(supplementary text and fig. S22) (*28*). This suggests that the increase in total number of genes in humans as compared with that in yeast yielded a system of higher complexity rather than more robustness through added redundancy. Nonessential human genes appear to frequently engage in synthetic lethal interactions. Our studies start to reveal an interconnected module of genetic interactions affecting the secretory pathway and link it to uncharacterized genes. The experimental strategy is applicable to various cellular processes and may help unravel the genetic network encoding a human cell.

REFERENCES AND NOTES

1. G. Giaever *et al.*, *Nature* **418**, 387–391 (2002).
2. A. H. Tong *et al.*, *Science* **294**, 2364–2368 (2001).
3. S. Mnaimneh *et al.*, *Cell* **118**, 31–44 (2004).
4. J. E. Carette *et al.*, *Nature* **477**, 340–343 (2011).
5. J. E. Carette *et al.*, *Nat. Biotechnol.* **29**, 542–546 (2011).
6. J. E. Carette *et al.*, *Science* **326**, 1231–1235 (2009).
7. L. T. Jae *et al.*, *Science* **340**, 479–483 (2013).
8. D. J. Kelleher, R. Gilmore, *Glycobiology* **16**, 47R–62R (2006).
9. A. Dumax-Vorzet, P. Roboti, S. High, *J. Cell Sci.* **126**, 2595–2606 (2013).
10. O. Parnas *et al.*, *Cell* **162**, 675–686 (2015).
11. G. Reiss, S. te Heesen, R. Gilmore, R. Zufferey, M. Aebi, *EMBO J.* **16**, 1164–1172 (1997).
12. A. H. Tong *et al.*, *Science* **303**, 808–813 (2004).
13. J. L. Hartman 4th, B. Garvik, L. Hartwell, *Science* **291**, 1001–1004 (2001).
14. C. Boone, H. Bussey, B. J. Andrews, *Nat. Rev. Genet.* **8**, 437–449 (2007).
15. M. Schuldiner *et al.*, *Cell* **123**, 507–519 (2005).
16. S. M. Nijman, S. H. Friend, *Science* **342**, 809–811 (2013).
17. J. Tischler, B. Lehner, A. G. Fraser, *Nat. Genet.* **40**, 390–391 (2008).
18. V. N. Subramaniam, F. Peter, R. Philp, S. H. Wong, W. Hong, *Science* **272**, 1161–1163 (1996).
19. D. Demaegd *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6859–6864 (2013).
20. E. L. Huttlin *et al.*, *Cell* **162**, 425–440 (2015).
21. A. Diao, D. Rahman, D. J. Pappin, J. Lucocq, M. Lowe, *J. Cell Biol.* **160**, 201–212 (2003).
22. W. L. Charng *et al.*, *PLOS Biol.* **12**, e1001777 (2014).
23. A. Rejman Lipinski *et al.*, *PLOS Pathog.* **5**, e1000615 (2009).
24. A. L. Greninger, G. M. Knudsen, M. Betegon, A. L. Burlingame, J. L. DeRisi, *MBio* **4**, e00098–e00013 (2013).
25. T. R. Graham, C. G. Burd, *Trends Cell Biol.* **21**, 113–121 (2011).
26. N. Y. Hsu *et al.*, *Cell* **141**, 799–811 (2010).
27. H. M. van der Schaar *et al.*, *Cell Res.* **22**, 1576–1592 (2012).
28. T. Wang *et al.*, *Science* **350**, 1096–1101 (2015).

GENOMICS

Identification and characterization of essential genes in the human genome

Tim Wang,[1,2,3,4] Kıvanç Birsoy,[1,2,3,4]* Nicholas W. Hughes,[3] Kevin M. Krupczak,[2,3,4] Yorick Post,[2,3,4] Jenny J. Wei,[1,2] Eric S. Lander,[1,3,5]†‡ David M. Sabatini[1,2,3,4,6]†‡

Large-scale genetic analysis of lethal phenotypes has elucidated the molecular underpinnings of many biological processes. Using the bacterial clustered regularly interspaced short palindromic repeats (CRISPR) system, we constructed a genome-wide single-guide RNA library to screen for genes required for proliferation and survival in a human cancer cell line. Our screen revealed the set of cell-essential genes, which was validated with an orthogonal gene-trap–based screen and comparison with yeast gene knockouts. This set is enriched for genes that encode components of fundamental pathways, are expressed at high levels, and contain few inactivating polymorphisms in the human population. We also uncovered a large group of uncharacterized genes involved in RNA processing, a number of whose products localize to the nucleolus. Last, screens in additional cell lines showed a high degree of overlap in gene essentiality but also revealed differences specific to each cell line and cancer type that reflect the developmental origin, oncogenic drivers, paralogous gene expression pattern, and chromosomal structure of each line. These results demonstrate the power of CRISPR-based screens and suggest a general strategy for identifying liabilities in cancer cells.

T he systematic identification of essential genes in microorganisms has provided critical insights into the molecular basis of many biological processes (*1*). Similar studies in human cells have been hindered by the lack of suitable tools. Moreover, little is known about how the set of cell-essential genes differs across cell types and genotypes. Differentially essential genes are likely to encode tissue-specific modulators of key cellular processes and important targets for cancer therapies. We used two independent approaches for inactivating genes at the DNA level to define the cell-essential genes of the human genome.

The first approach uses the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9–based gene editing system, which has emerged as a powerful tool to engineer the genomes of cultured cells and whole organisms (*2*, *3*). We and others have shown that lentiviral single-guide RNA (sgRNA) libraries can enable pooled loss-of-function screens and have used the technology to uncover mediators of drug resistance and pathogen toxicity (*4*–*6*). To systematically identify cell-essential genes, we constructed a library, which was optimized for high cleavage activity, and performed a proliferation-based screen in the near-haploid human KBM7 chronic myelogenous leukemia (CML) cell line (Fig. 1, table S1, and supplementary text S1).

