nature cancer

ARTICLES

https://doi.org/10.1038/s43018-022-00360-7

Check for updates

Phosphate dysregulation via the XPR1–KIDINS220 protein complex is a therapeutic vulnerability in ovarian cancer

Daniel P. Bondeson[1,6], Brenton R. Paolella[1,5,6], Adhana Asfaw[1], Michael V. Rothberg[1], Thomas A. Skipper[1], Carly Langan[1], Gabriel Mesa[1], Alfredo Gonzalez[1], Lauren E. Surface[2,3], Kentaro Ito[1], Mariya Kazachkova[1], William N. Colgan[1], Allison Warren[1], Joshua M. Dempster[1], John M. Krill-Burger[1], Maria Ericsson[3], Andrew A. Tang[1], Iris Fung[1], Emily S. Chambers[1], Mai Abdusamad[1], Nancy Dumont[1], John G. Doench[1], Federica Piccioni[1,5], David E. Root[1], Jesse Boehm[1], William C. Hahn[1,3,4], Michael Mannstadt[2,3], James M. McFarland[1], Francisca Vazquez[1,7] and Todd R. Golub[1,3,4,7] ✉

Despite advances in precision medicine, the clinical prospects for patients with ovarian and uterine cancers have not substantially improved. Here, we analyzed genome-scale CRISPR–Cas9 loss-of-function screens across 851 human cancer cell lines and found that frequent overexpression of *SLC34A2*—encoding a phosphate importer—is correlated with sensitivity to loss of the phosphate exporter XPR1, both in vitro and in vivo. In patient-derived tumor samples, we observed frequent PAX8-dependent overexpression of *SLC34A2*, *XPR1* copy number amplifications and *XPR1* messenger RNA overexpression. Mechanistically, in *SLC34A2*-high cancer cell lines, genetic or pharmacologic inhibition of XPR1-dependent phosphate efflux leads to the toxic accumulation of intracellular phosphate. Finally, we show that XPR1 requires the novel partner protein KIDINS220 for proper cellular localization and activity, and that disruption of this protein complex results in acidic "vacuolar" structures preceding cell death. These data point to the XPR1–KIDINS220 complex and phosphate dysregulation as a therapeutic vulnerability in ovarian cancer.

An emerging paradigm in cancer medicine is the tailoring of a therapeutic strategy to the specific molecular profile of a patient's tumor. Despite remarkable advances in 'precision medicine' in other cancer types, outcomes for ovarian and uterine cancers have not improved substantially in the past 20 years[1–3]. Accordingly, ovarian and uterine cancers remain among the most deadly, globally killing over 300,000 women in 2020 alone[4]. New therapeutic strategies are needed.

We and others have demonstrated that novel cancer vulnerabilities can be discovered from genome-scale, loss-of-function cell viability screens[5]. From these data, both biological insights[6,7] and therapeutic hypotheses[8,9] can be developed. Here, we systematically analyze CRISPR–Cas9 loss-of-function screens across 851 human cancer cell lines to identify novel candidate therapeutic targets in ovarian cancer. We identified an unexpected synthetic lethal relationship between increased expression of the phosphate importer SLC34A2 and loss of the phosphate exporter XPR1. We provide compelling evidence for the therapeutic development of XPR1 inhibitors through extensive in vitro and in vivo validation, analysis of patient samples and proof-of-concept pharmacologic inhibition. The relationship between increased phosphate import and reliance on phosphate export suggests that intracellular phosphate accumulation is toxic to cancer cells, and we leverage coessentiality data to elucidate XPR1 as a member of a phosphate efflux protein complex also containing KIDINS220. Together, these data highlight the power of functional genomics screens to identify compelling therapeutic targets and elucidate their biological function.

Results

Loss of XPR1 is toxic to gynecological cancers that express SLC34A2. To identify novel therapeutic targets for ovarian and uterine cancers, we analyzed genome-scale, pooled CRISPR–Cas9 loss-of-viability screens in 851 genomically characterized human cancer cell lines as part of the Cancer Dependency Map[10–12]. We focused on genes that, when inactivated, selectively lead to loss of viability in ovarian or uterine cancer cell lines, since a broad killing pattern is more likely to represent mechanisms that would be poorly tolerated if pharmacologically inhibited. This analysis (Fig. 1a) yielded 'selective dependencies' such as the transcription factor *PAX8* (ref. [8]), a known lineage-restricted transcription factor. This analysis also revealed that inactivation of the phosphate exporter *XPR1* has a cell-killing pattern that is highly selective and enriched in ovarian and uterine cancers (Extended Data Fig. 1a). XPR1 is a transmembrane protein[13] and the only phosphate exporter annotated in the human genome[14].

[1]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [2]Endocrine Unit, Massachusetts General Hospital, Boston, MA, USA. [3]Harvard Medical School, Boston, MA, USA. [4]Departments of Pediatric and Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. [5]Present address: Merck Research Laboratories, Cambridge, MA, USA. [6]These authors contributed equally: Daniel P Bondeson, Brenton R Paolella. [7]These authors jointly supervised this work: Francisca Vazquez, Todd R Golub. ✉e-mail: golub@broadinstitute.org

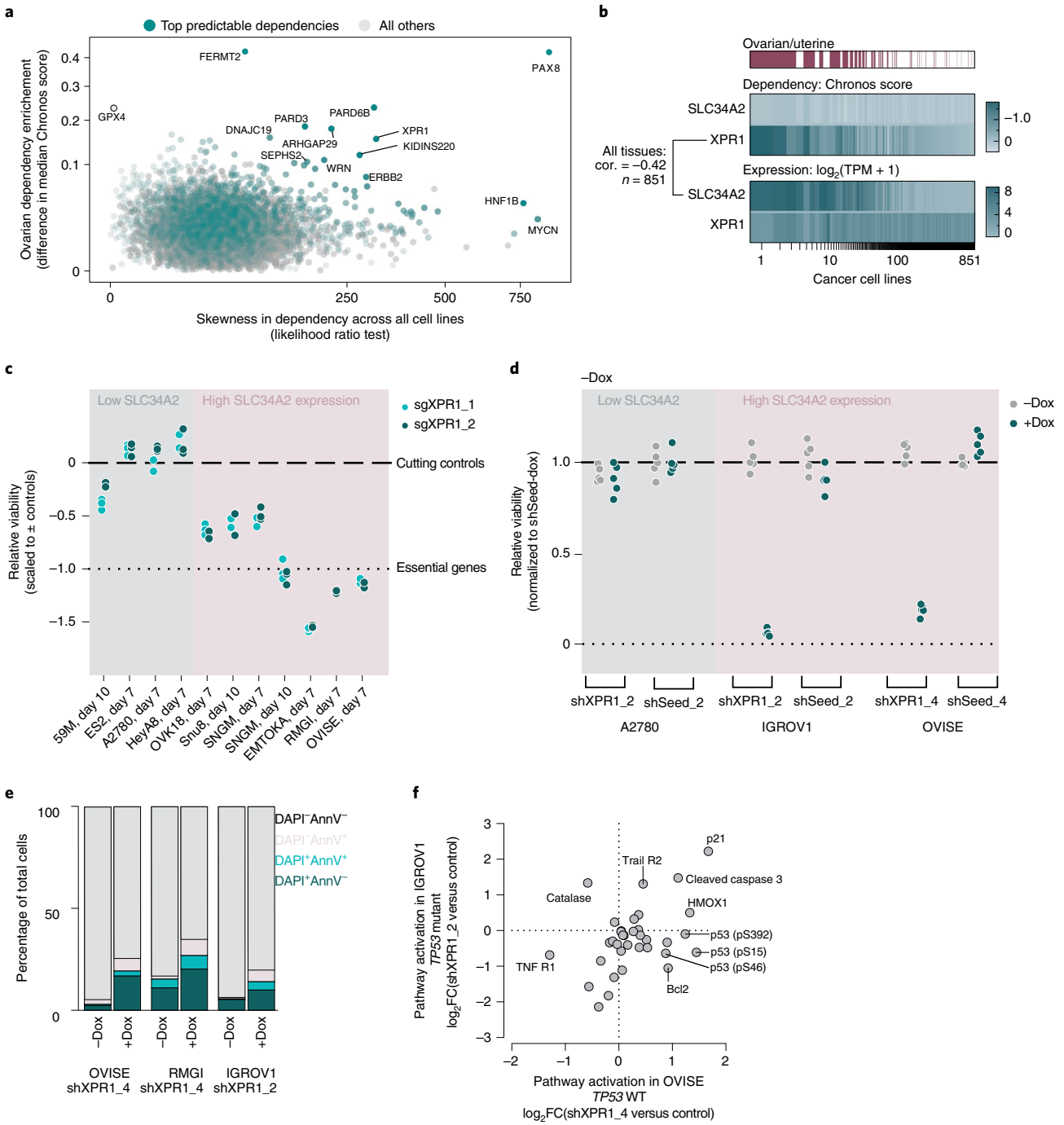


Fig. 1 | Functional genomics identifies *XPR1* loss as a cancer vulnerability in SLC34A2-high ovarian and uterine cancers. a, For the >18,000 genes tested in CRISPR–Cas9 loss-of-viability screens, the selectivity of the killing profile across all 851 cell lines (*x* axis, likelihood ratio test; Methods) and the enrichment of that gene's dependency (Chronos score) in ovarian and uterine cancers (*y* axis) are plotted. The top 5% most predictable dependencies are highlighted in teal, where a random forest model using the genomic and molecular features of cancer cell lines can predict the strength of dependency. **b**, Heatmap indicating *XPR1* and *SLC34A2* expression ($\log_2$(TPM + 1)) and dependency (CERES) values across all cell lines, ranked approximately by decreasing dependency on *XPR1*. The Pearson correlation (cor.) coefficient across all 851 cell lines is indicated. **c**, Across a panel of ovarian and uterine cancer cell lines, viability effects after inactivation of *XPR1* were evaluated by comparison with negative control sgRNA and sgRNA targeting pan-essential genes ($n = 3$ independent transductions representative of at least $n = 2$ independent experiments). Note that A2780 is not considered to be ovarian cancer despite its historical annotation. Data are scaled such that a value of 0 represents the viability effect of CRISPR–Cas9 genome editing and -1 represents loss of an essential gene. High *SLC34A2* expression indicates mRNA expression >3 TPM. **d**, Viability assessment 7 days after suppression of *XPR1* using the indicated shRNA or seed-matched controls (shSeed; $n = 5$ technical replicates representative of at least $n = 3$ independent experiments per cell line). **e**, Six days after induction of shXPR1 in the indicated cell lines, cells were stained with DAPI to distinguish live and dead cells, and with Annexin V (AnnV) to distinguish non- and pre-apoptotic cells ($n = 1$ flow cytometric analyses of at least 10,000 cells, representative of $n = 2$ independent experiments). Right, quantitation of percentage of cells in the indicated quadrants. **f**, Analysis of cell death pathways in OVISE and IGROV1 5 days after suppression of *XPR1* by shRNA, using protein arrays ($n = 1$). Note that OVISE has wild-type (WT) *TP53* while IGROV1 has an inactivating mutation in *TP53*.

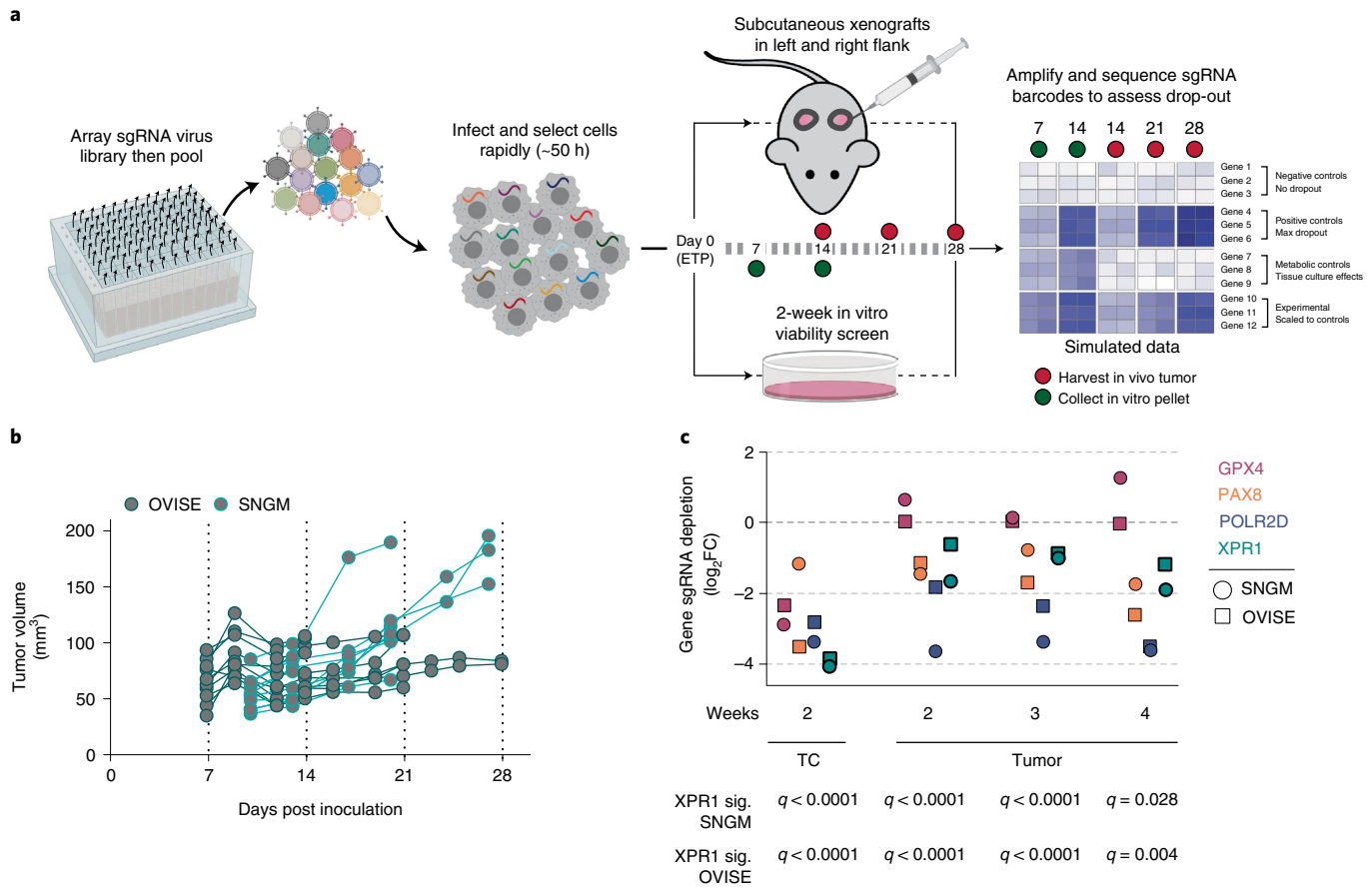


Fig. 2 | *XPR1* inactivation prevents tumor formation in vivo. a, Experimental design for in vivo competition assays. Using a rapid infection-and-selection protocol, pooled sgRNA can be introduced via lentivirus into cancer cell lines and inoculated as subcutaneous xenografts, and the effect of gene inactivation can then be evaluated in an environment more physiologically relevant than tissue culture (TC). **b**, Following rapid infection with pooled sgRNA, 8 million SNGM or OVISE cells were grown in TC and in parallel were inoculated as subcutaneous xenografts and allowed to grow. Tumor tissue was harvested at the indicated time points. **c**, Evaluation of sgRNA targeting *XPR1* and other cancer vulnerabilities in a tumor formation competition assay, as described in **a** for OVISE (squares) and SNGM (circles) cancer cell lines ($n = 2$–3 independent tumors derived from the same transduction per cell line per time point). GPX4, glutathione peroxidase 4, a metabolic dependency reliant on the amount of peroxidated lipids in the metabolic environment of cancer cells; PAX8, paired box 8, a transcription factor dependency in many ovarian cancer cell lines; POLR2D, RNA polymerase II subunit D, a pan-essential gene used as a positive control. Bottom, the significance (sig.) of depletion of three sgRNAs targeting *XPR1* relative to seven control sgRNAs was calculated via *t*-test and corrected for multiple comparisons with the Holm–Sidak method.

We next pursued the molecular basis of the selective dependency on *XPR1*. Using >100,000 molecular features of cancer cell lines[15], we built multivariate models—potential 'biomarkers' of response—to predict *XPR1* dependency[16,17]. Remarkably, the feature that most robustly predicted XPR1 dependency was expression of the phosphate importer *SLC34A2* (Fig. 1b; Pearson coefficient = –0.42 in all cell lines). *SLC34A2* overexpression in ovarian cancer is well documented[18,19] and was highly correlated with *XPR1* dependency in cell lines from the ovarian clear cell, high-grade serous and endometrial adenocarcinoma lineages (Extended Data Fig. 1b,c).

We initially validated the pooled screening results in a total of seven *SLC34A2*-high- and four -low-expressing cancer cell lines, confirming that *XPR1* is a selective and strong dependency in the context of *SLC34A2* overexpression (Fig. 1c,d and Extended Data Fig. 2). Loss of *XPR1* profoundly slows cell growth and leads to an increase in growth-inhibitory and proapoptotic markers (Fig. 1e,f and Extended Data Fig. 2e,f). We next assessed *XPR1* dependency in a CRISPR–Cas9-based tumor formation competition assay with 74 single-guide RNAs (sgRNA), and observed XPR1 sgRNA depletion in *SLC34A2*-high tumors (Fig. 2 and Extended Data Fig. 3). In contrast, sgRNAs targeting other metabolic dependencies, such as the ferroptosis regulator *GPX4*, were depleted in vitro but not in vivo, as previously reported[20,21]. These results indicate that *XPR1* dependency is retained in vivo.

Evidence of phosphate dysregulation in primary patient samples. To extend the relevance of *XPR1* and *SLC34A2* beyond cell lines, we evaluated the relationship between *XPR1* and *SLC34A2* in primary patient samples from The Cancer Genome Atlas (TCGA)[22,23] and normal samples from the Genotype-Tissue Expression project (GTEx)[24–26]. Ovarian tumors, on average, expressed *SLC34A2* at levels 16-fold higher than normal fallopian tube epithelium ($q = 0.006$), which is thought to be the cell of origin of ovarian cancers[27–30] (Fig. 3a). Uterine cancers similarly overexpressed *SLC34A2* relative to normal tissue. Interestingly, ovarian and uterine cancers were among the few tissues with high levels of *SLC34A2* expression (Extended Data Fig. 4a). We hypothesized that *SLC34A2* expression may be maintained at high levels in ovarian cancer because of its regulation by the transcription factor *PAX8*, the expression of which is elevated in ovarian cancer and which is required for ovarian cancer cell survival[8,31–33]. In support of this, we found a strong correlation between *PAX8* and *SLC34A2* expression in patient samples

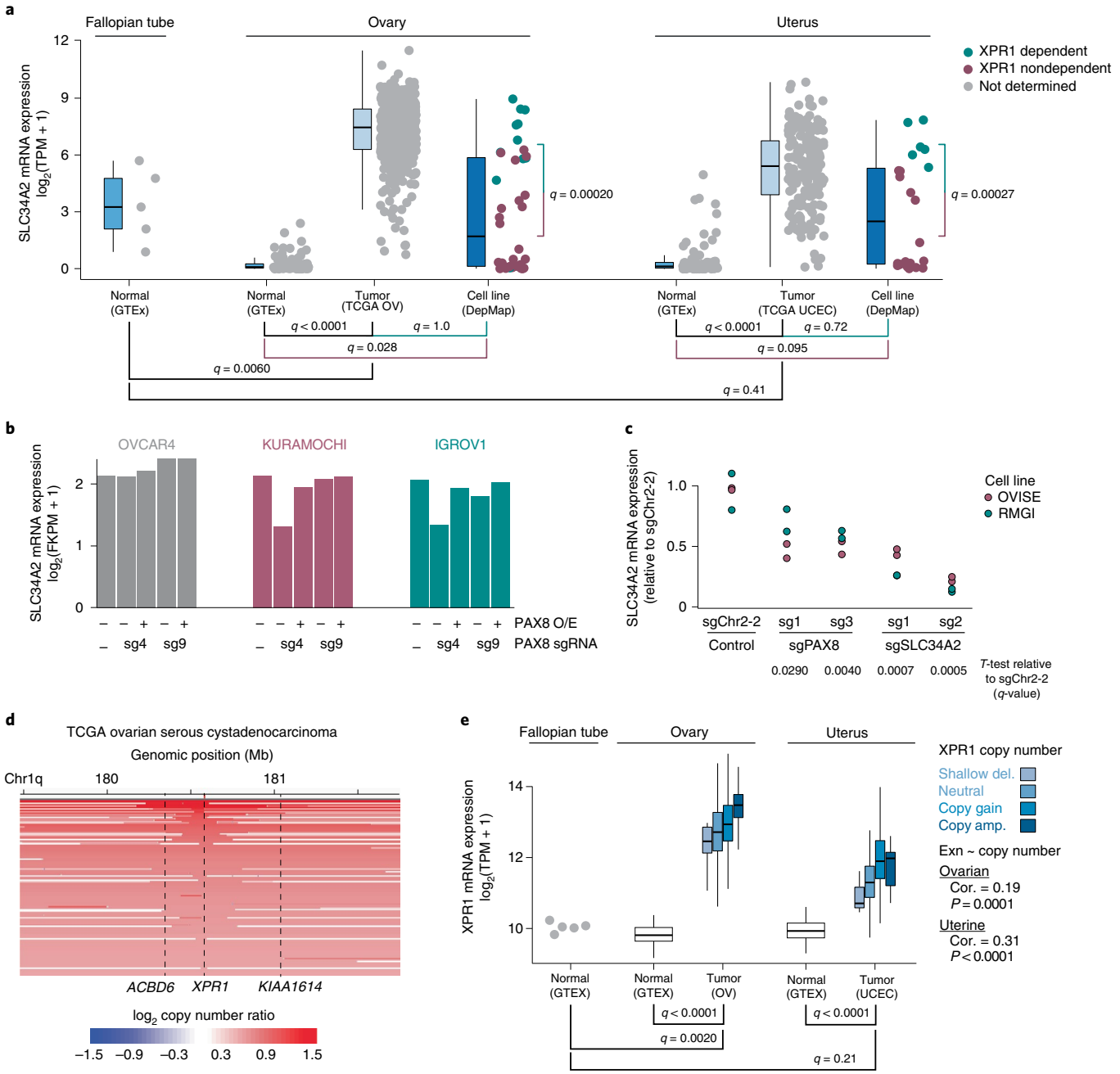


Fig. 3 | Expression of *XPR1* and *SLC34A2* in patient samples indicates cancer-specific phosphate dysregulation caused by lineage survival transcription factor *PAX8*. **a**, *SLC34A2* is expressed in ovarian and uterine tumor samples at levels sufficient to predict dependency on *XPR1*. RNA expression values for *SLC34A2* were compared within the indicated lineages. *XPR1* dependency status is indicated by color where applicable (CERES <−0.5). *Q*-values indicate the likelihood of the indicated populations having the same level of *SLC34A2* expression according to two-sided Wilcoxon ranked-sum test with Bonferroni correction for multiple comparisons. Boxplots are drawn indicating the first and third quartiles, and whiskers span to the largest value within 1.5× interquartile range. **b**, Expression of *SLC34A2* was measured using RNA-seq ($n = 1$) after stable overexpression of *PAX8* as indicated (PAX8 O/E), and induction of a *PAX8*-targeting (sg4) or control (sg9) sgRNA and dCas9-KRAB. **c**, Seven days after transduction with the indicated sgRNA ($n = 2$, separate transductions), RNA was extracted, converted to cDNA and the expression of *SLC34A2* measured using RT–PCR. Significance was assessed by comparison of expression relative to sgChr2-2 across two cell lines with a one-tailed *t*-test, and corrected for multiple comparisons using the Bonferroni method. Data are representative of $n = 2$ independent experiments. **d**, *XPR1* copy number heatmap for a ~2.5-Mb region of chromosome 1 indicating *XPR1* amplification in TCGA serous ovarian cancer[22]. Each patient sample is represented by a horizontal line. Red and blue indicate copy gain and loss, respectively. Dashed vertical lines represent the location of indicated genes. Data are a subset of 489 samples, rank ordered by highest copy gain to indicate both focal and chromosome arm variants. **e**, *XPR1* mRNA expression is increased in ovarian and uterine cancer. *XPR1* mRNA expression values from the same sources as in **a** are compared for the indicated tissues, with TCGA OV and TCGA UCEC color coded by *XPR1* copy number status as determined by GISTIC analysis. Boxplots are drawn as in **a**. Statistical differences between tissues were determined as in **a**. Correlation of *XPR1* copy number and expression was performed using Spearman's ranked correlation test. See Methods for exact *n* values. Del, deletion; amp, amplification; exn, expression; FKPM, fragments per kilobase million.

