

of infection, the number of EphA2<sup>low</sup> infected cells decreased over time (Fig. 2A). This difference could not be accounted for by division rates, because we observed lower levels of host cell division among EphA2<sup>low</sup> cells. Thus, our results may in fact underestimate the impact of EphA2 on infected cell survival (fig. S5). When we infected *EphA2*<sup>(-/-)</sup> and wild-type mice with  $10^5$  *P. yoelii* sporozoites, we observed a large decrease in liver-stage burden after 42 hours in *EphA2*<sup>(-/-)</sup> mice (Fig. 2B). *EphA2*<sup>(-/-)</sup> mice also exhibited a delay in the onset of blood-stage infection by 1 to 3 days (Fig. 2C). Thus, without EphA2, the host is far less susceptible to productive parasite liver infection.

The parasitophorous vacuole membrane (PVM) is critical for liver-stage development. One liver-stage PVM-resident protein, UIS4, is highly expressed after invasion when it is exported to the PVM (11), making it a useful marker. We constructed a *P. yoelii* parasite line, *PyUIS4-Myc*, which expressed a UIS4-Myc fusion protein driven by the endogenous UIS4 promoter (Fig. 2D). This allowed us to monitor PVM prevalence (UIS4<sup>pos</sup>) in infected cells by flow cytometry. Most of the UIS4<sup>pos</sup> infected host cells were in the EphA2<sup>high</sup> category (Fig. 2E). Similarly, the level of EphA2 expression was higher in UIS4<sup>pos</sup> infected cells than in UIS4<sup>neg</sup> infected cells (Fig. 2F). Thus, sporozoites not only preferentially entered EphA2<sup>high</sup> cells, but invasion accompanied by PVM formation was far more effective in these cells. UIS4<sup>neg</sup> infected hepatocytes suffered a higher frequency of cell death (Fig. 2G).

Two members of the 6-Cys family of parasite proteins (12, 13), P52 and P36, are expressed in sporozoites, are important for the invasion of hepatocytes (14–16), and are critical for PVM formation (14). In mouse livers, parasites without P52 or P36 were almost entirely eliminated within 3 hours after infection (fig. S6). We tested whether the lack of P52 and P36 phenocopies the lack of host EphA2 and found that *p52<sup>-/-</sup>/p36<sup>-/-</sup>* *P. yoelii* sporozoites exhibited a reduced preference for EphA2<sup>high</sup> cells (Fig. 3A). The related 6-Cys protein PI2 shows structural similarity to the mammalian ligand for EphA2, EphrinA1 (10).

We showed that an interaction in the extracellular region of EphA2 was required for sporozoite entry using an EphA2-blocking antibody (Fig. 1K). Therefore, we next asked whether the presence of P36 and P52 was required for the antibody to block sporozoite entry. The EphA2 antibody blocked infection for wild-type *P. yoelii* sporozoites, but *p52<sup>-/-</sup>/p36<sup>-/-</sup>* sporozoite entry was not affected (Fig. 3B). These data suggest that P36 or P52 engages EphA2 at the point of host cell invasion. We next tested whether P52 or P36 could directly impede the interaction between EphrinA1 and EphA2 on the hepatocyte surface, which results in EphA2 activation. When we added EphrinA1 in the presence of P36 to Hepal-6 cells, P36 blocked the activation of EphA2 (Fig. 3, C and D). P52, however, did not block EphrinA1-mediated activation of EphA2 (Fig. 3, C and D). To determine whether the interaction between EphA2 and P36 also occurs in human parasites, we assessed levels of EphA2 in *P. falciparum*

wild-type or *p52<sup>-/-</sup>/p36<sup>-/-</sup>/sap1<sup>-/-</sup>* parasite-infected HC-04 cells. The P52-P36-deficient *P. falciparum* sporozoites exhibited partially reduced selectivity for EphA2<sup>high</sup> HC-04 cells compared with *P. falciparum* wild-type sporozoites (Fig. 3E). Thus, P36 engages EphA2 but does not trigger its activation in rodent and human parasites.

We have shown that both host EphA2 and parasite 6-Cys proteins have a role in sporozoite invasion of hepatocytes and the establishment of the growth-permissive intracellular niche. Without either component, the parasite can still enter hepatocytes, but it does so without a PVM, which can result in death of the infected hepatocyte. The convergence of infection-permissive phenotypes is best explained by an interaction between parasite P36 and hepatocyte EphA2 when the PVM is formed. This role for EphA2 in hepatocyte infection does not preclude the possibility that additional hepatocyte receptors may be critical for infection. Interventional strategies aimed at either EphA2 or sporozoite 6-Cys proteins might block parasite infection before the onset of clinical malaria.

