Haploid Genetic Screens in Human Cells Identify Host Factors Used by Pathogens

Jan E. Carette,[1] Carla P. Guimaraes,[1] Malini Varadarajan,[1] Annie S. Park,[1] Irene Wuethrich,[1] Alzbeta Godarova,[1] Maciej Kotecki,[2] Brent H. Cochran,[2] Eric Spooner,[1] Hidde L. Ploegh,[1,3] Thijn R. Brummelkamp[1]*

Loss-of-function genetic screens in model organisms have elucidated numerous biological processes, but the diploid genome of mammalian cells has precluded large-scale gene disruption. We used insertional mutagenesis to develop a screening method to generate null alleles in a human cell line haploid for all chromosomes except chromosome 8. Using this approach, we identified host factors essential for infection with influenza and genes encoding important elements of the biosynthetic pathway of diphthamide, which are required for the cytotoxic effects of diphtheria toxin and exotoxin A. We also identified genes needed for the action of cytolethal distending toxin, including a cell-surface protein that interacts with the toxin. This approach has both conceptual and practical parallels with genetic approaches in haploid yeast.

Identification of gene products that play a role in human disease drives much of today's biomedical research. Classical genetics with induced mutations, as pioneered by Muller in 1927 (*1*), is the most powerful unbiased approach to elucidate the genetic components that underlie biological processes. The study of cultured human cells allows the recapitulation of many essential elements of human disease. However, the inability to efficiently generate and recover bi-allelic mutants in human diploid cells limits the contribution of mutagenesis-based genetics to the understanding of human disease.

The identification of cellular genes exploited by bacteria and viruses is essential to elucidate the mechanisms by which these pathogens cause disease (*2*–*4*). Bacterial toxins contribute greatly to pathogenicity of the microbes that produce them. Identification of host proteins involved in toxin cytotoxicity should help to identify targets for therapeutic intervention in diseases caused by bacteria, many of which now show increased resistance to conventional antibiotics. Likewise, an understanding of how viruses depend on host proteins to enter the cell, replicate their genome, and spread may accelerate the development of antiviral drugs. Influenza virus remains a threat to human health, causing several hundred thousands of deaths annually and many more in the course of a pandemic (*5*). The rapid spread of new strains of influenza A [for instance, avian (H5N1) and swine (H1N1) influenza] and the emergence of drug-resistant influenza strains (*6*) limit the effectiveness of vaccines and current antiviral therapeutics. Thus, we developed a method for genetic screens in human cells and isolated genes required for the action of several bacterial toxins and influenza viruses.

Development of an approach for haploid genetic screens in human cells. To facilitate mutagenesis-based genetic approaches in human cells, we used a derivative of the 7 KBM7 chronic myeloid leukemia (CML) cell line with a haploid karyotype except for chromosome 8 (Fig. 1A) (*7*, *8*). In this cell line of hematopoietic origin, gene inactivation should allow the generation of null mutants for most nonessential genes. We chose to inactivate genes with the use of insertional mutagenesis because this approach is highly mutagenic in a variety of organisms, and the integrated DNA sequences provide a convenient molecular tag to identify the disrupted gene. We used gene-trap retroviruses that contain a strong adenoviral splice acceptor site and a marker gene (green fluorescent protein or puromycin-resistance gene) in reverse orientation of the retroviral backbone. To examine whether gene-trap insertions were indeed mutagenic, we performed a pilot screen to isolate mutants resistant to the nucleotide analog 6-thioguanine (6-TG), converted to a toxic metabolite by the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT). The gene-trap virus was titrated to obtain a single viral integration in the majority of the infected cells. Cell lines resistant to 6-TG were recovered, and five independent mutants carried insertions in intron 1 of the X-linked HPRT gene (Fig. 1B and fig. S1A). We next performed two genetic screens to target autosomal genes. KBM7 cells are sensitive to the tumor necrosis factor–related apoptosis-inducing ligand (TRAIL) and to inhibition of the BCR-ABL oncogenic fusion protein by the kinase inhibitor Gleevec (Novartis, Basel, Switzerland). Gene trap–mutagenized KBM7 cells were exposed to either TRAIL or Gleevec and