(Extended Data Fig. 4b,c). Furthermore, CRISPR-mediated inactivation or transcriptional repression of *PAX8* led to a significant loss of *SLC34A2* expression, as previously reported[32–34] (Fig. 3b,c and Extended Data Fig. 4d,e).

In contrast to the more restricted pattern of expression of *SLC34A2* and *PAX8*, *XPR1* is widely expressed in both normal and cancer tissues (Extended Data Fig. 4f). Nevertheless, we found strong evidence for positive selection of *XPR1* copy number amplifications and enhanced mRNA expression in ovarian and uterine cancer, consistent with its dependency in these tissues (Fig. 3d,e). In ovarian cancer these amplifications were often focal, involving only the *XPR1* gene (Fig. 3d; $q = 0.0015$ (ref. [35])) whereas in uterine cancer, broader and less significant amplifications were observed (Extended Data Fig. 4g; $q = 0.568$). *XPR1* mRNA expression levels were correlated with *XPR1* copy number alterations, but other mechanisms probably also contribute to *XPR1* mRNA expression (Fig. 3e). Thus, the dysregulated expression of the essential transcription factor PAX8 results in enhanced expression of *SLC34A2*, thereby creating dependency on *XPR1*.

XPR1 loss causes toxic phosphate accumulation. To further understand the mechanism by which *XPR1* loss of function results in cancer cell death, we performed a genome-wide CRISPR–Cas9 rescue screen to determine which genes, when inactivated, were capable of rescuing XPR1-mediated loss of viability (Fig. 4a and Extended Data Fig. 5)[36]. Remarkably, the top-rescuing gene among >18,000 tested was also the top predictive biomarker: *SLC34A2*. We further demonstrated that *SLC34A2* is both necessary and sufficient to confer *XPR1* dependency in three ovarian and uterine cell lines (Fig. 4b).

The observation that high expression of the phosphate importer *SLC34A2* is required for loss of cell viability after inactivation of the phosphate exporter *XPR1* led us to hypothesize that accumulation of intracellular phosphate is toxic to ovarian and uterine cancer cells (Fig. 4c). XPR1 is the only known phosphate exporter in humans[14], suggesting that, in the context of increased phosphate import, XPR1-dependent phosphate efflux would be in higher demand, consistent with frequent copy number amplifications observed in patient samples (Fig. 3e).

Although the extracellular availability of phosphate in typical tissue culture medium far exceeds what is physiologically relevant, we found no correlation between the phosphate content of growth medium and *XPR1* dependency across the Cancer Dependency Map dataset (Extended Data Fig. 6a). Furthermore, *XPR1* dependency was retained when cells were adapted to growth medium with near-physiological phosphate concentrations (reduced by ~90% from 72.8 to 7.8 mg dl^{-1}; Extended Data Fig. 6b,c), indicating that *XPR1* dependency is not an artifact of high concentrations of extracellular phosphate.

Consistent with the phosphate accumulation hypothesis, we observed two-to-four-fold increased intracellular phosphate following XPR1 suppression (Fig. 4d). These large fluctuations in intracellular phosphate co-occur with loss of cell viability (Extended Data Fig. 6d). To understand the cellular response to phosphate accumulation, we used single-cell RNA sequencing (RNA-seq) of 2,501 cells across eight ovarian and uterine cancer cell lines at an early time point following *XPR1* inactivation (Fig. 4e and Extended Data Fig. 7a–g). The resulting transcriptional signature reflected cellular attempts to restore phosphate homeostasis, including the upregulation of FGF23. This critical phosphate homeostatic hormone is typically expressed in osteogenic bone cells, and its expression in ovarian cancer cells—although not represented at the protein level (Extended Data Fig. 7h)—is consistent with sensing of elevated phosphate[37]. We also observed downregulation of phosphate importers at both the mRNA (Fig. 4e) and protein level (Extended Data Figs. 2c and 7i) following *XPR1* inactivation or suppression. Consistent with this, *XPR1* inactivation led to a 60% decrease in phosphate uptake (Fig. 4f). This compensatory mechanism is also observed in XPR1-nondependent cancer cell lines with known regulatory mechanisms for SLC34A2 (for example, in lung[38,39]; Extended Data Figs. 1a,b and 7j). These data suggest a potential protective mechanism downregulating phosphate uptake in response to toxic levels of inorganic phosphate[40–42].

Phosphate efflux activity of XPR1 is required for cancer cell survival. We next confirmed that the phosphate efflux function of XPR1 is critical for cell survival. Expression of a naturally occurring hypomorphic *XPR1* mutation (L218S), associated with a rare brain calcification disorder[42,43], failed to rescue endogenous *XPR1* inactivation whereas WT *XPR1* fully restored cell viability (Fig. 5a and Extended Data Fig. 8). In addition, we pharmacologically inhibited XPR1 using a previously reported protein inhibitor (XRBD; Fig. 5b,c) and found that its cancer cell line growth-inhibitory effects paralleled inactivation of *XPR1* (Fig. 5d). Together, these results indicate that inhibition of the phosphate efflux capacity of XPR1 in *SLC34A2*-high cells is sufficient for loss of cancer cell viability.

The activity of XPR1 is entirely dependent on a novel partner protein, KIDINS220. To gain further insight into the mechanism by which XPR1 regulates phosphate homeostasis, we analyzed the Cancer Dependency Map for genes with dependency profiles highly correlated to *XPR1*. These 'codependencies' often indicate proteins that are part of the same protein complex[7,44,45]. Of the ~18,000 genes analyzed, *XPR1* dependency is most strongly correlated with that of *KIDINS220*, a gene with no known connection to phosphate homeostasis[46–48] (Pearson correlation = 0.81; Fig. 6a and Extended Data Fig. 9a). Given the strength of this correlation, we extensively validated KIDINS220 dependency (Extended Data Fig. 9b,c) and hypothesized that KIDINS220 might be part of an XPR1 phosphate export complex.

In support of an XPR1–KIDINS220 protein complex, protein interaction databases indicate that XPR1 and KIDINS220 interact with each other (Extended Data Fig. 9d). Further, their gene expression is highly correlated across diverse tissues (Extended Data Fig. 9e), suggesting cofunction and coregulation. To confirm this interaction, we performed coimmunoprecipitation experiments and found that XPR1 and KIDINS220 indeed interact with each other and with several other partner proteins (Fig. 6b,c and Extended Data Fig. 9f,g). Native XPR1–KIDINS220 is consistent with a highly oligomerized protein complex (Fig. 6d). We mapped the XPR1–KIDINS220 interaction to the C terminus of XPR1 containing the EXS domain (Fig. 6b and Extended Data Fig. 9g), an evolutionarily conserved domain known to be required for XPR1 trafficking between the Golgi apparatus and plasma membrane to achieve phosphate efflux[49–52]. In contrast, the N-terminal SPX domain of XPR1, which has been implicated in phosphate efflux and regulation[41,42], was neither necessary nor sufficient to bind KIDINS220 (Fig. 6b).

Further supporting an XPR1–KIDINS220 protein complex, we found dramatically decreased KIDINS220 protein levels following *XPR1* genetic inactivation or suppression, but not after inhibition by XRBD (Extended Data Figs. 2c, 7i and 9h). In addition, *KIDINS220* inactivation decreased XPR1 cell surface localization and dramatically changed the sedimentation pattern of XPR1, indicating that native localization and conformation of XPR1 requires KIDINS220 (Fig. 6d–f and Extended Data Fig. 9h). Finally, we measured phosphate efflux directly and found that inactivation of either *XPR1* or *KIDINS220* impaired it to a similar degree (Fig. 6g), and resulted in increased intracellular phosphate (Extended Data Fig. 9i). These results, taken together, indicate that phosphate efflux is achieved by the XPR1–KIDINS220 protein complex and that loss of either complex member leads to a disruption in the phosphate efflux required for cancer cell survival.

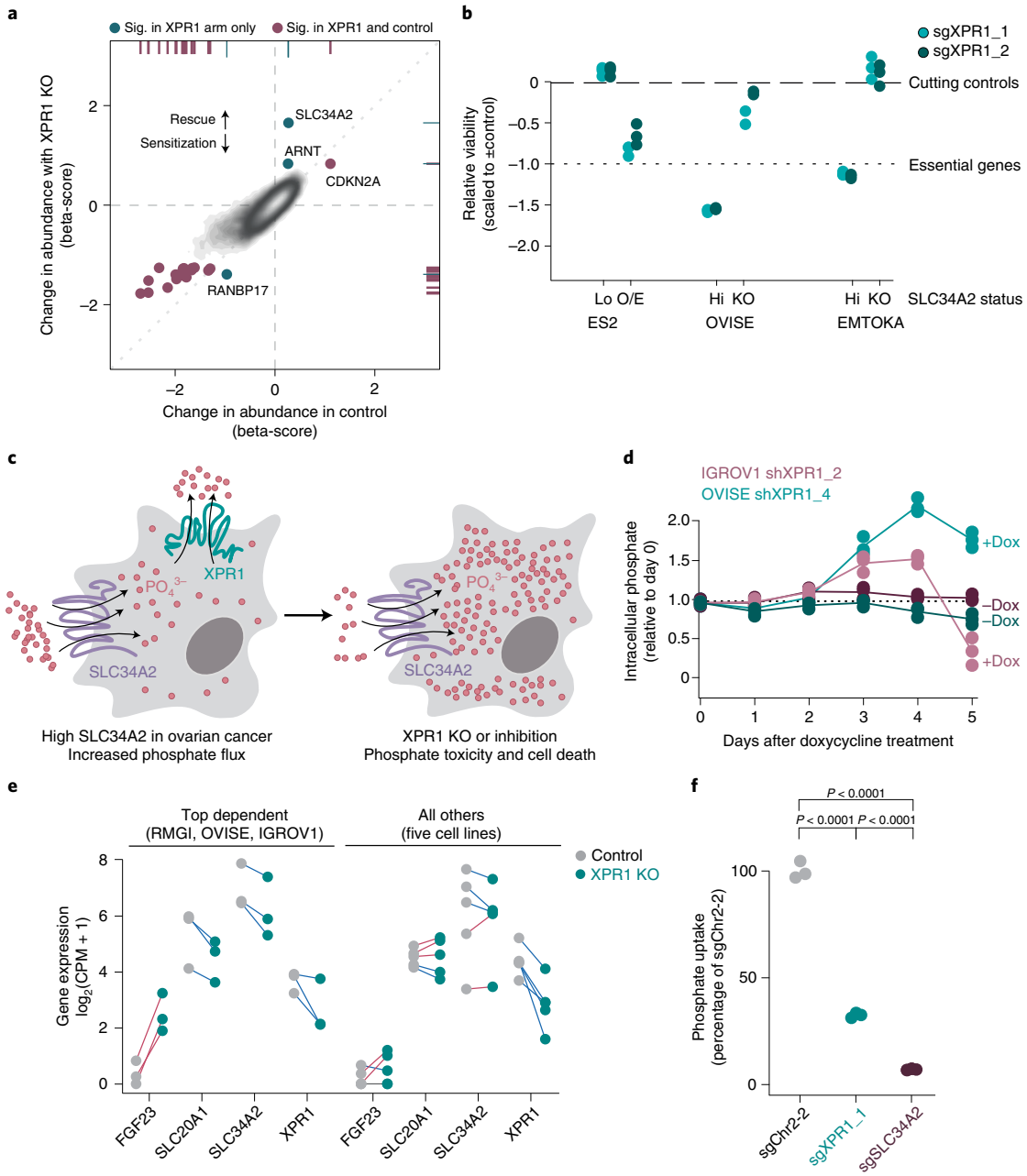


Fig. 4 | *XPR1* inactivation in *SLC34A2*-high ovarian cancer causes loss of cell viability via dysregulation of intracellular phosphate homeostasis. **a**, Identification of rescue genes that protect OVISE ovarian cancer cells from XPR1 dependency. Beta-scores (determined by MAGeCK MLE) represent the change in representation for each gene from the initial library to the final time point for either the control condition (*x* axis) or in combination with XPR1 inactivation (*y* axis). See Methods for full experimental and analytical details. $n=1$ transduction per cell line, expanded and cultured as $n=2$ independent cultures. **b**, SLC34A2 status of normally XPR1-resistant (ES2, SLC34A2-low (lo)) or XPR1-sensitive (EMTOKA and OVISE, SLC34A2-high (hi)) cell lines was modified by overexpression (O/E) or inactivation (KO) of *SLC34A2*, and *XPR1* dependency was evaluated as in Fig. 1c ($n=3$ separate transductions, representative of at least $n=2$ independent experiments per cell line). **c**, Because of their relative directionalities of phosphate transport, we hypothesize that XPR1 perturbation is toxic because of intracellular phosphate accumulation in SLC34A2-high ovarian and uterine cancers. **d**, At various time points after treatment with doxycycline and induction of shRNA, intracellular phosphate was measured in OVISE and IGROV1 cell lines ($n=3$ separate measurements per condition, representative of at least $n=3$ separate experiments). **e**, A pool of eight cancer cell lines was transduced with lentivirus to inactivate *XPR1*, and 4 days later cells were subjected to 10× single-cell transcriptomic measurement ($n=1$ transduction). The measured transcriptional change (relative to control sgRNA infection) in the indicated genes is plotted on the left for the three cell lines with the largest and most correlated transcriptional change (Extended Data Fig. 7), and for the other five on the right. Blue and red lines connect cell lines displaying decreased or reduced expression, respectively, upon *XPR1* inactivation. **f**, XPR1 perturbation causes compensatory inhibition of phosphate uptake, measured by incubation of OVISE ovarian cancer cells in medium supplemented with $^{32}PO_4^{3-}$ phosphate for 30 min before washing off excess medium and cell lysis ($n=1$ transduction measured in technical duplicate, representative of $n=2$ independent experiments). Significance was assessed by one-way analysis of variance and corrected for multiple comparisons using Bonferroni's method.

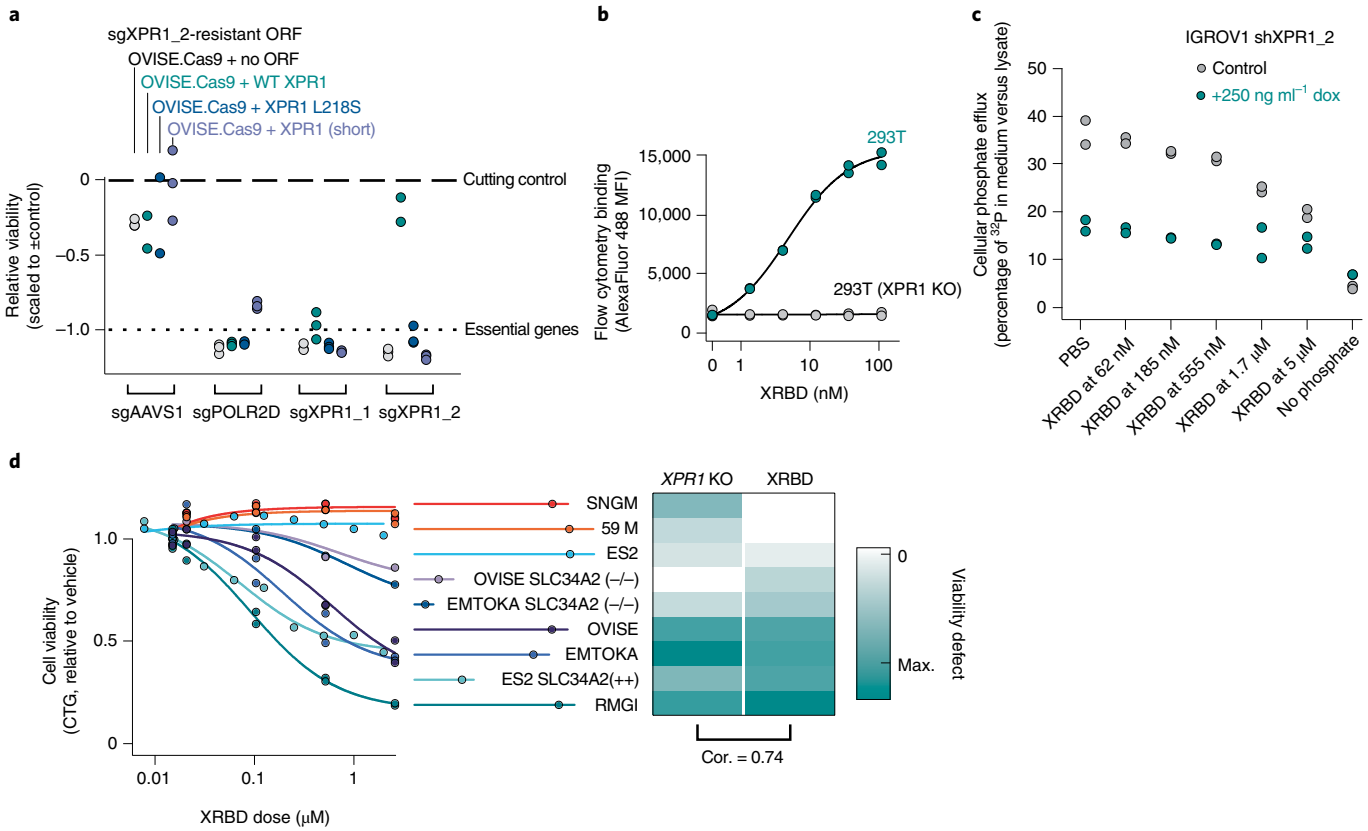


Fig. 5 | Phosphate efflux activity by XPR1 is required for SLC34A2-high cancer cell survival. a, The indicated *XPR1* ORFs were tested for their ability to rescue inactivation of endogenous *XPR1* ($n=3$ independent transductions, representative of $n=2$ independent experiments). The L218S mutation in *XPR1* has previously been shown to have only ~50% of the phosphate efflux function of WT *XPR1* (see main text). **b**, Purified XRBD (the soluble receptor-binding domain of the NZB strain of xenotropic and polytropic murine leukemia virus) was incubated with 293 T cells, or 293 T cells with *XPR1* inactivation, for 30 min at the indicated concentrations, washed and stained with AlexaFluor 488-conjugated anti-mouse secondary to detect the Fc tag on XRBD ($n=2$ flow cytometric analysis of at least 10,000 cells, representative of $n=2$ independent experiments). **c**, Phosphate efflux was measured in the presence of XRBD or *XPR1* suppression in IGROV1 mixed-lineage ovarian cancer cell lines. IGROV1 were treated with doxycline (where indicated) to induce expression of shXPR1_2, 3 days before evaluation of phosphate efflux by loading cells with ^{32}P-labeled phosphate, washing them to remove excess ^{32}P and measuring the percentage of total ^{32}P in the conditioned medium after 30 min ($n=2$ technical replicates representative of $n=3$ independent experiments). Where indicated, XRBD was added to cells during both the ^{32}P-loading and efflux portions of the experiment. Note that medium without phosphate does not stimulate phosphate efflux, and was used as a control. **d**, Treatment of various cancer cell lines with the XPR1 inhibitor XRBD (the soluble receptor-binding domain of the NZB strain of xenotropic and polytropic murine leukemia virus). Left, cells were incubated for 5 days with the indicated concentrations of XRBD and viability was assessed by Cell Titer Glo (CTG, $n=2$ independent treatments, representative of at least $n=2$ experiments per cell line.). Right, heatmap comparison of the viability defect following *XPR1* inactivation (XPR1 KO) or XRBD treatment (5 μM dose), with Pearson correlation coefficient indicated. MFI, mean fluorescence intensity.