The unusual karyotype of these cells also allows for an independent method of genetic screening. In this approach, null mutants are generated at random through retroviral gene-trap mutagenesis, selected for a phenotype, and monitored by sequencing the viral integration sites to pinpoint the causal genes (*7*). Positive selection–based screens by use of this method have identified genes underlying processes such as epigenetic silencing and viral infection (*7*–*9*). We extended this technique by developing a strategy for negative selection and conducted a screen for cell-essential genes (Fig. 1 and supplementary text S2).

For both methods, we computed a score for each gene that reflects the fitness cost imposed by inactivation of the gene. We defined the CRISPR score (CS) as the average $\log_2$ fold-change in the abundance of all sgRNAs targeting a given gene, with replicate experiments showing a high degree of reproducibility [correlation coefficient $(r) = 0.90$] (Fig. 2A, fig. S1A, and table S2). Of the 18,166 genes targeted by the library, 1878 scored as essential for optimal proliferation in our screen, although this precise number depends on the cutoff chosen (Fig. 2A and tables S2 and S3). Overall, this fraction represents ~10% of genes within our data set or roughly 9.2% of the entire genome (many of the genes not targeted by our library encode olfactory receptors that are unlikely to be cell-essential). Gene products that act in a non-cell-autonomous manner are not expected to score as essential in this pooled setting (fig. S1B).

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge MA 02139, USA. [2]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [3]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA. [4]David H. Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02139, USA. [5]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. [6]Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Present address: Laboratory of Metabolic Regulation and Genetics, The Rockefeller University, New York, NY 10065, USA.
†These authors contributed equally to this work.
‡Corresponding author. E-mail: lander@broadinstitute.org (E.S.L.); sabatini@wi.mit.edu (D.M.S.).

We defined the gene-trap score (GTS) as the fraction of insertions in a given gene occurring in the inactivating orientation. Because the accuracy of this score depends on the depth of insertional coverage, we set a requirement on the minimum number ($n = 65$) of antisense inserts in a gene needed for inclusion in our analysis by measuring the concordance between replicate experiments (Fig. 2B; fig. S1, C and D; table S4; and supplementary text S2). For the 7370 genes on the haploid chromosomes that exceeded this threshold, the GTS was well-correlated with the CS and with results from a copublished study that used a similar gene-trap approach ($r = 0.68$) (Fig. 2C; fig. S1, E and F; and supplementary text S3). The strong correspondence between the overlapping sets of cell-essential genes defined by the two methods provides support for the accuracy of the CRISPR scores for the full set of 18,166 genes.

The two methods differed with respect to the diploid chromosome 8. Whereas the gene-trap screen failed to detect any cell-essential genes on this chromosome, the CRISPR screen uncovered a similar proportion of cell-essential genes on all the autosomes (Fig. 2, A to C). These observations indicate that (i) the vast majority of cell-essential genes are haplosufficient and (ii) biallelic inactivation occurs at high frequency in our CRISPR screen (4).

To assess the accuracy of our scores with other measures of gene essentiality, we relied on functional profiling experiments conducted in yeast *Saccharomyces cerevisiae* as a benchmark (1, 10). Specifically, we ranked genes common to all data sets by their scores in each data set—CRISPR, gene trap, scores from similar loss-of-function RNA interference (RNAi) screens (11), and as a naïve proxy for gene essentiality, gene expression levels determined by means of RNA sequencing (RNA-Seq)—and compared these rankings with the essentiality of yeast homologs. The CRISPR and gene-trap methods had significantly stronger correlations with the yeast results than did the RNAi screens or gene expression, which performed similarly to each other (both methods, $P < 10^{-4}$, permutation test) (Fig. 2D). On the basis of additional comparisons with yeast gene essen-

tiality, we also found that (i) our new optimized sgRNA library gave better results than those from screens using older unoptimized libraries (4, 5) and (ii) the coverage of this library (~10 constructs per gene) approaches saturation, as evidenced by down-sampling (decreasing the coverage by randomly eliminating subsets of data) (fig. S2, A and B). Together, our results suggest that scores from the CRISPR and gene-trap screens both provide accurate measures of the cell-essentiality of human genes.

