlander, M., Arun, B., Loman, N., Schmutzler, R.K.*, et al.* (2010). Oral
714 poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2
715 mutations and advanced breast cancer: a proof-of-concept trial. Lancet *376*, 235-244.

716 Uhlen, M., Zhang, C., Lee, S., Sjostedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R.,
717 Arif, M., Liu, Z., Edfors, F.*, et al.* (2017). A pathology atlas of the human cancer
718 transcriptome. Science *357*.

719 Ulmke, P.A., Xie, Y., Sokpor, G., Pham, L., Shomroni, O., Berulava, T., Rosenbusch,
720 J., Basu, U., Fischer, A., Nguyen, H.P.*, et al.* (2021). Post-transcriptional regulation by
721 the exosome complex is required for cell survival and forebrain development via
722 repression of P53 signaling. Development *148*.

723 Wang, L., Song, C., Wang, N., Li, S., Liu, Q., Sun, Z., Wang, K., Yu, S.C., and Yang,
724 Q. (2020). NADP modulates RNA m(6)A methylation and adipogenesis via enhancing
725 FTO activity. Nat Chem Biol *16*, 1394-1402.

726 Wasmuth, E.V., Kurt, J., and Lima, C.D. (2014). Structure of an Rrp6-RNA exosome
727 complex bound to poly(A) RNA. Nature *511*, 435-439.

728 Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott,
729 K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas
730 Pan-Cancer analysis project. Nature genetics *45*, 1113-1120.

731 Yang, Q., Deng, X., Lu, B., Cameron, M., Fearns, C., Patricelli, M.P., Yates, J.R., 3rd,
732 Gray, N.S., and Lee, J.D. (2010). Pharmacological inhibition of BMK1 suppresses
733 tumor growth through promyelocytic leukemia protein. Cancer Cell *18*, 258-267.

734 Zhang, J., Wu, T., Simon, J., Takada, M., Saito, R., Fan, C., Liu, X.D., Jonasch, E.,
735 Xie, L., Chen, X.*, et al.* (2018). VHL substrate transcription factor ZHX2 as an
736 oncogenic driver in clear cell renal cell carcinoma. Science *361*, 290-295.

737 Zhang, Y., Ng, K.S., Kucherlapati, M., Chen, F., Liu, Y., Tsang, Y.H., Velasco, G.D.,
738 Kang, J.J., Akbani, R., and Hadjipanayis, A. (2017). A Pan-Cancer Proteogenomic
739 Atlas of PI3K/AKT/mTOR Pathway Alterations. Cancer Cell *31*, 820.

740 Zhou, S.L., Zhou, Z.J., Hu, Z.Q., Song, C.L., Luo, Y.J., Luo, C.B., Xin, H.Y., Yang,

741   X.R., Shi, Y.H., Wang, Z*., et al.* (2019). Genomic sequencing identifies WNK2 as a

742   driver in hepatocellular carcinoma and a risk factor for early recurrence. J Hepatol *71*,

743   1152-1163.

Figure legends

Figure 1. Identification of candidate ESMs in KIRC by statistical analyses. (**A**) Schematic of this study. (**B**) Illustration of base substitutions. (**C**) Schematic showing the archaeal and eukaryotic exosome complexes viewed from the top. (**D**) Summary statistics for the 6 types of c-substitutions in KIRC. (**E**) Scatter plots showing the correlation between the rank of mutation and gene expression. Each plot represents one KIRC sample. The orange dashed line shows the best fit for visualization. *P* values were calculated by spearman's rank correlation. (**F**) Volcano plots of *p* and r values calculated by spearman correlation analyses. Each plot represents one gene. The top 1% of genes were taken as candidates and marked in red. (**G**) Box plots showing ASNS expression in the high and low C>A/G>T mutation groups. The expression was normalized to TBP. (**H**) Volcano plots showing the *p* and fold change (FC) values calculated by the two-tailed student's *t*-test. Each plot represents one gene. FC was calculated by the formula: FC = the mean gene expression in the high group/that in the low group. The top 1% genes were taken as candidates and marked in red. (**I**) C-substitution mutation frequencies in 5 types of major cancers. (**J, K**) Mutation spectra of the TP53 gene in BRCA (**J**) and VHL gene in KIRC (**L**). (**L**) Kaplan-Meier (KM) analyses of OS between VHL C>A/G>T and C>T/G>A mutation groups. The median OSs in the C>A/G>T and C>T/G>A groups were 72.95 and 108.91 months, respectively. The *p* value was obtained from the log-rank test. (**M**) Venn diagram showing the overlap of the candidate genes identified by three types of statistical analyses as noted.

Figure 1
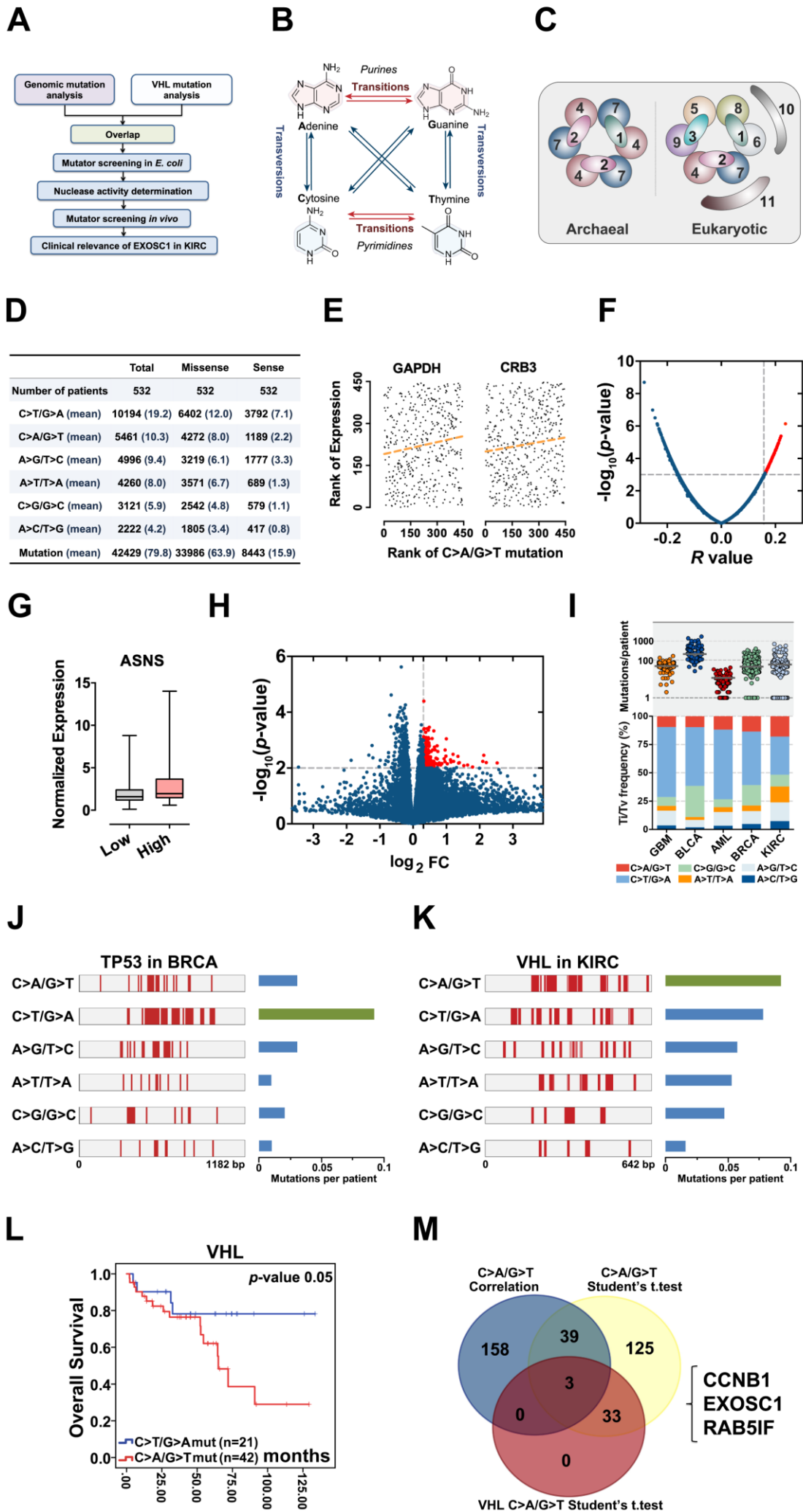
Figure 2. EXOSC1 promotes mutations in *E. coli*. (**A**) Western blot showing His-tagged protein levels in *E. coli*. (**B, C**) The frequencies of the Rif^R mutation in the *E. coli* cells expressing candidate ESMs (**B**) or exosome family members (**C**). The vector and AID were used as negative and positive controls, respectively. Each plot represents the mutational frequency of an independent overnight culture (n=30). Median mutational frequency of the gene is noted. (**D**) Frequencies of the Rif^R mutation in the *E. coli* cells treated with and without IPTG (n=30). (**E, F**) Representative images of *E. coli* cells treated with (**E**) and without (**F**) IPTG. (**G**) The mutational distribution in 25 independent Rif^R colonies transformed by vector or EXOSC1. (**H, I**) Summary of the c-substitutions in Rif^R colonies transformed by vector (**H**) and EXOSC1 (**I**). (**J**) The mutational frequencies of each c-substitution in Rif^R colonies. The *p* value was calculated by fisher's exact-test.
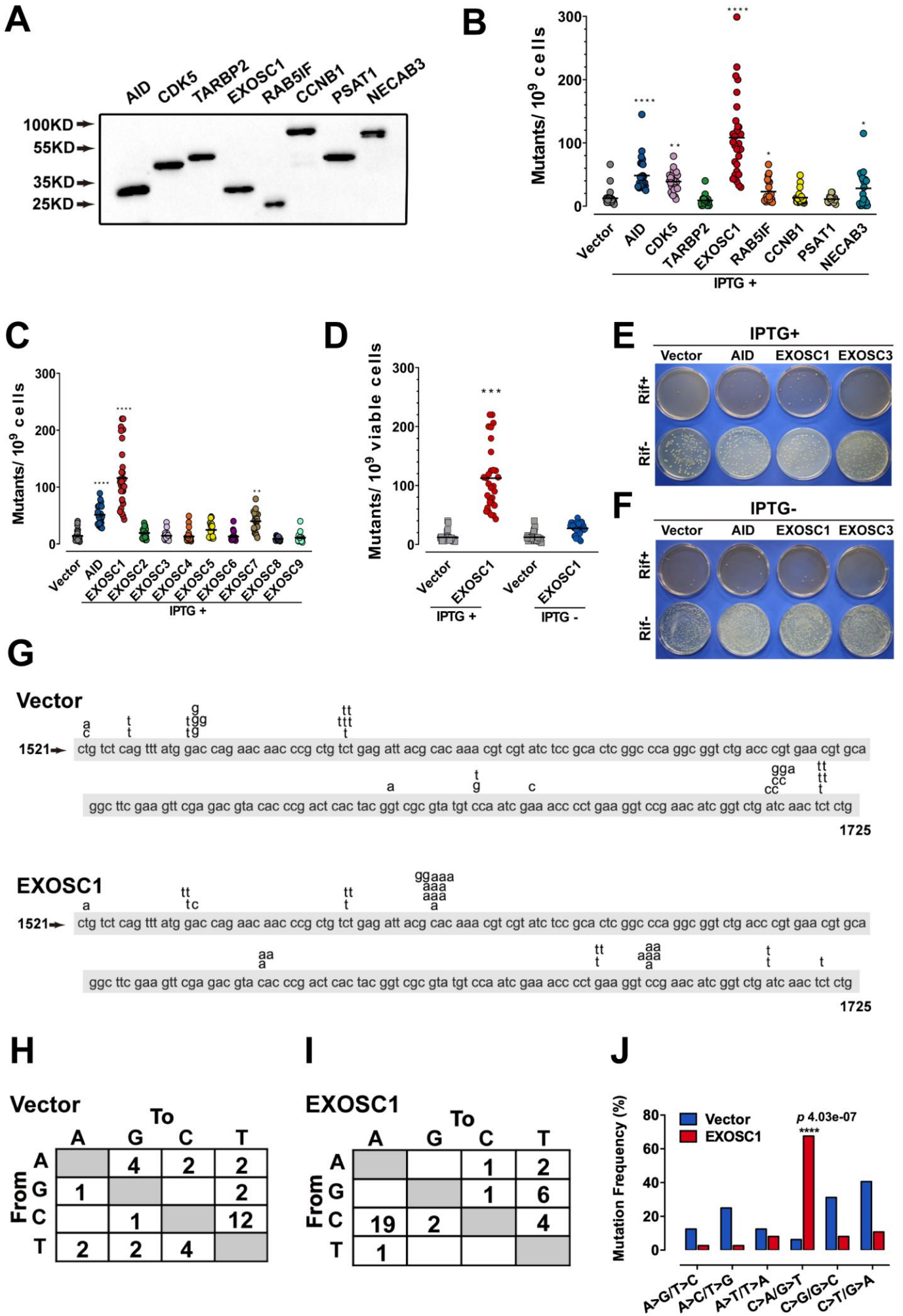
Figure 2