## REFERENCES AND NOTES

1. S. H. Kappe, A. M. Vaughan, J. A. Boddey, A. F. Cowman, *Science* **328**, 862–866 (2010).
2. A. Coppi et al., *Cell Host Microbe* **2**, 316–327 (2007).
3. U. Frevert et al., *J. Exp. Med.* **177**, 1287–1298 (1993).
4. O. Silvie et al., *Nat. Med.* **9**, 93–96 (2003).
5. S. Yalaoui et al., *Cell Host Microbe* **4**, 283–292 (2008).
6. C. D. Rodrigues et al., *Cell Host Microbe* **4**, 271–282 (2008).
7. L. S. Austin, A. Kaushansky, S. H. Kappe, *Cell. Microbiol.* **16**, 784–795 (2014).
8. A. Kaushansky et al., *Infect. Immun.* **83**, 39–47 (2015).
9. E. M. Lisabeth, G. Falivelli, E. B. Pasquale, *Cold Spring Harb. Perspect. Biol.* **5**, a009159 (2013).
10. S. A. Arredondo et al., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6692–6697 (2012).
11. A. K. Mueller et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3022–3027 (2005).
12. T. Annoura et al., *FASEB J.* **28**, 2158–2170 (2014).
13. M. R. van Dijk et al., *PLOS Pathog.* **6**, e1000853 (2010).
14. M. Labaied et al., *Infect. Immun.* **75**, 3758–3768 (2007).
15. M. R. van Dijk et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12194–12199 (2005).
16. K. M. VanBuskirk et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13004–13009 (2009).

## ACKNOWLEDGMENTS

We are grateful to W. Betz for mosquito and sporozoite production. We thank the vivarium staff of the Center for Infectious Disease Research for their work with mice. This work was funded by a NIH Research Project Grant (R01) to S.H.I.K. and A.K. (grant no. 1R01GM101183-01A1). A.K. is also a recipient of a Transition to Independence Award (award no. 1K99AI111785-01A1), which partially funded this work. A.K., N.A., A.N.D., V.V., N.D., H.K., and L.S.A. performed the experiments. A.K., D.N.S., and S.H.I.K. supervised the research. A.K. and S.H.I.K. wrote the paper. A provisional U.S. patent application (application no. 62/110018) has been filed by the Center for Infectious Disease Research, covering interventions that exploit the described host-parasite interaction. The data are included in the main manuscript and in the supplementary materials.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/350/4264/suppl/DC1](http://www.sciencemag.org/content/350/4264/suppl/DC1)  
Materials and Methods  
Figs. S1 to S7  
Table S1

28 August 2015; accepted 13 October 2015  
10.1126/science.aad3318

## GENOMICS

# Gene essentiality and synthetic lethality in haploid human cells

Vincent A. Blomen,<sup>1\*</sup> Peter Májek,<sup>2\*</sup> Lucas T. Jae,<sup>1\*</sup> Johannes W. Bigenzahn,<sup>2</sup> Joppe Nieuwenhuis,<sup>1</sup> Jacqueline Staring,<sup>1</sup> Roberto Sacco,<sup>2</sup> Ferdy R. van Diemen,<sup>1</sup> Nadine Olk,<sup>2</sup> Alexey Stukalov,<sup>2</sup> Caleb Marceau,<sup>3</sup> Hans Janssen,<sup>1</sup> Jan E. Carette,<sup>3</sup> Keiryn L. Bennett,<sup>2</sup> Jacques Colinge,<sup>2,4†</sup> Giulio Superti-Furga,<sup>2,5†</sup> Thijn R. Brummelkamp<sup>1,2,6†</sup>

Although the genes essential for life have been identified in less complex model organisms, their elucidation in human cells has been hindered by technical barriers. We used extensive mutagenesis in haploid human cells to identify approximately 2000 genes required for optimal fitness under culture conditions. To study the principles of genetic interactions in human cells, we created a synthetic lethality network focused on the secretory pathway based exclusively on mutations. This revealed a genetic cross-talk governing Golgi homeostasis, an additional subunit of the human oligosaccharyltransferase complex, and a phosphatidylinositol 4-kinase β adaptor hijacked by viruses. The synthetic lethality map parallels observations made in yeast and projects a route forward to reveal genetic networks in diverse aspects of human cell biology.