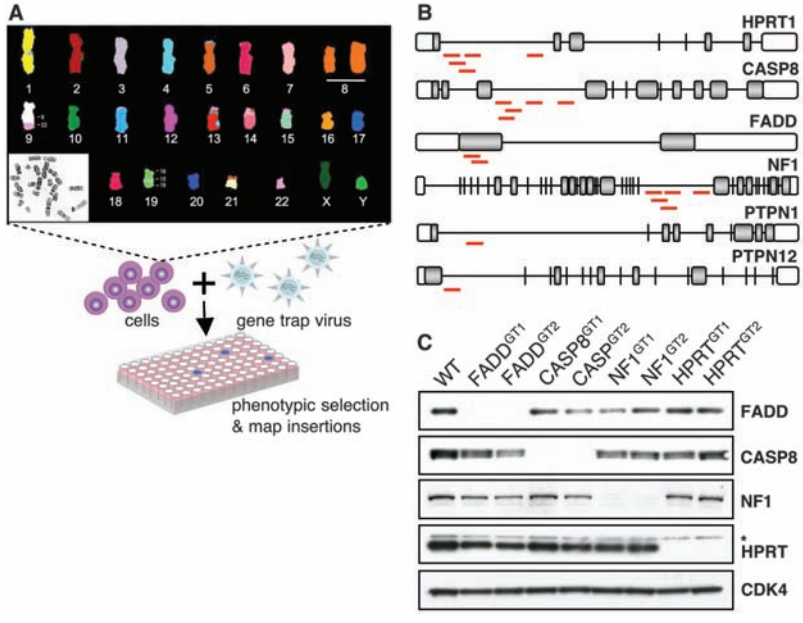

Fig. 1. Haploid genetic screens. (**A**) Twenty-four–color spectral karyotype of near-haploid KBM7 cells and schematic outline of gene trap–mutagenesis screens. (**B**) Schematic outline of the gene-trap insertion sites (red lines) in cells exposed to 6-TG, TRAIL, or Gleevec. (**C**) Immunoblot analysis of FADD, CASP8, NF1, and HPRT expression levels in clones that contain independent gene-trap insertions in the respective loci. Mutant alleles are labeled with GT in superscript notation, and an unspecific background band is indicated with an asterisk. CDK4 was used as a loading control.

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [2]Department of Physiology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA.

*To whom correspondence should be addressed. E-mail: brummelkamp@wi.mit.edu

resistant mutants were recovered. Five TRAIL-resistant clones showed independent insertions in caspase 8 (CASP8) and two independent insertions in Fas-associated death domain protein (FADD), genes known to be required for TRAIL-induced apoptosis (Fig. 1B and fig. S1A) (*9*). Resistance to TRAIL was confirmed in these mutants (fig. S1B). Five independent Gleevec-resistant mutants contained insertions in neurofibromin 1 (NF1) and one in protein tyrosine phosphatase-N1 (PTPN1); both genes are known to play an important role in the response of CML cells to Gleevec (*10*). One insertion was found in PTPN12, a tyrosine phosphatase that interacts with c-abl and negatively regulates its activity (*11*). Thus, PTPN12 is critical for Gleevec sensitivity. All insertions were in the same transcriptional orientation as the target gene, and immunoblot analysis of HPRT, FADD, CASP8, and NF1 mutant cells failed to detect the corresponding gene products (Fig. 1C). The haploid background of KBM7 thus enables the generation of mutant alleles for autosomal genes and pinpoints genes involved in the biological processes under study.

Identification of host factors required for cytolethal distending toxin. Because many pathogenic agents such as bacterial toxins or viruses readily kill the cells they target, a large-scale production of knockout alleles for human genes may enable the identification of host factors essential for pathogenesis, such as enzymes that create structures recognized by toxins or viruses or the receptors themselves. Several pathogenic bacteria (such as *Escherichia coli*, *Shigella dysenteriae*, *Actinobacillus actinomycetemcomitans*, *Campylobacter jejuni*, *Helicobacter* spp., *Salmonella typhi*, and *Haemophilus ducreyi*) secrete potent bacterial toxins named cytolethal distending toxins (CDTs). The DNAse I–like CdtB subunit of these toxins enters the nucleus and causes cytotoxicity by inducing DNA breaks (*12*, *13*). The membrane receptor(s) and other essential host genes involved in the entry or action of CDTs are unknown. KBM7 cells responded to *E. coli*–derived CDT by undergoing an arrest in the G_2/M phase of the cell cycle (Fig. 2A) that precedes cell death. Mutagenized KBM7 cells were treated with CDT and resistant clones were isolated. Eleven independent insertions in sphingomyelin synthase 1 (SGMS1) and three insertions in TMEM181, a gene that encodes a putative G protein–coupled receptor (*14*), were recovered (Fig. 2B and fig. S2A). SGMS1 and TMEM181 mutants were resistant to CDT, a phenotype reverted by complementing the mutant cells with the corresponding cDNAs (Fig. 2C and fig. S2B). The SGMS1 mutation reduced levels of sphingomyelin, as verified by treatment of cells with lysenin, a sphingomyelin-specific pore-forming toxin (fig. S3A). Sphingomyelin is a key component of lipid rafts: Depletion of SGMS1 activity disturbs lipid-raft function and prevents receptor clustering (*15*), a trait of possible relevance for CDT binding and/or entry. Extraction of the lipid-raft component cholesterol from the plasma membrane abolishes CDT binding (*16*).