Essential genes should be under strong purifying selection and should thus show greater evolutionary constraint than that of nonessential genes (12). Consistent with this expectation, the essential genes found in our screens were more broadly retained across species, showed higher levels of conservation between closely related species, and contain fewer inactivating polymorphisms within the human species, as compared with their dispensable counterparts (Fig. 2, E to G). Essential genes also tend to have higher expression and encode proteins that engage in more protein-protein interactions (13–15). These
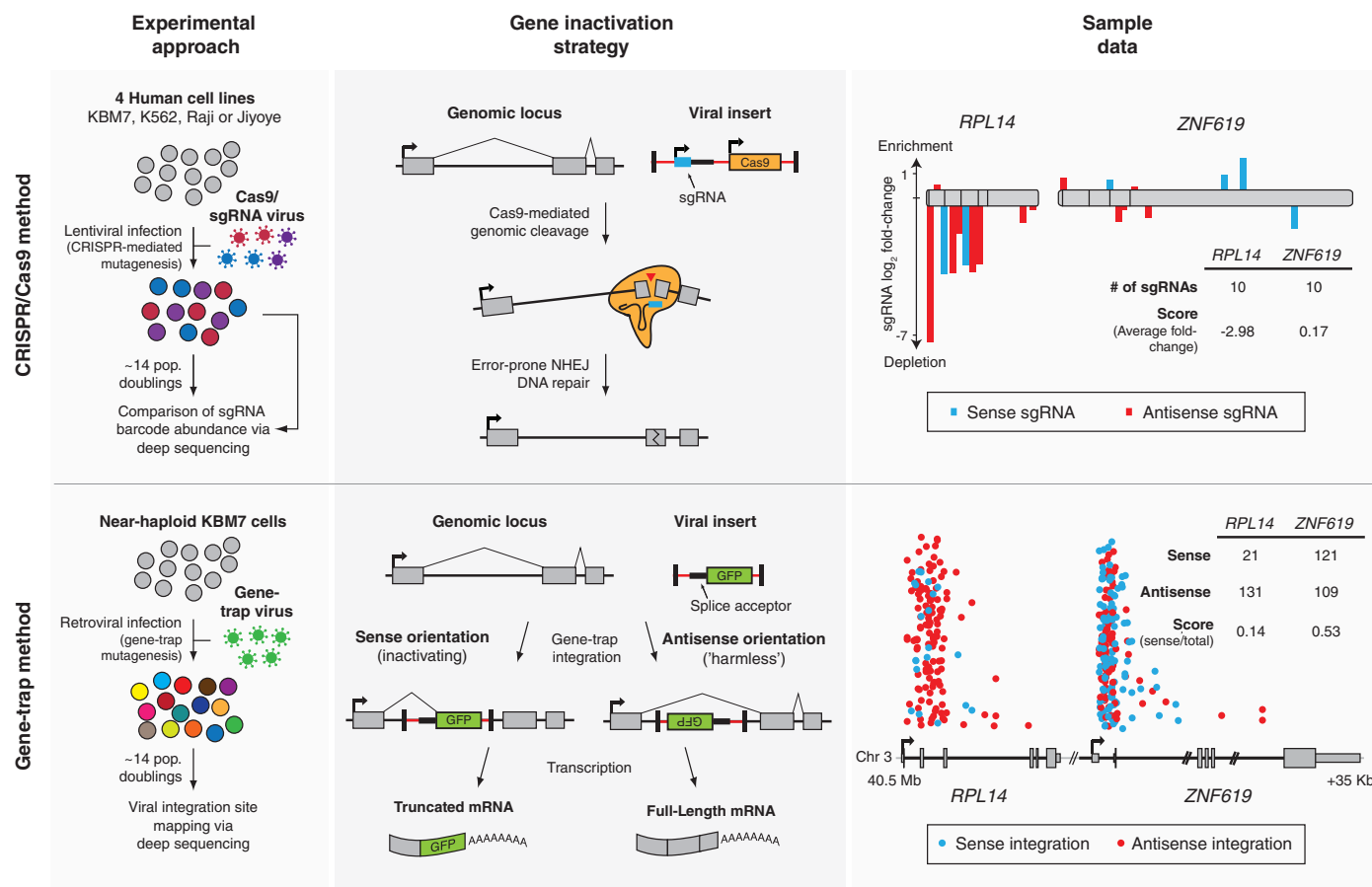

Fig. 1. Two approaches for genetic screening in human cells. (**Top**) CRISPR/Cas9 method. Cells are transduced with a genome-wide sgRNA lentiviral library. Gene inactivation via Cas9-mediated genomic cleavage is directed by the 20–base pair (bp) sequence at the 5′ end of the sgRNA. Cells bearing sgRNAs targeting essential genes are depleted in the final population. (**Bottom**) Gene-trap method. KBM7 cells are transduced with a gene-trap retrovirus that integrates in an inactivating or "harmless" orientation at random genomic loci. Essential genes contain fewer insertions in the inactivating orientation. Sample data for two neighboring genes—*RPL14*, encoding an essential ribosomal protein, and *ZNF619*, encoding a dispensable zinc finger protein—are displayed. For CRISPR/Cas9, sgRNAs are plotted according to their target position along each gene, with the height of each bar indicating the level of depletion. Boxes indicate individual exons. For gene trap, the intronic insertion sites in each gene are plotted according to their orientation and genomic position. The height of each point is randomized.

patterns were also observed in our CRISPR data set (Fig. 2, H and I).

In *S. cerevisiae*, genes with paralogous copies in the genome show a lower degree of essentiality, presumably because of at least partial functional overlap (*16*). Surprisingly, meta-analysis of knockout mouse collections has suggested that there is no such correlation in mammals (*17*, *18*). However, others have challenged this interpretation because the genes analyzed were far from a random sample (*19*). Using the results from our genome-wide screens, we revisited this question and observed that genes with paralogs are indeed less likely to be essential, which is consistent with the idea that paralogs can pro-

vide functional redundancy at the cellular level (Fig. 2J).

To examine the functions of the cell-essential genes, we used gene set enrichment analysis (GSEA) and found strong enrichment for many fundamental biological processes, such as DNA replication, RNA transcription, and mRNA translation (fig. S3A) (*20*). Whereas most of the genes could be assigned to such well-defined pathways, no function has been ascribed to ~330 of the cell-essential genes (18%) (Fig. 3A). For this set of uncharacterized genes, an analysis of the domains within their encoded gene products and comparisons with proteomic data sets from organellar purifications revealed substantial enrichment in

proteins found in the nucleolus and those containing domains associated with RNA processing (fig. S3, B and C) (*21*).