**S**ingle-cell organisms can often tolerate inactivating mutations in the majority of genes (1–3), but it is unclear whether human cells require more essential genes because of increased complexity, or fewer because of added redundancy. To study this, we used mutagenesis in the near-haploid chronic myeloid leu-

kemia (CML) cell line KBM7 (karyotype 25, XY, +8, Ph+), and its nonhematopoietic derivative HAP1, which is haploid for all chromosomes (fig. S1A) (4). More than 34.3 million and 65.9 million gene-trap integrations were identified in KBM7 and HAP1 cells, respectively. The gene-trap vector we used was unidirectional by design (fig. S1B),

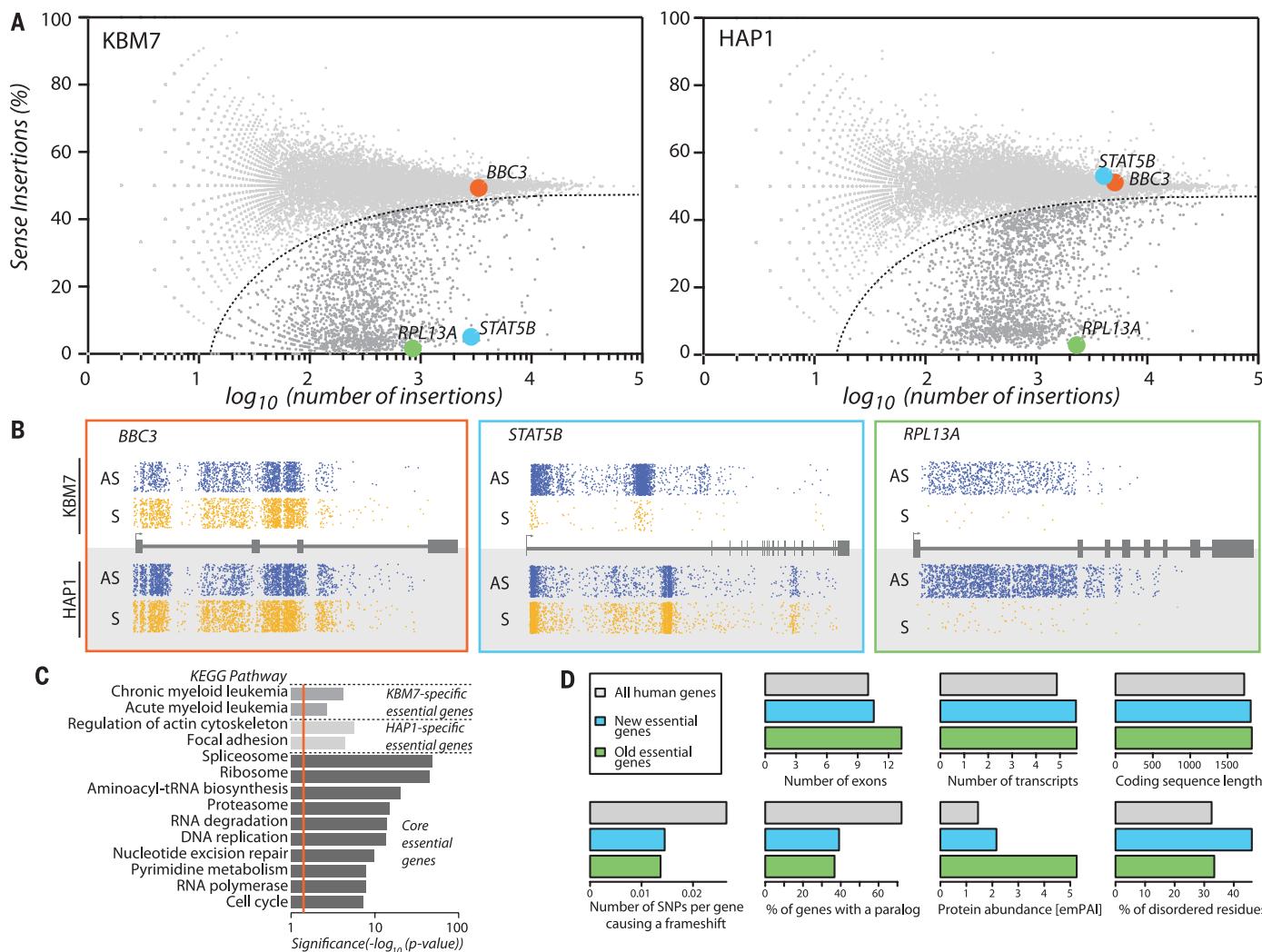
and for most genes, the number of intronic insertions in the sense direction was similar to that in the antisense direction (for example, pro-apoptotic factor *BBC3*) (Fig. 1, A and B) (5–7). For a fraction of genes, however, disruptive muta-

tions were underrepresented, which is indicative of impaired fitness: Some genes (such as *STAT5B*) appeared essential in one cell type (fig. S1C), whereas others appeared essential in both (such as *RPL13A*) (Fig. 1, A and B, and tables S1 to S3).