Figure 3. EXOSC1 cleaves ssDNA. (**A**) Coomassie blue staining of *in vitro* purified EXOSC1 protein. (**B**) Schematic of synthetic DNA substrates. (**C**) *In vitro* cleavage assays of EXOSC1 using generic ssDNA, dsDNA and DNA-RNA hybrid as substrates. The assays were performed in the presence of 1 μM oligonucleotides, 1 μM EXOSC1, 70 mM KCl, 700 μM MgCl$_2$, 1 mM DTT, and 20 mM Tris-HCl pH 7.0. After incubation at 37 ℃ for 4 h, resultant samples were analyzed by 15% polyacrylamide TBE-urea gels. (**D**) Cleavage assays of EXOSC1 in the presence or absence of EXOSC3, EXOSC6, EXOSC7 and EXOSC8 using generic ssDNA as substrates. (**E**) Cleavage assays in the presence of the components as noted. (**F**) Cleavage assays of EXOSC1 at the pH as noted. (**G**) Time course cleavage assays of EXOSC1 using generic ssDNA as substrates. (**H**) Rate curve of EXOSC1 cleavage at 37 ℃ and pH 7.0.
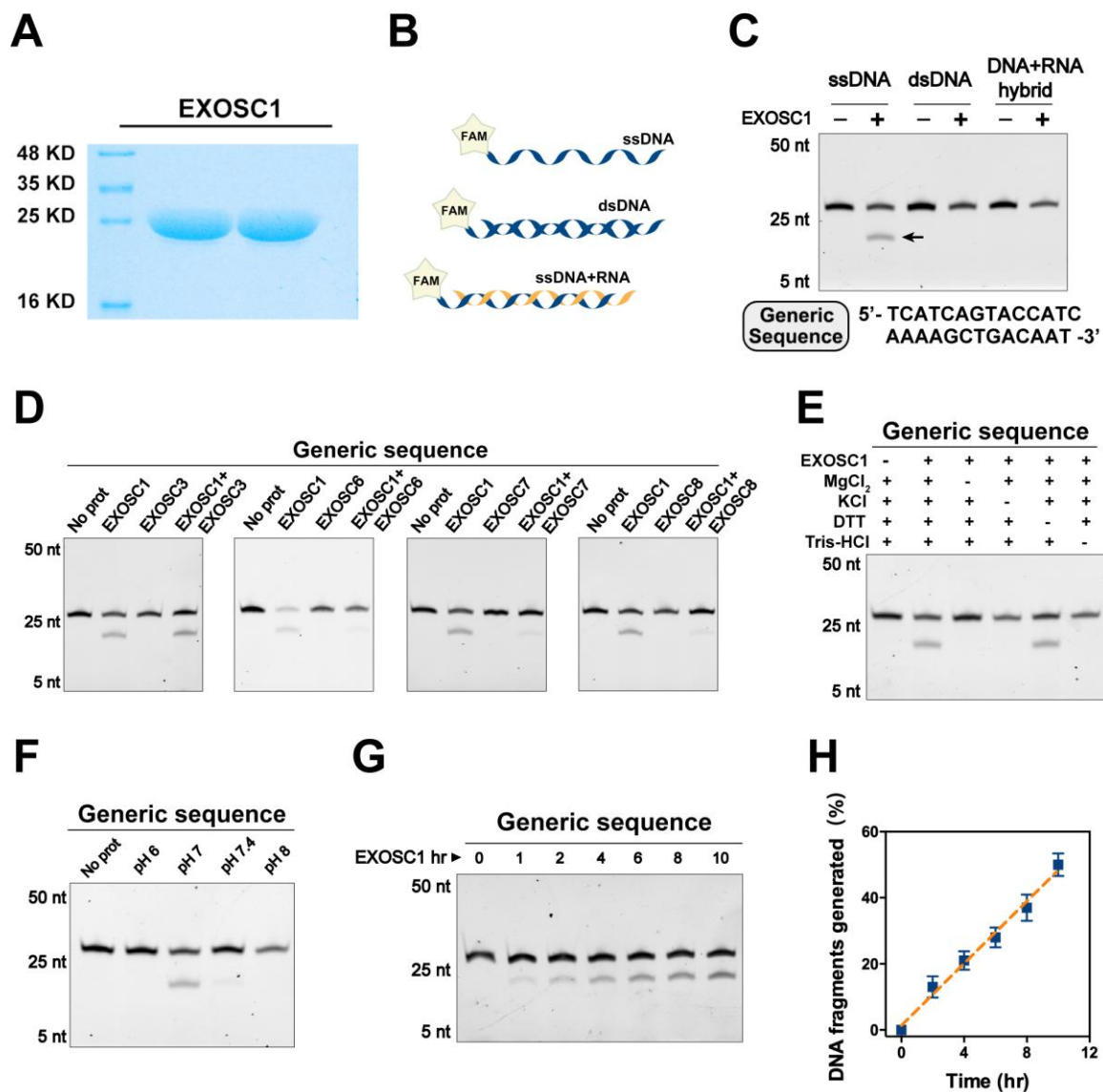
Figure 3

Figure 4. EXOSC1 prefers to cleave C sites in ssDNA. (**A, B**) Cleavage assay of EXOSC1 using unbiased DNA, dsDNA and DNA-RNA hybrid as substrates. (**C, D**) Cleavage assays of EXOSC1, EXOSC3, EXOSC7 and EXOSC8 using unbiased ssDNA as substrates. (**E**) Mass spectrometry (MS) analyses of the resultant mixtures described in (**B**). C, G, A and T were detected by MS using nucleoside to base ion mass transitions of 228.1 to 112.2 (C), 268.1 to 152.1 (G), 252.2 to 136.1 (A), and 243.1 to 127.2 (T) (Supplementary Figure 3E). Standard curves were generated by a serial dilution of C, G, A and T (Supplementary Figure 3F). Free C, G, A and T contained in the reaction mixtures were quantified by standard curves. (**F**) Schematic showing that the consequence of C>A mutation in the coding strand is distinct from that in the template strand. (**G**) Correlation between EXOSC1 expression and the c-substitution mutation as noted in KIRC. Each line represents one best fit for visualization. *P* values were from spearman's rank correlation. *P* and r values of C>A mutations were 0.0001 and 0.19, respectively; C>A/G>T: $p = 0.0006$, r = 0.17; G>T: $p = 0.0271$, r = 0.10; total mutations: $p = 0.0594$, r = 0.09; C/G>G/C: $p = 0.5730$, r = 0.03. (**H**) Student's *t*-tests analyses of the expression difference of EXOSC1 between the high and low mutation groups. FC = the mean gene expression in the high group/that in the low group.