We characterized three such genes—*C16orf80*, *C3orf17*, and *C9orf114*—whose mRNA expression patterns across the Cancer Cell Line Encyclopedia (CCLE) were correlated with that of genes involved in RNA processing (Fig. 3B). We validated the essentiality of these genes in short-term proliferation assays and detected localization of their products to the nucleus (C16orf80) or nucleolus (C3orf17 and C9orf114) (Fig. 3, C and D) (*22*). Additionally, mass spectrometric analyses of anti-FLAG-immunoprecipitates prepared from KBM7 cells expressing FLAG-tagged C16orf80,
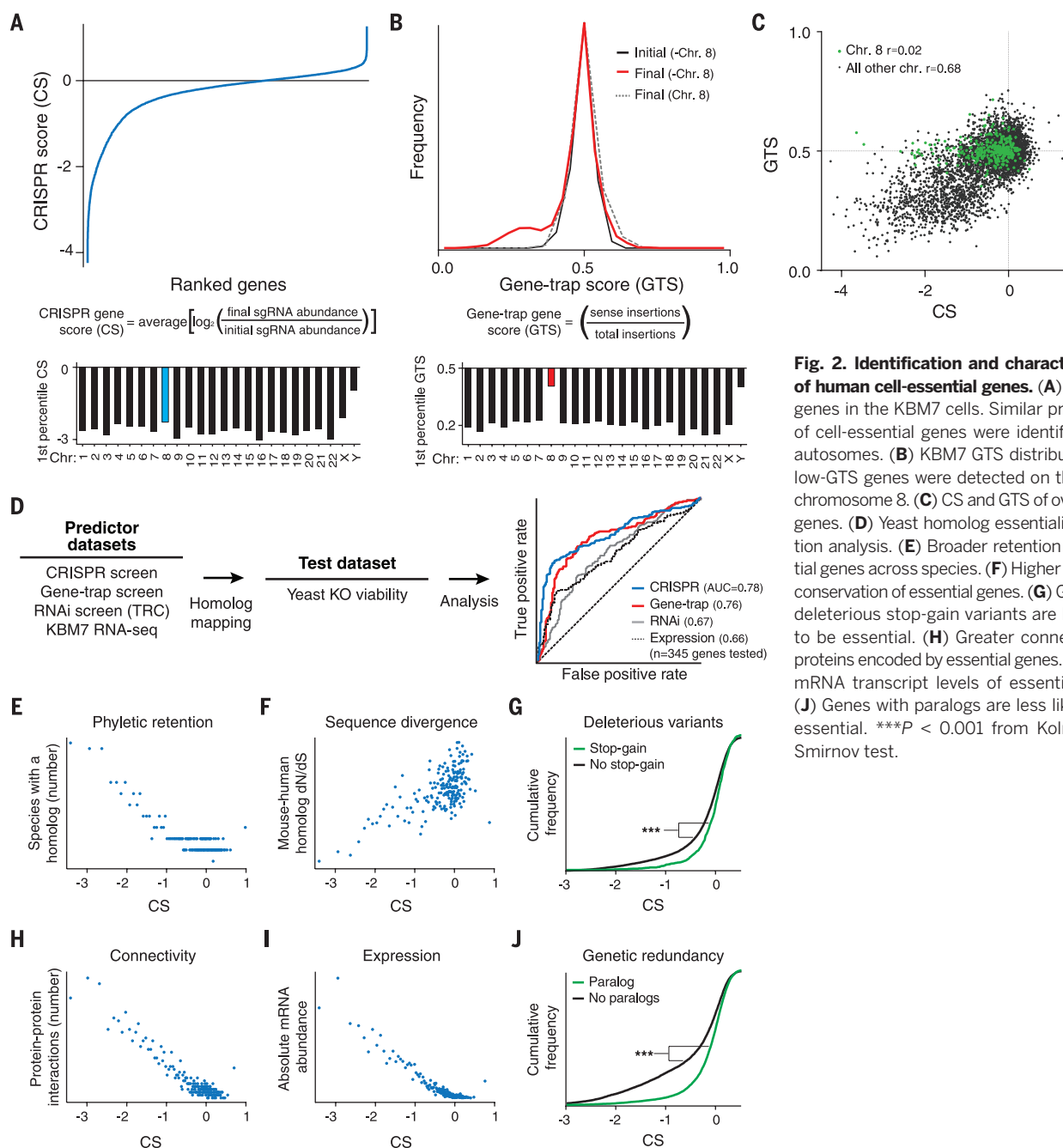

$$\text{CRISPR gene score (CS)} = \text{average}\left[\log_2\left(\frac{\text{final sgRNA abundance}}{\text{initial sgRNA abundance}}\right)\right]$$

$$\text{Gene-trap gene score (GTS)} = \left(\frac{\text{sense insertions}}{\text{total insertions}}\right)$$

Fig. 2. Identification and characterization of human cell-essential genes. (**A**) CSs of all genes in the KBM7 cells. Similar proportions of cell-essential genes were identified on all autosomes. (**B**) KBM7 GTS distributions. No low-GTS genes were detected on the diploid chromosome 8. (**C**) CS and GTS of overlapping genes. (**D**) Yeast homolog essentiality prediction analysis. (**E**) Broader retention of essential genes across species. (**F**) Higher sequence conservation of essential genes. (**G**) Genes with deleterious stop-gain variants are less likely to be essential. (**H**) Greater connectivity of proteins encoded by essential genes. (**I**) Higher mRNA transcript levels of essential genes. (**J**) Genes with paralogs are less likely to be essential. ***$P < 0.001$ from Kolmogorov-Smirnov test.

C3orf17, and C9orf114 revealed interactions with multiple subunits of the spliceosome, ribonuclease (RNase) P/MRP, and H/ACA small nucleolar ribonucleoprotein (snoRNP) complexes, respectively (Fig. 3E). These results implicate *C16orf80* in splicing, which is consistent with its association with mRNAs; *C3orf17* in ribosomal RNA/tRNA processing; and *C9orf114* in RNA modification (*23*). More broadly, our results indicate that the molecular components of many critical cellular processes, especially RNA processing, have yet to be fully defined in mammalian cells.

To determine how the set of essential genes differs among cell lines, we screened another CML cell line (K562) and two Burkitt's lymphoma cell lines (Raji and Jiyoye) using the CRISPR system (tables S2 and S3). Overall, the sets of essential genes in the four cell lines showed a high degree of overlap (Fig. 4A). Out of these

four cell lines, the KBM7 CRISPR results showed the highest correlation with the KBM7 gene-trap data set, suggesting that the few differences observed are likely to be biologically meaningful (fig. S4A).

We focused first on genes found to be essential in only one of the four cell lines. The Raji, Jiyoye, and KBM7 cell lines had 6, 7, and 19 such genes, respectively (fig. S4B and table S5). One example was *DDX3Y*, which resides in the nonpseudoautosomal region of the Y chromosome and was required only in Raji cells (Fig. 4B). Its X-linked paralog, *DDX3X*, was essential in KBM7 and K562 cells (Fig. 4E). Both genes encode DEAD-box helicases that likely have similar cellular functions (*24*). Thus, the dependence on one paralog might reflect functional absence of the other paralog. Indeed, DNA sequencing of *DDX3X* in Raji cells revealed hemizygous muta-

tions in the 5′-splice site of intron 8 that resulted in the production of a truncated mRNA transcript (Fig. 4, C and D). Conversely, *DDX3Y* was not expressed in KBM7 cells and was not present in K562 cells, which are of female origin (Fig. 4E). Introduction of wild-type *DDX3X* cDNA into Raji cells fully rescued the proliferation defect resulting from *DDX3Y* loss, indicating that the paralogous genes are essential and functionally overlapping (Fig. 4F). Essential paralogous gene pairs, involved in glucose metabolism (*HK1/2* and *SLC2A1/3*) and cell-cycle regulation (*CDK4/6*), were also observed in the Jiyoye line (fig. S4C and supplementary text S4). Vulnerabilities due to the loss of a paralogous partner may serve as targets for highly personalized antitumor therapies (*25*).