In KBM7 cells, 2054 genes (table S1 and figs. S2 and S3) and in HAP1, 2181 genes (table S2 and figs. S2 and S3) appeared to be needed for viability or optimal fitness under the experimental growth conditions (referred to as “essential,” although the approach does not distinguish between the two). The 1734 genes identified in both cell lines were designated as “core essentialome” (table S3). In KBM7 cells, genes on chromosome 8 (present in two copies) tolerated disruptive mutations, underscoring the specificity of the approach (fig. S1D). Furthermore, nearly all subunits of the proteasome were identified as es-

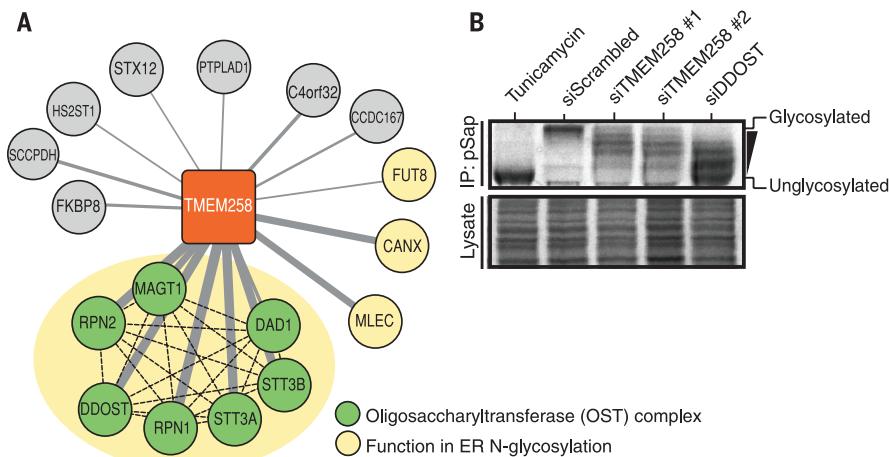
sential (fig. S4). In general, essential genes are overrepresented in categories such as translation or transcription but not signaling (Fig. 1C and figs. S5 and S6).

Many genes required for fitness in yeast were also essential in human cells. Exceptions were largely explained by paralogs in the human genome or by yeast-specific requirements (fig. S7A and table S4) (7). We estimated the evolutionary age of essential genes and found that 77% emerged in premetazoans (“old” essential genes) (fig. S7B). Essential genes had fewer paralogs and higher protein abundance and contained fewer single-nucleotide polymorphisms (SNPs) predicted to impair function (Fig. 1D). Proteins encoded by essential genes displayed more protein-protein interactions (fig. S8, A to D), and these occurred more frequently with other essential proteins (49.8%) (fig. S9A) and within the same functional



**Fig. 1. Identification of genes required for fitness in KBM7 and HAP1 cells through insertional mutagenesis.** (A) Distinct gene-trap insertions were mapped in KBM7 and HAP1 cells, and their orientation relative to the affected genes was counted. Per gene, the percentage of sense orientation gene-trap insertions (y axis) and the total number of insertions in a particular gene (x axis) are plotted. (B) Gene-trap insertions identified in the sense (S, yellow) or antisense orientation (AS, blue) in a nonessential gene