TMEM181 mutants remained fully sensitive to lysenin (fig. S3, A and B), suggesting that their resistance is acquired by means other than sphingomyelin depletion. Because TMEM181 is present at the cell surface (*17*) and a receptor for CDT must localize to the plasma membrane, we tested whether CDT bound to TMEM181. Flag-tagged CDT was adsorbed onto anti-Flag beads and incubated with cell lysates prepared from wild-type (WT) KBM7 cells and from KBM7 cells that express hemagglutinin (HA)–tagged TMEM181. Immunoblot analysis showed robust binding of TMEM181 to CDT (Fig. 2D). When TMEM181 was overexpressed by retroviral transduction in NIH3T3, U2OS, and Hela cells, it sensitized them to CDT intoxication (Fig. 2E and fig. S4), suggesting that TMEM181 expression levels are rate limiting for intoxication. Thus, CDT may bind to the cell-surface protein TMEM181, an event both required and rate limiting for intoxication, and then enter the cell through sphingomyelin-dependent, lipid-raft–mediated endocytosis, followed by nuclear entry and cleavage of cellular DNA. However, these results do not rule out the possibilities that TMEM181 is part of a complex that constitutes a functional receptor or plays a role in trafficking of a receptor-toxin complex.

Isolation of host factors essential for influenza virus infectivity. We next isolated mutant cells that were resistant to influenza virus A (PR/8/34; H1N1). Proviral-host junction sequencing revealed two independent insertions in cytidine monophosphate *N*-acetylneuraminic acid synthase (CMAS), encoding the enzyme responsible for activation of NeuAc to CMP-NeuAc, the glycosyl donor used in sialic acid–containing glycoconjugate synthesis. These structures are the receptors on influenza-susceptible cells recognized by the influenza hemagglutinin. We recovered three independent insertions in SLC35A2 (Fig. 3A and fig. S5A), a gene whose product transports uridine 5′-diphosphate–galactose from the cytoplasm to the Golgi, where it serves as a glycosyl donor (*18*) important for the generation of glycans to be modified with sialic acids. To determine whether mutant cells could be infected by influenza, we exposed cells to the virus and

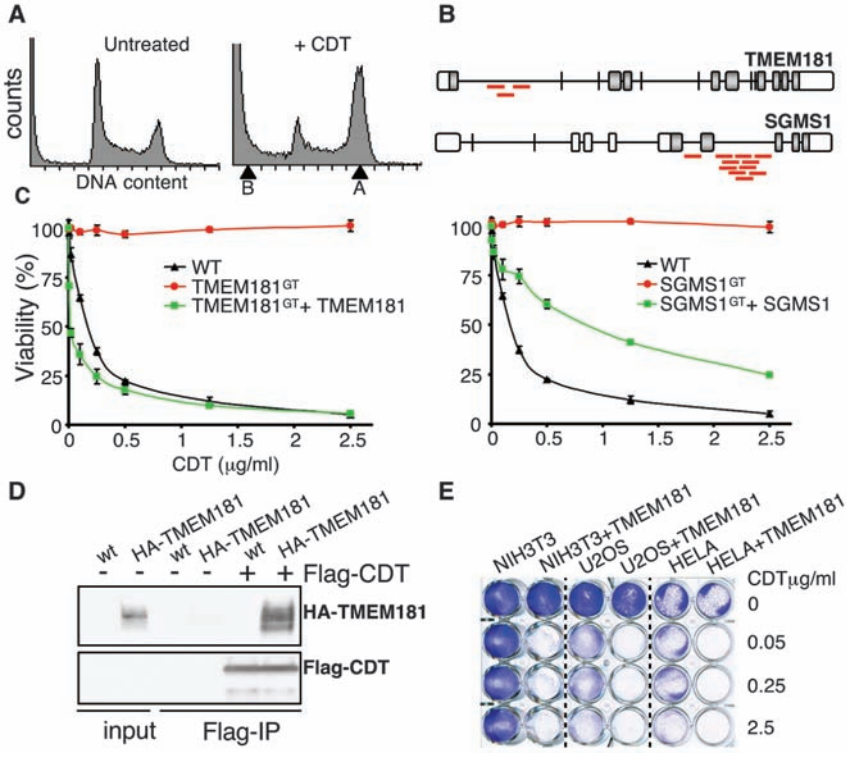

Fig. 2. Host factors for CDT. (**A**) Flow cytometric analysis of control KBM7 cells (left) and KBM7 cells after exposure to CDT purified from *E. coli* (right). Exposure of cells to CDT results in an increase of cells in the G_2/M phase of the cell cycle (arrow A) and cell death (arrow B). (**B**) Schematic outline of the insertion sites (red lines) in mutant cells unresponsive to CDT. (**C**) Resistance of TMEM181 mutant cells and SGMS1 mutant cells to CDT. Mutant cells reconstituted with the respective cDNAs re-acquire toxin sensitivity. Results are presented as mean values ± SD (error bars) (n = 3 biological replicates). (**D**) Flag-tagged CDT was bound to immobilized anti-Flag antibody and incubated with KBM7 cell lysates or lysates of cells expressing HA-TMEM181. Bound proteins were eluted and subjected to immunoblot analysis. Anti-Flag beads without bound CDT served as a control. (**E**) NIH3T3, U2OS, and HELA cells infected with a TMEM181-expressing retrovirus were treated with increasing amounts of CDT. After 5 days, adherent cells were stained with crystal violet.

stained for influenza nucleoprotein 12 hours after infection. As expected, KBM7 cells showed high levels of infection (~95% infection), whereas CMAS and SLC35A2 mutant cells showed near-complete resistance to virus infection (<0.01% infection) (Fig. 3B and fig. S5D). Absence of CMAS and SLC35A2 expression in the mutants was verified by reverse transcription polymerase chain reaction (RT-PCR) or immunoblot analysis (fig. S5, B and C). Transduction with cDNAs encoding the disrupted genes fully restored susceptibility to influenza infection (Fig. 3B), indicating that the observed resistance is attributable to the mutated loci. Although the KBM7 genome has not been screened at saturation for resistance to influenza, the transporter (SLC35A2) and enzyme (CMAS) identified here could lead to the development of antiviral therapies for influenza.