A striking feature of XPR1- or KIDINS220-mediated loss of cell viability is the formation of large, cytoplasmic, vacuolar structures preceding loss of cell viability (Fig. 7 and Supplementary Video 1). Colocalization with the acidic dye LysoTracker and the lysosomal marker LAMP1 (Fig. 7b,c) suggested they may be related to the lysosomal system. Ultrastructural analysis by transmission electron microscopy (TEM) found these structures to be bound by a double membrane that was often fenestrated (Fig. 7d,e). Although they lack the electron-dense appearance typical of lysosomes, we did note their fusion with lysosomes (Fig. 7d,e).

Discussion

This study highlights a previously unappreciated strategy used to kill cancer cells: the disruption of phosphate homeostatic mechanisms that are normally tightly regulated. Interestingly, antibody–drug conjugates targeting the phosphate importer SLC34A2 (ref. [53]) have been explored for the past decade, but this approach exploits SLC34A2 simply as a biomarker of ovarian cancer rather than serving as a mechanism to disrupt phosphate homeostasis. Furthermore, these strategies may have alternate mechanisms of resistance and/or on-target toxicity due to the physiological roles for SLC34A2 in intestinal[54] and lung[38,39] tissues. Nevertheless, the expression and regulation of SLC34A2 within fallopian tube epithelial cells—the probable cell of origin of ovarian and uterine cancers—has not been extensively studied. We observed elevated *SLC34A2* expression in normal fallopian tube samples (Fig. 3a), and hypothesize that SLC34A2 may play a similar role in the lung regarding uptake of inorganic phosphate derived from surfactant metabolism.

In the context of cancer, it is not clear whether enhanced *SLC34A2* expression is required for ovarian and uterine carcinogenesis or survival in vivo. Our data clearly indicate that *SLC34A2* is not required for cancer cell survival in vitro (Figs. 1b and 4a,b) and that these cell lines display a profound ability to downregulate *SLC34A2* in response to XPR1 inhibition (Extended Data Figs. 2c and 7i), suggesting that strong overexpression of *SLC34A2* is not necessarily required for ovarian cancer cell survival. We hypothesize

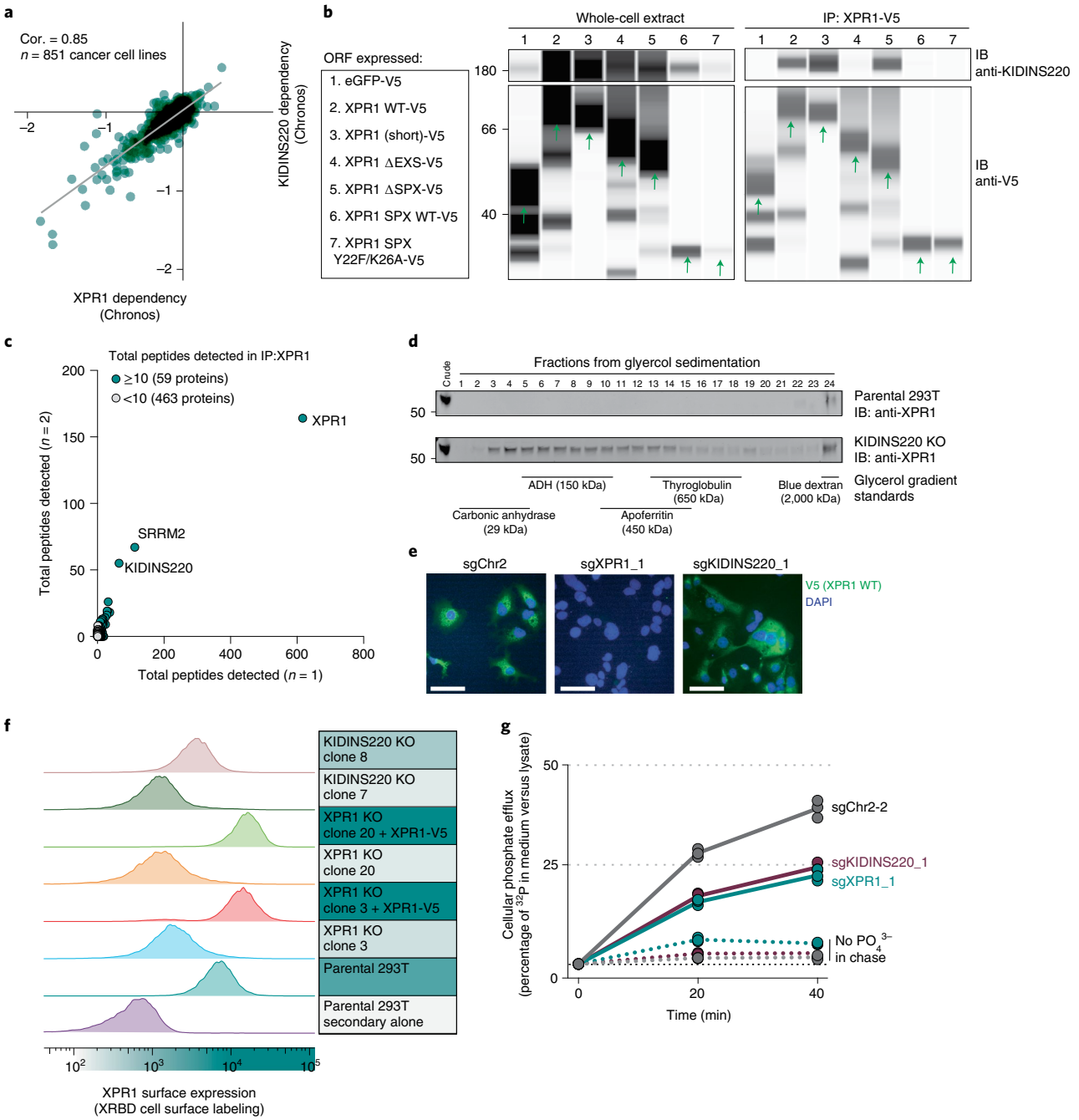


Fig. 6 | KIDINS220 is a critical component of the phosphate efflux protein complex. a, Across 851 cancer cell lines, the viability defects of *XPR1* and *KIDINS220* inactivation in each cell line were plotted and the Pearson correlation is indicated. A Chronos value of –1 is the median viability defect of inactivating pan-essential genes in a given cell line. **b**, The interaction between the V5-tagged XPR1 mutant and KIDINS220 was evaluated using coimmunoprecipitation. XPR1 WT corresponds to isoform NM_004736, while XPR1 (Short) corresponds to isoform NM_001135669. Green arrows indicate expected molecular weight; $n=1$ experiment representative of $n=3$ independent transfections. **c**, After XPR1-V5 immunoprecipitation, interacting proteins were identified using in-gel tryptic digestion followed by mass spectrometry. The *x* and *y* axes show the total number of peptides per protein detected specifically in XPR1 immunoprecipitation for $n=2$ independent transfections and immunoprecipitations. Higher-abundance proteins (more than ten peptides detected in IP:XPR1) are highlighted in teal. **d**, Glycerol gradient sedimentation analysis of XPR1-containing native protein complexes with or without *KIDINS220* inactivation. The crude lysate of the indicated cell lines was layered onto 10–30% glycerol gradients and centrifuged to fractionate protein complexes by molecular weight, followed by immunoblot analysis ($n=1$ centrifugation representative of $n=3$ independent experiments). The elution profile of protein standards is indicated below the immunoblot. **e**, Localization of XPR1-V5 proteins after inactivation of *KIDINS220*. Scale bars, 100 µm ($n=1$ technical replicate of $n=2$ independent experiments). **f**, Evaluation of XPR1 cell surface localization after KIDINS220 inactivation. $n=1$ flow cytometric analysis of 10,000 cells, representative of $n=4$ independent experiments. **g**, Cellular phosphate efflux after KIDINS220 inactivation. Three days after genetic inactivation of *XPR1* or *KIDINS220*, cellular phosphate efflux was assessed. Cells were loaded with ^{32}P-labeled phosphate, washed extensively to remove excess ^{32}P and then phosphate efflux was measured at the indicated times by isolation of conditioned medium and cellular lysates ($n=3$ technical replicates of the same transduction, representative of $n=3$ independent transductions). Phosphate efflux is calculated as the percentage of ^{32}P in the conditioned medium relative to total ^{32}P measured for that sample. Note that medium without phosphate does not stimulate phosphate efflux, and was used as a control.

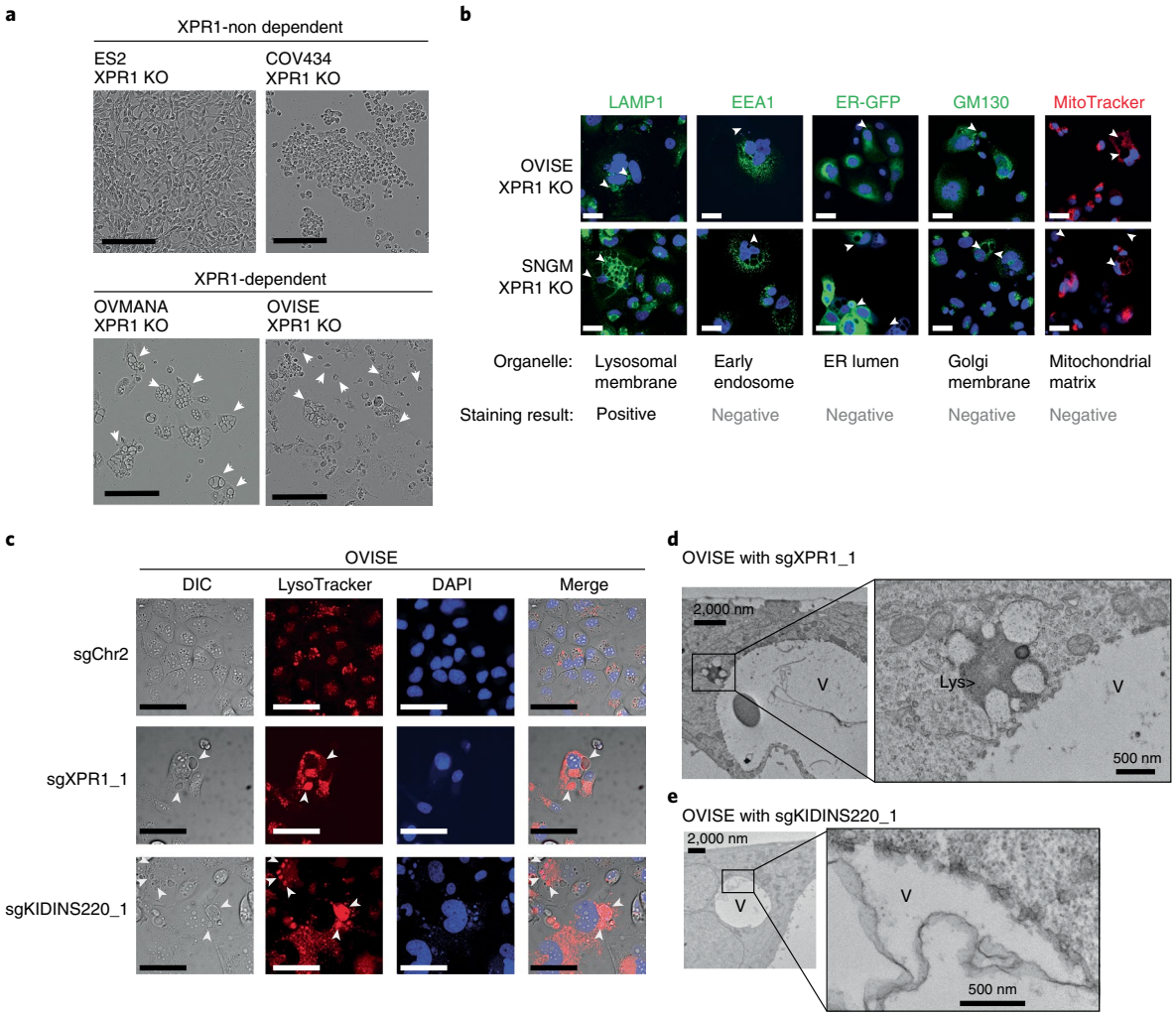


Fig. 7 | Vacuole structures precede loss of cell viability and are not derived from many common organelles. a, Phase-contrast images of vacuole-like phenotype 4–5 days after *XPR1* inactivation. Arrowheads indicate the locations of vacuole-like structures. Scale bars, 200 µm. Data are representative of $n=2$ independent transductions. **b**, Six days after infection with lentivirus encoding sgXPR1_2, OVISE and SNGM cell lines were stained and imaged using the indicated dyes and stains. Arrowheads indicate the locations of vacuole structures by phase contrast (not pictured). Positive staining was observed only for the lysosomal dye LAMP1. Scale bars, 100 µm. Data are representative of $n=2$ independent transductions. **c**, The acidic dye LysoTracker was used to stain live cells 5 days after inactivation of XPR1. Scale bars, 100 µm. Data are representative of $n=2$ independent transductions. **d**, Transmission electron micrographs of vacuole-like structures (v) and lysosomes (lys) in OVISE cancer cells after *XPR1* inactivation ($n=1$ experiment). **e**, As in **d**, but with *KIDINS220* inactivation ($n=1$ experiment).

that elevated activity of the lineage survival transcription factor PAX8 drives *SLC34A2* expression, and that PAX8-driven *SLC34A2* expression is sufficient to engender XPR1 sensitivity. Further work is needed to elucidate the many mechanisms by which *SLC34A2* is regulated in both normal and cancer cells.

This study also highlights fundamental gaps in our understanding of how intracellular phosphate is sensed, regulated and stored. Although excessive phosphate has previously been shown to be toxic[55,56], the exact mechanism of this toxicity is unknown. Our transcriptional profiling experiments failed to identify a clear 'phosphate stress response', and yet XPR1 perturbation causes a profound delay in cell growth. Along with cell cycle arrest, we also observed large, acidic, vacuolar structures (Fig. 7). Whether these structures reflect a compensatory cell survival mechanism involving the sequestration of potentially toxic phospho-metabolites[57,58], or are themselves the cause of cell death, remains to be determined.

Finally, this study uses the power of large-scale functional genomics to elucidate biological processes: we show that the phosphate efflux activity of XPR1 is entirely dependent on KIDINS220. The *KIDINS220* gene essentiality profile (Fig. 6a) is both selective and highly correlated with *XPR1*, suggesting a specific role for these two proteins in phosphate efflux as opposed to a more general or pleiotropic role. We show that the XPR1–KIDINS220 protein complex is probably oligomeric (Fig. 6d) and trafficks between multiple subcellular compartments, an activity previously attributed to both proteins[45,50]. Interestingly, we found that *XPR1* inactivation or suppression invariably caused loss of KIDINS220 protein; nevertheless, KIDINS220 protein loss is not required for cell death because the XRBD phosphate efflux inhibitor did not decrease KIDINS220 protein levels (Extended Data Fig. 7i). The exact mechanism of the XPR1–KIDINS220 phosphate efflux complex requires further study, as does reconciliation of cellular phosphate efflux with other activities attributed to KIDINS220, such as neurotrophin signaling[47,48] and genetic associations between *XPR1*, *KIDINS220* and various diseases[59].

Together, this study establishes the XPR1–KIDINS220 protein complex as a previously unrecognized therapeutic target in ovarian

and uterine cancer. Moreover, the work highlights disruption of phosphate homeostasis as a potential new anticancer strategy.

Methods

In vivo studies were carried out under the Institutional Animal Care and Use Committee (IACUC) of the Broad Institute under animal protocol no. 0194-01-18.

Statistics and reproducibility. In general, tissue culture experiments were conducted with cells grown in parallel vessels to evaluate the reproducibility of numeric data (for example, the two wells treated with an identical dose of XRBD in Fig. 5d). In most cases, experiments were always conducted at least twice to confirm effect size; in most situations only representative experiments are shown. This is the case for data from each cell line in Figs. 1c–e, 3c, 4b,d,f, 5, 6b,c (both replicates are shown), 6d–g and 7, and in Extended Data Figs. 2b–f, 4d, 5b, 6c,d, 7h–j, 8 and 9b,c,g,i. The experiment presented in Extended Data Fig. 2a was conducted once only. The cell lines presented in Fig. 6f and Extended Data Fig. 9h are representative of at least two clones derived by single-cell isolation. For the in vivo competition assay (Fig. 2), cells were infected with the small library of sgRNA only once, and parallel cultures or different subcutaneous xenografts were treated as 'technical replicates'. Sample sizes were chosen (eight mice, two tumors per mouse) to account for the penetrance of tumor development, to evaluate multiple time points and to have at least two animals per time point. Those animals with the largest tumor burden were euthanized at the indicated time points (that is, no randomization was employed). No statistical method was used to predetermine sample size. For cell death pathway profiling after *XPR1* suppression (Fig. 1f), RNA-seq after PAX8 suppression (Fig. 3b), genome-scale modifier screen of XPR1 dependency (Fig. 4a) and MixSeq transcriptional profiling (Fig. 4e), experiments were conducted once only. Statistical significance testing is discussed in more detail in Methods .

Genetic dependency data. The dependency data used in this manuscript were derived from the Public Avana 21Q2 dataset, consisting of dependency data for 18,025 genes across 859 cancer cell lines from 26 lineages. Expression data from the Cancer Cell Line Encyclopedia (CCLE) were also used. These data are available online, at https://depmap.org/portal/download/all/. In Fig. 1a, selectivity[60] and predictability[12,17] were determined as previously reported. 'Highly predictable' genes are indicated if the Pearson correlation coefficient between experimental data and the top predictive model is >0.4. The median dependency for each gene in ovarian/uterine cancers ($n=62$) was subtracted from that in all other cancer cell lines ($n=671$) to calculate ovarian/uterine genetic dependency enrichment on the *y* axis of Fig. 1a. In Fig. 1b and Extended Data Fig. 1b, correlation of SLC34A2 expression and XPR1 dependency was performed using a two-tailed Pearson correlation test. To compare XPR1 codependencies (Extended Data Fig. 9a), a two-tailed Pearson correlation test was performed for XPR1 versus all other genes ($k=18,025$ genes, $n=859$ cell lines, although some cell line–gene pairs are not represented), and *P* values are reported after correction for multiple comparisons with the Benjamini–Hochberg method.

Cell lines. Cancer cell lines ES2, HeyA8, A2780, 59 M, SNU8, OVK18, SNGM, OVISE, EMTOKA, IGROV1, OVCAR4, KURAMOCHI, RMGI, COV413a, JHOS4, HEC6 and JHUEM1 were collected by the CCLE before distribution for our use. The sources of the aforementioned cell lines can be found at DepMap.org. All cell lines were adapted to growth in RPMI 1640 (Corning) + 10% fetal bovine serum (FBS) before use. All cell lines are routinely validated using short-tandem repeat profiling.

sgRNA sequences. The negative control guides sgChr2 and sgAAVS1 were designed to cleave a gene desert and an intronic region in *PPP1R12C*, respectively, to control for the effects of DNA double-strand breaks. sgLacZ targets a sequence not found in the human genome. Positive-control sgRNA targets common essential splicing factors (SF3B1), ribosomal subunits (POLR2D) or kinesin motor proteins (KIF11). The 20-base-pair (bp) targeting sequences can be found in Supplementary Table 2.

Lentiviral production. Lentiviral production was performed using HEK293T cells as described on the Broad Institute Genetic Perturbation Web Portal (https://portals.broadinstitute.org/gpp/public/).

Plasmids, overexpression constructs and site-directed mutagenesis. Open reading frames (ORFs) of the following genes were obtained from a genome-scale library of annotated genes[61]. SLC34A2 (NM_006424) was isolated from this library in pDONR223 and was transferred into the expression vector pLX-TRC313 (similar to Addgene, no. 118017) using gateway cloning. The resultant construct has a C-terminal V5 tag and, after stable integration into cell lines using lentiviral infection, the proper protein product with a V5 tag was detected using immunoblot (not shown). XPR1 constructs (both isoforms, NM_004736 and NM_001135669) were obtained in a similar way. Only the NM_004736 isoform was observed using isoform-agnostic PCR primers and complementary DNA generated from the OVISE cancer cell line. Mutations were introduced with PCR-based methods, either the Q5 Site-directed Mutagenesis kit (NEB, no. E0554S) for large deletions or the QuickChange II XL (Agilent, no. 200521) for point mutations, and were confirmed using Sanger sequencing. The XPR1 ORFs were then transferred to pLX-TRC313 or the same expression vector with a weaker, PGK promoter. For coimmunoprecipitation experiments in HEK293T cells, the stronger promoter XPR1 mutants were used to maximize expression levels. The weaker promoter construct was used for stable expression in ovarian cancer cell lines and in immunofluorescence and mutant-rescue experiments.