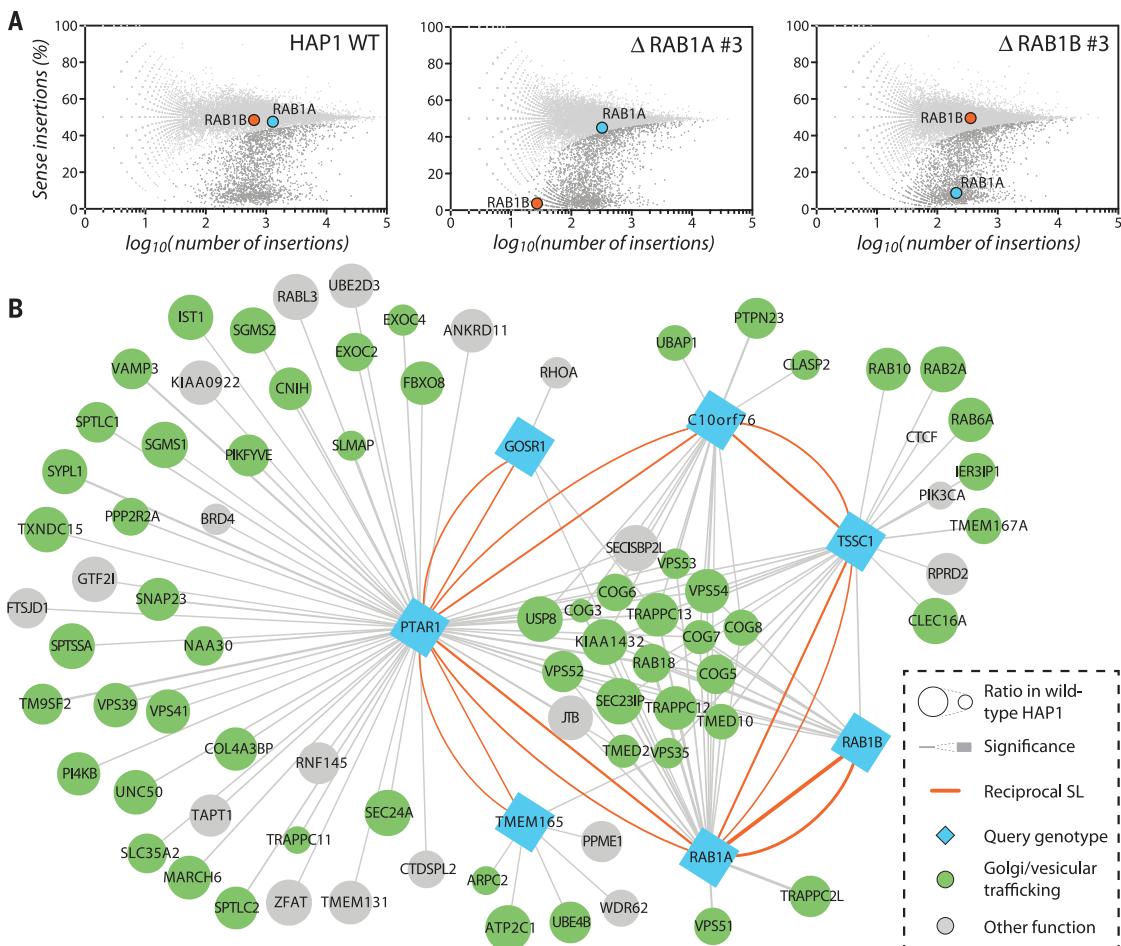
(*BBC3*), a gene essential only in KBM7 cells (*STAT5B*), and a gene essential in both cell lines (*RPL13A*). (C) KEGG pathway enrichment analysis of essential genes shared between or specific to KBM7 or HAP1 cells. (D) Properties of new and old essential genes compared with the human genome. Averages for the sets are displayed, except for protein abundance, in which median Exponentially Modified Protein Abundance Index (emPAI) values are shown.



**Fig. 2. The essential gene *TMEM258* encodes a component of the OST complex.** (A) High-confidence protein-protein interactions associated with *TMEM258*. Green proteins indicate members of the OST complex. Dashed lines indicate the OST complex subnetwork. (B) Effects of depletion of *TMEM258* with small interfering RNAs on the glycosylation of endogenous prosaposin. Cells were pulsed with  $^{35}\text{S}$ -methionine/cysteine, lysed, and subjected to immunoprecipitation by using antibodies to prosaposin. Precipitated proteins were detected by means of phosphorimaging, and hypoglycosylated prosaposin species are indicated. Tunicamycin treatment and depletion of the established OST subunit *sDDOST* served as positive controls.

category (fig. S5B). Remarkably, the products of “new” essential genes are more often connected with old rather than other new essential gene products, suggesting that they largely function within ancient molecular machineries (fig. S9, B and C).