Identification of host factors for ADP-ribosylating toxins. Diphtheria and anthrax toxins are AB toxins composed of a cell-binding moiety (B) and an active (A) subunit that targets a host function to increase virulence. We have a detailed molecular understanding of how diphtheria toxin enters the cell and induces cell death (*19*, *20*). Can haploid genetic screens identify previously unidentified components in this well-characterized host/pathogen interaction? We screened mutagenized cells with diphtheria or anthrax toxin. Because native anthrax toxin is not cytotoxic for KBM7 cells, we exposed cells to the cell-binding component of anthrax toxin–protective antigen (PA) and anthrax lethal factor (LFN) fused to the catalytic domain of diphtheria toxin (LFN-DTA) (*21*). Resistant mutants were classified as either being resistant to anthrax toxin (PA-LFN-DTA) (class I), resistant to diphtheria toxin (class II), or resistant to both (class III). Mutants in the known anthrax toxin receptor (ANTXR2) (*22*) were recovered with 10 independent insertions and with 12 insertions for the known diphtheria toxin receptor [heparin-binding EGF-like growth factor (HB-EGF)] (Fig. 4A and fig. S6A) (*23*). The third class of mutants included genes involved in diphthamide biosynthesis [DPH1, DPH2, and DPH5; see (*20*)] and a previously uncharacterized gene named WDR85 (Fig. 4A and fig. 6A). All of these insertions were in the same transcriptional orientation as the mutated gene and were thus predicted to impair gene function. In the WDR85 mutant (hereafter referred to as WDR85^{GT}), no WDR85 transcripts were observed, as determined by RT-PCR (Fig. 4B). The resistance of WDR85^{GT} was readily complemented by transduction with WDR85 cDNA, which restored the sensitivity of WDR85^{GT} cells to diphtheria toxin and anthrax toxin (PA-LFN-DTA) (Fig. 4C and fig. S6B). WDR85^{GT} cells were also resistant to *Pseudomonas* exotoxin A, another adenosine diphosphate (ADP)–ribosylating toxin with a similar mode of action as diphtheria toxin (Fig. 4D). Although native anthrax toxin is not lethal to most cell types, including KBM7, its cellular entry and activity can be probed by monitoring cleavage of its cellular target MEK-3. WDR85^{GT} cells were still responsive to the native anthrax toxin, because the extent of proteolytic cleavage of MEK-3 was comparable for WDR85^{GT} and WT cells (fig. S7A), suggesting that toxin entry was normal in WDR85^{GT} cells.

WDR85 is part of the dipthamide biosynthetic pathway. Given the resistance of WDR85 mutant cells to different bacterial toxins, we further explored the mechanism by which WDR85 conferred sensitivity to toxin-mediated cell death. Diphtheria toxin, LFN-DTA, and exotoxin A inhibit host translation through ADP ribosylation of elongation factor 2 (EF2) and thus cause cell death (*19*). ADP ribosylation occurs on diphthamide, a posttranslationally modified histidine uniquely present in EF2 and conserved among all eukaryotes. As WDR85 was not required for toxin entry, we investigated EF2 ribosylation in response to diphtheria toxin. In cell lysates derived from WDR85^{GT} cells, EF2 ADP ribosylation was impaired and could be restored by re-expression of a WDR85 cDNA (fig. S7B). EF2 fused to a streptavidin-binding peptide (SBP) purified from WDR85^{GT} cells was also a poor substrate for ADP ribosylation in vitro (Fig. 5A). Impaired ADP ribosylation is thus an inherent property of EF2 derived from WDR85^{GT} cells and is not due to the presence or absence of other factors present in cell lysates. Diphthamide biosynthesis is the result of stepwise posttranslational modification of His^{175} (fig. S7D), the proteins responsible for which are known (*20*, *24*, *25*). The second step involves the trimethylation of "intermediate" EF2 by the methyltransferase DPH5, with *S*-adenosylmethionine as the methyl donor (*26*). To investigate if this methylation step was affected by the loss of WDR85, we purified intermediate EF2 from DPH5 null cells and performed in vitro methylation assays in cell lysates. Efficient methylation of intermediate EF2 by WT and WDR85^{GT} cell lysates suggested that WDR85 is not required for the second step of diphthamide biosynthesis (Fig. 5B). Next, we purified EF2 from WDR85^{GT} cells and used LC/MS/MS (liquid chromatography–tandem mass spectrometry) to monitor the relevant modifications of His^{175}. Modifications of His^{175} predict an increase in mass by +143 (diphthamide), +142 (diphthine), and +101 (the intermediate) mass units for those peptides that carry the modified His residue. SBP-tagged EF2 isolated from WDR85^{GT} showed a mass consistent with the presence of unmodified His^{175} (Fig. 5C), whereas modifications of EF2 purified from WT and DPH5 mutant cells showed a mass that was expected for the presence of diphthamide and intermediate, respectively (fig. S8, A and B). The absence of modified histidine in EF2 suggests that WDR85 plays a role in the first step in diphthamide biosynthesis.