Figure 4
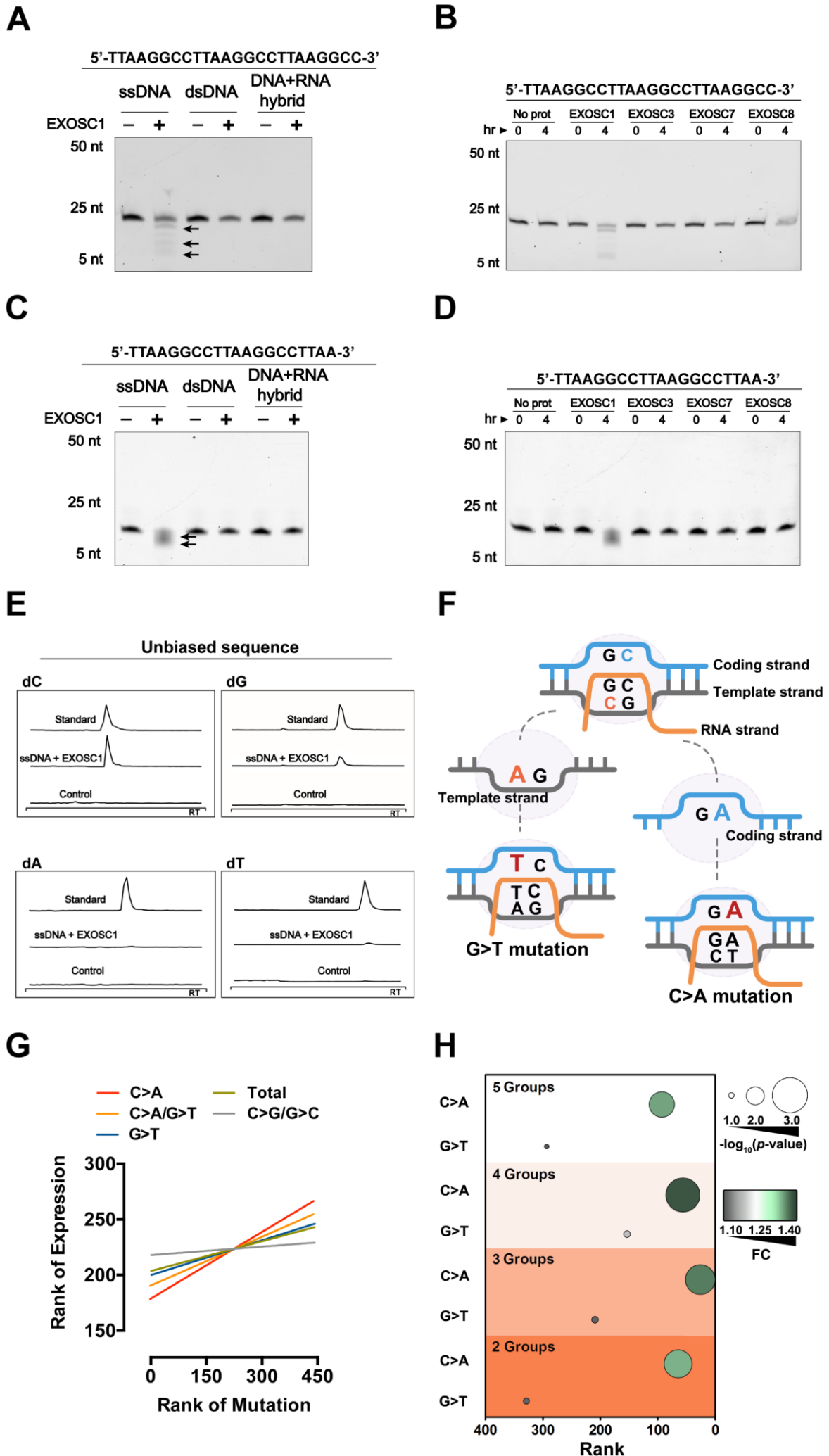
Figure 5. EXOSC1 enhances DNA damages and mutations in KIRC Cells. (**A**) Representative fluorescent images of γ-H2AX foci in 769-P cells transfected with control (pCDH, empty vector) or pCDH-Flag EXOSC1 plasmids for 2 days. Scale bar = 10 μm. (**B**) Percentage of cells with more than 20 γ-H2AX foci in the KIRC cells transfected with control or pCDH-Flag EXOSC1 plasmids for 2 days. (**C**) Percentage of cells with more than 20 γ-H2AX foci in the KIRC cells infected with lentivirus encoding shRNAi control (pLKO scramble) or pLKO sh-EXOSC1 for 2 days. (**D**) Representative images of DNA comets in 769-P cells transfected with control or pCDH-Flag EXOSC1 plasmids for 2 days. (**E**) Comet tail moment of the 769-P and TUHR14TKB cells transfected with control or pCDH-Flag EXOSC1 plasmids for 2 days. (**F**) 3D-PCR and subsequent sequencing analyses of the VHL mutations in the TUHR14TKB cells stably expressing control (vector) or EXOSC1 (pCDH-Flag EXOSC1, EXOSC1-OE). (**G**) 3D-PCR analyses of VHL in TUHR14TKB cells stably expressing control (pLKO.1 vector) or shRNA against EXOSC1 (pLKO shEXOSC1-1, EXOSC1-KD). (**H**) 3D-PCR analyses of VHL in TUHR14TKB cells stably expressing control, EXOSC1-OE and/or shRNA against XRCC1 (pMKO.1 shXRCC1-1, XRCC1-KD). (**I**) 3D-PCR analyses of VHL in TUHR14TKB cells stably expressing control (pLKO.1 vector), shRNA against EXOSC1 (pLKO shEXOSC1-1, EXOSC1-KD) and/or shRNA against XRCC1 (pMKO.1 shXRCC1-1, XRCC1-KD). (**J**) Schematic showing the subcutaneous xenograft tumor models. The 5 x 10^6 control, EXOSC1-OE and EXOSC1-KD 769-P cells were implanted subcutaneously. After 2 weeks, the resultant tumors were analyzed by 3D-PCR. (**K**) 3D-PCR analyses of VHL in the tumors described in (**J**).

Figure 5
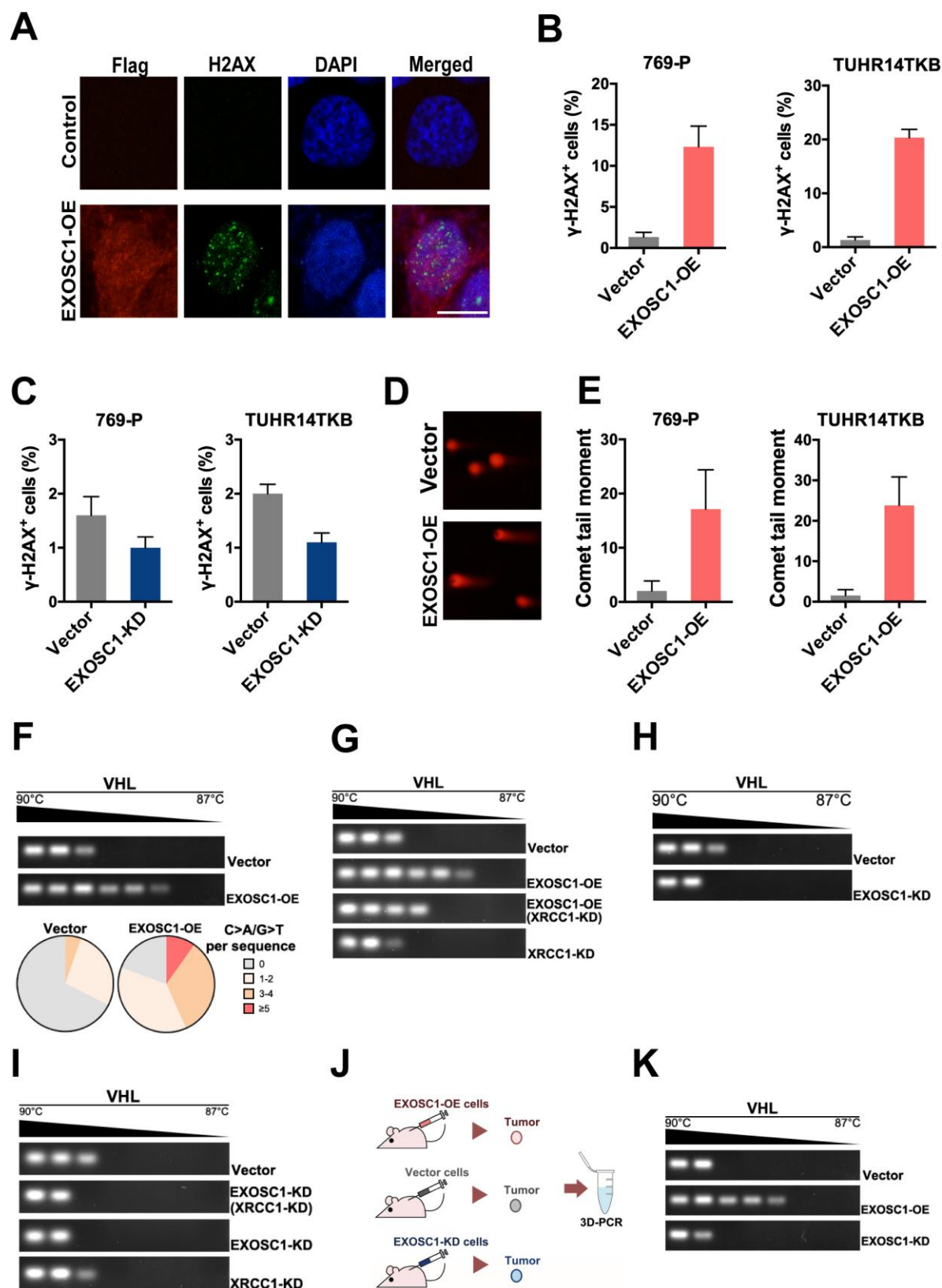
Figure 6. EXOSC1 sensitizes KIRC cells to PARP inhibitor. (**A, B**) KM analyses of DFS (**A**) and OS (**B**) in KIRC patients with different EXOSC1 levels. *P* values were obtained from the log-rank test. (**C**) KM analysis of OS in KIRC patients with different EXOSC1 expression levels and VHL mutations. (**D**) Colony formation of control and EXOSC1-OE TUHR14TKB cells treated with 600 nM niraparib. (**E, F**) Clonogenic survival of control and EXOSC1-OE TUHR14TKB (**E**) and 769-P (**F**) cells treated with niraparib. (**G, I**) Tumor volumes of 769-P (**G**) and Caki-2 (**I**) xenografts treated with or without niraparib. n = 4 groups; n = 6 mice/group; ± SEM. (**H, J**) Body weight change percentage of 769-P (**H**) and Caki-2(**J**) xenografts treated as described in (**G**).