To identify proteins interacting with products of 18 uncharacterized essential genes, we used tandem affinity purification coupled to mass spectrometry (fig. S10). Interactors were frequently essential proteins (52.4%,  $P < 2.5 \times 10^{-36}$ , hypergeometric test) involved in processes such as splicing, translation, and trafficking (fig. S11 and table S5). The small transmembrane protein *TMEM258* associated with components of the conserved oligosaccharyltransferase (OST) complex (Fig. 2A and fig. S12A) that are essential for protein N-glycosylation (8). *TMEM258* localized to the endoplasmic reticulum (fig. S12B), and depletion (fig. S12, C and D) impaired OST catalytic activity as monitored by means of hypoglycosylation of prosaposin (Fig. 2B) (9). This also rationalizes the observed clustering of *TMEM258* with OST complex subunits in a recent genetic screen (10). Thus, *TMEM258* constitutes a subunit of the



**Fig. 3. Synthetic lethality network generated based on mutations.** (A) Essentiality of *RAB1A* and *RAB1B* in wild-type HAP1 cells and cells deficient for *RAB1A* or *RAB1B*. (B) Genetic interaction network indicating synthetic lethal/sick interactions that were identified by scoring genes for fitness reduction in three nuclease-generated knockout clones per genotype. This revealed an interconnected network, with many genes that could be functionally assigned to the secretory pathway (labeled in green). Reciprocal interactions, scored in either query genotype, are indicated by orange edges. Edge thickness reflects the effect size of the interaction (compared with wild-type cells).

human OST complex, and although homology searches (fig. S12E) do not identify a yeast ortholog, TMEM258 may relate to the similarly sized yeast transmembrane protein OST5 (11).

Whereas most genes appear nonessential, their function may be buffered by other genes so that only simultaneous disruption is lethal (12–15). The frequency of such synthetic lethal interactions between human genes is debated and challenging to address experimentally (16, 17). We studied the small guanosine triphosphatases (GTPases) RAB1A and RAB1B by creating individual knockout lines and assessing the genes needed for fitness in these backgrounds (Fig. 3A and fig. S13A). Whereas neither *RAB1A* nor *RAB1B* were essential in wild-type cells, *RAB1A* became indispensable in *RAB1B* knockout cells and vice versa (Fig. 3A and fig. S13B). To explore the breadth of synthetic lethality, we probed the secretory pathway using three independent knockout cell lines (fig. S14) for *RAB1A*, *RAB1B*, *GOSR1* [a subunit of the Golgi soluble *N*-ethylmaleimide-sensitive factor attachment protein (SNAP) receptor] (18), and *TMEM165* (a Golgi-resident  $\text{Ca}^{2+}/\text{H}^+$  antiporter whose deficiency impairs glycosylation) (Fig. 3B, figs. S15 and S16, and table S6) (19). Most of their genetic interactions impinged on the secretory pathway (Fig. 3B and table S7), and many were found synthetic lethal with *PTARI*. Synthetic lethality screens in *PTARI*-deficient cells confirmed these genetic interactions and additionally identified the uncharacterized gene *C10orf76* (Fig. 3B and fig. S17A). Validation by use of *C10orf76* as query gene confirmed synthetic lethality with *PTARI* and (reciprocally) identified