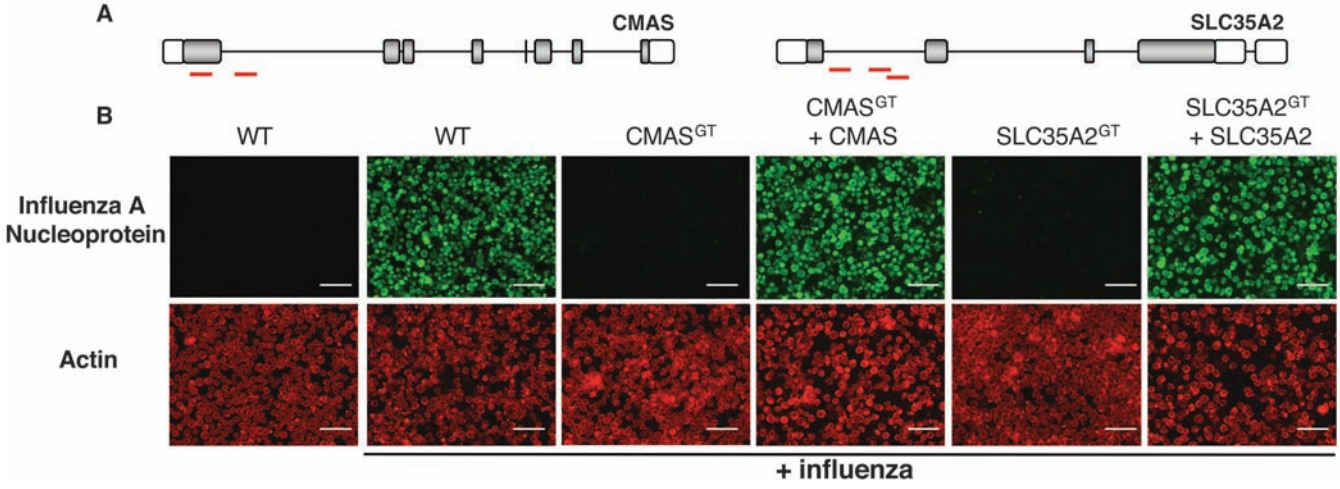

Fig. 3. Cellular genes required for influenza infection. (**A**) Schematic outline of the identified insertion sites (red lines) in the mutated genes. (**B**) Cells were exposed to influenza virus and stained 12 hours later with antibodies directed against influenza A nucleoprotein. Mutant cells reconstituted with cDNAs that correspond to the mutated gene products re-acquire virus sensitivity. Scale bars, 50 μm.

In the course of purification of EF2 from WDR85^{GT} cells, we detected a protein that strongly interacted with EF2 (Fig. 5D). Mass spectrometry and immunoblot analysis identified this protein as DPH5 (Fig. 5E and fig. S7C). WDR85 lacks homology to known proteins involved in diphthamide biosynthesis but does contain WD40 repeats, often involved in protein/protein interactions. Thus, WDR85 may serve as a scaffold to coordinate the association (or dissociation) of enzymatic complexes required for the stepwise biosynthesis of diphthamide.