In some cases, cell line–specific essentiality of paralogous genes did not reflect differential expression. For example, the transcription factors *GATA1* and *GATA2* are expressed in both K562 and KBM7 cells, but the first is specifically essential in K562 cells and the second in KBM7 cells (fig. S4D). These master regulators are known to promote proliferation and survival during distinct developmental stages in the hematopoietic lineage; *GATA1* is required for the survival of erythroid progenitors, and *GATA2* is required for the maintenance and proliferation of immature hematopoietic progenitors (*26*). These two cell types likely correspond to the cells of origin of the two CML lines (*27, 28*). We also identified similar instances of genes required for cell line–specific functions such as nuclear factor κB (NF-κB) pathway regulation and homology-directed DNA repair in Raji and KBM7 cells, respectively (fig. S4, E and F, and supplementary text S5).

Whereas the other three cell lines showed only a few cell line–specific essential genes, the K562 cell line was an outlier, with 63 such genes (fig. S4B). Oddly, these genes showed no discernible common functions, and many were not even expressed in the K562 cell line. (Additionally, a few encoded secreted factors whose loss would not be expected to be lethal in a pooled screen.) This mystery was resolved when we examined the chromosomal location of the genes: The majority (39 of the 63 genes) reside near 9q34 and 22q11. These two regions are translocated to produce the *BCR-ABL* oncogene and are present in a high-copy tandem amplification in K562 cells (fig. S5, A and B) (*29*). All 61 contiguous genes within the amplicon on 22q11 scored as essential, suggesting that sgRNA-mediated cleavage in this repeated region induces cytotoxicity in a manner unrelated to the function of the target genes themselves (Fig. 4G and fig. S5F). Indeed, two sgRNAs targeting nongenic sites within the amplicon were toxic to K562 but not KBM7 cells. They increased the abundance of phosphorylated histone H2AX (γH2AX), a marker of DNA damage, and induced erythroid differentiation, which occurs upon DNA damage in this cell line (Fig. 4, H to J, and fig. S5C) (*30*). We obtained similar results in another cell line, HEL, which contains a highly amplified
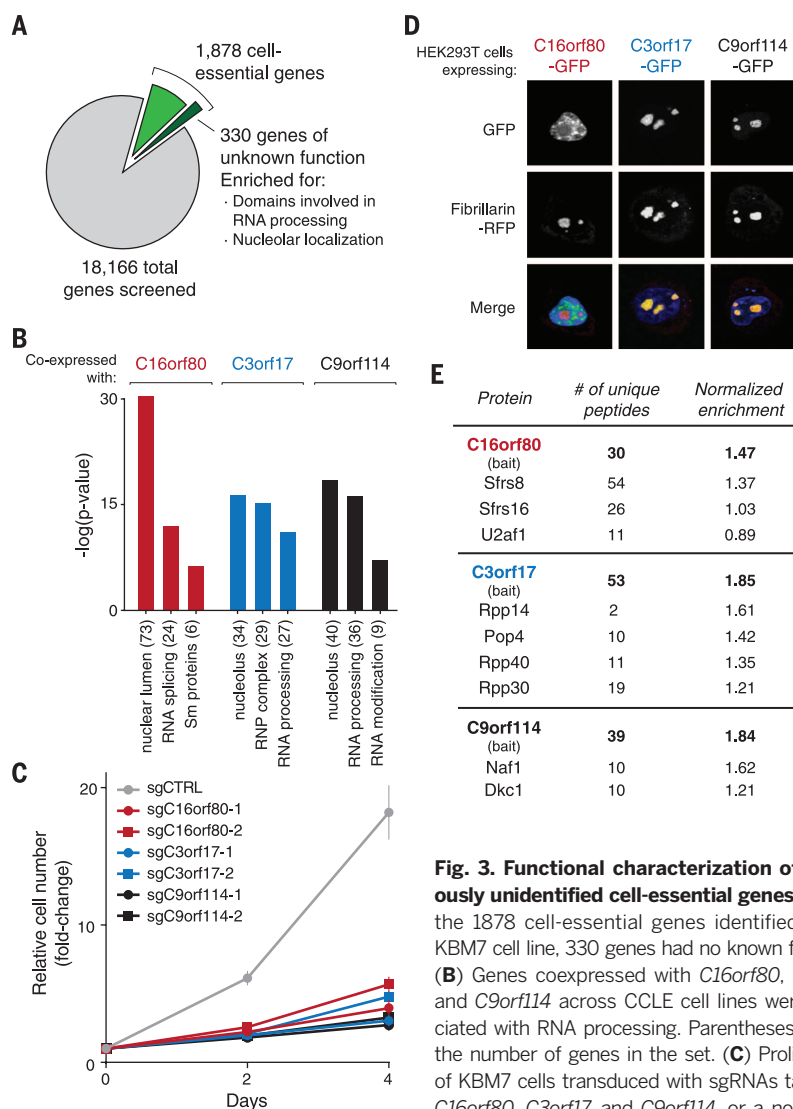