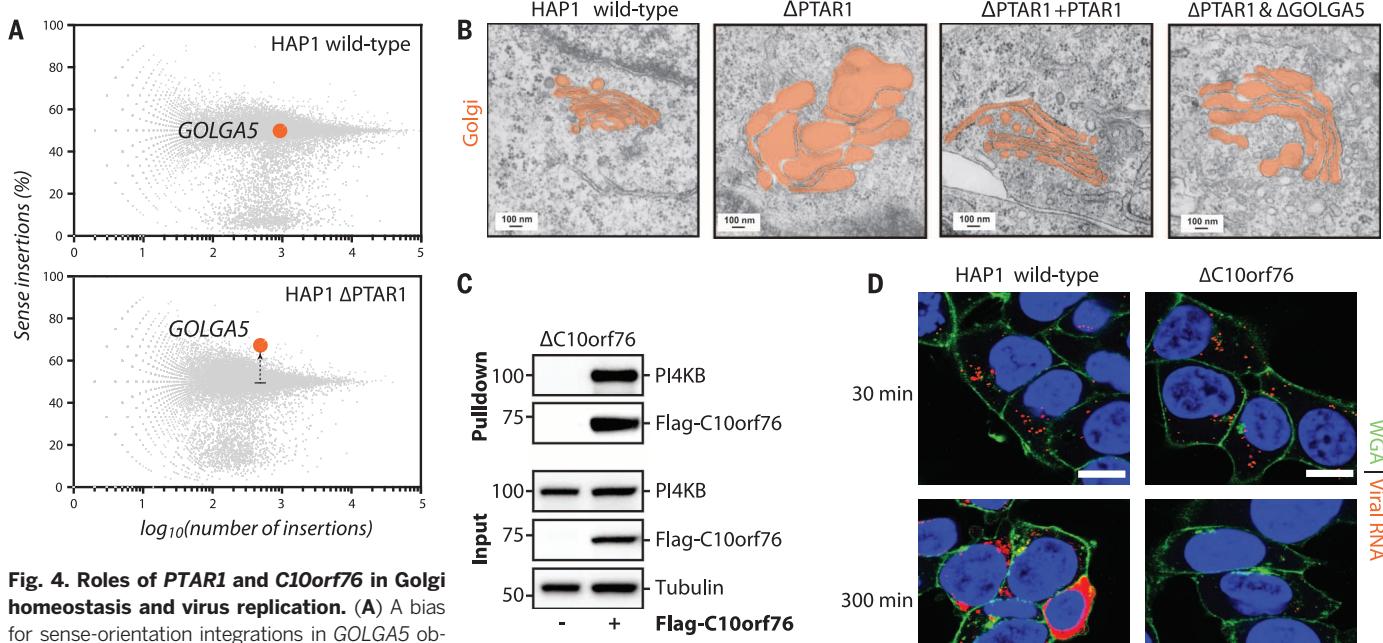
*TSSC1*, which was recently reported to interact with the Golgi-associated retrograde protein complex (GARP) (Fig. 3B) (20). The human genes we studied display on average ~20 synthetic lethal interactions, a number comparable with that in yeast (12), although this varies between genes, with *PTARI* (causing a fitness defect when deleted alone) having close to 60 interactions (fig. S17B). This illustrates that synthetic lethal interactions can be identified and validated by using reciprocal haploid screens and that, similarly to yeast, interactions frequently occur between genes whose products act in related processes (fig. S17B) (13, 16). However, we acknowledge a caveat that this approach cannot readily distinguish between synthetic lethal or synthetic “sick” interactions.

The impaired growth of *PTARI*-deficient cells (table S2) was suppressed by loss of the Golgi factor *GOLGA5* (Fig. 4A and fig. S18) (21). *PTARI*-deficient cells had an abnormally dilated Golgi morphology (fig. S19A), which was partially corrected by codeletion of *GOLGA5* (Fig. 4B and fig. S19B). Functionally, *PTARI* deficiency impaired glycosylation (fig. S19C) (7), possibly owing to dysregulation of RAB proteins (22). Indeed, *PTARI*-deficient cells showed attenuated geranylgeranylation of RAB1A and RAB1B (fig. S19D). Partial correction of the Golgi morphology in cells lacking both *PTARI* and *GOLGA5* could relate to the effect of *GOLGA5*, itself a RAB effector, on Golgi fragmentation (21, 23). Thus, the interaction map reveals *PTARI* and *GOLGA5* as opposing handles tuning Golgi morphology and homeostasis.

Genetic analysis suggested a link between the unstudied gene *C10orf76* and *PI4KB*, which were

both synthetic lethal with *PTARI* (Fig. 3B). A host factor screen with coxsackievirus A10 also identified *C10orf76* as well as *PI4KB* (fig. S20, A and B), and a proteomics survey (24) suggested association between *C10orf76* and *PI4KB*. We confirmed this interaction in immunoprecipitation experiments with cells expressing FLAG-tagged *C10orf76* (Fig. 4C). Phosphatidylinositol 4-kinase  $\beta$  (*PI4KB*) regulates abundance of phosphatidylinositol 4-phosphate [*PI(4)P*] (25) and has a role in genome replication of various RNA viruses, including coxsackieviruses (26). Infection studies confirmed that cells in which *C10orf76* was knocked out were particularly resistant to coxsackievirus A10 (fig. S20C). Although virus entry occurred normally, replication of viral RNA was decreased in cells in which *C10orf76* was knocked out (Fig. 4D and fig. S20D). Enteroviruses hijack *PI4KB* activity to construct “replication factories,” which were abundant in wild-type cells but rare in *C10orf76*-deficient cells (fig. S20E). Amounts of *PI(4)P* were decreased in these cells, and Golgi retention of *PI4KB* after chemical inhibition (27) was largely dependent on *C10orf76*, which also localized to this compartment under these conditions (fig. S21, A and B). Thus, *C10orf76* is a *PI4KB*-associated factor hijacked by specific picornaviruses for replication.