Immunohistochemistry. Cell pellets were fixed using paraformaldyhde (PFA) and then paraffin embedded. Immunohistochemistry was performed on the Leica Bond RX automated staining platform using the Leica Biosystems Refine Detection Kit. The antibody for SLC34A2/MX35 (Creative Biolabs, no. TAB-467MZ, recombinant) was run at 1:400 dilution with citrate antigen retrieval.

CRISPR viability assays. CRISPR viability assays were performed in 96-well plates with cells seeded at a low density to allow for logarithmic growth throughout the entire assay. For 7-day assays, cells were seeded and infected with lentivirus expressing the sgRNA in pXPR-BRD003 on day 0. The next day, infection medium was replaced with 100 µl of medium. On day 7 postinfection, viability was evaluated by the addition of 25 µl per well of Cell Titer Glo (Promega) reagent and luminescence measured. Infection efficiency was determined by comparing the viability of cells with and without puromycin after infection, and assays were repeated if <80% of cells were infected with every sgRNA. The data were normalized such that the cutting-control sgRNA (targeting Chr2-2 and AAVS1) was 0 and positive-control sgRNA (targeting the common essential genes KIF11, SF3B1 and POLR2D) was −1.0. For 10-day assays, infections were carried out in six-well plates. Three days postinfection, cells were extracted and seeded into replicate 96-well plates. On days 3, 7 and 10 postinfection, viability was evaluated by the addition of 50 µl per well of Cell Titer Glo (Promega) reagent and monitoring of luminescence. Fold change (FC) viability was calculated by comparison of days 7 or 10 to day 3, and data were normalized as above.

Short hairpin RNA sequences. Short hairpin RNA sequences for XPR1 were selected from project DRIVE's subgenome-scale shRNA library[60] using DEMETER2 estimates for on- and off-target seed effects[62,63]. A detailed protocol for selection of shRNA using these datasets is available online (https://protocols.io/view/shrna-selection-and-quality-control-for-cancer-tar-bfmnjk5e). Doxycycline (dox)-inducible shRNA expression was accomplished by cloning these sequences into the pRSITEP-U6Tet-(shRNA)-EF1-TetRep-2A-Puro vector (Cellecta, no. SVSHU6TEP-L). shRNA seed-matched negative control sequences[62] were generated by substitution of complementary bp sequences into positions 9–11 bp of the target shRNA using a web-based tool (https://web.archive.org/web/20180605134130/;http://rnai.nih.gov/haystack/C911Calc2.html).

shXPR1 and seed-matched control sequences were rigorously tested for both on-target XPR1 suppression and off-target cell viability effects. Off-target cell viability effects were determined when a given shSeed control sequence did not knock down XPR1 but produced strong loss of cell viability, regardless of whether a cell line was predicted to be XPR1 dependent or nondependent. shRNA target sequences are provided below.

Antibodies. A full list of antibodies and their respective dilutions is included in the Reporting summary accompanying this paper.

Protein analysis of cell lysates by immunoblotting. For protein analyses, cells were grown in six-well dishes and harvested by washing with PBS and incubation with 0.5 ml of TrypLE until all cells had lifted, followed by dilution to 1.5 ml with PBS. Cells were then centrifuged and washed once with PBS and lysed with radioimmunoprecipitation assay buffer (150 mM NaCl, 1.0% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris pH 8.0) supplemented with cOmplete, Mini Protease and Phosphatase Inhibitor Cocktail Tablets (Roche). Cell extracts were cleared by spinning at 15,000 r.p.m. for 10 min at 4°C. Protein content was quantified by bicinchoninic acid (BCA) analysis, with the addition of 4× lithium dodecyl sulfate reducing sample buffer. We found, by immunoblot, that boiling of lysates led to a loss of XPR1 protein, and so they were incubated at 37 °C for 30 min to denature proteins. For standard immunoblots (Extended Data Figs. 5b and 6e), equal amounts of protein (typically 25 µg) were resolved by SDS–polyacrylamide gel electrophoresis (SDS–PAGE), transferred to nitrocellulose and incubated with the indicated primary and secondary antibodies to visualize protein levels. Images were obtained using a LI-COR Odyssey CLx system.

For quantitative determination of cellular proteins (Fig. 6b), lysates were analyzed using Protein Simple, an automated capillary-based protein separation and immunoblotting assay. Lysates were prepared as above and diluted to 0.1 mg ml^{-1} with sample buffer and 40 mM DTT before loading 3 µl of sample onto each plate.

Foci formation. Cells stably expressing doxycycline-inducible short hairpins against XPR1 or the corresponding seed controls were plated at three different densities (18,000, 12,000 and 6,000 cells per well) to determine optimal seeding

density. Six replicates of each cell line were plated in 24-well plates, and half of the wells were treated with 0.5 μg ml^{-1} doxycycline. Media and doxycycline were refreshed every 2–3 days until untreated wells reached 100% confluence. Cells were then washed with 1× PBS and fixed with 4% PFA in 1× PBS for 15 min. Fixation was quenched with deionized water and cells were then stained with 0.1% crystal violet in deionized water for 20 min. Cells were then washed with deionized water to remove residual crystal violet and left to dry overnight. To quantify crystal violet staining, 10% acetic acid was incubated in each well for 20 min, diluted 1:3 with deionized water and replated in quadruplicate in a 96-well plate. Absorbance was measured at 590 nm.

Time-course analysis of cell growth. To measure cell growth over time after *XPR1* suppression, cells were seeded in a 96-well plate, treated with doxycycline to induce shXPR1 and imaged every 6 h (Essen Incucyte S3). Images were quantified using built-in software.

Proteome Profiler for cell death markers. To evaluate markers associated with different cell death pathways, cells were treated with doxycycline for 5 days to induce shXPR1, washed with PBS, lysed and processed according to the instructions provided for the Proteome Profiler Human Apoptosis Array Kit (R&D Systems, no. ARY009). The resulting dot blot array was quantified using ImageJ, and values were normalized to the –shXPR1 condition per cell line.

Annexin and DAPI staining. To evaluate the extent of cell death, cells were treated for 6 days with doxycycline to induce shXPR1, then both nonattached and adherent cells were collected, pooled and stained for 40 min at room temperature with DAPI, washed with Annexin Staining Buffer and incubated for 15 min at room temperature with Annexin V-APC (ThermoFisher, no. A35110). Cells were then washed, and >10,000 individual cells analyzed by flow cytometry. Data were analyzed using FlowJo, gating against smaller cellular debris as well as events consistent with more than one cell per droplet.

In vivo sgRNA competition assay. These studies were used under the IACUC of the Broad Institute under animal protocol no. 0194-01-18. A detailed protocol for the tumor formation competition assay is available online (https://www.protocols.io/view/in-vivo-nanopool-pooled-sgrna-competition-assays-t-bfsbjnan). A small lentiviral library of 74 total sgRNAs (including nontargeting, negative controls targeting gene deserts or introns, positive controls targeting pan-essential genes and experimental sgRNA) was made in an arrayed format and then pooled. To minimize time in tissue culture, we optimized infection and puromycin conditions to achieve roughly 30–50% infection efficiency and nearly 100% selection 50 h after infection. OVISE cells were selected with 8 μg ml^{-1} puromycin while SNGM cells were selected with 4 μg ml^{-1} puromycin.

Fifty hours after infection, cells were lifted, counted and diluted in 50% matrigel to a final concentration of 8 million cells 100 μl^{-1}. Some cells were frozen to determine the early representation of the library. For in vitro experiments, cells were replated and grown under standard conditions for 2 weeks. In total, this study used 26 female, 7-week-old, Rag1$^{-/-}$ Il2rg$^{-/-}$ mice obtained from The Jackson Laboratories. For in vivo experiments, mice were anesthetized under isoflurane gas and two bilateral subcutaneous xenografts inoculated in each of five mice (ten tumors per experiment). Tumors were measured twice weekly with calipers, and tumor volumes calculated using the following formula: 3.14/6 × (width^2 × length). Animal body weights were recorded once or twice weekly during the course of all studies. Mice were euthanized and tumors harvested on days 14, 21 and 28 after inoculation, ensuring they had not reached endpoint (tumor size >2,500 mm^3, no visual distress and no evidence of ascites development). After harvesting and weighing, tumors were flash-frozen in liquid nitrogen until required for genomic DNA isolation.

At the end of the study (28 days postinoculation, 30 days postinfection), tumors were thawed and minced and genomic DNA extracted from all samples using Qiagen DNeasy Blood and Tissue kits. Sample barcodes were sequenced by Illumina Next-Generation sequencing then deconvoluted with Broad Genetic Perturbation Platform's PoolQ software for sgRNA read counts. Data from all animals included in the study are provided.

Statistical analysis for in vivo competition assay. The gene-level effect was determined by comparison of normalized sgRNA read counts at the early time point (2 days after infection, the day of inoculation) with those at the indicated time points. If normalized read counts at the early time points were significantly different for a particular sgRNA, that sgRNA was not included in downstream analyses (an indication that the plasmid had not produced lentivirus). sgRNA read counts were normalized such that the seven cutting-control sgRNAs (representing viability effects from CRISPR–Cas9 genome editing) had a median depletion of 0. Next the median fold change of all sgRNAs targeting a particular gene was calculated, and only those genes with twofold depletion or greater relative to control sgRNA are reported per replicate in Extended Data Fig. 3. The different time-point replicates were median averaged and are reported in Fig. 2c. The statistical significance of XPR1 depletion was calculated by comparison of the gene-level depletion of XPR1 (three different sgRNAs per replicate) with cutting-control sgRNA (seven different sgRNAs per replicate), using a *t*-test to compare all replicates at each time point. The test was conducted in GraphPad Prism 8.0 with Holm–Sidak multiple comparisons correction. Corrected *P* values (*q*-values) are reported in Fig. 2c. Test statistics are as follows: OVISE TC 2 weeks ($n = 2$, t-value = 29.6, 14 degrees of freedom); OVISE tumor 2 weeks ($n = 4$, t-value = 8.9, 30 degrees of freedom); OVISE tumor 3 weeks ($n = 4$, t-value = 7.1, 30 degrees of freedom); OVISE tumor 4 weeks ($n = 2$, t-value = 4.6, 30 degrees of freedom); SNGM TC 2 weeks ($n = 2$, t-value = 19.1, 14 degrees of freedom); SNGM tumor 2 weeks (n = 2, t-value = 5.6, 16 degrees of freedom); SNGM tumor 3 weeks ($n = 4$, t-value = 2.3, 28 degrees of freedom); SNGM tumor 4 weeks ($n = 4$, t-value = 5.0, 22 degrees of freedom).

Comparing expression of genes across normal and tumor tissues. We compiled transcripts per million (TPM) gene expression data for normal fallopian tube (GTEx, $n = 5$), normal ovary (GTEx, $n = 88$), normal uterus (GTEx, $n = 78$), ovarian cancer (TCGA OV, $n = 426$) and uterine cancer (TCGA UCEC, $n = 238$) from the TOIL RSEM $\log_2$(TPM + 0.01) data at Xena Browser (https://xenabrowser.net/) and then converted the data to $\log_2$(TPM + 1). RNA-seq gene expression for ovarian cancer cell lines (CCLE, $n = 40$), and uterine cancer cell lines (CCLE, $n = 22$) was downloaded from the CCLE (https://depmap.org) as $\log_2$(TPM + 1). Because most TCGA ovarian and uterine samples have relatively high purity (>80%)[64], we used these data directly for the following comparisons. In Fig. 3a,e and Extended Data Fig. 4c,f, boxplots were drawn using the 'geom_boxplot' command in the R package ggplot2, such that boxes span the first and third quartiles of values with the median indicated by a line; whiskers extend 1.5× interquartile range, with outliers beyond this range excluded. In Fig. 3a,e and Extended Data Fig. 4c, the Wilcoxon ranked-sum test was employed using the R package rstatix to compare the distribution of expression of the indicated gene between the indicated tissues, and *P* values were corrected for multiple comparisons using Bonferroni's method. In Extended Data Fig. 4a a pairwise Wilcoxon ranked-sum test was used to compare the expression of SLC34A2 in each tissue relative to all other tissues, and *P* values were corrected for multiple comparisons using Bonferroni's method and are reported on the *y* axis. The difference in median SLC34A2 expression (tissue expression across all tissues) is plotted on the *x* axis of Extended Data Fig. 4a. In Extended Data Fig. 4b, the correlation between SLC34A2 and PAX8 mRNA was tested using a two-tailed Pearson correlation test across the indicated tissues ($n = 897$). In Extended Data Fig. 9e, the correlation between XPR1 and KIDINS220 mRNA was tested using a two-tailed Pearson correlation test for the indicated tissue groups (all 60 tissues, $n = 17{,}194$; top 15 correlated, $n = 2{,}799$; and for all other tissues, listed from top to bottom, $n = 337$, 55, 172, 173, 36, 47, 182, 520, 66, 182, 496, 154, 119, 181 and 79, respectively).

PCR with reverse transcription. After the indicated perturbations, cells were lysed and RNA extracted using Qiazol and phenol-chloroform. Total RNA was determined spectrophotometrically, and normalized amounts of RNA converted to cDNA using the iScript kit (Bio-Rad, no.1708890). Diluted cDNA was then mixed with gene-specific primers and Power Sybr Green Master Mix (Thermo Fisher, no. 4367659) and analyzed on a Quant Studio 7 PCR with reverse transcription (RT–PCR) instrument. A full list of RT–PCR primers used can be found in Supplementary Table 3. The quality of RNA extraction was evaluated by comparison of CT values of cDNA samples with control samples treated in the same way but without the addition of reverse transcriptase. Gene expression values were corrected for loading with a housekeeping gene (*VCL*). In Fig. 3c, SLC34A2 gene expression values were compared using a one-tailed *t*-test corrected for multiple comparisons with the Bonferroni method. Test statistics are as follows: sgPAX8_1 ($n = 4$, t-value = 3.54, degrees of fredom = 5.46); sgPAX8_3 ($n = 4$, t-value = 5.66, degrees of freedom = 5.19); sgSLC34A2_1 ($n = 4$, t-value = 7.24, degrees of freedom = 5.94); sgSLC34A2_2 ($n = 4$, t-value = 11.45, degrees of freedom = 4.20).

Analysis of XPR1 copy number in TCGA. To evaluate the frequency of XPR1 amplification, we evaluated precomputed GISTIC2 (ref. [35]) analyses for recurrent copy number alterations in ovarian and uterine TCGA datasets[22,23]. To compare the expression of XPR1 with its copy number status, XPR1 copy number thresholds—as determined by GISTIC—were downloaded from CBIoPortal[65,66] and then matched to the corresponding TCGA samples. In total, 410 ovarian cancer and 171 uterine cancer samples were included in this analysis. In Fig. 3d and Extended Data Fig. 4g, each patient sample is represented by a horizontal line; red indicates copy gain and blue indicates copy loss, and dashed vertical lines represent the locations of indicated genes. The samples are rank ordered by highest copy gain, to indicate both focal and chromosome arm variants. Figure 3e, shows testing of whether there was a significantly nonzero correlation between XPR1 copy number and XPR1 mRNA expression using a two-tailed Spearman correlation test, and reports Spearman's rho and *P* value. Also in Fig. 3e, the expression of XPR1 between tissue categories was compared using a Wilcoxon ranked-sum test with Bonferroni correction.

Modifier screen. The anchor modifier screen was performed as described previously[36]. OVISE cells stably expressing sgRNA targeting Chr2-2, XPR1_1 or XPR1-2 in the lentiviral guide-only vector pXPR-BRD016 were infected with the

Brunello all-in-one vector (pXPR-BRD023), in a format such that each cell received a maximum of one sgRNA. Twenty-four hours postinfection, cells were split into two replicates and treated with 2 μg ml^{-1} puromycin to select for those with stable integration. Every 3–4 days, cells were trypsinized and replated to maintain a minimum representation of 500× library representations per replicate. At 15 days postinfection, all cells were collected and counted. The final representations for each replicate were as follows: 533× and 533× for sgChr2-2, 287× and 231× for sgXPR1_1 and 81× and 83× for sgXPR1_2. Genomic DNA was then extracted from each cell line and sgRNA barcodes were amplified by PCR.

Modifier screen analysis. The representation of each sgRNA in each arm of the experiment was determined by next-generation sequencing, and compared with plasmid DNA representation of the entire library. Using MAGeCK MLE[67], this change in representation was converted to gene-level beta-values, representing the variability of each gene relative to all other genes. A permutation test (ten different permutations) on sgRNA labels was used to empirically determine *P* values, which were then corrected using the Benjamini–Hochberg procedure. Any gene with false discovery rate (FDR) <0.1 was considered a potential hit in that screen. SLC34A2 was the only statistically significant 'rescue' gene (that is, beta-value >0) for both sgXPR1_1 and sgXPR1_2 arms that was not significant in the control arm. There were several genes with beta-values significantly <0. In comparison of beta-values for this list of potential 'sensitizer' genes between the sgXPR1_1/2 arms and the sgChr2-2 arm, we note that all of these are significantly depleted across every arm, with little difference in beta-value. This indicates that these are essential genes, and their depletion observed in every arm of the experiment is probably not due to a relationship with XPR1.

Comparison of XPR1 dependency and tissue culture medium. Information on growth media used is available for cell lines at depmap.org/portal. Because each cell line is grown in a mixture of several medium types (for example, 90% RPMI 1640 + 10% FBS), we estimated the concentration of phosphate using a weighted average of phosphate concentrations of each component (Supplementary Table 4 and ref. [68]). The XPR1 dependency score was then compared to the concentration of phosphate using a one-tailed Pearson correlation test, in which we expected to find that more dependent cell lines were grown in higher concentrations of phosphate, and exact *P* values and sample sizes are reported in Extended Data Fig. 6a.

Cell competition assays in low-phosphate media. A competition experiment was designed using isogenic cell lines expressing either Firefly or Renilla luciferase; the Firefly luciferase-expressing version of the cell line was also engineered to express Cas9.

For the competition assay in low phosphate, luciferase-expressing cells were adapted from normal growth media (RPMI 1640 + 10% FBS) to low-phosphate media for 7 days. The concentration of phosphate in these media was empirically determined as the minimum required to support sustained cell viability for 3 weeks in culture. Low-phosphate media had a 9:1 ratio of (RPMI 1640 without L-glutamine phosphate + 10% FBS; MP Biomedicals, no. 09162975)/normal growth medium (RPMI 1640 + 10% FBS). We estimate that 100% FBS contains 5.3 mg dl^{-1} inorganic phosphate, although this value is probably variable[68]. Therefore, this low-phosphate RPMI would have a concentration of 7.8 mg dl^{-1}.

After cells were adapted to low-phosphate media for 7 days, Renilla luciferase-expressing cells were mixed at a 1:1 ratio with Firefly luciferase + Cas9-expressing cells and infected with the indicated sgRNAs on day 8. Cells were selected with puromycin 24 h after infection. On day 11, cells were split with half replated to propagate the cultured mixtures while the other half was subjected to Dual-Glo Luciferase Assay (Promega) to set the baseline signal. The final luciferase assay was performed on day 16 after initial culture in low-phosphate media.

XRBD protein purification. We thank J. L. Battini for providing the sequence for the XRBD-mFc construct published previously[14]. The plasmid-encoding XRBD (strain NZB) was synthesized and cloned into pcDNA3.4 by GeneArt Gene Synthesis (GENEART, Thermo Fisher Scientific). XRBD protein was expressed in CHO cells and purified using Protein A affinity chromatography (GENEART).

XRBD flow cytometry. Cells were extracted from culture vessels using TrypLE Express (ThermoFisher, no.12604013) and then diluted in PBS + 2% FBS. A total of 300,000 cells in 50 μl were transferred to a U-bottom. 96-well plate followed by the addition of 50 μl of XRBD staining solution. In Fig. 5b testing of the indicated doses is shown while Fig. 6f shows the use of XRBD at 100 nM. Cells were incubated at 37 °C for 40 min, washed once and then incubated with an anti-mouse secondary antibody conjugated to AlexaFluor 488 (ThermoFisher, no. A-11004). Cells were then incubated on ice for 40 min, washed four times in PBS + 2% FBS and analyzed on a CytoFlex LX instrument. At least 10,000 single-cell events were recorded for each condition.

XRBD viability assays. To assess viability defects after exposure to XRBD, dilutions of XRBD were made in PBS and then added to cells plated 24 h previously in 96-well plates. Five days after the addition of XRBD, cellular viability was assessed by Cell Titer Glo.

Measurement of intracellular phosphate. To measure intracellular phosphate, cells were plated in six-well plates, as described above, to perturb XPR1, KIDINS220 or SLC34A2. After the indicated times, cells were washed three times with tris-buffered saline (TBS) to remove residual phosphate from media. Cells were then lysed in 1% NP40, 50 mM Tris pH 7.5 and protease inhibitors, and cellular debris cleared by centrifugation. The cell lysate, or a dilution series of a phosphate standard, was then diluted in water to 50 μl in a 96-well microtiter plate and 10 μl Malachite Green Reagent A (R&D Systems, no. DY966) was added according to the manufacturer's instructions. After 10 min, 10 μl of Malachite Green Reagent B was added to the samples and the absorbance at 620 nm read immediately. If any sample was not within the linear range of the assay, that sample was iteratively rediluted and reanalyzed. Intracellular phosphate levels were calculated by dividing the interpolated phosphate concentration by protein concentration (determined by BCA assay) in each sample. In Fig. 4d and Extended Data Figs. 6d and 9i, technical triplicates representative of at least three experiments are displayed.