Figure 6
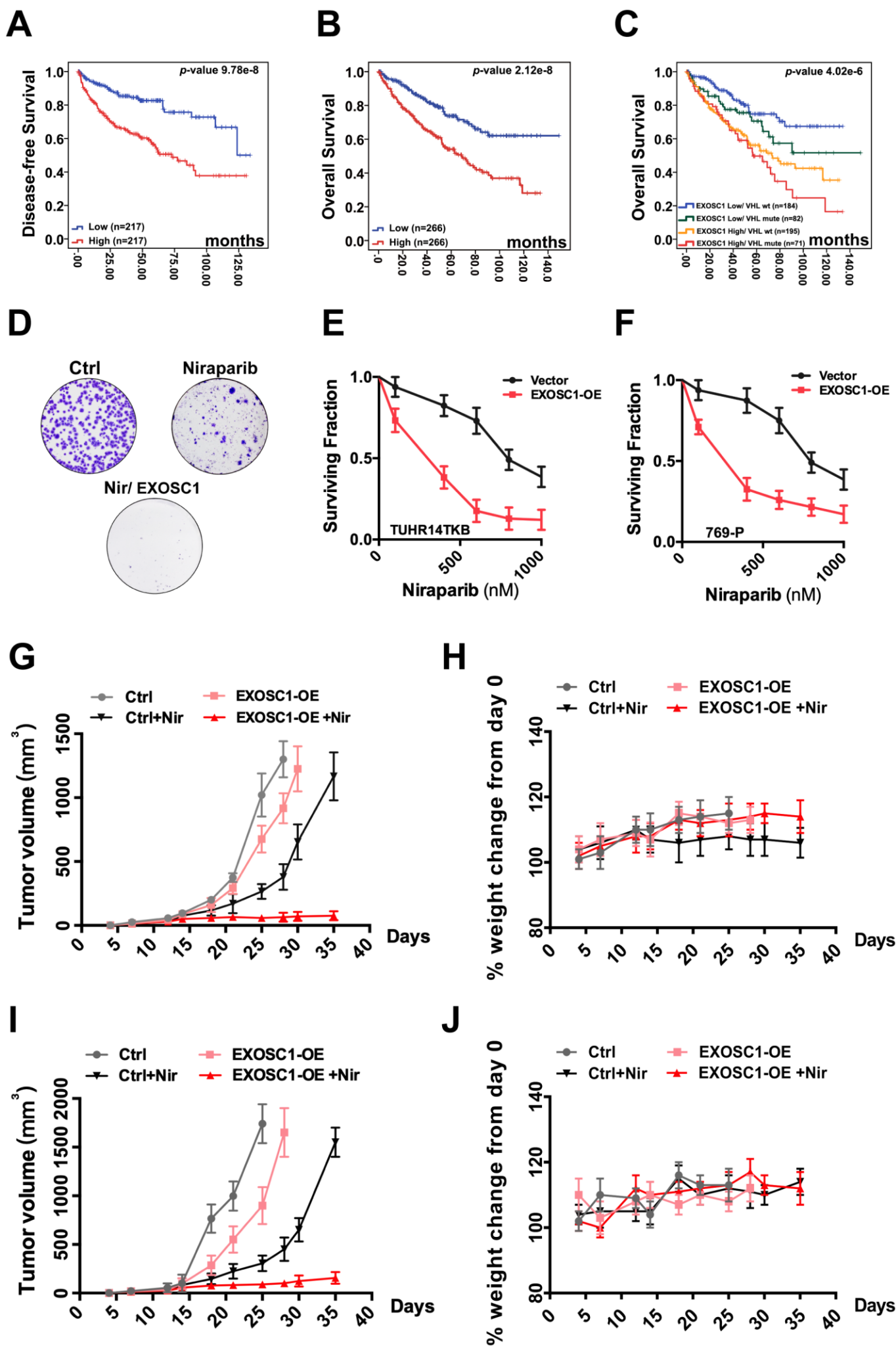
Supplementary Figure Legends

Figure 1—figure supplement 1. Identification of candidate ESMs in KIRC by statistical analyses. (**A** and **B**) Mutation spectra of TP53 (**A**) and VHL (**B**) normalized by the nucleotide abundance.

Figure 1—figure supplement 1
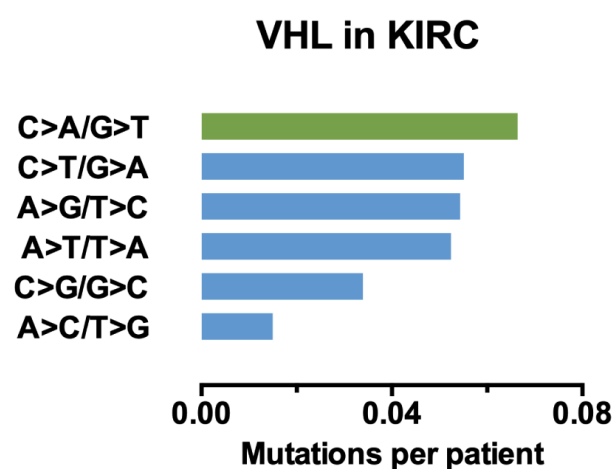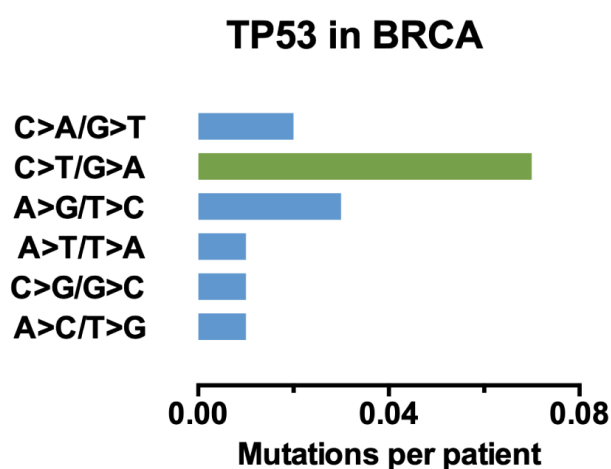
Figure 2—figure supplement 1. EXOSC1 promotes mutations in *E. coli*. (**A**) Setup plot showing the overlap numbers of the candidate genes of the 6 types of c-substitutions. More details can be found in Supplementary Table S4. (**B**) Western blot showing the His-tagged protein levels of the exosome complex members in the *E. coli* cells. (**C**) Growth rates of control and EXOSC1-transformed cells in the presence of IPTG. n = 5. (**D**) Frequencies of c-substitutions in Rif^R colonies were normalized to the nucleotide abundances. The *p* value was calculated by fisher's exact test.