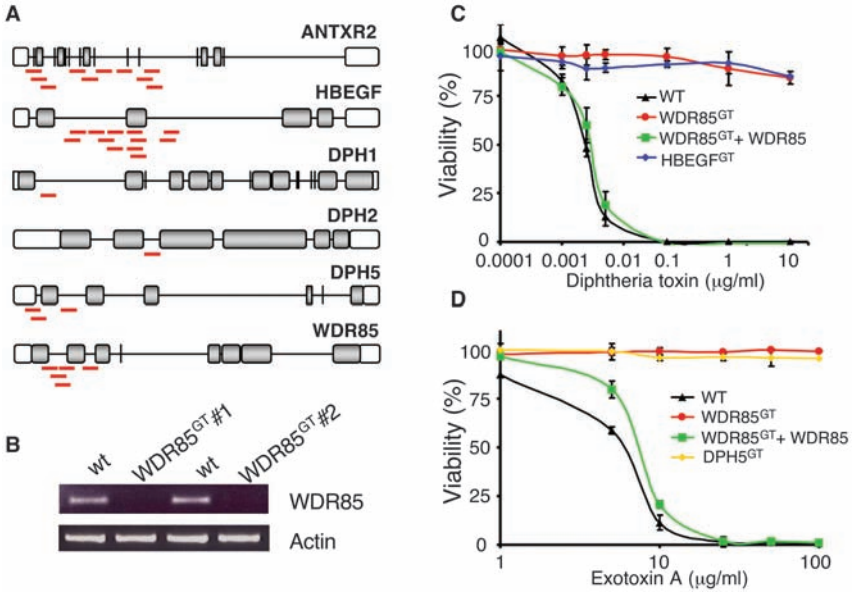

Fig. 4. Identification of loci that confer resistance to ADP-ribosylating bacterial toxins. (**A**) Schematic outline of the insertion sites (red lines) in the mutated genes. (**B**) RT-PCR for WDR85 shows undetectable WDR85 mRNA levels in independent clones with gene-trap insertions in the WDR85 locus. (**C**) Resistance of WDR85^{GT} cells to diphtheria toxin. Error bars indicate SD. (**D**) Resistance of WDR85^{GT} cells to exotoxin A. Identified clones with mutations in HB-EGF and DPH5 served as insensitive controls for these respective toxins, and WDR85^{GT} cells reconstituted with a WDR85 cDNA (WDR85^{GT} + WDR85) re-acquired sensitivity to the toxins. Results are presented as mean values ± SD (n = 3).

WDR85 is a conserved protein with homology to yeast YBR246W (fig. S9A). We used a database containing fitness profiles of deletion strains of all nonessential yeast genes under 1144 chemical conditions to cluster genes with similar profiles to YBR246W (*27*). The top 10 genes that phenoclustered with YBR246W by homozygous cosensitivity included DPH2 and DPH5 (fig. S9, B and C). The only gene annotation terms we found enriched concerned diphthamide biosynthesis [P value = 9×10^{-4} (fig. S9C)]. To directly test whether YBR246W is involved in diphthamide biosynthesis, we undertook ribosylation assays in protein extracts derived from WT yeast or yeast strains deleted for YKL191W (DPH2) or YBR246W. Deficiency of YKL191W and YBR246W both impaired ADP-ribosyl acceptor activity of EF2 in yeast (Fig. 5F). Thus, the role of WDR85 in diphthamide biosynthesis is conserved in eukaryotes, and the proposed scaffolding role may be the main function of WDR85 in cells. WDR85 therefore represents a host gene involved in the first step in diphthamide biosynthesis, despite previous suggestions that all proteins involved in this complex posttranslational modification were known (*20*).

Perspective. One of the main strengths of yeast as a genetic tool is the relative ease with which recessive mutations can be recognized and characterized at its haploid life stage (*28*), a trait absent from the commonly studied higher eukaryotes. The approach described here will allow similar types of screens for the human genome.

The discovery of RNA interference has enabled targeted reduction of gene expression in diploid cultured mammalian and insect cells, which

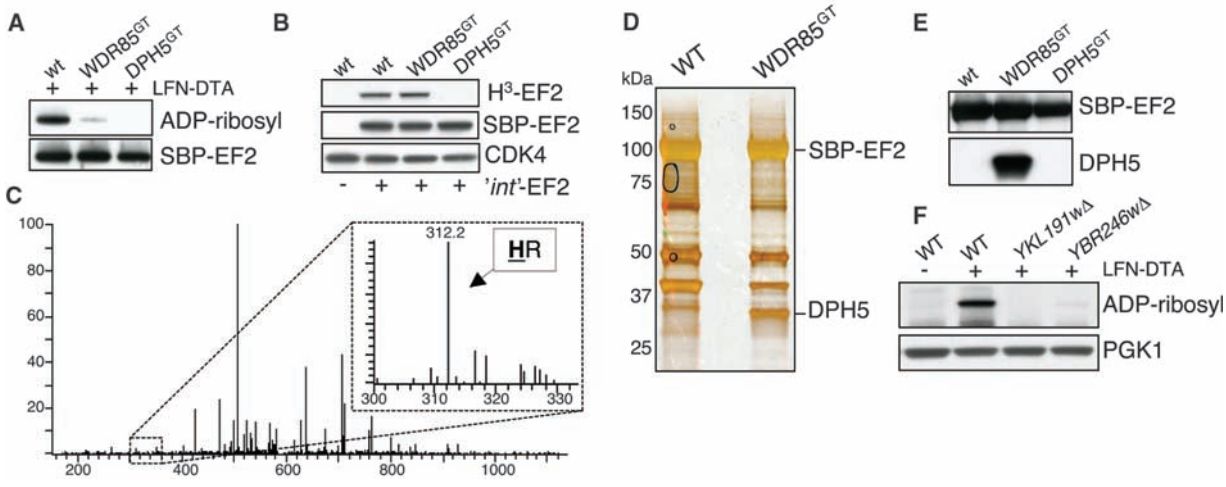

Fig. 5. WDR85 is involved in diphthamide biosynthesis. (**A**) In vitro ADP ribosylation of SBP-tagged EF2 purified from WT, WDR85, and DPH5 mutant cells by LFN-DTA in the presence of nicotinamide adenine dinucleotide (NAD)–biotin. Streptavidin–horseradish peroxidase (HRP) was used to detect ADP ribosylation, and total EF2 was detected by immunoblot analysis. (**B**) Methylation of intermediate EF2 by WT, WDR85, and DPH5 mutant cell lysates. SBP-tagged intermediate EF2 was purified from DPH5 mutant cells and incubated in lysates derived from the indicated genotypes in the presence of [methyl-^{3}H] adenosylmethionine (Ado-S-Me) as methyl donor. The amount of supplied intermediate EF2 was detected by immunoblot analysis, with CDK4 as a loading control. (**C**) MS/MS spectra of a tryptic peptide derived from SBP-tagged EF2 purified from WDR85 mutant cells. Peptide fragments characteristic for unmodified His^{715} are indicated. (**D**) Silver stain of SBP-EF2 purified from WT and WDR85-deficient cells. kDa, kilodaltons. (**E**) Immunoblot analysis of SBP-EF2 pull-down from WT, WDR85-, and DPH5-deficient cells probed with an antibody directed against DPH5. (**F**) Protein extracts from WT, YKL191W-, and YBR246W-deficient *S. cerevisiae* strains were incubated with LFN-DTA in the presence of NAD-biotin. Streptavidin-HRP was used to detect ADP ribosylation, and PGK1 was used as loading control.

opened the door to large-scale screens. At the same time, limitations of this approach are increasingly apparent, such as the induction of off-target effects that complicate genome-wide screens in particular (*29*, *30*) and the inability to completely switch off gene expression. When similar small interfering RNA screens are conducted independently in mammalian cells, the lack of concordance between them is an additional complicating factor (*31*, *32*). Finally, mammals are rather robust in their tolerance to partial loss of gene function: Haploinsufficiency appears to be the exception rather than the rule, because inactivation of one gene copy, as in heterozygous knockout mice, rarely leads to severe phenotypes.

Although we have focused on host-pathogen biology, similar screens could in principle be applied to any phenotype that can be recognized in a population of mutant cells, such as modulation of a genetically encoded reporter. In the future, haploid genetic screens could be used to generate comprehensive compendia of host factors that are used by different pathogens and may yield new strategies to combat infectious disease. In conclusion, the haploid genetic screens described here expand mutagenesis-based screens in model organisms by providing a window on disease-associated molecular networks that can be studied in cultured human cells.

References and Notes

1. H. J. Muller, *Science* **66**, 84 (1927).
2. A. L. Brass *et al.*, *Science* **319**, 921 (2008); published online 10 January 2008 (10.1126/science.1152725).
3. J. A. Philips, E. J. Rubin, N. Perrimon, *Science* **309**, 1251 (2005); published online 14 July 2005 (10.1126/science.1116006).
4. L. Hao *et al.*, *Nature* **454**, 890 (2008).
5. R. Salomon, R. G. Webster, *Cell* **136**, 402 (2009).
6. A. Moscona, *N. Engl. J. Med.* **360**, 953 (2009).
7. M. Kotecki, P. S. Reddy, B. H. Cochran, *Exp. Cell Res.* **252**, 273 (1999).
8. Materials and methods are available as supporting material on *Science* Online.
9. S. Nagata, *Cell* **88**, 355 (1997).
10. B. Luo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20380 (2008).
11. F. Cong *et al.*, *Mol. Cell* **6**, 1413 (2000).
12. M. Lara-Tejero, J. E. Galán, *Science* **290**, 354 (2000).
13. D. Nesic, Y. Hsu, C. E. Stebbins, *Nature* **429**, 429 (2004).
14. M. Wistrand, L. Kall, E. L. L. Sonnhammer, *Protein Sci.* **15**, 509 (2006).
15. M. Miyaji *et al.*, *J. Exp. Med.* **202**, 249 (2005).
16. L. Guerra *et al.*, *Cell. Microbiol.* **7**, 921 (2005).
17. B. Wollscheid *et al.*, *Nat. Biotechnol.* **27**, 378 (2009).
18. H. Sprong *et al.*, *Mol. Biol. Cell* **14**, 3482 (2003).
19. R. J. Collier, *Toxicon* **39**, 1793 (2001).
20. S. Liu, G. T. Milne, J. G. Kuremsky, G. R. Fink, S. H. Leppla, *Mol. Cell. Biol.* **24**, 9487 (2004).
21. J. C. Milne, S. R. Blanke, P. C. Hanna, R. J. Collier, *Mol. Microbiol.* **15**, 661 (1995).
22. H. M. Scobie, G. J. A. Rainey, K. A. Bradley, J. A. T. Young, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5170 (2003).
23. J. G. Naglich, J. E. Metherall, D. W. Russell, L. Eidels, *Cell* **69**, 1051 (1992).
24. L. C. Mattheakis, W. H. Shen, R. J. Collier, *Mol. Cell. Biol.* **12**, 4026 (1992).
25. S. Liu, S. H. Leppla, *Mol. Cell* **12**, 603 (2003).
26. J. Y. Chen, J. W. Bodley, *J. Biol. Chem.* **263**, 11692 (1988).
27. M. E. Hillenmeyer *et al.*, *Science* **320**, 362 (2008).
28. S. L. Forsburg, *Nat. Rev. Genet.* **2**, 659 (2001).
29. Y. Ma, A. Creanga, L. Lum, P. A. Beachy, *Nature* **443**, 359 (2006).
30. C. J. Echeverri *et al.*, *Nat. Methods* **3**, 777 (2006).
31. S. P. Goff, *Cell* **135**, 417 (2008).
32. F. D. Bushman *et al.*, *PLoS Pathog.* **5**, e1000437 (2009).

33. We thank D. Sabatini, S. Nijman, J. Roix, and J. Pruszak for discussion and critical review of the manuscript; C. Y. Wu and G. Fink for yeast deletion strains; J. Kaper for the CDT expression plasmid; J. Collier, R. Moon, and M. Wernig for plasmids; and E. Guillen for help with influenza infections. C.P.G. has a fellowship from Fundacao Ciencia Tecnologia, Portugal. T.R.B. was funded by the Kimmel Foundation and the Whitehead Institute Fellows Program. The Whitehead Institute has filed a patent on the application of gene-trap mutagenesis in haploid or near-haploid cells to identify human genes that affect cell phenotypes, including host factors used by pathogens.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1231/DC1
Materials and Methods
Figs. S1 to S9
References

10 July 2009; accepted 5 October 2009
10.1126/science.1178955

Proteome Organization in a Genome-Reduced Bacterium

Sebastian Kühner,[1]* Vera van Noort,[1]* Matthew J. Betts,[1] Alejandra Leo-Macias,[1] Claire Batisse,[1] Michaela Rode,[1] Takuji Yamada,[1] Tobias Maier,[2] Samuel Bader,[1] Pedro Beltran-Alvarez,[1] Daniel Castaño-Diez,[1] Wei-Hua Chen,[1] Damien Devos,[1] Marc Güell,[2] Tomas Norambuena,[3] Ines Racke,[1] Vladimir Rybin,[1] Alexander Schmidt,[4] Eva Yus,[2] Ruedi Aebersold,[4] Richard Herrmann,[5] Bettina Böttcher,[1]† Achilleas S. Frangakis,[1] Robert B. Russell,[1] Luis Serrano,[2,6] Peer Bork,[1]‡ Anne-Claude Gavin[1]‡

The genome of *Mycoplasma pneumoniae* is among the smallest found in self-replicating organisms. To study the basic principles of bacterial proteome organization, we used tandem affinity purification–mass spectrometry (TAP-MS) in a proteome-wide screen. The analysis revealed 62 homomultimeric and 116 heteromultimeric soluble protein complexes, of which the majority are novel. About a third of the heteromultimeric complexes show higher levels of proteome organization, including assembly into larger, multiprotein complex entities, suggesting sequential steps in biological processes, and extensive sharing of components, implying protein multifunctionality. Incorporation of structural models for 484 proteins, single-particle electron microscopy, and cellular electron tomograms provided supporting structural details for this proteome organization. The data set provides a blueprint of the minimal cellular machinery required for life.

Biological function arises in part from the concerted actions of interacting proteins that assemble into protein complexes and networks. Protein complexes are the first level of cellular proteome organization: functional and structural units, often termed molecular machines, that participate in all major cellular processes. Complexes are also highly dynamic in the sense that their organization and composition vary in time and space (*1*), and they interact to form higher level networks; this property is central to whole-cell functioning. However, general rules concerning protein complex assembly and dynamics remain elusive.

The combination of affinity purification with mass spectrometry (MS) (*2*) has been applied to several organisms to provide a growing repertoire of molecular machines. Genome-wide screens in *Saccharomyces cerevisiae* (*3*–*5*) captured discrete, dynamic proteome organization and revealed higher-order assemblies with direct connections between complexes and frequent sharing of common components. To date these exhaustive analyses have been applied only in yeast. In bacteria, genome-wide yeast two-hybrid analyses have been reported (*6*, *7*), but only a few biochemical analyses on selected sets of complexes are available (*8*–*11*). The understanding of proteome organization in these organisms concerns thus the binary interaction networks.

Here, we report a genome-scale analysis of protein complexes in the bacterium *Mycoplasma pneumoniae*, a human pathogen that causes atypical

[1]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [2]Centro Regulacion Genomica–Universidad Pompeu Fabra, Dr Aiguader 88, 08003 Barcelona, Spain. [3]Pontificia Universidad Catolica de Chile, Alameda 340, Santiago, Chile. [4]ETH (Eidgenössische Technische Hochschule) Zürich, Wolfgang-Pauli-Strasse 16, 8093 Zürich, Switzerland; Faculty of Science, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland, and Institute for Systems Biology, Seattle, WA 98013, USA. [5]ZMBH (Zentrum für Molekulare Biologie der Universität Heidelberg), Im Neuenheimer Feld 282, 69120 Heidelberg, Germany. [6]ICREA (Institució Catalana de Recerca i Estudis Avançats), 08010 Barcelona, Spain.

*These authors contributed equally to this work.
†Present address: University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JR.
‡To whom correspondence should be addressed. E-mail: gavin@embl.de (A.-C.G.); bork@embl.de (P.B.)