Immunofluorescence and fixed cell-compatible dyes. Cells stably expressing V5-tagged XPR1 and KIDINS constructs were plated at a density of 10,000–20,000 per well in Nunc Lab-Tek II CC2 eight-well chamber slides (ThermoFisher). When investigating localization following knockdown, cells were plated in chamber slides after 5 days of lentiviral transduction of sgRNA targeting XPR1, KIDINS220 or a noncoding portion of chromosome 2. The following day, cells were washed with 1× PBS (Corning) and fixed in 4% PFA (Electron Microscopy Sciences) in 1× PBS for 15 min. Cells were washed twice with 1× PBS to quench fixation, permeabilized with 0.1% Triton X-100 in 1× PBS for 15 min and then blocked in 1% bovine serum albumin (BSA) in 1× PBS for 1 h. Cells were probed overnight at 4 °C with primary antibody diluted in 0.1% BSA in 1× PBS. The following day, cells were washed three times with 0.1% Triton in 1× PBS and stained for 1 h at 25 °C with AlexaFluor conjugated secondary antibodies (Molecular Probes, ThermoFisher) diluted in the blocking buffer according to the table below. The wells were then washed three times with 1× PBS and counterstained with DAPI in 1× PBS at 2 μg ml^{-1} for 20 min. The wells were washed twice with deionized water, and cells were mounted in ProLong Gold AntiFade Mountant (ThermoFisher).

For determination of the organelle source of vacuole-like structures, the same general immunofluorescent staining protocol was used as above with the following changes. SNGM and OVISE Cas9 stable cell lines were plated at 10,000–20,000 cells per well on μ-Slide eight-well coated chamber slides (IBIDI, no. 80826) and simultaneously infected with lentivirus-expressing sgRNAs. The next day, transduced cells were selected with 2 μg ml^{-1} puromycin for 24 h, removed from puro selection and fixed 6 days after infection.

For endoplasmic reticulum labeling, cells were transduced with 24 μl of CellLight ER-GFP, BacMam 2.0 in 200 μl of culture medium 24 h before imaging (ThermoFisher, no. C10590), with no permeabilization step performed. For mitochondrial imaging, cells were treated with 100 nM MitoTracker Red CMXRos (ThermoFisher, no. M7512) in serum-free RPMI for 30 min at 37 °C, then MitoTracker dye medium was replaced with normal growth medium (RPMI with 10% FBS) and incubated for 1 h at 37 °C before fixation. All other antibody-based stains were treated as described above. Information on the antibodies used and their concentrations for staining are provided above.

Multiplexed transcriptional profiling. Multiplexed transcriptional profiling (MixSeq[69]) was performed using custom pools of ovarian and uterine cancer cell lines. Cancer cell lines were pooled (five to seven cell lines per minipool) based on doubling time, and then frozen. To initiate the experiment, cells were thawed and plated in 12-well dishes. The next day, virus-encoding mixtures of sgRNAs (sgLacZ/sgChr2-2 or sgXPR1_1/sgXPR1_2) were prepared under conditions in which each cell received both sgRNAs to increase the penetrance of inactivation. Cells were treated with 2 μg ml^{-1} puromycin 24 h after infection. Four days after infection, cells were extracted with TrypLE, spun down, resuspended in cell-staining buffer (PBS + 2%BSA + 0.02% Tween) and counted.

Perturbations were multiplexed for 10X sequencing using Cell Hashing[70]. Equal numbers of each minipool were then pooled, blocked with FcX blocking buffer (BioLegend) for 10 min on ice and incubated with hash-antibodies (TotalSeq, BioLegend) for 30 min on ice. Cells were then washed three times with cell-staining buffer and resuspended in Cell Capture buffer (PBS + 0.04% BSA), filtered through a 40-μm filter and diluted to ~1,500 cells μl^{-1}. The detailed protocol can be found online (https://www.protocols.io/view/cell-hashing-zn9f5h6). Approximately 40,000 cells were then loaded onto a 10X Chromium controller using v.3 single-cell 3' reagent chemistries. Library preparation and next-generation sequencing were performed as previously reported[70].

Sequencing data were processed using 10X Cell Ranger software (v.3, hg19 reference genome) run with the Cumulus cloud-based analysis framework[71]. Single-nucleotide polymorphism (SNP)-based cell line classification and quality control was performed according to the methods described in ref. [72]. In brief, for each cell line the allelic fractions across a predefined 100,000-SNP reference set were estimated from bulk RNA-seq data using Freebayes[72]. A logistic regression model was then used to estimate the likelihood of the observed SNP reads for an individual cell having come from each cell line given the allelic fractions across

the SNP reference set. A similar model was used to detect doublets, where allelic fractions were modeled as a mixture of allelic fractions from two different cell lines. After SNP-based classification, low-quality cells and doublets were removed according to a set of stringent filters. These include the proportion of unique molecular identifiers from mitochondrial genes being 0.25–0.01, the number of reads at reference SNP sites being >50 and the likelihood that a cell matches a particular cell line being much higher than the likelihood that it matches a different cell line or doublet.

MixSeq single-cell data were analyzed using the Seurat R package v.3 (ref. [73]). Data were first normalized and scaled using the functions NormalizeData and ScaleData (default parameters). To generate uniform manifold approximation and projection embedding (Extended Data Fig. 7d), the top 5,000 most variable genes were selected using the function FindVariableFeatures, principal components were computed using the function RunPCA and embedding was generated using the function RunUMAP with 20 principal components, ten nearest neighbors and a minimum distance of 0.3 (default parameters otherwise). Cell cycle phase classification was performed with the function CellCycleScoring, using S- and G2M-phase marker gene sets reported previously. Differences in the proportion of cycling cells (ΔG0/G1) (Extended Data Fig. 7e) was calculated for each cell line as the fraction of cells in G0/G1 in the XPR1 sgRNA condition minus that in the control condition.

MixSeq differential expression analysis was performed using the pipeline limma-voom[74,75]. First, per-cell normalization factors were calculated using the TMMwzp method from the edgeR R package[76]. Then, counts per million (CPM) were converted to log(CPM) and the mean–variance relationship estimated using the voom function from the limma R package[74,75]. To identify each cell line's transcriptional response to XPR1 knockout (KO), removing the effect of cell cycle (Extended Data Fig. 7e), a linear model was used with the S.Score and G2M.Score values from Seurat CellCycleScoring as covariates to regress out the effect of cell cycle. The top 500 differentially expressed genes displayed in the heatmap were identified based on the average log(FC) across the eight cell lines. To identify the average transcriptional response within the three highly correlated cell lines (RMGI, IGROV1 and OVISE) and the five less correlated (JHUEM1, OVCAR4, COV413A, JHOS4 and HEC6; Extended Data Fig. 7f,g) a similar linear model was used, but cell line identity was added as a covariate to account for differences in baseline gene expression between cell lines. *P* values were derived from empirical Bayes-moderated *t*-statistics, and FDR adjusted *q*-values were obtained using the Benjamini–Hochberg method.

FGF23 ELISA. To measure levels of secreted FGF23 from cancer cell lines, the indicated cell lines were plated in six-well dishes and treated with doxycycline for 4 days to induce shRNA-targeting XPR1. The conditioned medium was then collected, centriguged to remove any debris and FGF23 levels were measured according to the manufacturer's instructions (R&D Systems, no. DY2604-05). Similar results were obtained using a second manufacturer's ELISA (Quidel Corp., no. 606600).

Phosphate uptake and efflux assays. To determine phosphate uptake or efflux, OVISE cells were infected in six-well plates with lentivrius-encoding, sgRNA-targeting XPR1, KIDINS220. Stable cell lines with SLC34A2 KO were also used, as were IGROV1 cells 3 days after doxycycline treatment to suppress *XPR1* using shRNA. The day before the experiment, cells were split into 96-well plates such that they would be confluent the following day. Cells were first 'pulsed' using low-phosphate RPMI 1640 (a 1:9 ratio of standard RPMI to no-phosphate RPMI; see the lower-phosphate assay, above) supplemented with 10 μci ml^{-1} $^{32}PO_4$ (Perkin Elmer, no. NEX053001MC) and incubated at room temperature for 30 min. Cells were then washed with no-phosphate RPMI 1640. To determine phosphate uptake and initial phosphate levels for time-course efflux experiments, cells were lysed with 1% Triton X-100 and the amount of intracellular ^{32}P measured using a liquid scintillation counter. For efflux time-course experiments, the 'chase' of phosphate efflux was then measured using high-phosphate RPMI (standard RPMI 1640). When incubated without phosphate in the medium (0 mM phosphate RPMI 1640), phosphate efflux was far lower (as previously reported[14]) and was used as a control. At each time point, the conditioned medium was taken, cells were washed three times with no-phosphate RPMI 1640 (to remove any radioactivity outside of the cells) and then lysed with 1% Triton X-100. Conditioned medium and cell lysates were analyzed for ^{32}P using a liquid scintillation counter. The extent of phosphate efflux was determined by dividing the ^{32}P level measured in the conditioned medium by total ^{32}P measured for that sample (in both cell lysates and the conditioned medium). For experiments using XRBD, protein was diluted in PBS and added to the medium during both the uptake/pulse and efflux/chase portions of the process.

High-throughput mass spectrometry analysis. High-throughput protein interaction databases BioPlex[77] and BioGRID[78] were used to search and download lists of primary physical interactors of both XPR1 and KIDINS220. Interactions were identified by affinity capture–mass spectrometry, where epitope tags on target 'bait' proteins act as affinity capture probes for identification of 'prey' interactor proteins. Common interactors were found by comparing the four gene lists then visualizing the overlap.

Coimmunoprecipitation of V5-tagged proteins. HEK293T cells were transiently transfected with 40 μg of pLX-TRC313 ORF vectors expressing XPR1, GFP or Luciferase for 24 h. Cells were then washed twice with TBS and lysed with 0.4% NP40, 50 mM Tris and 150 mM NaCl, supplemented with Halt Protease Inhibitor Cocktail (ThermoScientific). Cell extracts were cleared by spinning at 15,000 r.p.m. for 15 min at 4 °C on a tabletop centrifuge, and quantified by BCA. Fifty microliters of V5-tagged immunomagnetic beads (MBL International, no. M167-11) was incubated with 1 mg of protein lysate overnight at 4 °C. The next day, beads were washed five times with 0.2% NP40, 50 mM Tris and 150 mM NaCl. Bound protein was then eluted by the addition of 10× sample buffer (Protein Simple) and incubation of beads for 30 min at 37 °C. Eluate was removed by beads and probed alongside whole-cell lysates by either Protein Simple Automated standard or immunoblotting.

Protein ID by mass spectrometry. Eluate from V5 coimmunoprecipitation was run on 4–12% PAGE, stained with Safe Stain (Invitrogen, no. LC6065) and cut into eight pieces spanning the entire mW range of the gel. These excised gel bands were cut into three pieces of approximately 1 mm then subjected to a modified in-gel trypsin digestion procedure[79]. Gel pieces were washed and dehydrated with acetonitrile for 10 min, followed by removal of acetonitrile. Samples were then completely dried in a speed-vac. Rehydration of gel pieces was performed with 50 mM ammonium bicarbonate solution containing 12.5 ng μl^{-1} modified sequencing-grade trypsin (Promega) at 4 °C. After 45 min, excess trypsin solution was removed and replaced with 50 mM ammonium bicarbonate solution till the gel pieces were just covered. Samples were then placed in a room at 37 °C overnight. Peptides were later extracted by removal of ammonium bicarbonate solution, followed by one wash with a solution containing 50% acetonitrile and 1% formic acid. Extracts were then dried in a speed-vac (~1 h) and stored at 4 °C until required for analysis. On the day of analysis, samples were reconstituted in 5–10 μl of high-performance liquid chromatography (HPLC) solvent A (2.5% acetonitrile, 0.1% formic acid). A nanoscale, reverse-phase HPLC capillary column was created by packing 2.6-μm C18 spherical silica beads into a fused silica capillary (100 μm inner diameter × ~30 cm length) with a flame-drawn tip. After equilibration of the column, each sample was loaded via a Famos autosampler (LC Packings) onto the column. A gradient was formed and peptides were eluted with increasing concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid).

As peptides eluted they were subjected to electrospray ionization and then entered into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher Scientific). Peptides were detected, isolated and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. Peptide sequences (and hence protein identity) were determined by matching protein databases with the acquired fragmentation pattern using the software program Sequest (Thermo Fisher Scientific)[80]. All databases include a reversed version of all sequences, and data were filtered to 1–2% peptide FDR. Proteins identified and displayed in Fig. 6c met the following criteria: fewer than ten peptides identified in the Luciferase immunoprecipitate (IP:Luc) IP/Luciferase and tenfold more peptides in IP:XPR1 than IP:Luc.

Live-cell LysoTracker staining of acidic organelles. Similar to the immunofluorescence methods described above, OVISE.Cas9 stable cells were plated at 10,000 per well in μ-Slide, eight-well coated chamber slides (IBIDI, no. 80826) and simultaneously infected with lentivirus-expressing sgRNAs. Twenty-four hours after infection, cells were selected with 2 ug ml^{-1} puromycin for 2 days and then the medium was replaced with phenol red-free RPMI 1640 supplemented with 10% FBS. Live cells were incubated with dyes and imaged 5 days after infection.

For dye staining, LysoTracker Red DND-99 (Invitrogen, no. L7528) was resuspended as 1 mM stock in DMSO. Cells were stained in phenol red-free RPMI growth medium and 50 nM LysoTracker for 45 min at 37 °C in the dark. The dye was then washed out once with dye-free growth medium and incubated with 1 μg ml^{-1} Hoechst 33342 (Invitrogen, no. H3570) in phenol red-free growth medium at 37 °C for 30 min, and retained in Hoechst 33342-containing medium during imaging. Cells were imaged immediately after on a Nikon Eclipse Ti microscope with a Yokogawa Life Sciences CSU-W1 spinning-disc confocal system.

Ultrastructural analysis by TEM. After infection with either cutting-control (sgChr2-2) or experimental sgRNA (sgXPR1_1 and sgKIDINS220_1), cell lines OVISE.Cas9 were grown for 5 days in six-well dishes. Cells were washed once with PBS and then fixed in 2.5% glutaraldehyde + 2.5% PFA in 100 mM sodium cacodylate buffer pH 7.4. Cells were then washed in 100 mM sodium cacodylate buffer pH 7.4, postfixed for 30 min in a solution of 1% osmium tetroxide and 1.5% potassium ferrocyanide, washed in water three times and incubated in 1% aqueous uranyl acetate for 30 min, followed by two washes in water and subsequent dehydration in graded alcohol (5 min each; 50, 70, 95% and twice in 100%).

Cells were removed from the dish in propylene oxide, pelleted at 3,000 r.p.m. for 3 min and infiltrated for 2–16 h in a 1:1 mixture of propylene oxide and TAAB Epon (TAAB Laboratories, https://taab.co.uk). Samples were subsequently embedded in TAAB Epon and polymerized at 60 °C for 48 h. Ultrathin sections (~60 nm) were cut on a Reichert Ultracut-S microtome, extracted onto copper

grids stained with lead citrate and examined in a JEOL 1200EX transmission electron microscope, and images recorded with an AMT 2k CCD camera.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Publicly available data used in this study include CRISPR–Cas9 loss of viability screens for 851 cancer cell lines[10–12], cancer cell line RNA-seq expression data[15], harmonized gene expression data for GTEx and TCGA datasets[24] and copy number alterations for ovarian adenocarcinoma and uterine corpus endometrial carcinoma[22,23]. RNA-seq data following PAX8 suppression and multiplexed transcriptional profiling are available from the corresponding author upon request. RNA-seq data following PAX8 knockdown are available at https://doi.org/10.6084/m9.figshare.19125671.v1. MixSeq data following XPR1 knockdown are available at https://doi.org/10.6084/m9.figshare.19125677.v1. Source data are provided with this paper.

Code availability

Computer code to reproduce these results is available from the corresponding authors upon request.

Received: 17 July 2020; Accepted: 4 March 2022;
Published online: 18 April 2022

References

1. Lord, C. J. & Ashworth, A. PARP inhibitors: synthetic lethality in the clinic. *Science* **355**, 1152–1158 (2017).
2. Lheureux, S., Braunstein, M. & Oza, A. M. Epithelial ovarian cancer: evolution of management in the era of precision medicine. *CA Cancer J. Clin.* **69**, 280–304 (2019).
3. Vaughan, S. et al. Rethinking ovarian cancer: recommendations for improving outcomes. *Nat. Rev. Cancer* **11**, 719–725 (2011).
4. Sung, H. et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249 (2021).
5. Boehm, J. S. et al. Cancer research needs a better map. *Nature* **589**, 514–516 (2021).
6. Costanzo, M. et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**, aaf1420 (2016).
7. Pan, J. et al. Sparse dictionary learning recovers pleiotropy from human cell fitness screens. *Cell Syst.* https://www.sciencedirect.com/science/article/pii/S2405471221004889 (2022).
8. Cheung, H. W. et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc. Natl Acad. Sci. USA* **108**, 12372–12377 (2011).
9. Price, C. et al. Genome-wide interrogation of human cancers identifies EGLN1 dependency in clear cell ovarian cancers. *Cancer Res.* **79**, 2564–2579 (2019).
10. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
11. DepMap. DepMap 21Q2 Public. *figshare* https://doi.org/10.6084/m9.figshare.14541774.v2 (2021).
12. Dempster, J. M., Rossen, J., Kazachkova, M. & Pan, J. Extracting biological insights from the Project Achilles genome-scale CRISPR screens in cancer cell lines. Preprint at *bioRxiv* https://doi.org/10.1101/720243 (2019).
13. Battini, J. L., Rasko, J. E. & Miller, A. D. A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction. *Proc. Natl Acad. Sci. USA* **96**, 1385–1390 (1999).
14. Giovannini, D., Touhami, J., Charnet, P., Sitbon, M. & Battini, J.-L. Inorganic phosphate export by the retrovirus receptor XPR1 in metazoans. *Cell Rep.* **3**, 1866–1873 (2013).
15. Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
16. Tsherniak, A. et al. Defining a cancer dependency map. *Cell* **170**, 564–576 (2017).
17. Dempster, J. M. et al. Gene expression has more power for predicting in vitro cancer cell vulnerabilities than genomics. Preprint at *bioRxiv* https://doi.org/10.1101/2020.02.21.959627 (2020).
18. Levan, K. et al. Immunohistochemical evaluation of epithelial ovarian carcinomas identifies three different expression patterns of the MX35 antigen, NaPi2b. *BMC Cancer* **17**, 303 (2017).
19. Yin, B. W. T. et al. Monoclonal antibody MX35 detects the membrane transporter NaPi2b (SLC34A2) in human carcinomas. *Cancer Immun.* **8**, 3 (2008).
20. Zou, Y. et al. A GPX4-dependent cancer cell state underlies the clear-cell morphology and confers sensitivity to ferroptosis. *Nat. Commun.* **10**, 1617 (2019).
21. Magtanong, L. et al. Exogenous monounsaturated fatty acids promote a ferroptosis-resistant cell state. *Cell Chem. Biol.* **26**, 420–432 (2019).
22. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
23. Cancer Genome Atlas Research Network et al. Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67–73 (2013).
24. Goldman, M. J. et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678 (2020).
25. Vivian, J. et al. Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* **35**, 314–316 (2017).
26. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
27. Köbel, M. et al. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. *PLoS Med.* **5**, e232 (2008).
28. Piek, J. M. et al. Dysplastic changes in prophylactically removed Fallopian tubes of women predisposed to developing ovarian cancer. *J. Pathol.* **195**, 451–456 (2001).
29. Jarboe, E. et al. Serous carcinogenesis in the fallopian tube: a descriptive classification. *Int. J. Gynecol. Pathol.* **27**, 1–9 (2008).
30. Hu, Z. et al. The repertoire of serous ovarian cancer non-genetic heterogeneity revealed by single-cell sequencing of normal fallopian tube epithelial cells. *Cancer Cell* **37**, 226–242 (2020).
31. Mittag, J., Winterhager, E., Bauer, K. & Grümmer, R. Congenital hypothyroid female pax8-deficient mice are infertile despite thyroid hormone replacement therapy. *Endocrinology* **148**, 719–725 (2007).
32. Elias, K. M. et al. Epigenetic remodeling regulates transcriptional changes between ovarian cancer and benign precursors. *JCI Insight* **1**, e87988 (2016).
33. Adler, E. K. et al. The PAX8 cistrome in epithelial ovarian cancer. *Oncotarget* **8**, 108316–108332 (2017).
34. Shi, K. et al. PAX8 regulon in human ovarian cancer links lineage dependency with epigenetic vulnerability to HDAC inhibitors. *eLife* **8**, e44306 (2019).
35. Mermel, C. H. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
36. DeWeirdt, P. C. et al. Genetic screens in isogenic mammalian cell lines without single cell cloning. *Nat. Commun.* **11**, 752 (2020).
37. Minisola, S. et al. Tumour-induced osteomalacia. *Nat. Rev. Dis. Primers* **3**, 17044 (2017).
38. Whitsett, J. A., Wert, S. E. & Weaver, T. E. Alveolar surfactant homeostasis and the pathogenesis of pulmonary disease. *Annu. Rev. Med.* **61**, 105–119 (2010).
39. Lituiev, D. S. & Kiyamova, R. G. Mutations in the gene of human type IIb sodium-phosphate cotransporter SLC34A2. *Zurich Open Repository and Archive* https://www.zora.uzh.ch/id/eprint/44008/1/5.pdf (2010).
40. Wilson, M. S., Jessen, H. J. & Saiardi, A. The inositol hexakisphosphate kinases IP6K1 and -2 regulate human cellular phosphate homeostasis, including XPR1-mediated phosphate export. *J. Biol. Chem.* **294**, 11597–11608 (2019).
41. Li, X. et al. Control of XPR1-dependent cellular phosphate efflux by InsP8 is an exemplar for functionally-exclusive inositol pyrophosphate signaling. *Proc. Natl Acad. Sci. USA* https://www.pnas.org/doi/10.1073/pnas.1908830117 (2020).
42. López-Sánchez, U. et al. Interplay between PFBC-associated SLC20A2 and XPR1 phosphate transporters requires inositol polyphosphates for control of cellular phosphate homeostasis. *J. Biol. Chem.* **295**, 9366–9378 (2020).
43. Legati, A. et al. Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. *Nat. Genet.* **47**, 579–581 (2015).
44. Wainberg, M. et al. A genome-wide almanac of co-essential modules assigns function to uncharacterized genes. *Nat. Genet.* **53**, 638–649 (2021).
45. Kim, E. et al. A network of human functional gene interactions from knockout fitness screens in cancer cells. *Life Sci. Alliance* **2**, e201800278 (2019).
46. Cabrera-Poch, N., Sánchez-Ruiloba, L., Rodríguez-Martínez, M. & Iglesias, T. Lipid raft disruption triggers protein kinase C and Src-dependent protein kinase D activation and Kidins220 phosphorylation in neuronal cells. *J. Biol. Chem.* **279**, 28592–28602 (2004).
47. Iglesias, T. et al. Identification and cloning of Kidins220, a novel neuronal substrate of protein kinase D. *J. Biol. Chem.* **275**, 40048–40056 (2000).
48. Arévalo, J. C., Yano, H., Teng, K. K. & Chao, M. V. A unique pathway for sustained neurotrophin signaling through an ankyrin-rich membrane-spanning protein. *EMBO J.* **23**, 2358–2368 (2004).
49. Kong, H., Boulter, J., Weber, J. L., Lai, C. & Chao, M. V. An evolutionarily conserved transmembrane protein that is a novel downstream target of neurotrophin and ephrin receptors. *J. Neurosci.* **21**, 176–185 (2001).
50. Wege, S. et al. The EXS domain of PHO1 participates in the response of shoots to phosphate deficiency via a root-to-shoot signal. *Plant Physiol.* **170**, 385–400 (2016).