Figure 2—figure supplement 1
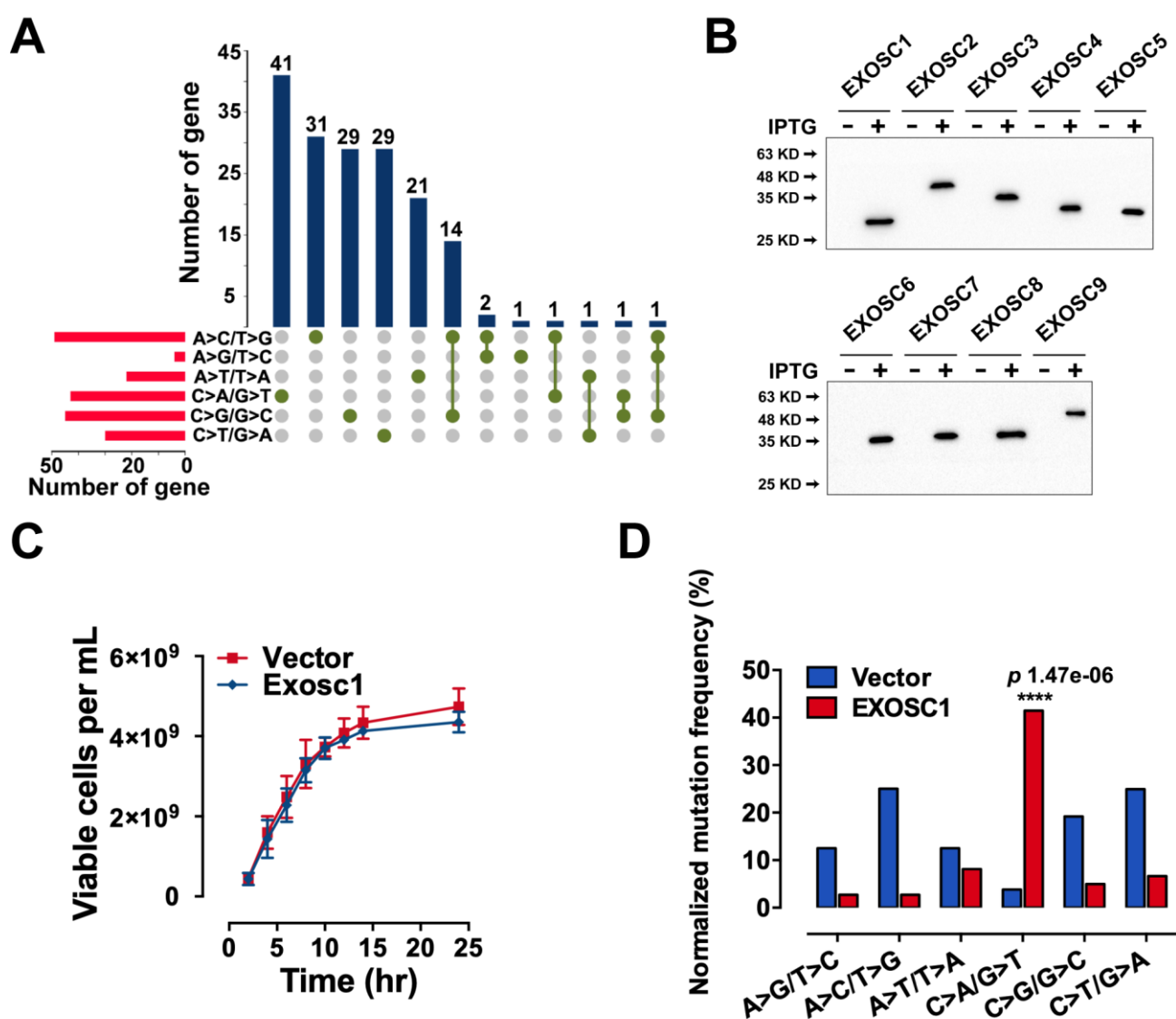
Figure 3—figure supplement 1. EXOSC1 cleaves ssDNA. (**A**) *In vitro* cleavage assays of EXOSC1 homologs (EXOSC2–EXOSC9) using generic ssDNA, dsDNA and DNA-RNA hybrid as substrates. (**B**) Cleavage assays of EXOSC1 in the presence/absence of EXOSC2, EXOSC4, EXOSC5 or EXOSC9.
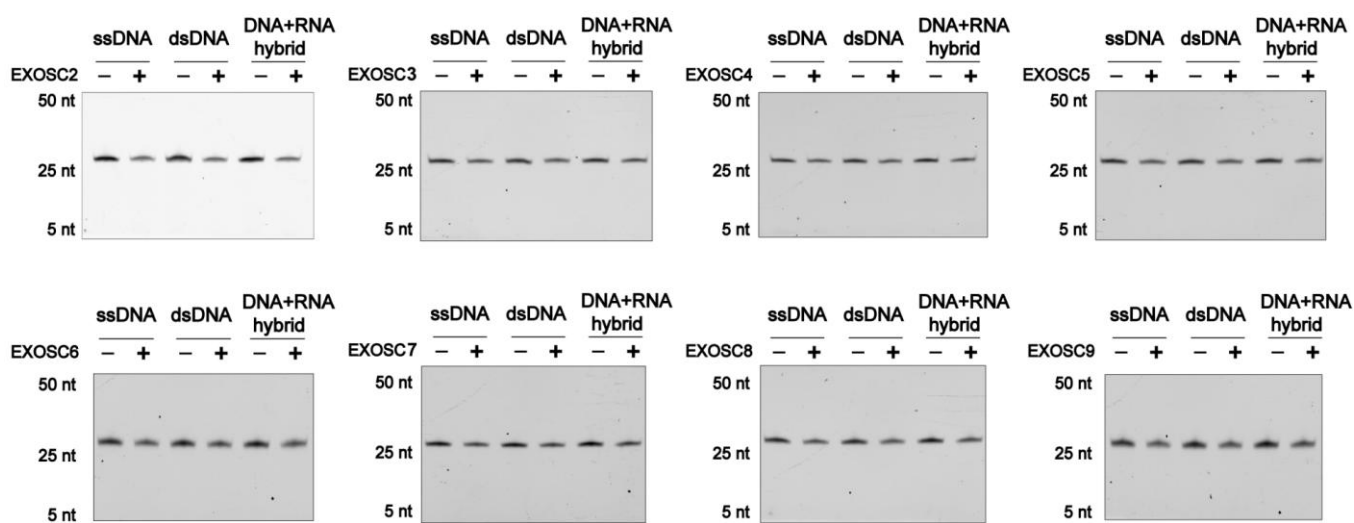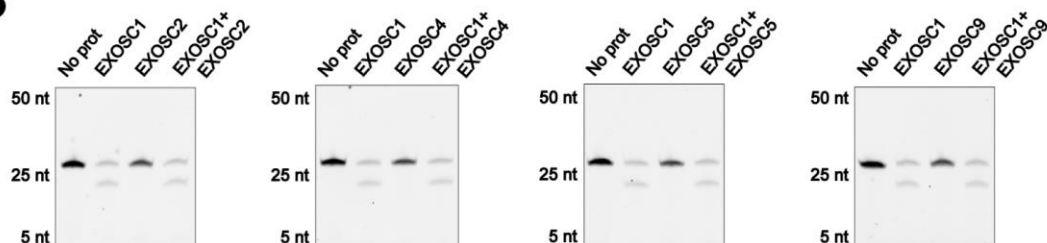
Figure 4—figure supplement 1. EXOSC1 prefers to cleave C sites in ssDNA. (**A**) Cleavage assays of EXOSC2-EXOSC9 using unbiased ssDNA, dsDNA and DNA-RNA hybrid as substrates. (**B**) Time course cleavage assays of EXOSC1 using unbiased ssDNA as substrates. (**C**) Base ion mass transitions for LC-MS/MS analyses of C, G, A and T. Free nucleosides were detected using nucleoside to base ion mass transitions of 228.1 to 112.2 (C), 268.1 to 152.1 (G), 252.2 to 136.1 (A), and 243.1 to 127.2 (T). (**D**) LC-MS/MS standard curves of nucleoside as noted. (**E** and **F**) Frequencies (**E**) and normalized frequencies (**F**) of 12 types of substitutions in Rif^R colonies. *P* values were calculated using fisher's exact test. (**G**) Correlation between EXOSC1 expression and the mutation as noted in KIRC. Each plot represents one KIRC sample. (**H**) Student's *t*-test analyses of EXOSC1 expression difference between the high and low mutation groups. The 532 KIRC patients were grouped into 2, 3, 4 or 5 groups to evaluate the impact of group number. (**I** and **J**) Summary statistics for the mutation as noted in KIRC. *P* values were calculated by the two-tailed student's *t*-test.