This study identifies ~2000 genes required for optimal fitness of cultured haploid human cells. Despite technical limitations, the identification of gene essentiality shows high concordance with the gene-trap and clustered regularly interspaced short palindromic repeats (CRISPR) data reported in the accompanying manuscript of Wang *et al.*



**Fig. 4. Roles of *PTARI* and *C10orf76* in Golgi homeostasis and virus replication.** (A) A bias for sense-orientation integrations in *GOLGA5* observed in *PTARI*-deficient HAP1 cells but not wild-type cells. (B) Electron micrographs of the Golgi apparatus (orange highlight) in the indicated genotypes. (C) Interaction of FLAG-tagged *C10orf76* with *PI4KB* in HAP1 cells detected with immunoprecipitation by using antibodies to FLAG. (D) Coxsackievirus A10 amplification in wild-type and *C10orf76*-deficient cells measured by means of single molecule fluorescent in situ hybridization (smFISH) to localize individual viral genomes (red). Intracellular viral RNA was first detected after 30 min. Increased RNA signal after 300 min indicates RNA replication.

(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. Non-essential 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

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

## ACKNOWLEDGMENTS

We thank T. Sixma, S. Nijman, J. Roix, J. Neefjes, and members of the Brummelkamp group for discussion and K. Sandhoff for antibodies directed against saposin. This work was supported by the Cancer Genomics Center (CGC.nl), Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)—VIDI grant 917.11.316, European Research Council (ERC) Starting Grant (ERC-2012-STG 309634) to T.R.B., the Austrian Academy of Sciences and the Advanced ERC grant i-FIVE and Austrian Science Fund grant (FWF SFB F4711) to G.S.-F., and the ARC Foundation grant PJA 20141201975 to J.C. Sequence data have been deposited at the National Center for Biotechnology Information Sequence Read Archive with accession number SRP058962.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/350/6264/1092/suppl/DC1](http://www.sciencemag.org/content/350/6264/1092/suppl/DC1)

Materials and Methods

Supplementary Text

Figs. S1 to S23

Tables S1 to S8

References (29–62)

28 September 2014; accepted 1 October 2015

Published online 15 October 2015

10.1126/science.aac7557

## GENOMICS

# Identification and characterization of essential genes in the human genome

Tim Wang,<sup>1,2,3,4</sup> Kivanç Birsoy,<sup>1,2,3,4\*</sup> Nicholas W. Hughes,<sup>3</sup> Kevin M. Krupeck,<sup>2,3,4</sup> Yorick Post,<sup>2,3,4</sup> Jenny J. Wei,<sup>1,2</sup> Eric S. Lander,<sup>1,3,5†‡</sup> David M. Sabatini<sup>1,2,3,4,6†‡</sup>

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.

The 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).

\*Department of Biology, Massachusetts Institute of Technology, Cambridge MA 02139, USA. <sup>†</sup>Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA.

<sup>‡</sup>Broad Institute of MIT and Harvard, 7 Cambridge Center,

Cambridge, MA 02142, USA. <sup>§</sup>David H. Koch Institute for

Integrative Cancer Research at MIT, Cambridge, MA 02139, USA. <sup>¶</sup>Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. <sup>||</sup>Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

<sup>\*</sup>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.).

## Gene essentiality and synthetic lethality in haploid human cells

Vincent A. Blomen, Peter Májek, Lucas T. Jae, Johannes W. Bigenzahn, Joppe Nieuwenhuis, Jacqueline Staring, Roberto Sacco, Ferdy R. van Diemen, Nadine Olk, Alexey Stukalov, Caleb Marceau, Hans Janssen, Jan E. Carette, Keiryn L. Bennett, Jacques Colinge, Giulio Superti-Furga and Thijn R. Brummelkamp

Science 350 (6264), 1092-1096.  
DOI: 10.1126/science.aac7557 originally published online October 15, 2015

### 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/350/6264/1092>

#### SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2015/10/14/science.aac7557.DC1>

#### RELATED CONTENT

<http://stke.sciencemag.org/content/sigtrans/8/404/ec345.abstract>  
<http://science.sciencemag.org/content/sci/350/6264/1028.full>  
<http://science.sciencemag.org/content/sci/350/6264/1096.full>

#### REFERENCES

This article cites 61 articles, 16 of which you can access for free  
<http://science.sciencemag.org/content/350/6264/1092#BIBL>

#### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2015, American Association for the Advancement of Science