Fig. 3. Functional characterization of previously unidentified cell-essential genes. (**A**) Of the 1878 cell-essential genes identified in the KBM7 cell line, 330 genes had no known function. (**B**) Genes coexpressed with *C16orf80*, *C3orf17*, and *C9orf114* across CCLE cell lines were associated with RNA processing. Parentheses denote the number of genes in the set. (**C**) Proliferation of KBM7 cells transduced with sgRNAs targeting *C16orf80*, *C3orf17*, and *C9orf114*, or a nontargeting control. Error bars denote SD (*n* = 4 experiments per group). (**D**) C16orf80 localized to the nucleus, and C3orf17 and C9orf114 localized to the nucleolus. Fibrillarin–red fluorescent protein (RFP) was used as a nucleolar marker. GFP, green fluorescent protein. (**E**) Multiple subunits of the spliceosome, RNase P/MRP, and H/ACA ribonucleoprotein complexes interact with C16orf80, C3orf17, and C9orf114, respectively.
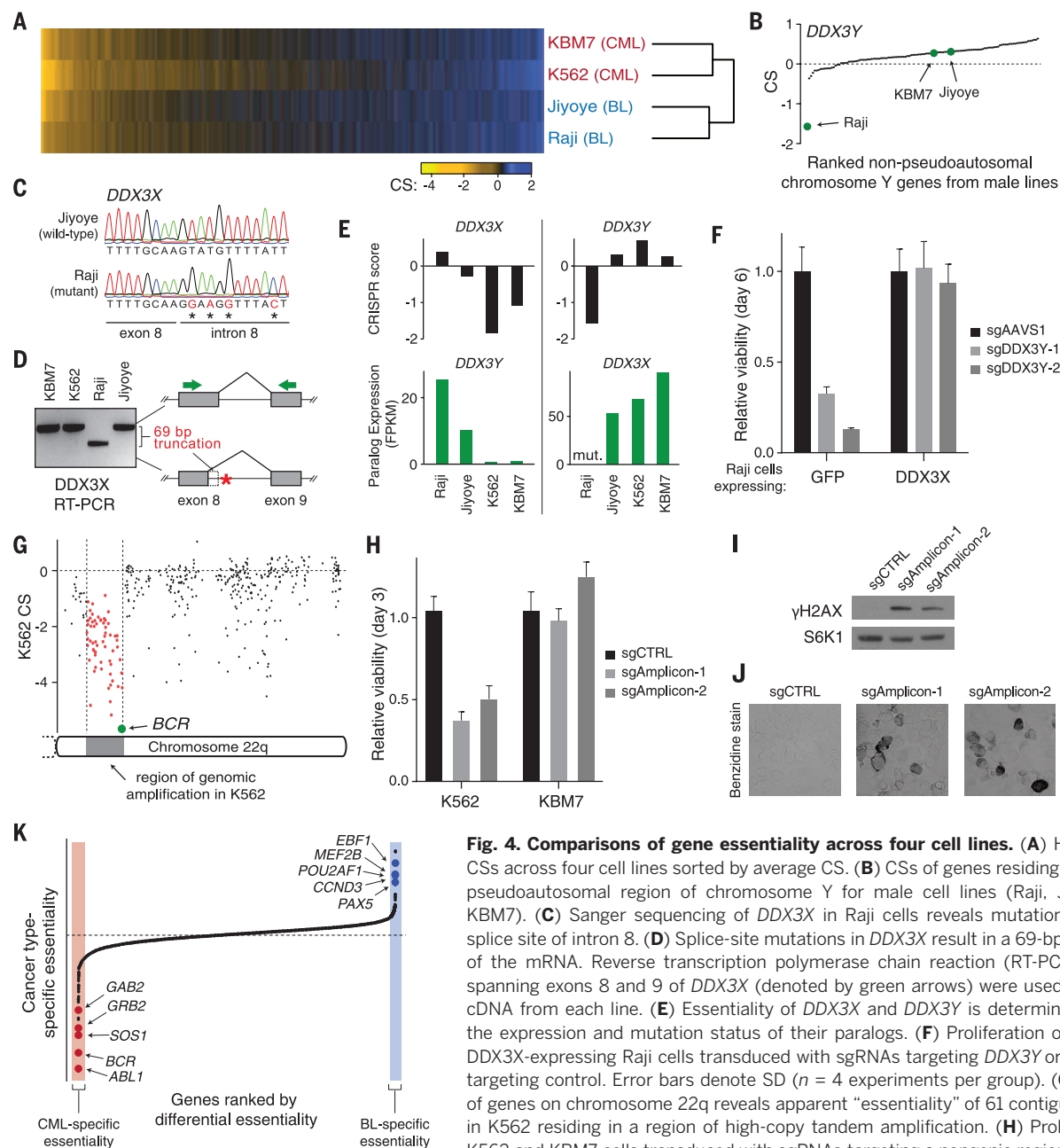
Figure 3E table:

Protein	# of unique peptides	Normalized enrichment
C16orf80 (bait)	**30**	**1.47**
Sfrs8	54	1.37
Sfrs16	26	1.03
U2af1	11	0.89
C3orf17 (bait)	**53**	**1.85**
Rpp14	2	1.61
Pop4	10	1.42
Rpp40	11	1.35
Rpp30	19	1.21
C9orf114 (bait)	**39**	**1.84**
Naf1	10	1.62
Dkc1	10	1.21

Fig. 4. Comparisons of gene essentiality across four cell lines. (**A**) Heatmap of CSs across four cell lines sorted by average CS. (**B**) CSs of genes residing in the non-pseudoautosomal region of chromosome Y for male cell lines (Raji, Jiyoye, and KBM7). (**C**) Sanger sequencing of *DDX3X* in Raji cells reveals mutations in the 5′ splice site of intron 8. (**D**) Splice-site mutations in *DDX3X* result in a 69-bp truncation of the mRNA. Reverse transcription polymerase chain reaction (RT-PCR) primers spanning exons 8 and 9 of *DDX3X* (denoted by green arrows) were used to amplify cDNA from each line. (**E**) Essentiality of *DDX3X* and *DDX3Y* is determined through the expression and mutation status of their paralogs. (**F**) Proliferation of GFP- and DDX3X-expressing Raji cells transduced with sgRNAs targeting *DDX3Y* or an AAVS1-targeting control. Error bars denote SD (*n* = 4 experiments per group). (**G**) Analysis of genes on chromosome 22q reveals apparent "essentiality" of 61 contiguous genes in K562 residing in a region of high-copy tandem amplification. (**H**) Proliferation of K562 and KBM7 cells transduced with sgRNAs targeting a nongenic region within the BCR-ABL amplicon in K562 cells or a nontargeting control. Error bars denote SD (*n* = 4 experiments per group). (**I**) γH2AX (phospho-S139 H2AX) immunoblot analysis of K562 transduced with sgRNAs as in (H). S6K1 was used as a loading control. (**J**) Cleavage within amplified region induced erythroid differentiation in K562 cells as assessed by means of 3,3′-dimethoxybenzidine hemoglobin staining. (**K**) Comparison of gene essentiality between the two cancer types reveals oncogenic drivers and lineage specifiers. Genes were ranked by the difference between the average CS of each cancer type.