Figure 4—figure supplement 1
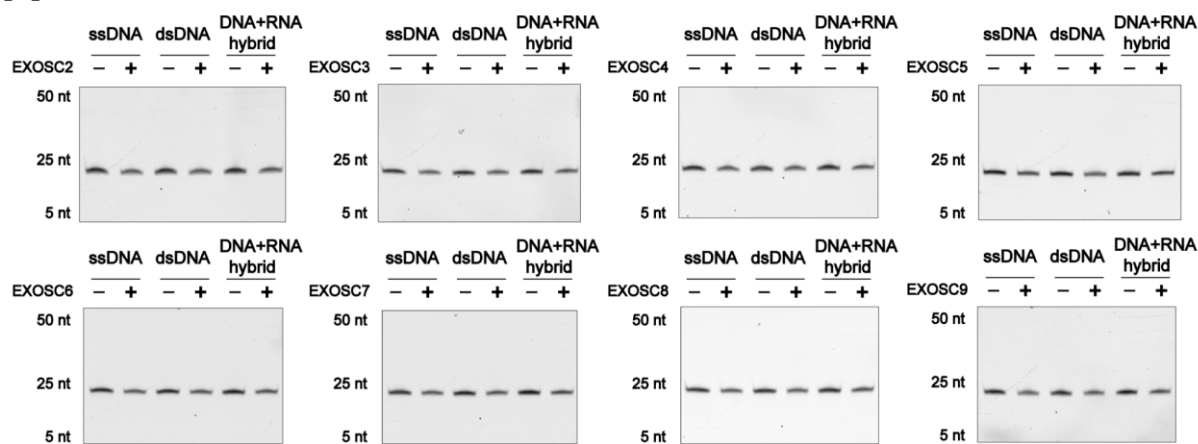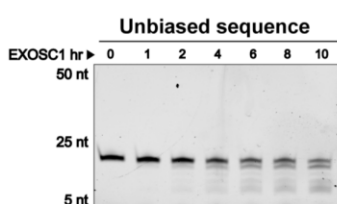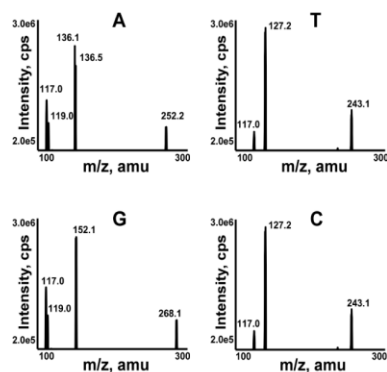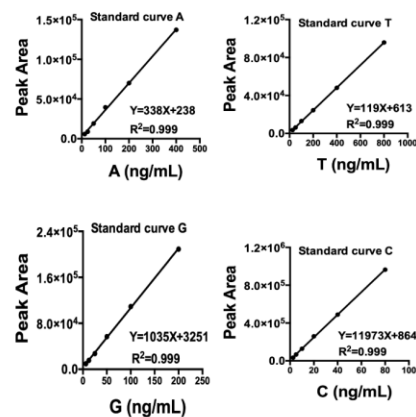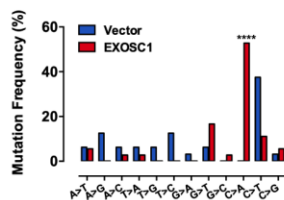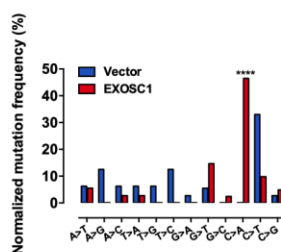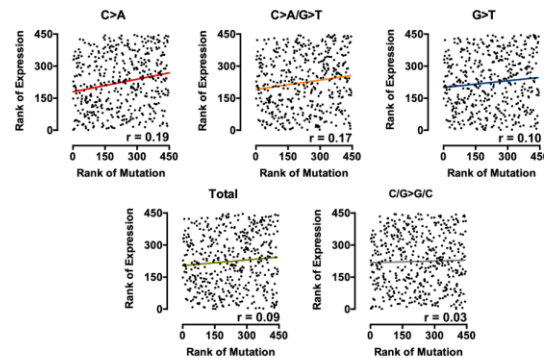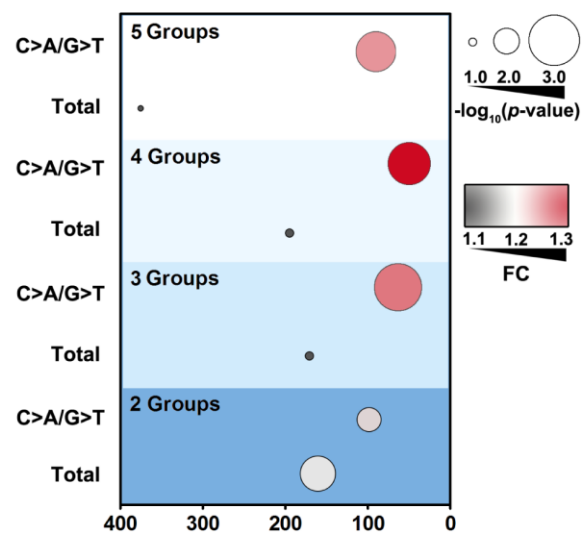
Figure 5—figure supplement 1. EXOSC1 enhances DNA damages and mutations in KIRC cells. (**A**) Western blot analyses of EXOSC1 in stable control, enhanced EXOSC1, and EXOSC1 knockdown cells. (**B**) Qantitative PCR analyses of EXOSC1 mRNA in the 769-P cells stably expressing shRNA (pLKO EXOSC1-1 or pLKO EXOSC1-2) against EXOSC1. (**C**) Representative fluorescence images of γ-H2AX foci in the 769-P cells stably expressing shRNA against control (empty vector) or EXOSC1 (pLKO EXOSC1-1, EXOSC1-KD). Scale bar = 10 μm. (**D** and **F**) Western blot analyses of EXOSC1 and XRCC1 protein levels in the stable cells as noted. (**E**) Qantitative PCR analyses of the mRNA in the cells stably expressing shRNA against EXOSC1 and/or XRCC1.

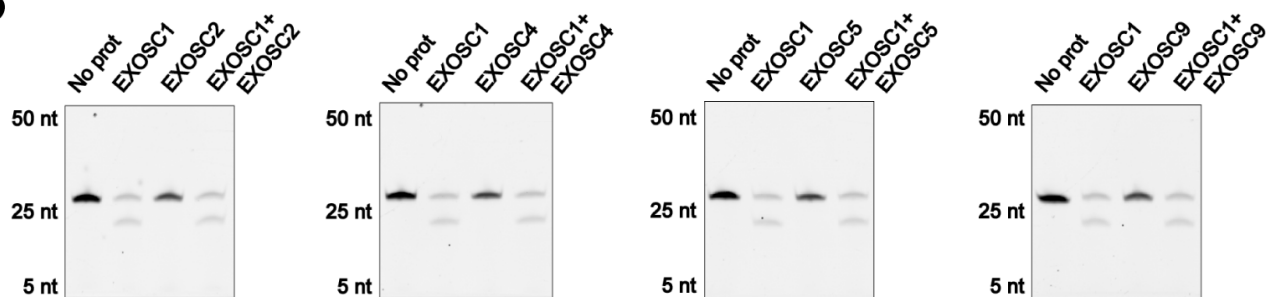

Figure 5—figure supplement 1
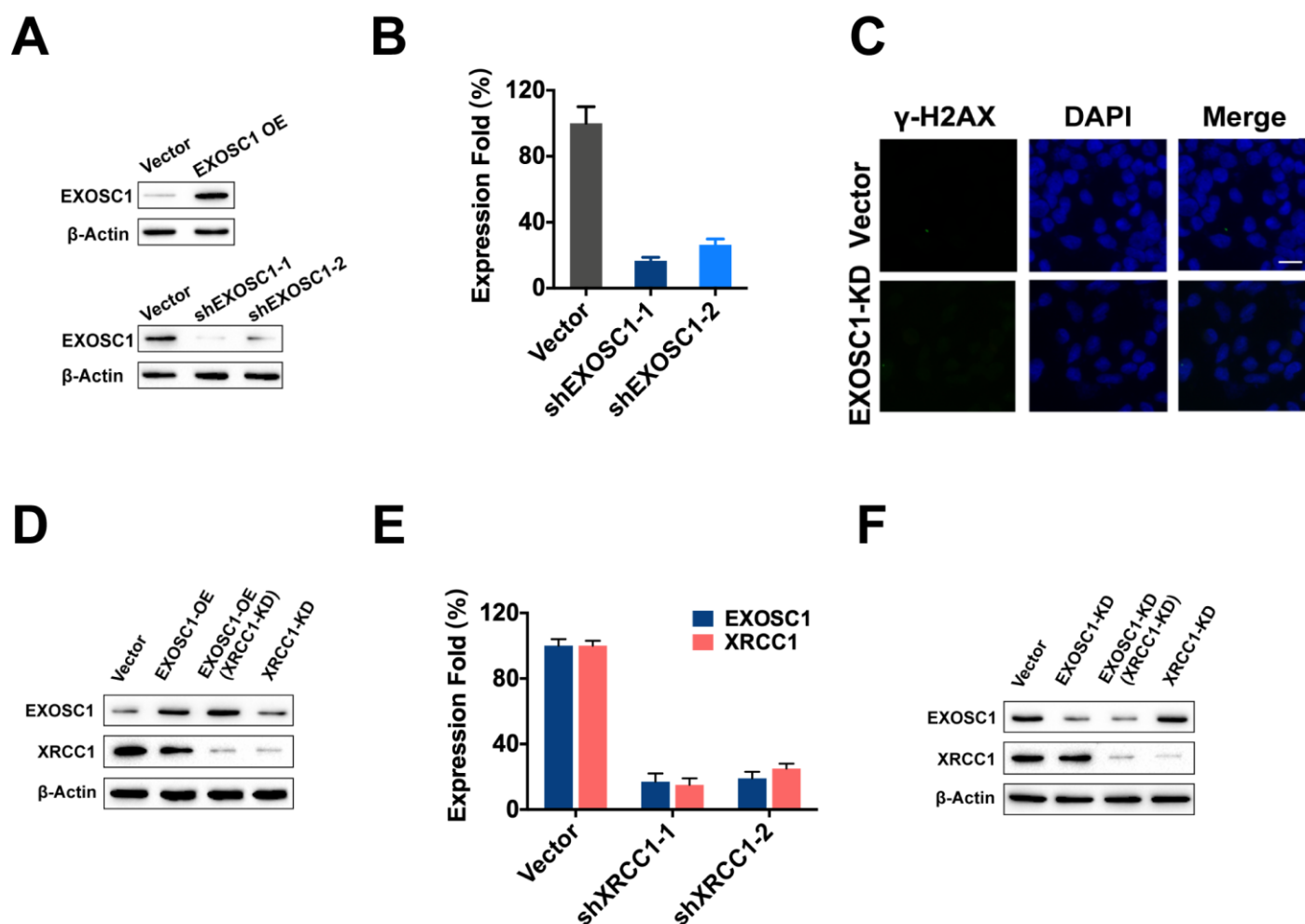
**Figure 6—figure supplement 1.** EXOSC1 sensitizes KIRC cells to PARP inhibitor. (**A**) The fragments per kilo base per million mapped reads (FPKM) of the noted gene in 532 KIRC. Each dot represents one KIRC patient. (**B**) Best separation KM analyses of OS in KIRC patients with different EXOSC1 levels. The median OS (the low vs high EXOSC1 group) = 106.40 vs 71.91 months. (**C**) Summary statistics for Figure 6C. (**D**) Clonogenic survival of control/EXOSC1-OE SNU-1272 and Caki-2 cells treated with niraparib.

Figure 6—figure supplement 1

Figure 1

Figure 2

Figure 3

A

EXOSC1

48 KD
35 KD
25 KD
16 KD

B

FAM ssDNA

FAM dsDNA

FAM ssDNA+RNA

C

ssDNA	dsDNA	DNA+RNA hybrid

EXOSC1   −   +   −   +   −   +

50 nt
25 nt
5 nt

Generic Sequence   5'- TCATCAGTACCATC AAAAGCTGACAAT -3'

D

Generic sequence

No prot, EXOSC1, EXOSC3, EXOSC1+EXOSC3, No prot, EXOSC1, EXOSC6, EXOSC1+EXOSC6, No prot, EXOSC1, EXOSC7, EXOSC1+EXOSC7, No prot, EXOSC1, EXOSC8, EXOSC1+EXOSC8

50 nt
25 nt
5 nt

E

Generic sequence

EXOSC1	−	+	+	+	+	+
MgCl$_2$	+	+	−	+	+	+
KCl	+	+	+	−	+	+
DTT	+	+	+	+	−	+
Tris-HCl	+	+	+	+	+	−

50 nt
25 nt
5 nt

F

Generic sequence

No prot, pH 6, pH 7, pH 7.4, pH 8

50 nt
25 nt
5 nt

G

Generic sequence

EXOSC1 hr ▶ 0 1 2 4 6 8 10

50 nt
25 nt
5 nt

H

DNA fragments generated (%)

60
40
20
0

0   4   8   12
Time (hr)

Figure 4

Figure 5

Figure 6

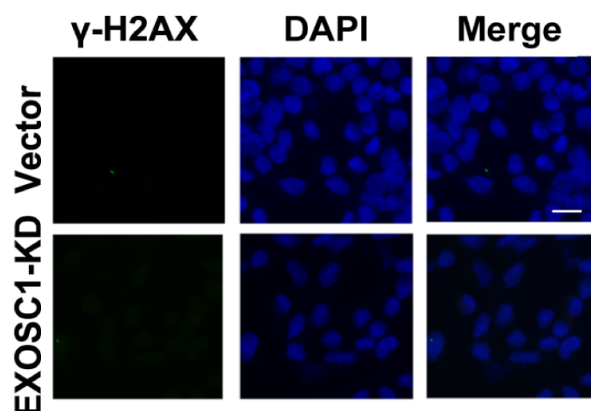# Figure 1—figure supplement 1

A

TP53 in BRCA
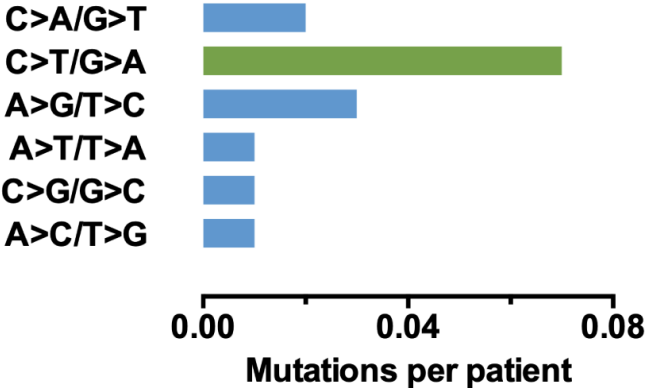
B

VHL in KIRC
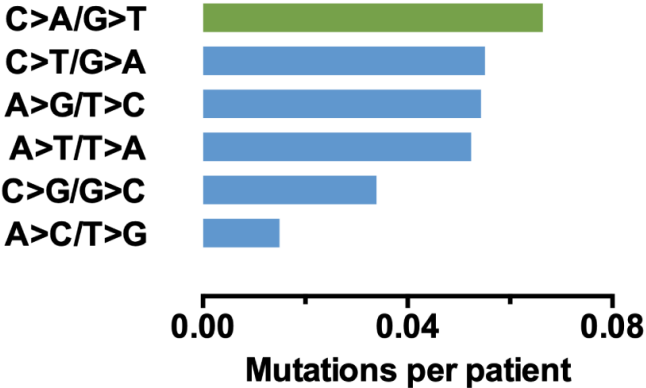
Figure 2—figure supplement 1

A
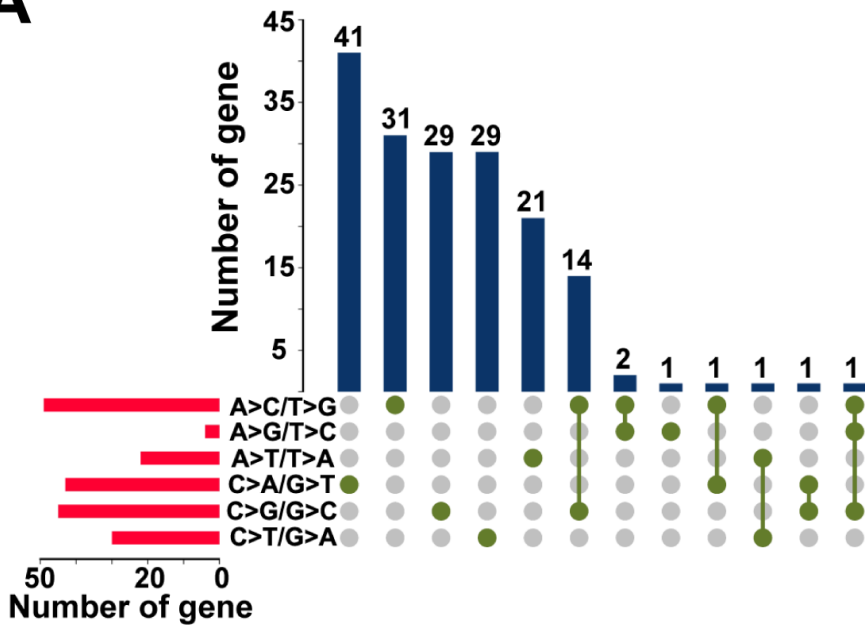
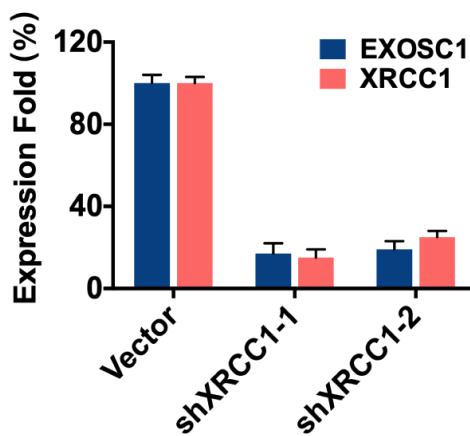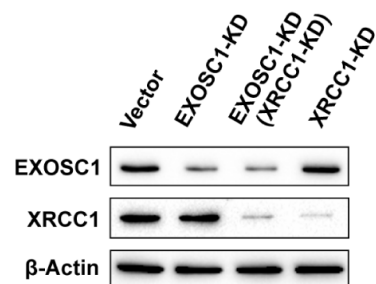
B
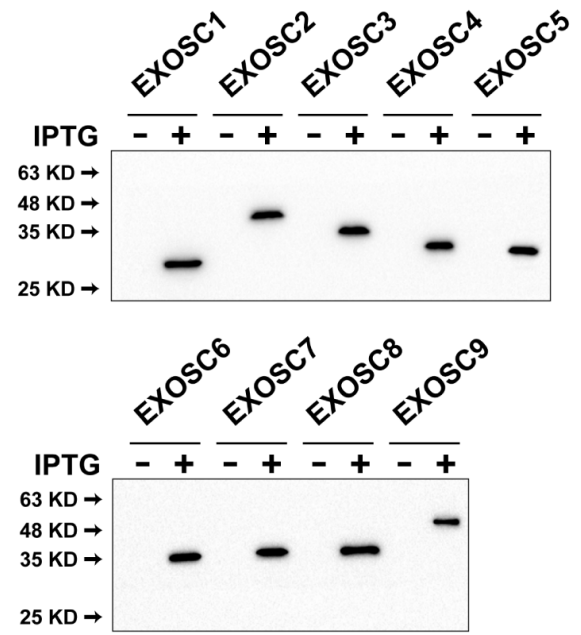
C
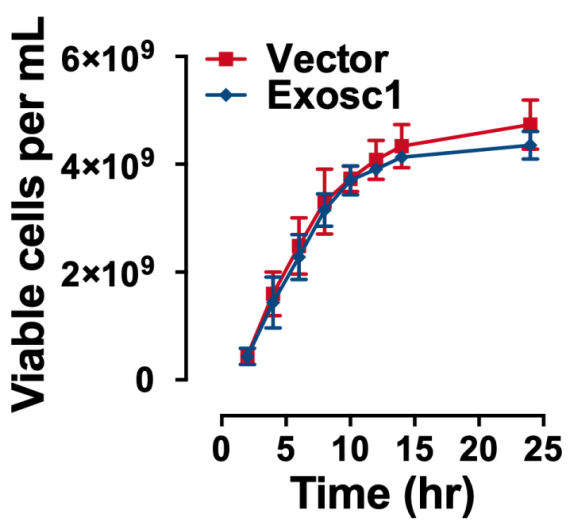
D
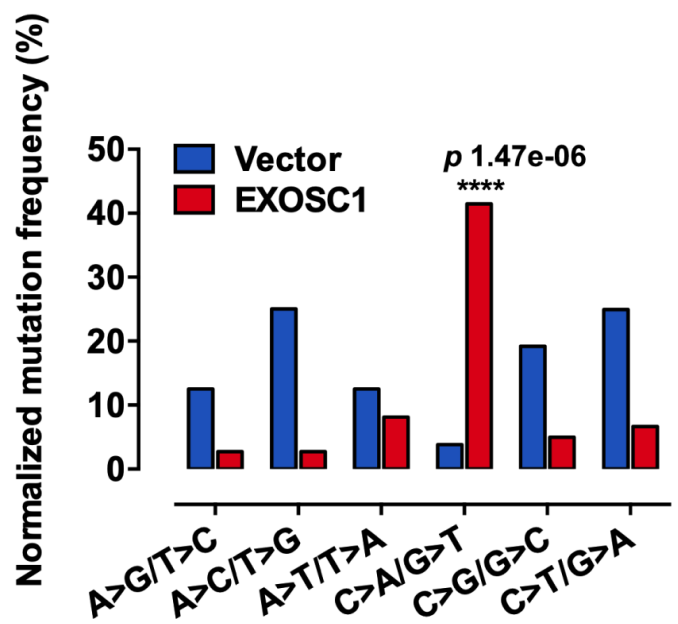
Figure 3—figure supplement 1