51. Arpat, A. B. et al. Functional expression of PHO1 to the Golgi and trans-Golgi network and its role in export of inorganic phosphate: PHO1 localization to the Golgi/trans-Golgi network. *Plant J.* **71**, 479–491 (2012).
52. Snyder, N. A. et al. H^+ and Pi byproducts of glycosylation affect Ca^{2+} homeostasis and are retrieved from the Golgi complex by homologs of TMEM165 and XPR1. *G3 (Bethesda)* **7**, 3913–3924 (2017).
53. Gerber, D. E. et al. Phase Ia study of anti-NaPi2b antibody-drug conjugate lifastuzumab vedotin DNIB0600A in patients with non-small cell lung cancer and platinum-resistant ovarian cancer. *Clin. Cancer Res.* **26**, 364–372 (2020).
54. Marks, J. The role of SLC34A2 in intestinal phosphate absorption and phosphate homeostasis. *Pflugers Arch.* **471**, 165–173 (2019).
55. He, P. et al. High phosphate actively induces cytotoxicity by rewiring pro-survival and pro-apoptotic signaling networks in HEK293 and HeLa cells. *FASEB J.* **35**, e20997 (*2021*).
56. Di Marco, G. S. et al. Increased inorganic phosphate induces human endothelial cell apoptosis in vitro. *Am. J. Physiol. Renal Physiol.* **294**, F1381–F1387 (2008).
57. Marks, M. S., Heijnen, H. F. G. & Raposo, G. Lysosome-related organelles: unusual compartments become mainstream. *Curr. Opin. Cell Biol.* **25**, 495–505 (2013).
58. Ruiz, F. A., Lea, C. R., Oldfield, E. & Docampo, R. Human platelet dense granules contain polyphosphate and are similar to acidocalcisomes of bacteria and unicellular eukaryotes. *J. Biol. Chem.* **279**, 44250–44257 (2004).
59. Josifova, D. J. et al. Heterozygous KIDINS220/ARMS nonsense variants cause spastic paraplegia, intellectual disability, nystagmus, and obesity. *Hum. Mol. Genet.* **25**, 2158–2167 (2016).
60. McDonald, E. R. III et al. Project DRIVE: a compendium of cancer dependencies and synthetic lethal relationships uncovered by large-scale, deep RNAi screening. *Cell* **170**, 577–592 (2017).
61. Yang, X. et al. A public genome-scale lentiviral expression library of human ORFs. *Nat. Methods* **8**, 659–661 (2011).
62. McFarland, J. M. et al. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat. Commun.* **9**, 4610 (2018).
63. Buehler, E., Chen, Y.-C. & Martin, S. C911: a bench-level control for sequence specific siRNA off-target effects. *PLoS ONE* **7**, e51942 (2012).
64. Aran, D., Sirota, M. & Butte, A. J. Systematic pan-cancer analysis of tumour purity. *Nat. Commun.* **6**, 8971 (2015).
65. Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
66. Gao, J. et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).
67. Li, W. et al. Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* **16**, 281 (2015).
68. Williams, S. N. et al. Phosphorus concentrations in blood, milk, feces, bone and selected fluids and tissues of growing heifers as affected by dietary phosphorus. *Livest. Res. Rural. Dev.* http://www.lrrd.org/lrrd3/2/florida4.htm (1991).
69. McFarland, J. M. et al. Multiplexed single-cell profiling of post-perturbation transcriptional responses to define cancer vulnerabilities and therapeutic mechanism of action. *Nat. Commun.* **11**, 4296 (2020).
70. Stoeckius, M. et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
71. Li, B. et al. Cumulus provides cloud-based data analysis for large-scale single-cell and single-nucleus RNA-seq. *Nat. Methods.* **17**, 793–798 (2020).
72. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at https://arxiv.org/abs/1207.3907 (2012).
73. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
74. Ritchie, M. E. et al. limma Powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
75. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
76. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
77. Huttlin, E. L. et al. Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* **184**, 3022–3040 (2021).
78. Oughtred, R. et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, D529–D541 (2019).
79. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858 (1996).
80. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass. Spectrom.* **5**, 976–989 (1994).

Acknowledgements

This work was funded in part by the Slim Initiative in Genomic Medicine for the Americas, a joint United States–Mexico project funded by the Carlos Slim Foundation (to T.R.G.), and grants from the National Cancer Institute (nos. CA242457 to T.R.G. and CA212229 to D.P.B.). We thank the Dana-Farber/Harvard Cancer Center in Boston, MA, for the use of the Specialized Histopathology Core, which provided histology and immunohistochemistry services and is supported in part by a NCI Cancer Center Support Grant (no. CA06516). We thank J. Barnett, B. Buckley, M. Veneskey, M. Cecilia-Saberi and D. Tennen for technical support.

Author contributions

D.P.B., B.R.P., F.V. and T.R.G. initiated the project and oversaw the research plan. J.K.B, D.P.B. and B.R.P. analyzed genetic dependency data under the supervision of W.C.H., D.E.R., J.B., F.V. and T.R.G. and with support from I.F. and E.S.C. M.V.R., A.A., T.A.S., B.R.P. and D.P.B. conducted viability experiments and immunoblotting. A.A., B.P. and D.P.B. conducted the genome-scale modifier screen with analysis support from M.K., J.M.D. and J.M.M. and supervision from J.M.D. In vivo experiments were conducted by A.G. and N.D. under the supervision of F.P. Intracellular phosphate assays were performed by M.V.R. and D.P.B. PAX8 RNA-seq experiments were conducted by K.I. with analytical support from W.N.C. D.P.B. analyzed GTEx, TCGA and CCLE expression datasets with supervision from J.M.M. Multiplexed transcriptional profiling was conducted by B.R.P. and D.P.B. and was analyzed by A.W. and W.N.C. Phosphate uptake and efflux assays were conducted by D.P.B. and L.E.S. with supervision from M.M. A.A. and D.P.B. conducted coimmunoprecipitation experiments. B.R.P. conducted cellular imaging studies. D.P.B. and M.E. conducted ultrastructural analysis. D.P.B., B.R.P., F.V. and T.R.G. wrote the manuscript, and all authors edited and approved the manuscript.

Competing interests

F.V. receives research funding from Novo Holdings. D.E.R. receives research funding from the Functional Genomics Consortium (Abbvie, Jannsen, Merck and Vir) and is a director of Addgene. T.R.G. previously received cash and/or equity compensation for consulting to GlaxoSmithKline, Sherlock Biosciences and FORMA Therapeutics, is currently a paid consultant for Anji Pharmaceuticals and receives research funding from Bayer HealthCare, Calico Life Sciences and Novo Holdings. W.C.H. is a consultant for ThermoFisher, Solasta, MPM Capital, iTeos, Frontier Medicines and Paraxel and is a Scientific Founder and serves on the Scientific Advisory Board for KSQ Therapeutics. All other authors declare no competing interests.

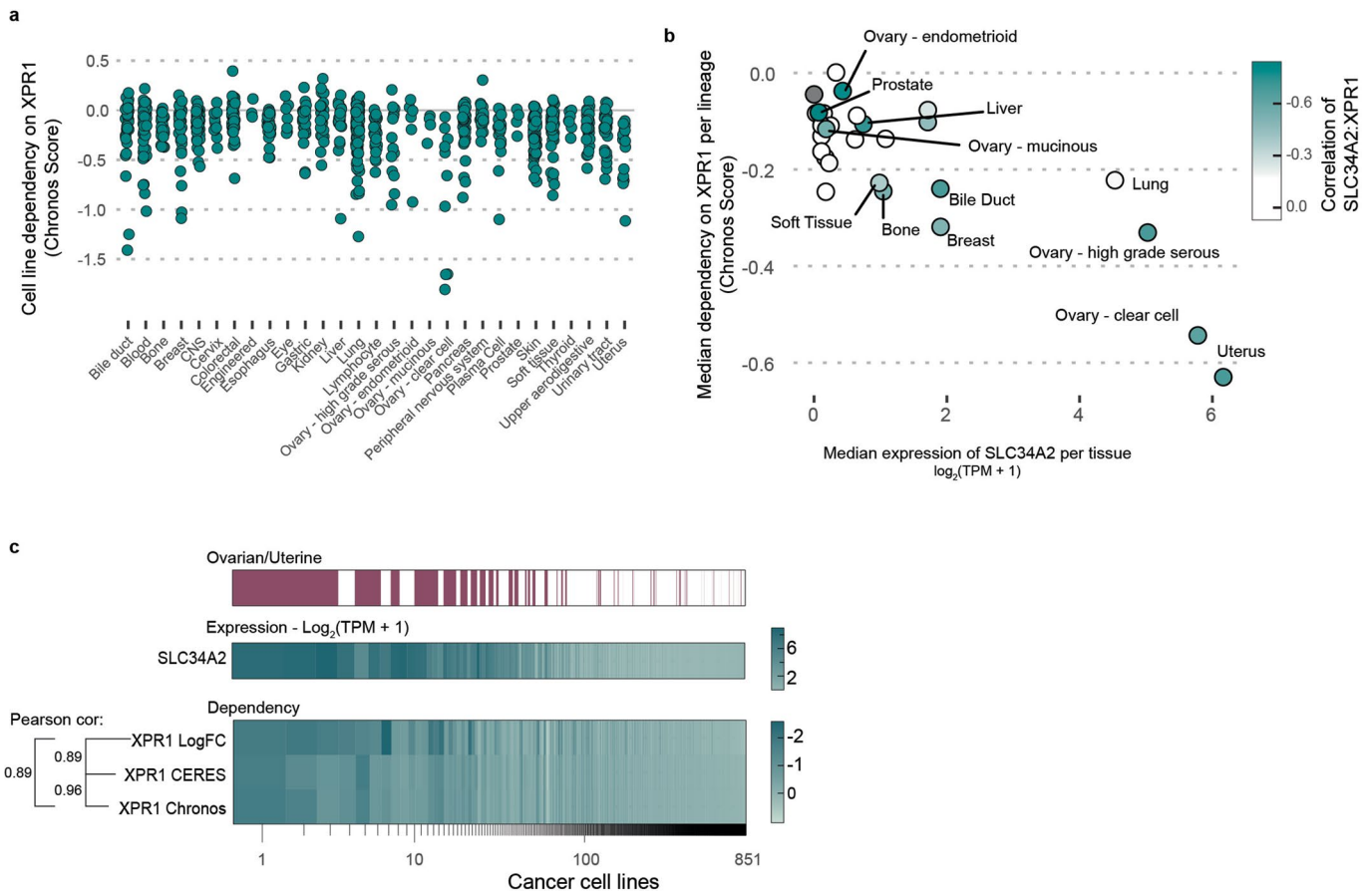
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s43018-022-00360-7.

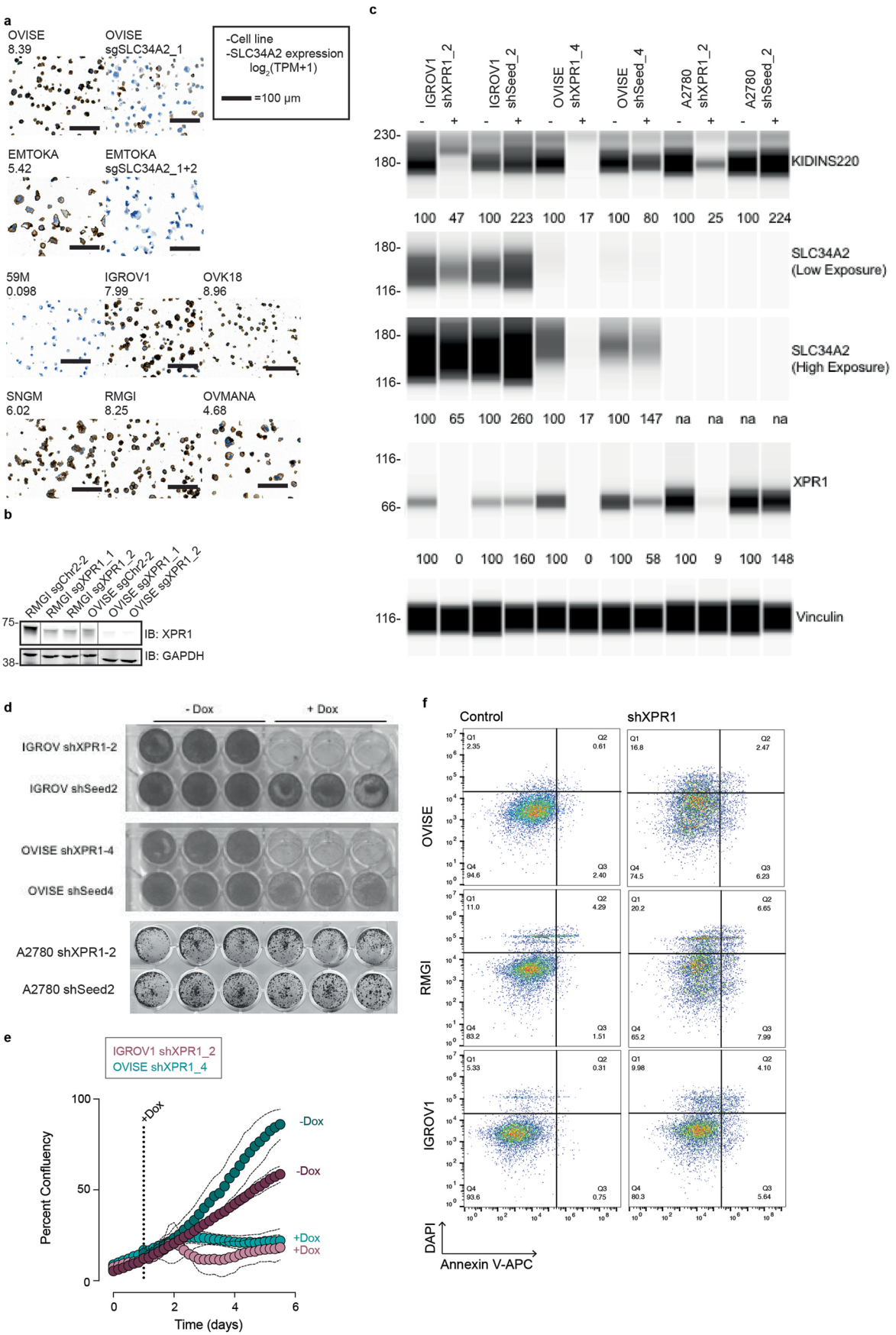
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43018-022-00360-7.

Correspondence and requests for materials should be addressed to Todd R. Golub.

Peer review information *Nature Cancer* thanks Ronny Drapkin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

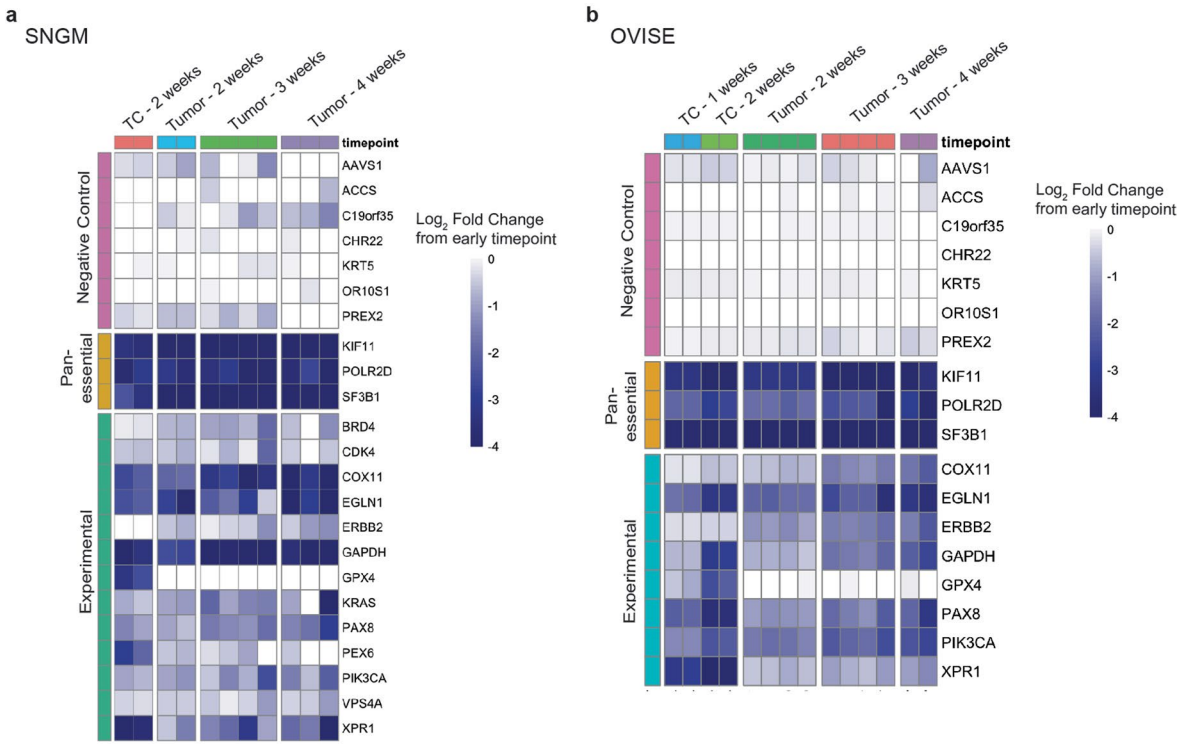


Extended Data Fig. 1 | XPR1 dependency is observed selectively in SLC34A2-high cancer cell lines. a) For every cell line profiled in the Cancer Dependency Map dataset (N = 851 cancer cell lines), the degree of XPR1 essentiality is plotted on the Y-axis. The Chronos score is a scaled value of gene essentiality, where 0 is the effect of CRISPR/Cas9 genome editing and -1 is the effect of inactivation of pan-essential genes. Note that the ovarian lineage is separated into cancer subtypes. **b**) For every tissue type, the 10 highest SLC34A2 expressing cell lines were analyzed for their median expression of SLC34A2 (X-axis) and dependency on XPR1 (Y-axis). Note that some lineages may have less than 10 cell lines. Color encodes the correlation of SLC34A2 expression and XPR1 dependency across all cell lines within that lineage. **c**) Comparison of analytical methods for CRISPR/Cas9 genome-scale loss of function screens.

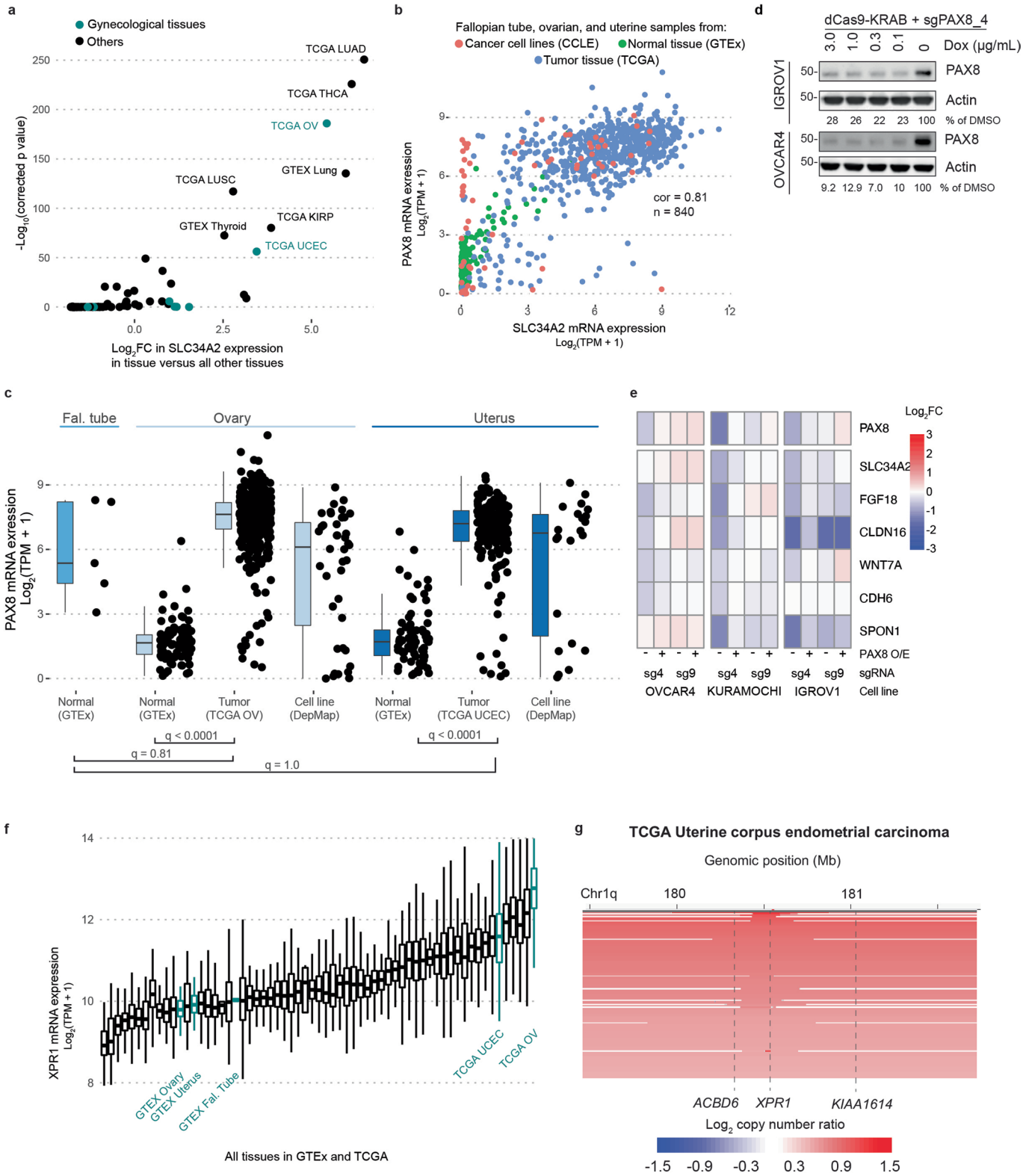


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Validation of SLC34A2 and XPR1 protein levels and viability defects upon shRNA induction. a) Validation of SLC34A2 in cell lines using immunohistochemistry. N = 1 experiment. **b**) Five days after viral transduction of the indicated sgRNA in the indicated cell lines stably expressing Cas9, cells were harvested and XPR1 levels were analyzed by immunoblotting. Note that irrelevant lanes were cropped out for clarity, but that this image represents a single blot at a single exposure. N = 1 technical replicate of at least N = 5 representative experiments. **c**) Three days after induction of shRNA, protein levels were measured in cellular lysates by protein simple. Protein levels normalized to vinculin and the untreated (-Dox) conditions are expressed below each band. Note that shXPR1 reagents effectively suppress XPR1 protein levels but shSeed reagents do not. N = 1 technical replicate of at least N = 5 representative experiments. **d**) Colony formation assay to measure the long-term (14 day) penetrance and viability effect of suppression of XPR1 using shRNA in IGROV1 and OVISE cells. N = 3 technical replicates of at least N = 2 representative experiments. **e**) Growth curves of SLC34A2-expressing cell lines after suppression of *XPR1*. In 96-well plates, confluency of the indicated cell lines was assessed every 4 hours. N = 3 technical replicates of at least N = 2 representative experiments. **f**) Six days after induction of shXPR1 in the indicated cell lines, cells were stained with DAPI to distinguish live and dead cells and Annexin V to distinguish non- and pre-apoptotic cells. N = 2 flow cytometric analyses of at least 10,000 cells, representative of N = 2 experiments.

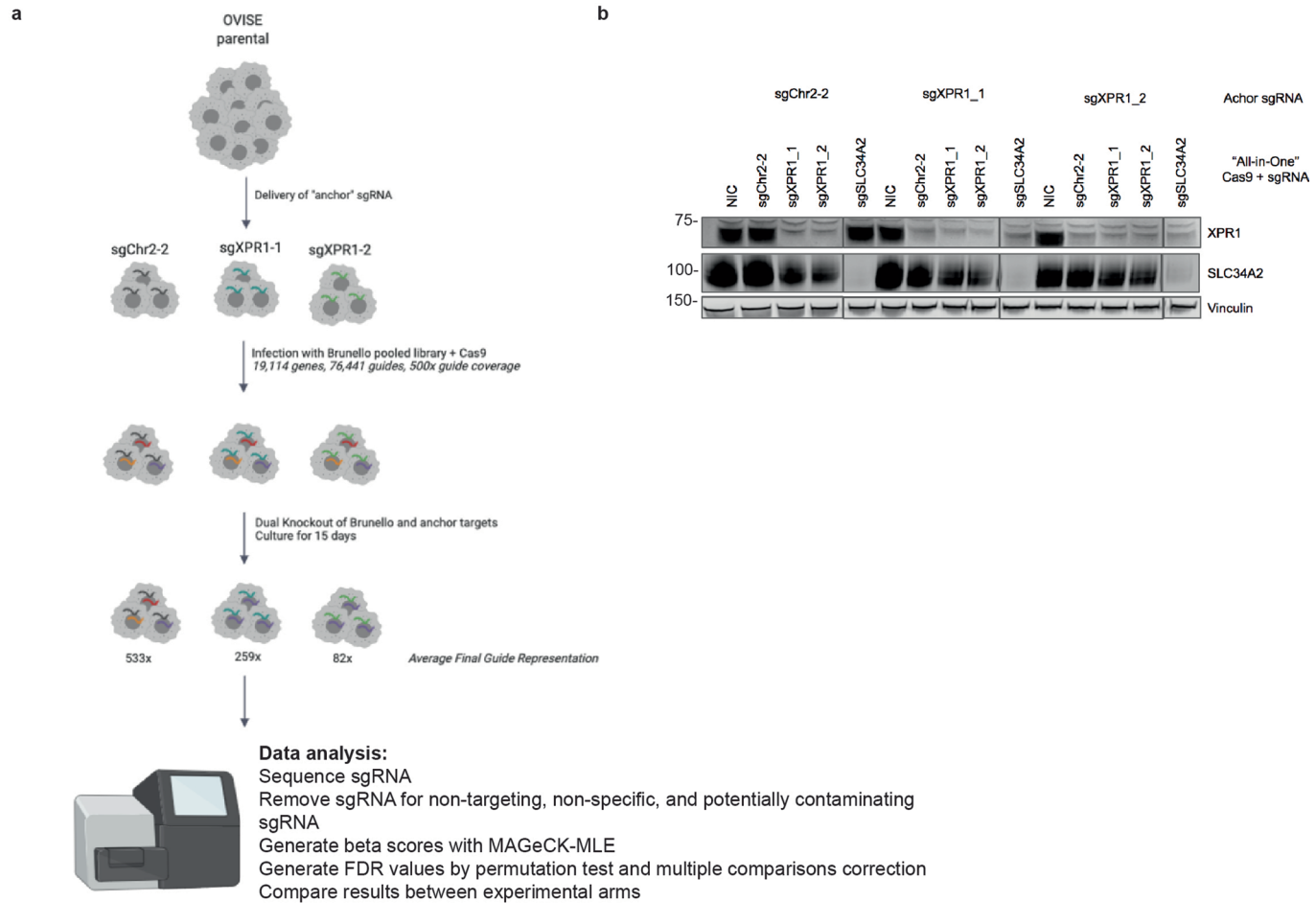


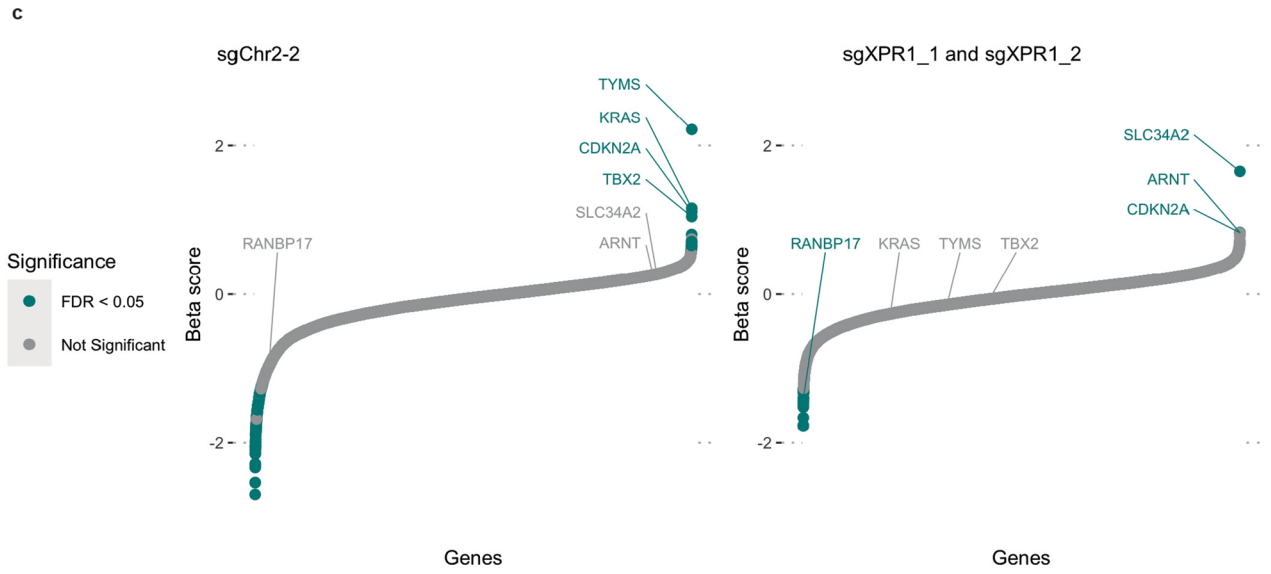
Extended Data Fig. 3 | *In vivo* CRISPR/Cas9 competition assays for target validation in mouse xenografts. a) sgRNA abundance in SNGM tumor xenografts was evaluated by PCR and next-generation sequencing analysis, and the fold change compared to the early time point is shown as a heatmap for all of the negative control genes as well as any gene with a > 4 fold change in abundance in any of the screens. Each tumor/replicate is shown as an individual column, N = 1 transduction. **b**) Same as in **d**, but with the OVISE cancer cell line.



Extended Data Fig. 4 | See next page for caption.

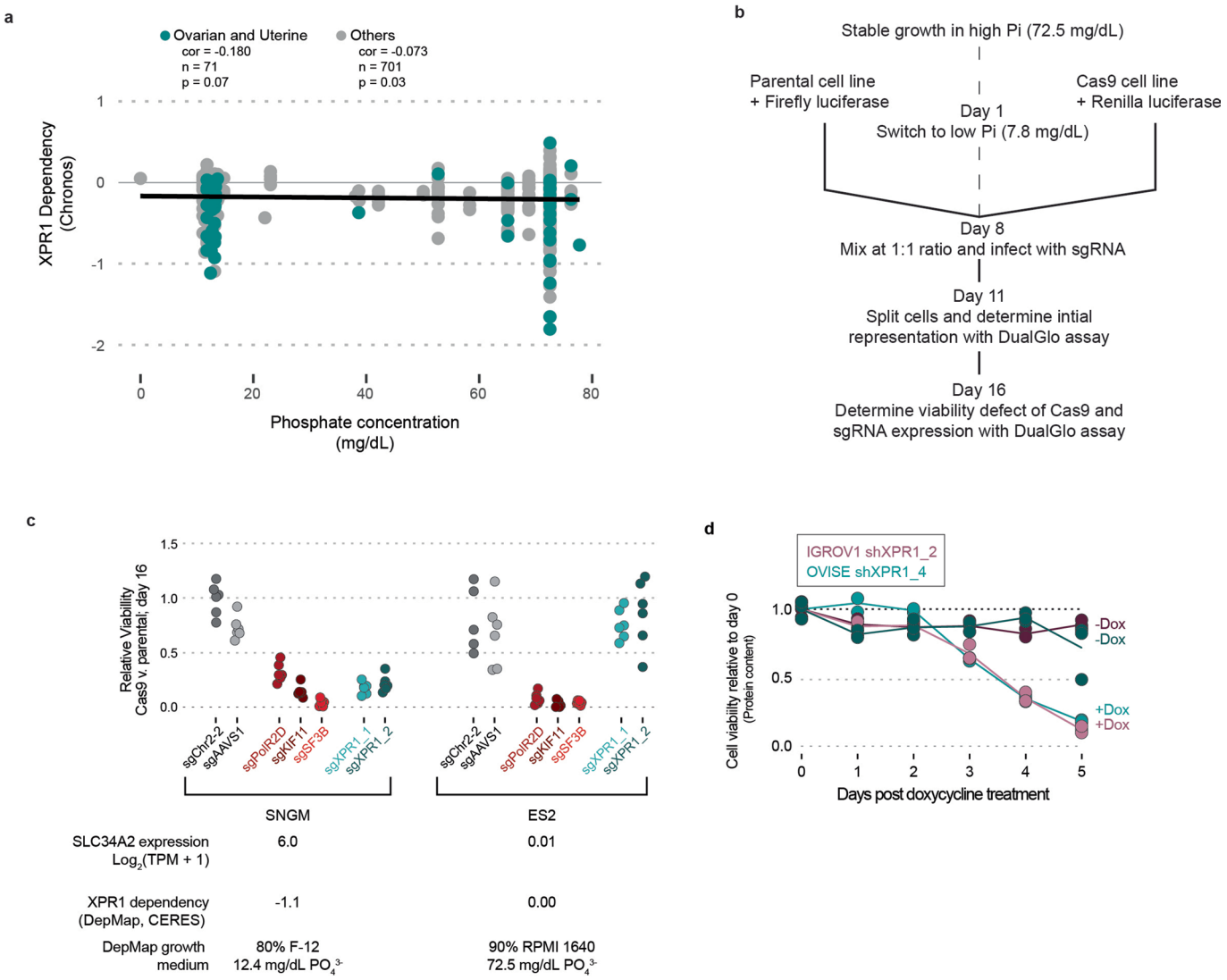
Extended Data Fig. 4 | *SLC34A2* in ovarian cancer is likely driven by PAX8. a) Using the combined GTEx, TCGA, and CCLE dataset, the differential expression of *SLC34A2* in each tissue relative to the average of all tissues is compared. The relevant gynecological tissues (fallopian tube, ovary, and uterus) are highlighted in teal. The false discovery rate (FDR) was calculated using a two-sided Wilcoxon ranked sum test comparing each group to the average expression across all groups and correcting for multiple comparisons using Bonferoni's method. The Cancer Genome Atlas abbreviations used include: LUAD = Lung adenocarcinoma; THCA = Thyroid carcinoma; KRP = Kidney renal cell papillary carcinoma; LUSC = Lung squamous cell carcinoma; OV = Ovarian serous cystadenocarcinoma; UCEC = Uterine corpus endometrial carcinoma. **b**) The expression of *PAX8* and *SLC34A2* mRNA in the indicated tissues is plotted. The pearson correlation within these samples is indicated. **c**) Expression of *PAX8* across the indicated tissues was compared as in Fig. 3a. See methods for exact N values. Boxplots are drawn indicating the first and third quartiles, and whiskers span to the largest value within 1.5x the interquartile range. **d**) Immunoblot validation of CRISPR-interference mediated suppression of PAX8. N = 1 technical replicate, representative of N = 2 independent experiments. **e**) Gene expression - relative to un-perturbed, parental cell lines profiled in parallel - of reported PAX8 target genes (see main text) after stable overexpression of WT-PAX8 ('PAX8 O/E') and/or induction of PAX8-target (sg4) or control (sg9) sgRNA and dCas9-KRAB. Data represents a single experiment. N = 1 replicate. **f**) XPR1 expression across all tissues in TCGA and GTEx, with ovarian and uterine tissues highlighted in teal. Boxplots are drawn as in **b. g**) *XPR1* copy number heatmap for a ~2.5 Mb region of chromosome 1 indicating XPR1 amplification in TCGA Uterine Corpus Endometrial Carcinoma[20]. Each patient sample is represented by a horizontal line. Red indicates copy gain and blue indicates copy loss. Data are a subset of the samples rank-ordered by highest copy gain to indicate both focal and chromosome arm-level gains.



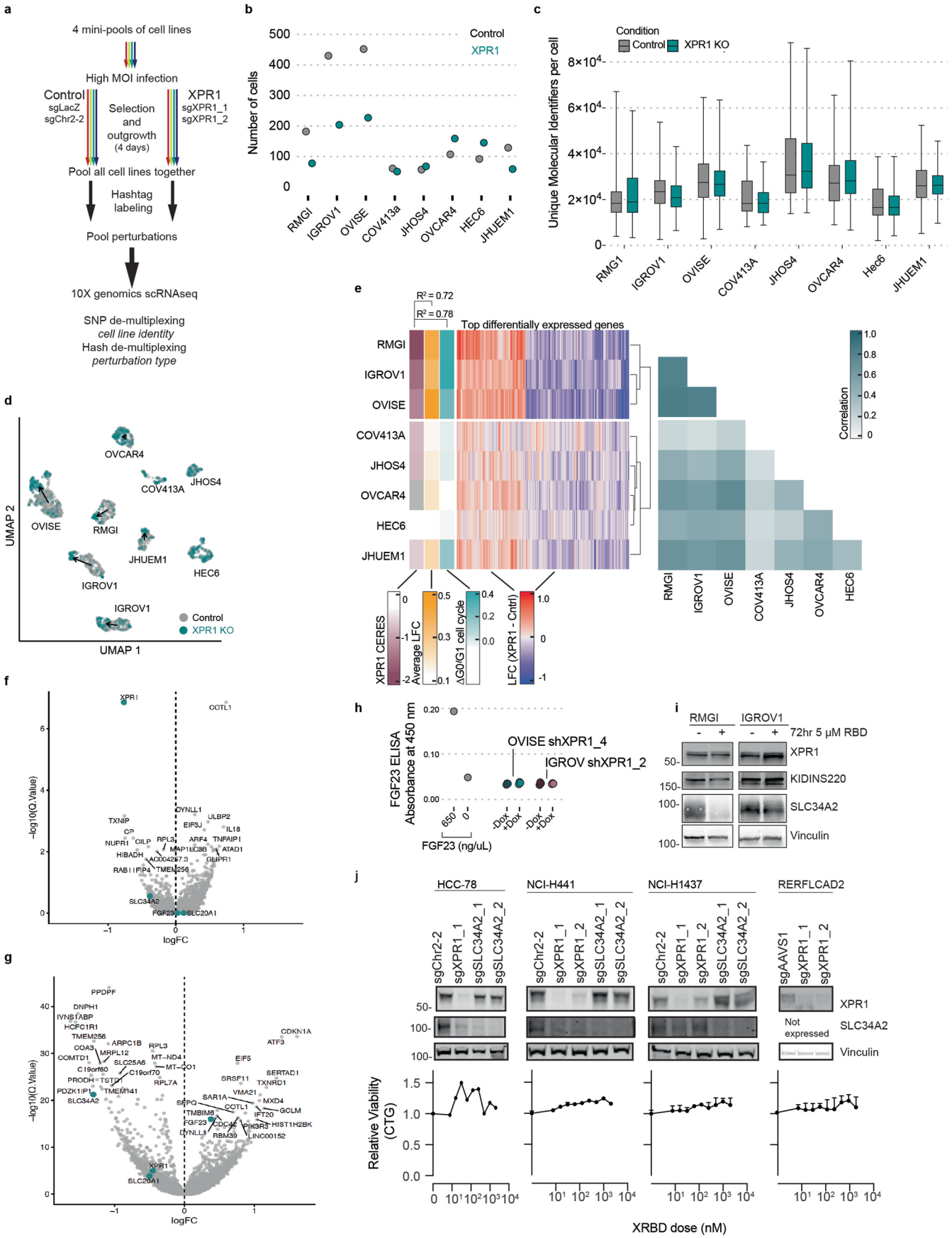


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | A genome-scale CRISPR/Cas9 screen validates the relationship between XPR1 dependency in the context of high expression of SLC34A2. a) Outline of the experimental method and analysis for a genome-scale, dual-knock-out modifier screen. OVISE (without Cas9 expression) is engineered to stably express sgRNA targeting XPR1 (the 'anchor' sgRNA). Upon introduction of 'all-in-one' lentivirus, containing both Cas9 ORF and a second sgRNA, both genes are simultaneously inactivated by Cas9. We used three achor sgRNA: one targeting a gene desert on chromosome 2 (sgChr2-2) and two targeting XPR1 (sgXPR1_1 and sgXPR1_2) and infecting the cells with the Brunello genome-scale sgRNA library. 15 days after infection, cells were harvested, genomic DNA was isolated, and sgRNA barcodes were quantified with next generation sequencing. See methods for full experimental and analytical details. **b**) Western confirmation of dual-knock-out of XPR1 and SLC34A2. The three cell lines used in the genome-scale screen were infected with 'all-in-one' lentivirus expressing control-, XPR1-, or SLC34A2-targeting sgRNA. Note that in the sgXPR1 'anchor' cell lines, XPR1 is suppressed with the control virus, indicating that the first infection provides XPR1-targeting sgRNA and the second infection provides Cas9 protein. NIC stands for 'no-infection control'. N = 1 technical replicate representative of N = 3 independent transductions. **c**) Arm-level results of the genome-scale modifier screen. See methods for full analysis details. Beta-scores represent the extent to which a gene was enriched or depleted relative to the initial plasmid representation. An XPR1-positive and control-neutral score represents a likely rescue gene (that is SLC34A2 and ARNT). XPR1-positive and control-positive scores represent genes with profound viability defects without specificity for XPR1 (for example RANBP17). N = 1 transduction per anchor condition, expanded and cultured as N = 2 independent replicates.

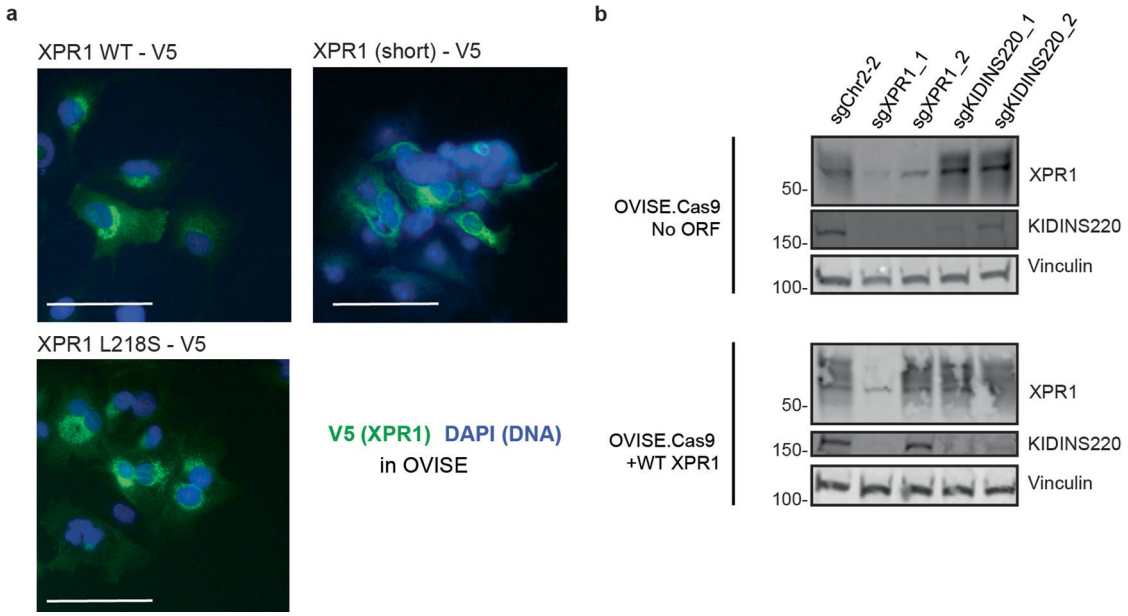


Extended Data Fig. 6 | The XPR1 dependency is not affected by phosphate levels in the tissue culture medium. a) The concentration of phosphate in the growth medium of DepMap cell lines does not determine the extent of XPR1 dependency. Concentrations of phosphate were estimated from manufacturer formulations (see methods) and the pearson correlation between growth medium phosphate and XPR1 dependency is indicated. **b**) Experimental procedure for manipulating tissue culture medium and assessing its effect on XPR1 dependency. The same parental cancer cell line was engineered to express firefly luciferase and Cas9, or Renilla luciferase alone. After a one-week adaptation to lowered phosphate, the two variants were mixed together and infected with sgRNA-encoding lentivirus. After selection for lentivirus-infected cell lines, the initial representation of Cas9:parental cells was determined by measuring the ratio of Firefly:Renilla luciferase using a DualGlo assay (Promega). One week after infection (Day 16 of the protocol), the extent to which genetic perturbation was detrimental to cell viability was determined using the DualGlo assay.**c**) The XPR1 dependency is maintained in a low phosphate medium condition. SNGM and ES2 were profiled in the assay outlined in panel **b**. Note that the CERES score - displayed below the plot - represents the viability defect of the cell line 21 days after knock-out of XPR1 and growth in the indicated growth medium. N = 5 technical replicates representative of N = 2 experiments. **d**) The viability of cells (as measured by total protein content) was measured in parallel with total phosphate as in Fig. 4d. N = 3 technical replicates representative of N = 4 independent experiments per cell line.

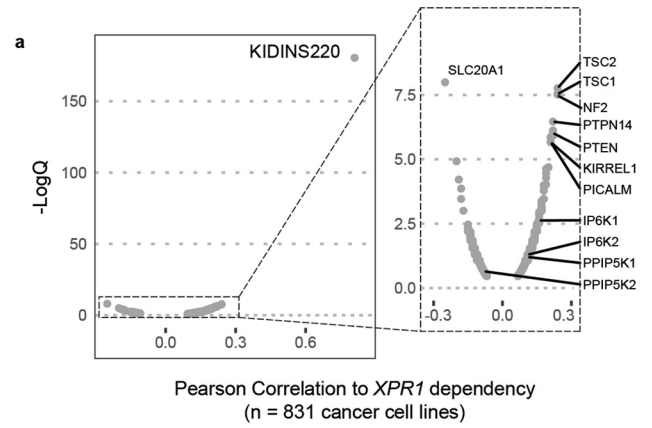


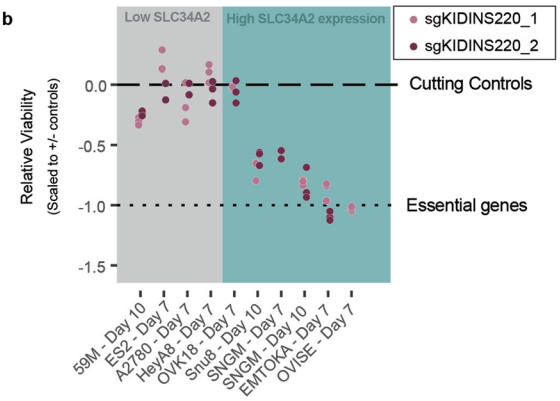
Extended Data Fig. 7 | See next page for caption.

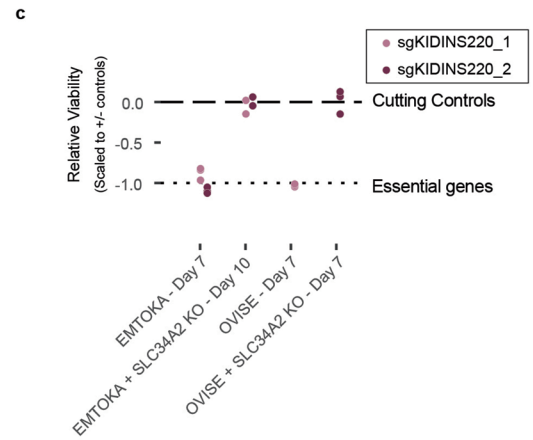
Extended Data Fig. 7 | Transcriptional profiling reveals a phosphate-related homeostatic response after XPR1 inactivation. a) Experimental workflow to determine the transcriptional profile of XPR1 inactivation across many different cancer cell lines. See methods for full details; N = 1 transduction event for panels **b-g**. **b**) The total number of cells per cell line de-multiplexed from the 10X scRNAseq library. **c**) The total number of unique transcripts measured for each cell, as measured by unique molecular identifiers (UMIs). Box plots represent the 1st – 3rd quartiles, with whiskers representing the minimum and the maximum. The exact N for each sample in **c** is indicated in panel **b. d**) UMAP projection of the 2,501 cells from the indicated cell lines (determined by SNP profiles) and perturbations (indicated by cell-surface antibody 'hash-tag' labeling). Arrows indicate the degree to which the average transcriptional profile changes between the control sgRNA and the sgXPR1 infection condition. **e**) Summary of transcriptional effects across cell lines after inactivation of *XPR1*. Middle, the log-fold change of the top 500 differentially expressed genes after regressing out the effect of cell cycle. Left, summary annotations for each cell line include XPR1 dependency (XPR1 CERES), the overall transcriptional change (average log2 fold-change), and the degree of cell cycle arrest observed in the single-cell data (ΔG0/G1). The pearson correlation of these values is shown above the heatmap. Right, the pearson correlation of the top 500 differentially expressed genes between each cell line. **f**) Differentially expressed genes - after correcting for cell cycle - in the less sensitive cell lines (COV413a, JHOS4, OVCAR4, HEC6, and JHUEM1). Significance was assessed by the limma-voom pipeline using a two-tailed statistical test (see methods). **g**) Same as in **f**, but for the highly correlated cell lines RMG1, IGROV1, and OVISE. **h**) Four days after induction of shXPR1_2 (IGROV1) or shXPR1_4 (OVISE) using doxycycline, the amount of secreted FGF23 was measured in the conditioned medium using ELISA. N = 2 technical replicates representative of N = 3 independent experiments. **i**) 72 hours after treatment with the XPR1 inhibitor XRBD, the indicated proteins were detected using immunoblot. N = 1 technical replicate representative of N = 2 independent experiments. **j**) Top, western blot analysis of SLC34A2 and XPR1 in the SLC34A2-high yet XPR1-insensitive lung cancer cell lines, five days after infection with lentivirus expressing the indicated sgRNA. Bottom, viability of the indicated cancer cell lines was assessed using Cell Titer Glo after a five day XRBD treatment to inhibit XPR1. Points represent the mean of N = 3 technical replicates; error bars represent standard error of the mean. Data are representative of N = 2 independent experiments.

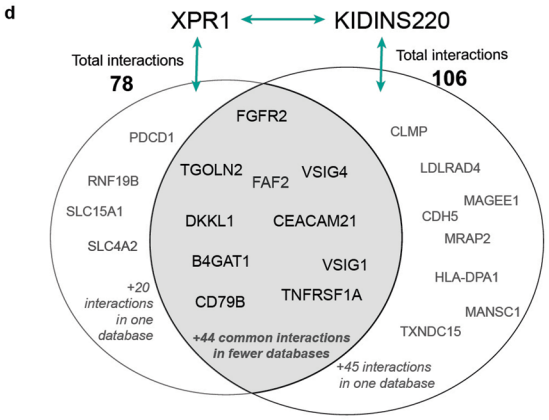


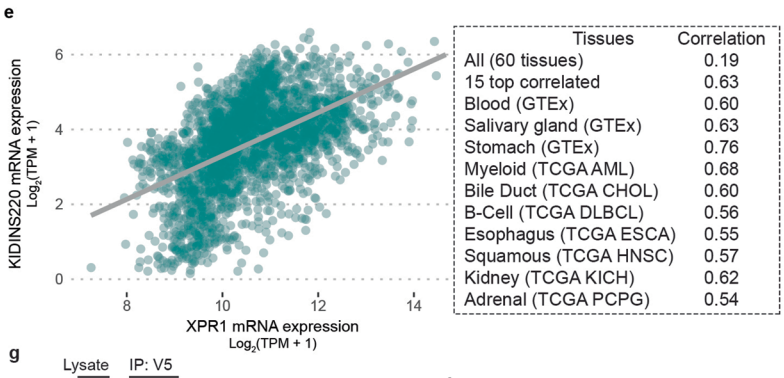
Extended Data Fig. 8 | Open-reading frames of *XPR1* resistant to CRISPR/Cas9-mediated gene editing. a) Immunofluorescent localization of XPR1 mutants using the V5 epitope tag. Left, WT XPR1 localizes to the secretory pathways as well as puncta within the cytoplasm. Middle, XPR1 (short) staining appears more diffuse. Note similar localization patterns between L218S and wildtype XPR1 alleles. Scale bar = 200 μm. N = 1 experiment representative of N = 2 independent transductions. **b**) Western blot validation of guide-resistant ORF. OVISE.Cas9 cells (parental, left, or overexpressing the WT XPR1 ORF, right, used in Fig. 3e) were infected with the indicated sgRNA and harvested 5 days after infection. The XPR1 ORF includes a mutation to block binding of sgXPR1_2 but not sgXPR1_1. Note the inactivation of both endogenous and overexpression ORF with sgXPR1_1 and only endogenous XPR1 with sgXPR1_2. N = 1 experiment representative of N = 2 independent transductions.



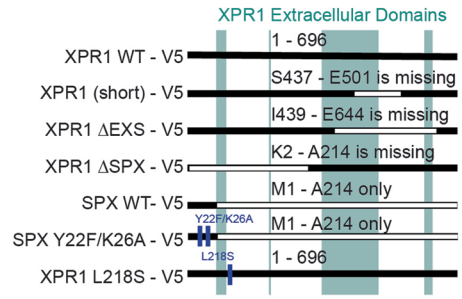


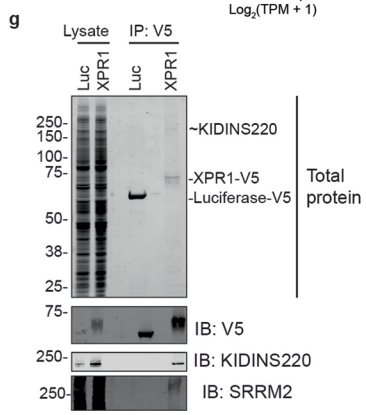


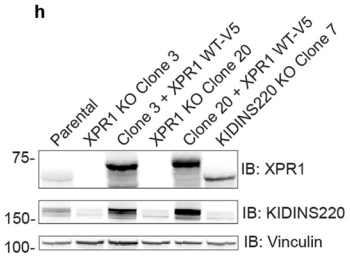


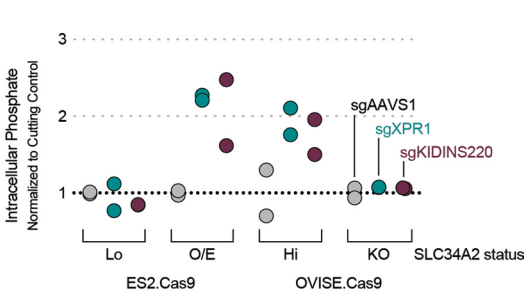


Tissues	Correlation
All (60 tissues)	0.19
15 top correlated	0.63
Blood (GTEx)	0.60
Salivary gland (GTEx)	0.63
Stomach (GTEx)	0.76
Myeloid (TCGA AML)	0.68
Bile Duct (TCGA CHOL)	0.60
B-Cell (TCGA DLBCL)	0.56
Esophagus (TCGA ESCA)	0.55
Squamous (TCGA HNSC)	0.57
Kidney (TCGA KICH)	0.62
Adrenal (TCGA PCPG)	0.54









Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | KIDINS220 is a unique partner protein of the XPR1 phosphate efflux complex. a) Genetic dependency correlations to XPR1 dependency across 851 cancer cell lines was assessed by pearson correlation test and corrected for multiple comparisons using the Benjamini-Hochberg method. Genes with significantly correlated dependency profiles are highlighted, as are proteins with known connection to XPR1 regulation. **b**) The viability defect of the indicated cancer cell lines after *KIDINS220* inactivation was evaluated as in Fig. 1c. N = 3 technical replicates representative of at least N = 2 independent transductions per cell line. **c**) *SLC34A2* was inactivated in EMTOKA and OVISE, and the KIDINS220 dependency was evaluated as in **b**. N = 3 technical replicates representative of at least N = 2 independent transductions per cell line.**d**) The interacting partners of XPR1 and KIDINS220 were downloaded from the BioGrid and Bioplex databases and compared. Proteins present in the XPR1 or KIDINS220 interactomes are highlighted as text.**e**) Left, the mRNA expression of XPR1 and KIDINS220 is shown for the fifteen tissues with the highest correlation in expression. The line represents linear regression for these samples (N = 2,799). Right, the Pearson correlation for those tissues, highlighting the diverse tissues in which there is a high correlation between XPR1 and KIDINS220 expression. **f**) Mutants of XPR1 used in this manuscript. XPR1 WT refers to the 696 amino acid protein produced by NM_004736 (the only isoform detected by RT-PCR of OVISE mRNA), while XPR1 (short) refers to the 631 amino acid product of NM_001135669. All constructs have C-terminal V5 tags for immunoprecipitation, western blotting, and immunofluorescent detection. **g**) XPR1 or Luciferase (Luc) were overexpressed in 293 T cells and immunoprecipitated using the V5 tag and analyzed by targeted immunoblot or for total protein. Proteins were extracted from this gel and identified using mass spectrometry, the results of which are shown in Fig. 6c. N = 1 replicate of N = 2 independent transfections. **h**) Cas9 + sgRNA targeting *XPR1* or *KIDINS220* were transfected into 293 T cells, and clones were isolated that lacked expression of the target proteins. For *XPR1* inactivated cells, the *XPR1* ORF was re-expressed, and the relative levels of the indicated proteins were detected by immunoblot. At least N = 2 clonal populations were profiled for each inactivation condition. **i**) Five days after infection with the indicated sgRNA targeting XPR1 or KIDINS220, free inorganic intracellular phosphate was determined as in Fig. 4d. N = 3 technical replicates of N = 3 independent transductions.

nature research

Corresponding author(s): Todd R Golub
Francisca Vazquez

Last updated by author(s): 1 February 2022

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	As described in the Methods section, the following software was used for data analysis: Graphpad Prism v. 8.1, the publicly available PoolQ software (v. 3.2.1) from the Broad Institute Genetic Perturbation Platform, the program MAGeCK-MLE (v. 0.5.9), the program FlowJo (v. 10.8.1), the 10x Cell Ranger software (v. 3.0.2), Freebayes (v. 0.2.3), the R packages RStatix (v. 0.7.0), ggplot2 (v. 3.3.3), Seurat (v. 4.0.5), limma-voom (v. 3.44.3), and EdgeR (3.30.3). Code to reproduce figures is available upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Publicly available data used in this study include CRISPR/Cas9 loss of viability screens for 851 cancer cell lines (https://doi.org/10.6084/m9.figshare.14541774.v2) cancer cell line RNAseq expression data (https://doi.org/10.6084/m9.figshare.14541774.v2), harmonized gene expression data for GTEx and TCGA datasets (https://doi.org/10.1101/326470), and copy number alterations for Ovarian adenocarcinoma and Uterine Corpus Endometrial Carcinoma (https://doi.org/10.1038/

nature12113 and https://doi.org/10.1038/nature10166). RNAseq data after PAX8 knockdown are available at https://doi.org/10.6084/m9.figshare.19125671.v1. MixSeq data after XPR1 knockdown are available at https://doi.org/10.6084/m9.figshare.19125677.v1.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the analysis of cancer cell line data, 859 cancer cell lines have robust gene expression and gene dependency data, and so all lines included within the analysis. In some cases, individual genes were not profiled (e.g. SLC34A2 expression and XPR1 dependency in 851 cell lines) and the adjustments in sample size are noted throughout the text. Where appropriate, all samples from the GTEx and TCGA projects were analyzed. Where appropriate, technical replicates (N=2 or 3) were included. For in vivo experiments, no statistical test was used to pre-determine sample sizes.
Data exclusions	No data was excluded from the analysis.
Replication	All experiments are representative of at least two successful technical replicates.
Randomization	Randomization was not relevant to this study. For the in vivo competition assays (Figure 2), animals with the highest tumor burden were euthanized at a given timepoint; thus, no randomization was appropriate.
Blinding	Blinding was not relevant to this study, since the biases of the researcher were unlikely to influence the results of the reported experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

All antibodies used in this study include:
XPR1
ProteinTech
14174-1-AP
Immunblot dilution of 1:2000

XPR1
Atlas Antibodies
HPA016557
Immunblot dilution of 1:2000
Protein Simple dilution of 1:37

KIDINS220
ProteinTech
21856-1-AP
Protein Simple dilution of 1:25
Immunblot dilution of 1:2000

SLC34A2
CST
#66445
Immunblot dilution of 1:2000
Protein Simple dilution of 1:50

SLC34A2 (MX35)
Creative Biolabs
TAB-467MZ
Immunohistochemistry dilution of 1:200

Vinculin
Sigma
#V9131
Immunblot dilution of 1:10000

Vinculin
Abcam
ab129002
Protein Simple dilution of 1:75

V5
CST
13202
Immunblot dilution of 1:2000

GM130
CST
12480
Immunofluorescence dilution of 1:800

LAMP1
CST
9091
Immunofluorescence dilution of 1:200

EEA1
CST
3288
Immunofluorescence dilution of 1:100

SSRRM2
Abcam
ab122719
Immunoblot dilution of 1:500

Anti-Rabbit HRP
ProteinSimple
DM-001
ProteinSimple dilution of 1:1

Anti-Rabbit Igg
LiCor
926-32211
Immunoblot dilution of 1:20,000

Anti-Mouse IgG
LiCor
926-68020
Immunoblot dilution of 1:20,000

Anti-Mouse 488
ThermoFisher
A-11004
Flow Cytometry dilution of 1:250

Validation

Each antibody (XPR1, KIDINS220, SLC34A2) was validated by detection of signal at the appropriate molecular weight with matching signal upon knockout and/or overexpression. The SRRM2 antibody only recognized samples in which an orthogonal method confirmed the precense of the protein (i.e. Mass Spectrometry). V5 antibody showed no signal in samples lacking V5-tagged proteins. GM130, LAMP1, and EEA1 were validated previously by the manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	All cancer cell lines (59M, ES2, A2780, HEYA8, OVK18, SNU8, SNGM, EMTOKA, RMGI, OVISE, KURAMOCHI, OVCAR4, HCC-78, NCI-H441, NCI-H1437, RERFLCAD2, JHUEM1, HEC6, and JHOS4 were collected by the Cancer Cell Line Encyclopedia prior to distribution for our use. The original sources are available at depmap.org. COV413A and HEK-293T were collected from ATCC.
Authentication	All cell lines were authenticated by exome sequencing and SNP profiling by the Cancer Cell Line Encyclopedia prior to distribution, and are only used for <15 passages to avoid cell line drift and potential cross contamination.
Mycoplasma contamination	All cell lines are routinely tested for mycoplasma, and only cell lines with negative results are used.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	In total, this study used 26 female, 7-week-old, Rag1-/- Il2rg-/- (NRG) mice obtained from The Jackson Laboratories.
Wild animals	This study did not use wild animals
Field-collected samples	This study did not employ field collected samples.
Ethics oversight	These studies were used under the Institutional Animal Care and Use Committee (IACUC) of the Broad Institute under animal protocol 0194- 01-18.

Note that full information on the approval of the study protocol must also be provided in the manuscript.