region surrounding the *JAK2* tyrosine kinase (fig. S5, D and E). Together, these findings indicate that lethality upon Cas9-mediated cutting may also reflect chromosome structure and therefore should be evaluated in light of copy-number information.

Last, we looked for consistent differences in essential genes between the two CML and two Burkitt's lymphoma lines. Such genes might represent attractive targets for antineoplastic therapies because their inhibition is less likely to be broadly cytotoxic. Overall, we identified 33 genes that were specifically essential in the CML lines

and 15 genes in the Burkitt's lymphoma lines (Fig. 4K and table S6). As a control, permuted comparisons—that is, a set containing of one CML and one Burkitt's line versus the complementary sets—showed roughly half as many "set-specific" essential genes (fig. S6, A to C).

In the CML lines, the top two genes were *BCR* and *ABL1*, which is consistent with the known essentiality of the BCR-ABL translocation product and the therapeutic effect of BCR-ABL inhibitors such as imatinib (*31*). Additional members of the BCR-ABL signaling pathway—*SOS1*,

GRB2, and *GAB2*—scored strongly as well (ranked 3, 4, and 7, respectively). Network analysis of the other top hits also uncovered several genes encoding assembly factors for the electron transport chain, as well as enzymes involved in folate-mediated one-carbon metabolism. These results suggest additional potential targets for CML therapy (table S6).

In the B cell–derived Burkitt's lymphoma cell lines, the top genes included three B cell–lineage transcription factors *EBF1*, *POU2AF1*, and *PAX5* (ranked 3, 6, and 8, respectively). Each of these

genes is the target of recurrent translocations in lymphoma (*32–34*). Enhancers of the corresponding three gene loci all show a high level of bromodomain containing 4 (BRD4) occupancy in Ly1 cells, a related diffuse large B cell lymphoma cell line, suggesting bromodomain inhibitors such as JQ1 as potential treatments (*35*). Other selectively essential genes included *MEF2B*—a transcriptional activator of *BCL6*—and *CCND3*, both of which are frequently mutated and implicated in the pathogenesis of various lymphomas (*36*). Intriguingly, the top two hits, *CHM* and *RPP25L*, do not appear to have specific roles in B cells; rather, their differential essentiality is likely explained by the lack of expression of their paralogs, *CHML* and *RPP25*, in both of the Burkitt's lymphoma cell lines studied (fig. S6D).

We used two complementary and concordant approaches, CRISPR and gene trap, to define the cell-essential genes in the human genome. Although the gene-trap method is suitable only for loss-of-function screening in rare haploid cell lines, the CRISPR method is broadly applicable. Extending our analysis across different cell lines and tumor types, we developed a framework to assess differential gene essentiality and identify potential drivers of the malignant state. The method can be readily applied to more cell lines per cancer type so as to eliminate idiosyncrasies particular to a given cell line and to more cancer types so as to systematically uncover tumor-specific liabilities that might be exploited for targeted therapies.

REFERENCES AND NOTES

1. G. Giaever et al., *Nature* **418**, 387–391 (2002).
2. L. Cong et al., *Science* **339**, 819–823 (2013).
3. H. Wang et al., *Cell* **153**, 910–918 (2013).
4. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, *Science* **343**, 80–84 (2014).
5. O. Shalem et al., *Science* **343**, 84–87 (2014).
6. H. Koike-Yusa, Y. Li, E. P. Tan, M. C. Velasco-Herrera, K. Yusa, *Nat. Biotechnol.* **32**, 267–273 (2014).
7. J. E. Carette et al., *Science* **326**, 1231–1235 (2009).
8. J. E. Carette et al., *Nature* **477**, 340–343 (2011).
9. I. A. Tchasovnikarova et al., *Science* **348**, 1481–1485 (2015).
10. J. P. Kastenmayer et al., *Genome Res.* **16**, 365–373 (2006).
11. G. S. Cowley, B. A. Weir, W. C. Hahn, *Sci. Data* 10.1038/data.2014.35 (2014).
12. A. C. Wilson, S. S. Carlson, T. J. White, *Annu. Rev. Biochem.* **46**, 573–639 (1977).
13. Y. Ishihama et al., *BMC Genomics* **9**, 102 (2008).
14. H. Jeong, S. P. Mason, A. L. Barabási, Z. N. Oltvai, *Nature* **411**, 41–42 (2001).
15. T. Hart, K. R. Brown, F. Sircoulomb, R. Rottapel, J. Moffat, *Mol. Sys. Biol.* 10.15252/msb.20145216 (2014).
16. Z. Gu et al., *Nature* **421**, 63–66 (2003).
17. H. Liang, W.-H. Li, *Trends Genet.* **23**, 375–378 (2007).
18. B.-Y. Liao, J. Zhang, *Trends Genet.* **23**, 378–381 (2007).
19. T. Makino, H. Hokamp, A. McLysaght, *Trends Genet.* **25**, 152–155 (2009).
20. A. Subramanian et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
21. Y. Ahmad, F.-M. Boisvert, E. Lundberg, M. Uhlen, A. I. Lamond, *Mol. Cell. Proteomics* **11**, 013680 (2012).
22. J. Barretina et al., *Nature* **483**, 603–607 (2012).
23. A. G. Baltz et al., *Mol. Cell* **46**, 674–690 (2012).
24. H. Sekiguchi, H. Iida, J. Fukumura, T. Nishimoto, *Exp. Cell Res.* **300**, 213–222 (2004).
25. F. L. Muller et al., *Nature* **488**, 337–342 (2012).
26. K. Ohneda, M. Yamamoto, *Acta Haematol.* **108**, 237–245 (2002).
27. L. C. Andersson, K. Nilsson, C. G. Gahmberg, *Int. J. Cancer* **23**, 143–147 (1979).
28. B. S. Andersson et al., *Leukemia* **9**, 2100–2108 (1995).
29. S. Q. Wu et al., *Leukemia* **9**, 858–862 (1995).
30. A. Constantinou, K. Kiguchi, E. Huberman, *Cancer Res.* **50**, 2618–2624 (1990).
31. B. Scappini et al., *Cancer* **100**, 1459–1471 (2004).
32. S. Iida et al., *Blood* **88**, 4110–4117 (1996).
33. H. Bouamar et al., *Blood* **122**, 726–733 (2013).
34. S. Galiègue-Zouitina et al., *C. R. Acad. Sci. III* **318**, 1125–1131 (1995).
35. B. Chapuy et al., *Cancer Cell* **24**, 777–790 (2013).
36. R. D. Morin et al., *Nature* **476**, 298–303 (2011).