A

B

Figure 4—figure supplement 1
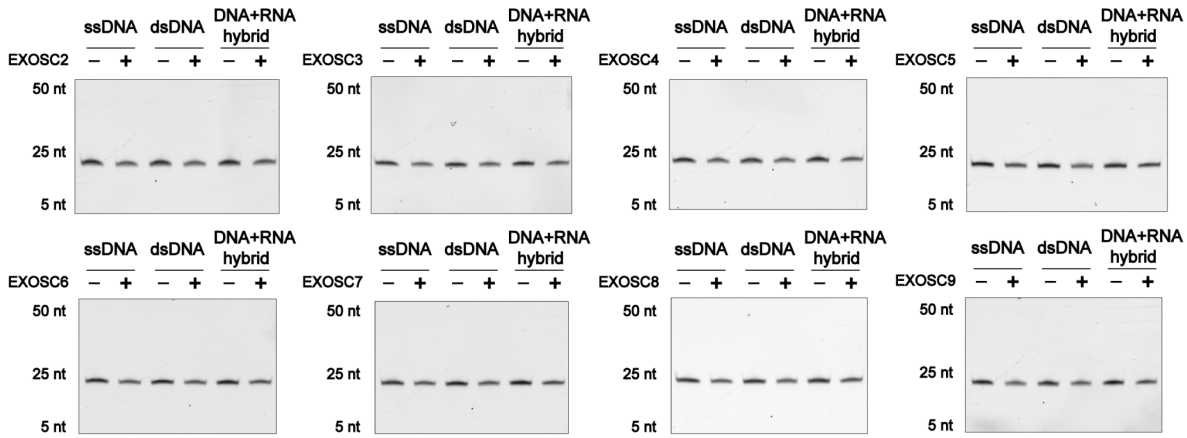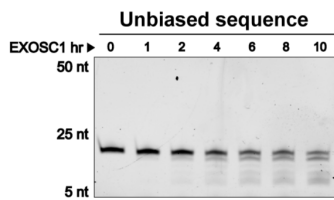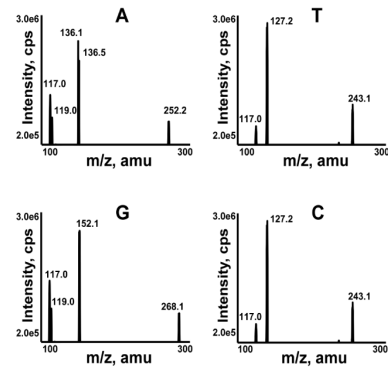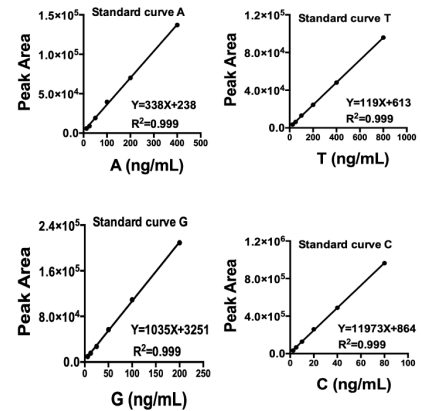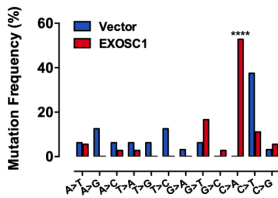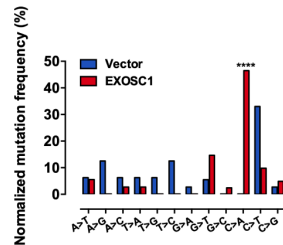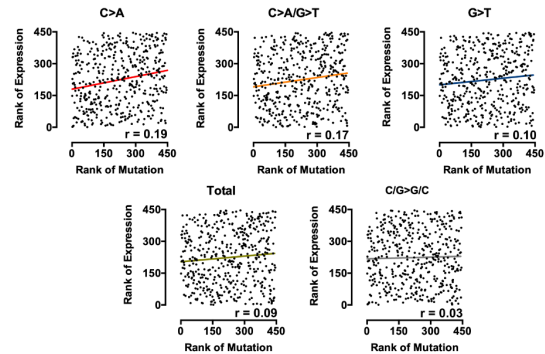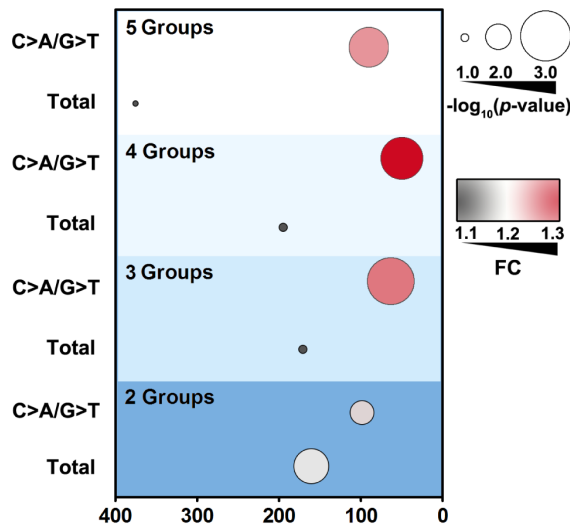
Figure 5—figure supplement 1

Figure 6—figure supplement 1

A

B

Best separation

p-value 9.26e-12

Low (n=306)
High (n=226)
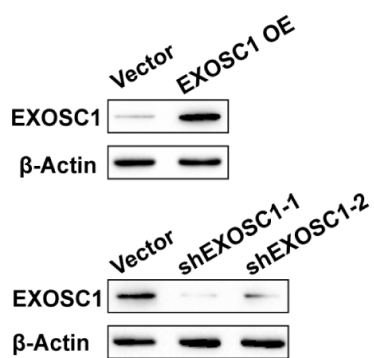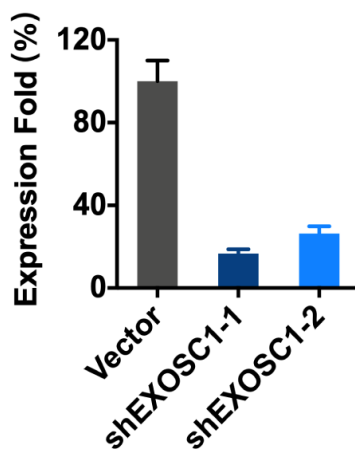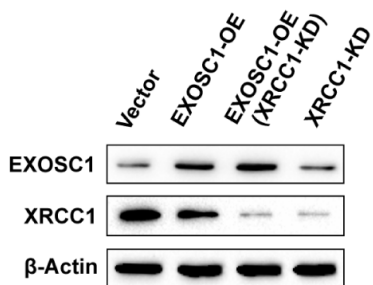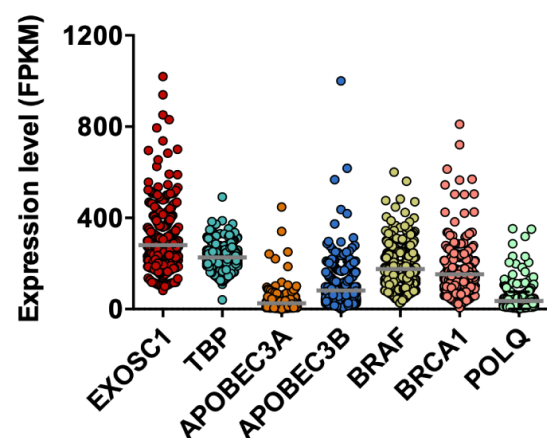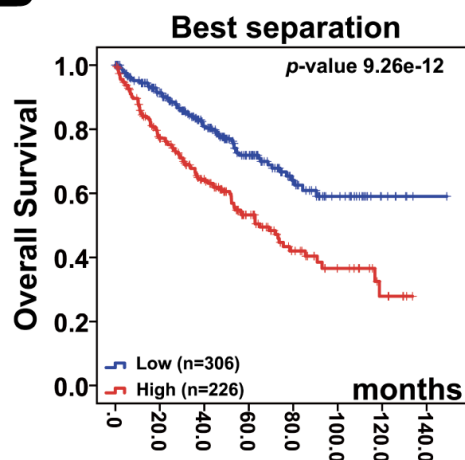
months

C

	EXOSC1- VHL wt	EXOSC1- VHL mut	EXOSC1+ VHL wt	EXOSC1+ VHL mut
Median Survival (Month)	104.47	98.46	75.66	65.19
Logrank Test P-Value	4.02e-6			

D