GENOME EDITING

Genome-wide inactivation of porcine endogenous retroviruses (PERVs)

Luhan Yang,[1,2,3]*† Marc Güell,[1,2,3]† Dong Niu,[1,4]† Haydy George,[1]† Emal Lesha,[1] Dennis Grishin,[1] John Aach,[1] Ellen Shrock,[1] Weihong Xu,[6] Jürgen Poci,[1] Rebeca Cortazio,[1] Robert A. Wilkinson,[5] Jay A. Fishman,[5] George Church[1,2,3]*

The shortage of organs for transplantation is a major barrier to the treatment of organ failure. Although porcine organs are considered promising, their use has been checked by concerns about the transmission of porcine endogenous retroviruses (PERVs) to humans. Here we describe the eradication of all PERVs in a porcine kidney epithelial cell line (PK15). We first determined the PK15 PERV copy number to be 62. Using CRISPR-Cas9, we disrupted all copies of the PERV *pol* gene and demonstrated a >1000-fold reduction in PERV transmission to human cells, using our engineered cells. Our study shows that CRISPR-Cas9 multiplexability can be as high as 62 and demonstrates the possibility that PERVs can be inactivated for clinical application of porcine-to-human xenotransplantation.

Pig genomes contain from a few to several dozen copies of PERV elements (*1*). Unlike other zoonotic pathogens, PERVs cannot be eliminated by biosecure breeding (*2*). Prior strategies for reducing the risk of PERV transmission to humans have included small interfering RNAs (RNAi), vaccines (*3–5*), and PERV elimination using zinc finger nucleases (*6*) or TAL effector nucleases (*7*), but these have had limited success. Here we report the successful use of the CRISPR-Cas9 RNA-guided nuclease system (*8–10*) to inactivate all copies of the PERV *pol* gene and effect a 1000-fold reduction of PERV infectivity of human cells.

To design Cas9 guide RNAs (gRNAs) that specifically target PERVs, we analyzed the sequences of publicly available PERVs and other endogenous retroviruses in pigs (methods). Using droplet digital polymerase chain reaction (PCR), we identified a distinct clade of PERV elements (Fig. 1A) and determined that there were 62 copies of PERVs in PK15 cells (a porcine kidney epithelial cell line) (Fig. 1B). We then designed two Cas9 gRNAs that targeted the highly conserved catalytic center (*11*) of the *pol* gene on PERVs (Fig. 1C and fig. S1). The *pol* gene product functions as a reverse transcriptase (RT) and is thus essential for viral replication and infection. We determined that these gRNAs targeted all PERVs but no other endogenous retrovirus or other sequences in the pig genome (methods).

Initial experiments showed inefficient PERV editing when Cas9 and the gRNAs were transiently transfected (fig. S2). Thus, we used a PiggyBac transposon (*12*) system to deliver a doxycycline-inducible Cas9 and the two gRNAs

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA, USA. [3]eGenesis Biosciences, Boston, MA 02115, USA. [4]College of Animal Sciences, Zhejiang University, Hangzhou 310058, China. [5]Transplant Infectious Disease and Compromised Host Program, Massachusetts General Hospital, Boston, MA 02115, USA. [6]Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.
*Corresponding author. E-mail: gchurch@genetics.med. harvard.edu (G.C.); luhan.yang@egenesisbio.com (L.Y.)
†These authors contributed equally to this work.

Science

Identification and characterization of essential genes in the human genome

Tim Wang, Kivanç Birsoy, Nicholas W. Hughes, Kevin M. Krupczak, Yorick Post, Jenny J. Wei, Eric S. Lander and David M. Sabatini

Zeroing in on essential human genes

More powerful genetic techniques are helping to define the list of genes required for the life of a human cell. Two papers used the CRISPR genome editing system and a gene trap method in haploid human cells to screen for essential genes (see the Perspective by Boone and Andrews). Wang *et al.*'s analysis of multiple cell lines indicates that it may be possible to find tumor-specific dependencies on particular genes. Blomen *et al.* investigate the phenomenon in which nonessential genes are required for fitness in the absence of another gene. Hence, complexity rather than robustness is the human strategy.

Science, this issue p. 1096 and p. 1092; see also p. 1028

ARTICLE TOOLS	http://science.sciencemag.org/content/350/6264/1096
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2015/10/14/science.aac7041.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/350/6264/1028.full http://science.sciencemag.org/content/sci/350/6264/1092.full
REFERENCES	This article cites 50 articles, 14 of which you can access for free http://science.sciencemag.org/content/350/6264/1096#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions