LETTER

doi:10.1038/nature22376

Genetic wiring maps of single-cell protein states reveal an off-switch for GPCR signalling

Markus Brockmann[1]*, Vincent A. Blomen[1]*, Joppe Nieuwenhuis[1], Elmer Stickel[1], Matthijs Raaben[1], Onno B. Bleijerveld[1], A. F. Maarten Altelaar[1,2], Lucas T. Jae[1]† & Thijn R. Brummelkamp[1,3,4]

As key executers of biological functions, the activity and abundance of proteins are subjected to extensive regulation. Deciphering the genetic architecture underlying this regulation is critical for understanding cellular signalling events and responses to environmental cues. Using random mutagenesis in haploid human cells, we apply a sensitive approach to directly couple genomic mutations to protein measurements in individual cells. Here we use this to examine a suite of cellular processes, such as transcriptional induction, regulation of protein abundance and splicing, signalling cascades (mitogen-activated protein kinase (MAPK), G-protein-coupled receptor (GPCR), protein kinase B (AKT), interferon, and Wingless and Int-related protein (WNT) pathways) and epigenetic modifications (histone crotonylation and methylation). This scalable, sequencing-based procedure elucidates the genetic landscapes that control protein states, identifying genes that cause very narrow phenotypic effects and genes that lead to broad phenotypic consequences. The resulting genetic wiring map identifies the E3-ligase substrate adaptor KCTD5 (ref. 1) as a negative regulator of the AKT pathway, a key signalling cascade frequently deregulated in cancer. KCTD5-deficient cells show elevated levels of phospho-AKT at S473 that could not be attributed to effects on canonical pathway components. To reveal the genetic requirements for this phenotype, we iteratively analysed the regulatory network linked to AKT activity in the knockout background. This genetic modifier screen exposes suppressors of the KCTD5 phenotype and mechanistically demonstrates that KCTD5 acts as an off-switch for GPCR signalling by triggering proteolysis of Gβγ heterodimers dissociated from the Gα subunit. Although biological networks have previously been constructed on the basis of gene expression[2,3], protein–protein associations[4–6], or genetic interaction profiles[7,8], we foresee that the approach described here will enable the generation of a comprehensive genetic wiring map for human cells on the basis of quantitative protein states.

Genetic perturbation approaches have been employed to study the wiring of human cells using (trans)gene expression[9–12] or cell viability as readouts[13]. However, many cellular events do not affect cell fitness or result in large transcriptional effects. Using protein states as readouts for phenotypes[12,14,15], we apply random mutagenesis in haploid cells to directly link genome mutations to protein phenotypes within the same cell.

Induction of the interferon-regulatory factor 1 (IRF1) by interferon-γ is under strong genetic control. We used this as a proof of concept to determine whether the genomic elements required for this phenotype could be revealed. A population of 10^8 gene-trap mutagenized human haploid HAP1 cells[13,16,17] were expanded, fixed, permeabilized, and stained using antibodies directed against IRF1 before fluorescence-activated cell sorting (FACS) (Fig. 1a). Two populations of 1.5×10^7 cells, corresponding to the cells with the highest and lowest IRF1 protein abundance, were isolated and used for mapping the gene-trap integrations. Because of the high complexity of the library, the effect of genes on a measured phenotype could be assessed by counting the number of unique genomic mutations in each channel rather than by measuring the abundance of few distinct mutants. Approximately 1.5×10^7 gene-trap integration events were mapped in both populations, resulting in approximately one cell accounting for each mapped mutation. This pointed out all described pathway components[18] including *STAT1*, for which a high number of disruptive mutations were identified in the 'low' population and few in the 'high' population (135 versus 2 independent mutations, false discovery rate-corrected P value $= 4.3 \times 10^{-37}$) and additional regulators (Fig. 1b, Extended Data Fig. 1a and Supplementary Table 1).

To identify genetic regulators affecting different protein fates, we used a series of specific antibodies to study protein phosphorylation (p38, ERK), splicing (XBP1), methylation (H3K27), crotonylation (H2AK119), glycosylation (LAMP1), and proteolysis (β-catenin) (Fig. 1c–i, Extended Data Fig. 1b and Supplementary Tables 2–8). In all cases, mutations in the gene encoding, or required for generation of, the antibody target were expectedly enriched in the 'low' cell population. For example, we identified the PRC2 complex (EED, EZH2, and SUZ12; Fig. 1f), which mediates trimethylation of H3K27 (ref. 19), and the acetyltransferase EP300 (Fig. 1g), which was shown to transfer a crotonyl mark on histones[20]. In addition to the genes required for the respective antibody targets, numerous factors known to affect the trait of interest were identified. These included a series of lysosomal storage disease genes that affect the abundance of the lysosomal marker LAMP1, which is in agreement with the increased number of enlarged lysosomes observed in cells derived from patients with lysosomal storage disease (Fig. 1h)[21]. In summary, genetic regulators affecting a variety of protein states in human cells could be identified.

In addition to previously identified regulators, these genetic surveys identified numerous loci affecting the protein phenotypes studied, suggesting that these traits could be controlled by extensive genetic networks. Assuming only expressed genes could affect a phenotype, we determined gene expression levels in HAP1 cells by RNA sequencing (RNA-seq) (Supplementary Table 9) and examined the identified regulators for each query phenotype. Importantly, the 25% lowest or non-expressed genes rarely affected phenotypes, suggesting few false positives among the identified genetic regulators (Extended Data Fig. 2). Next, we classified the identified regulators across all query phenotypes. This showed that some regulators affect a single measured trait whereas others display a broader phenotypic range (Fig. 2a). Examining functionally related genes linked to multiple phenotypes, we noticed similar behaviour across screens (Fig. 2a). We observed that the most strongly enriched Reactome term for genes connected

[1]Netherlands Cancer Institute, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands. [2]Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Padualaan 8, 3584CH Utrecht, The Netherlands. [3]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria. [4]Cancergenomics.nl, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands. †Present address: Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Straße 25, 81377 Munich, Germany.
*These authors contributed equally to this work.

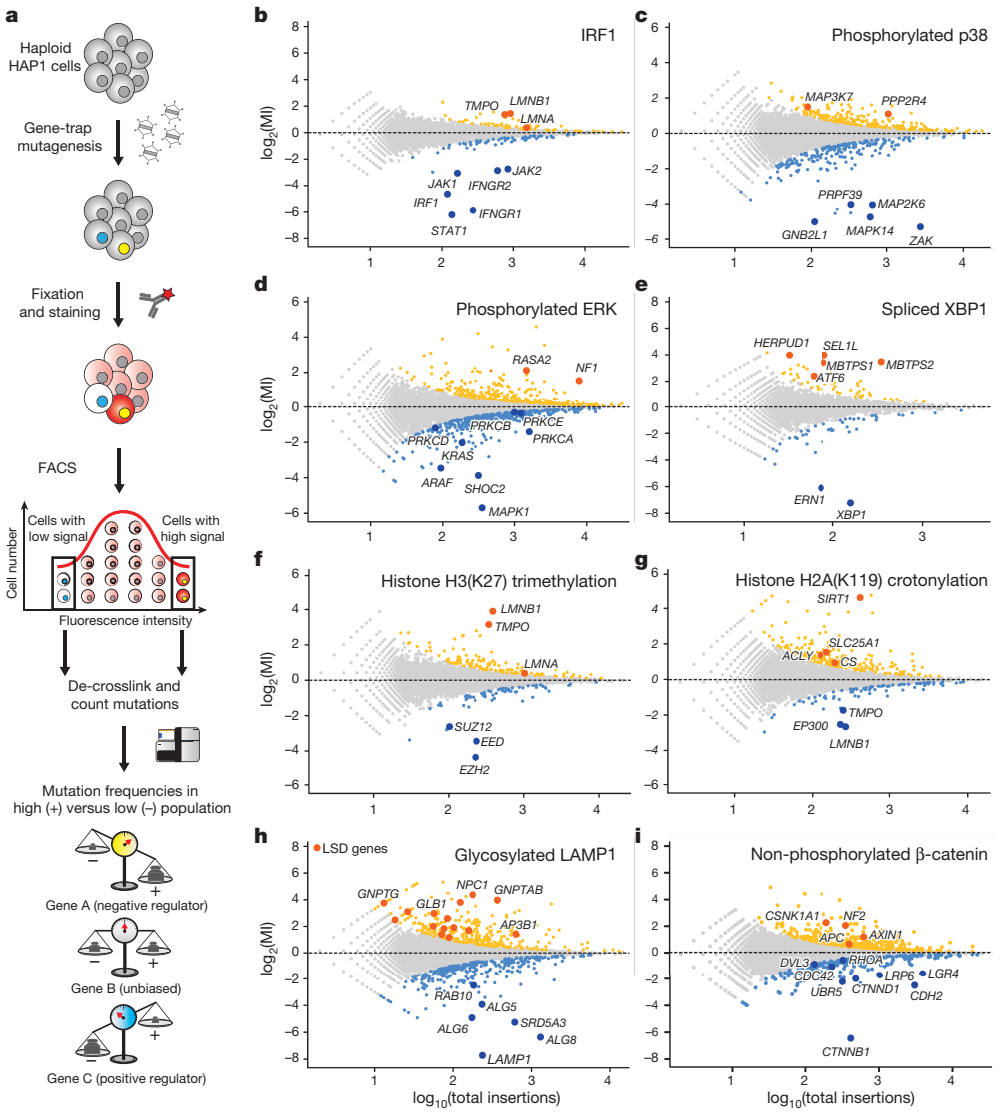


Figure 1 | Genetic wiring maps for protein phenotypes measured in cultured human cells. a, Mutagenesis-based approach in haploid cells to study protein phenotypes (see Supplementary Table 16). **b–i**, Genetic screens using antibodies for the indicated targets. Per gene (dots), the frequency of mutations in the 'high' channel divided by the frequency of mutations in the 'low' channel is plotted as the mutation index (MI, *y* axis) against the total number of mutations assigned to the gene (*x* axis). Genes enriched in either channel (two-sided Fisher's exact test, false discovery rate-corrected $P \leq 0.05$) are coloured (yellow, negative regulators; blue, positive regulators) and selected regulators are labelled; **b**, induced IRF1, **c**, phosphorylated p38(Thr180/Tyr182); **d**, phosphorylated ERK(Thr202/Tyr204); **e**, spliced XBP1; **f**, trimethylated histone H3(K27); **g**, crotonylated histone H2A(K119); **h**, glycosylated LAMP1 (LSD, lysosomal storage disease); **i**, non-phosphorylated (active) β-catenin (Ser33/37/Thr41). Interactive graphs are available at https://phenosaurus.nki.nl.

to more than one phenotype was chromatin biology (Fig. 2b), as seen for hubs in genetic interaction networks[7], and more connected genes engaged more frequently in protein–protein interactions or affected cell fitness (Extended Data Fig. 3). Conversely, Reactome terms related to the specific query trait were identified for most regulators connected to only one trait (Fig. 2b). Finally, we assigned human genes in groups of approximately 1,000 on the basis of expression levels in HAP1 cells. We observed that transcription levels of a gene related to its likelihood of being found as a regulator (Extended Data Fig. 4). Together, these results suggest that protein outputs are under complex genetic control, that this regulation is disproportionally elicited by the most highly expressed genes, and that regulators can be classified by their phenotypic range.

The identification of genes affecting cellular phenotypes opens the possibility of expanding known biological pathways. The AKT cascade affects many aspects of cell biology such as proliferation, cell survival, and metabolism[22]. Expectedly, many known pathway components were readily identified in a genetic screen measuring AKT phosphorylation at S473 (Fig. 2c, Extended Data Fig. 5 and Supplementary Table 10). Notably, the cullin E3 ligase adaptor KCTD5 (ref. 1) (but no other KCTD family member; Extended Data Fig. 6) was identified as a potent negative regulator together with CUL3 itself, and their observed phenotypes could be verified in different human cell lines using immunoblot analyses (Extended Data Fig. 7a, b). Because KCTD5-deficient cells displayed elevated phospho-AKT (pAKT) levels that could not be explained by effects on the examined canonical pathway components (Extended Data Fig. 7c), we applied a genetic suppressor approach to dissect the underlying mechanism (Fig. 2d) by using the abundance of pAKT as a readout in KCTD5-deficient HAP1 cells (Fig. 2e and Supplementary Table 11). As expected, *KCTD5* was not identified, but importantly neither was *CUL3*, indicating that the effect of CUL3 on pAKT requires KCTD5. Notably, when regulators of pAKT levels were compared between wild-type and KCTD5-deficient cells, a set of genes was identified that represented potent drivers of pAKT levels in KCTD5-deficient but not wild-type cells (Fig. 2c, e, f). These corresponded to G-protein β and γ subunits (*GNB1*, *GNB2*, and *GNG5*), their chaperone *PDCL*[23],

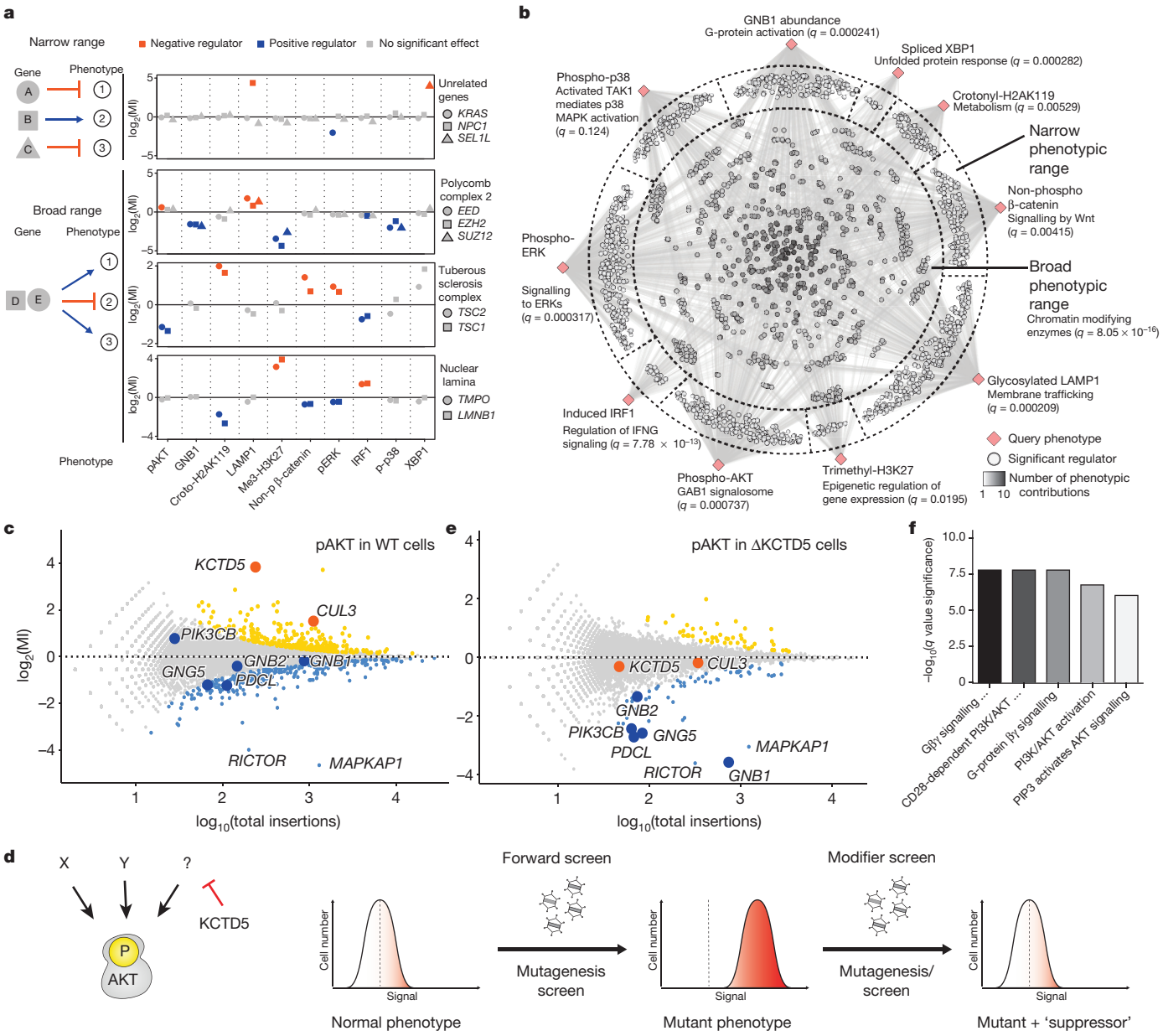


Figure 2 | Protein phenotypes are regulated by extensive genetic networks and can be influenced by suppressor interactions. a, Mutation index of selected genes is plotted per screen (red, negative regulator; blue, positive regulator; grey, not significant). **b**, Network depicting query phenotypes (orange diamonds) connected to identified regulators (nodes). Unique regulators are grouped in the outer rim (indicating selected Reactome terms), while genes connected to multiple phenotypes are grouped towards the centre. **c**, Genetic screen for pAKT. **d**, Genetic suppressor approach to study KCTD5 mechanism. **e**, Screen for pAKT in ΔKCTD5 cells. **f**, Top Reactome terms identified specifically in KCTD5-deficient cells. PIP3, phosphatidylinositol-3,4,5-trisphosphate.

required to generate Gβγ dimers, and their effector PI3KCB[24,25]. Activation of PI3KCB by Gβγ is known to activate AKT signalling. This genetic modifier screen suggests that the activation of AKT by Gβγ is suppressed by KCTD5/CUL3.

We then used label-free quantitative proteomics to compare wild-type and KCTD5-deficient cells to identify substrates of KCTD5/CUL3. Hypothesizing that KCTD5 substrates are ubiquitinated and subsequently degraded, we searched for potential substrates and identified 70 proteins (out of a total of 3,701 quantified proteins; Supplementary Table 12) with increased abundance in KCTD5-deficient cells (Fig. 3a and Supplementary Table 12). In parallel, we used the ubiquitin remnant motif antibody (di-Gly) to purify and quantify ubiquitinated proteins, yielding 217 downregulated ubiquitination sites (out of a total of 2,417 quantified sites) (Fig. 3b and Supplementary Table 12). Comparison of both datasets yielded only four proteins that showed fewer di-Gly remnants in KCTD5- deficient cells and increased protein levels. Remarkably, three out of four hits encoded Gβγ subunits (*GNB1*, *GNB2*, and *GNG5*) (Fig. 3c and Supplementary Table 13). Western blot analysis confirmed that both KCTD5 and CUL3 knockout cells contained elevated levels of GNB1 and GNG5 (Fig. 3d and Extended Data Fig. 8a). Together with the genetic suppressor screen, these data suggest a model in which KCTD5 limits pAKT activation through ubiquitin-mediated proteolysis of Gβγ.

To determine the circumstances under which KCTD5 mediates degradation of Gβγ, we designed a comparative genetic strategy: using our FACS-based screening approach, we identified genes affecting GNB1 abundance both in wild-type and in KCTD5-deficient HAP1 cells and searched for genotype-specific regulators. *GNB1* was strongly enriched for disruptive mutations in the 'low' channel and KCTD5 stood out as a strong negative regulator in the wild-type screen, independently

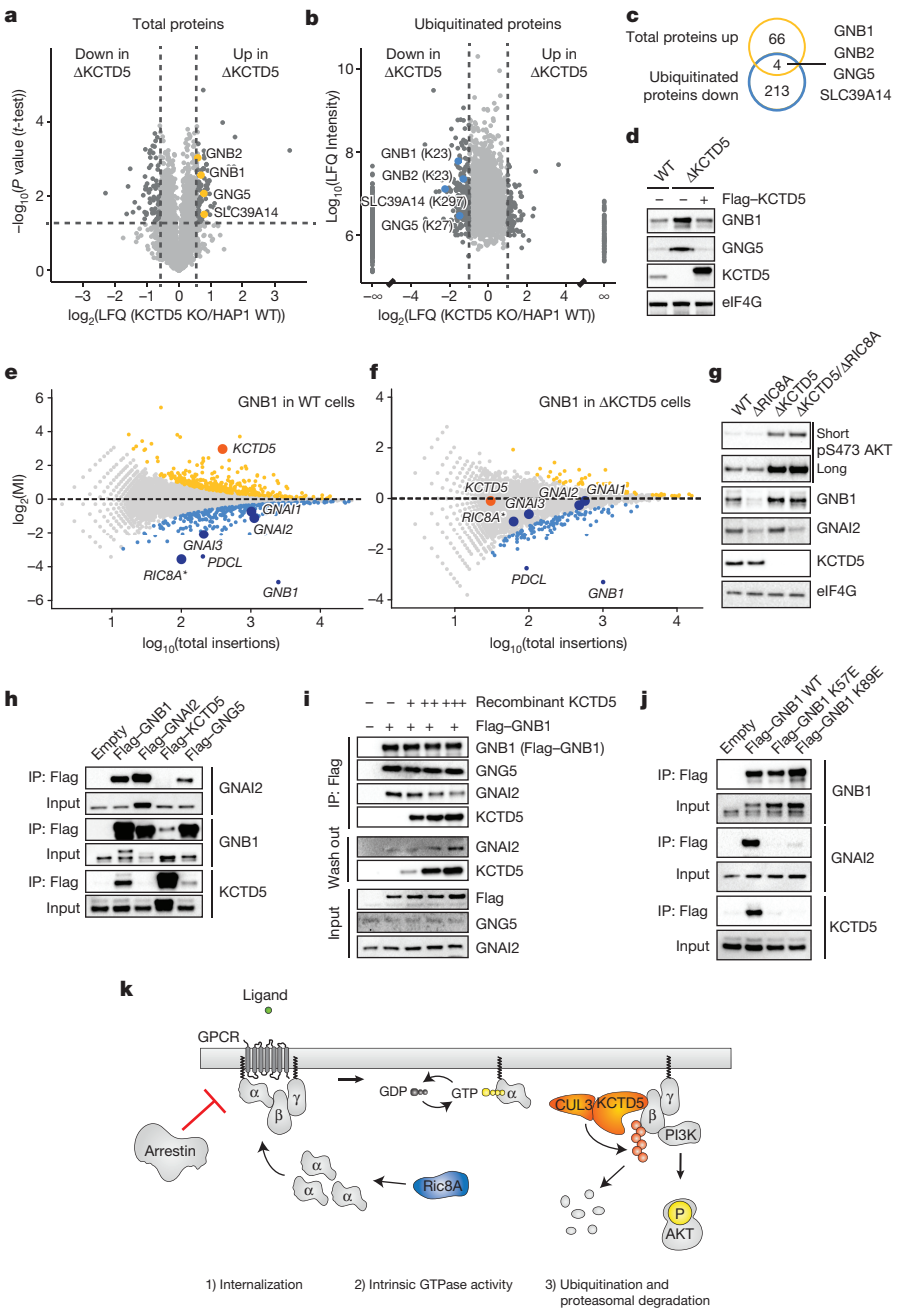


Figure 3 | KCTD5 acts as off-switch for GPCR Gβγ signalling. **a**, **b**, Label-free mass spectrometry identifies Gβγ subunits with increased abundance (**a**) and decreased ubiquitination (**b**) in Δ*KCTD5* cells. **c**, Intersection of data from **a** and **b**. **d**, Lysates of wild-type, KCTD5-deficient, and reconstituted HEK293T cells analysed by immunoblotting. **e**, **f**, Screens for GNB1 abundance in wild-type (**e**) and KCTD5-deficient HAP1 cells (**f**). Selected differentially acting genes are labelled. For RIC8A*, transcript uc001lof.3 was used (Extended Data Fig. 8b). **g**, Lysates of indicated HAP1 genotypes analysed by immunoblotting. **h**, HEK293T cells expressing the indicated genes were subjected to Flag immunoprecipitation (IP) and immunoblotting. **i**, KCTD5-deficient HEK293T cells expressing Flag–GNB1 were lysed, immobilized, and treated with increasing amounts of recombinant KCTD5. Input samples, supernatant (wash out), and precipitates were subjected to immunoblot analysis. **j**, HEK293T cells expressing the indicated genes were lysed, immunoprecipitated, and analysed as before. **k**, Model of Gβγ regulation by KCTD5.

confirming that it acts as an important regulator of GNB1 abundance (Fig. 3e and Supplementary Table 14). Notably, intersection of the wild-type and Δ*KCTD5* datasets revealed that the Gα subunits *GNAI1*, *GNAI2*, *GNAI3*, and *RIC8A* (encoding a protein required for Gα subunit abundance)[26] specifically regulated GNB1 levels in wild-type cells (Fig. 3f and Supplementary Table 15). This could be confirmed by western blot analysis using cells deficient for RIC8A in the absence or presence of KCTD5 (Fig. 3g). Because loss of Gα subunits decreases GNB1 levels in a KCTD5-dependent manner, it is possible that KCTD5 specifically targets those Gβγ dimers dissociated from Gα subunits for degradation. Indeed, KCTD5 associated with Gβγ subunits but could not be detected in complex with purified Gα subunits (Fig. 3h). Moreover, the addition of recombinant KCDT5 protein to immobilized Gαβγ resulted in binding of KCTD5 to the Gβγ and simultaneous dissociation of Gα, demonstrating that Gα and KCTD5 bind to Gβγ dimers competitively (Fig. 3i). Finally, mutations in Gβ subunits have been observed in a variety of solid cancers and myeloid malignancies[27]. These oncogenic Gβ subunits potently trigger AKT phosphorylation because of their inability to interact with Gα[27]. Intriguingly, those oncogenic GNB1 point mutations also attenuated

KCTD5 binding (Fig. 3j). We propose that this attenuated binding contributes to the oncogenic activity of GNB1 point mutants, since liberation of wild-type Gβγ through Gα depletion (as in *RIC8A* knockout cells) did not increase AKT phosphorylation when KCTD5 was present (Fig. 3g).

GPCRs are the largest receptor family mediating many physiological responses. Their activity is tightly controlled and reset upon activation. Although the activity of Gβγ dimers can be neutralized through re-association with Gα subunits, we here present an unexpected proteolysis-based mechanism, which irreversibly removes dissociated Gβγ dimers, limiting activation of their effector molecule PI3K (Fig. 3k). Although other KCTD proteins have been linked to GPCR signalling before[28], these were associated with specific receptor sub-complexes and did not trigger proteolysis because of their inability to interact with CUL3[29,30].

By examining a variety of protein-related phenotypes, we reveal extensive underlying genetic networks enriched for highly transcribed genes. The identification of genetic suppressors in modifier screens can readily provide mechanistic insights into the investigated phenotypes. Together, these genetic explorations applicable to any specific antibody may facilitate the generation of a phenotypic map for human cells.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 December 2016; accepted 13 April 2017.
Published online 31 May 2017.

1. Bayón, Y. *et al.* KCTD5, a putative substrate adaptor for cullin3 ubiquitin ligases. *FEBS J.* **275,** 3900–3910 (2008).
2. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95,** 14863–14868 (1998).
3. Hughes, T. R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102,** 109–126 (2000).
4. Huttlin, E. L. *et al.* The BioPlex network: a systematic exploration of the human interactome resource. *Cell* **162,** 425–440 (2015).
5. Havugimana, P. C. *et al.* A census of human soluble protein complexes. *Cell* **150,** 1068–1081 (2012).
6. Hein, M. Y. *et al.* A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163,** 712–723 (2015).
7. Costanzo, M. *et al.* A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353,** aaf1420 (2016).
8. Tong, A. H. Y. *et al.* Global mapping of the yeast genetic interaction network. *Science* **303,** 808–813 (2004).
9. Parts, L. Genome-wide mapping of cellular traits using yeast. *Yeast* **31,** 197–205 (2014).
10. Jaitin, D. A. *et al.* Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* **167,** 1883–1896.e15 (2016).
11. Dixit, A. *et al.* Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic resource. *Cell* **167,** 1853–1866.e17 (2016).
12. Parnas, O. *et al.* A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162,** 675–686 (2015).
13. Blomen, V. A. *et al.* Gene essentiality and synthetic lethality in haploid human cells. *Science* **350,** 1092–1096 (2015).
14. Albert, F. W., Treusch, S., Shockley, A. H., Bloom, J. S. & Kruglyak, L. Genetics of single-cell protein abundance variation in large yeast populations. *Nature* **506,** 494–497 (2014).
15. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
16. Carette, J. E. *et al.* Haploid genetic screens in human cells identify host factors used by pathogens. *Science* **326,** 1231–1235 (2009).
17. Carette, J. E. *et al.* Ebola virus entry requires the cholesterol transporter Niemann–Pick C1. *Nature* **477,** 340–343 (2011).
18. Platanias, L. C. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat. Rev. Immunol.* **5,** 375–386 (2005).
19. Lund, A. H. & van Lohuizen, M. Polycomb complexes and silencing mechanisms. *Curr. Opin. Cell Biol.* **16,** 239–246 (2004).
20. Sabari, B. R. *et al.* Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol. Cell* **58,** 203–215 (2015).
21. Meikle, P. J. *et al.* Diagnosis of lysosomal storage disorders: evaluation of lysosome-associated membrane protein LAMP-1 as a diagnostic marker. *Clin. Chem.* **43,** 1325–1335 (1997).
22. Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307,** 1098–1101 (2005).
23. Lukov, G. L., Hu, T., McLaughlin, J. N., Hamm, H. E. & Willardson, B. M. Phosducin-like protein acts as a molecular chaperone for G protein βγ dimer assembly. *EMBO J.* **24,** 1965–1975 (2005).
24. Stephens, L. *et al.* A novel phosphoinositide 3 kinase activity in myeloid-derived cells is activated by G protein βγ subunits. *Cell* **77,** 83–93 (1994).
25. Leopoldt, D. *et al.* Gβγ stimulates phosphoinositide 3-kinase-γ by direct interaction with two domains of the catalytic p110 subunit. *J. Biol. Chem.* **273,** 7024–7029 (1998).
26. Oner, S. S. *et al.* Regulation of the G-protein regulatory-$G\alpha_i$ signaling complex by nonreceptor guanine nucleotide exchange factors. *J. Biol. Chem.* **288,** 3003–3015 (2013).
27. Yoda, A. *et al.* Mutations in G protein β subunits promote transformation and kinase inhibitor resistance. *Nat. Med.* **21,** 71–75 (2015).
28. Schwenk, J. *et al.* Native GA_BA_B receptors are heteromultimers with a family of auxiliary subunits. *Nature* **465,** 231–235 (2010).
29. Smaldone, G. *et al.* Cullin 3 recognition is not a universal property among KCTD proteins. *PLoS ONE* **10,** e0126808 (2015).
30. Turecek, R. *et al.* Auxiliary GABAB receptor subunits uncouple G protein βγ subunits from effector channels to induce desensitization. *Neuron* **82,** 1032–1044 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Goedhart, L. Wessels, B. van Steensel, S. Nijman, and members of the Brummelkamp, Perrakis, and Sixma laboratories for discussions. We thank R. Spaapen for providing *CUL3* knockout cells, P. Celie and M. Stadnik for assistance with the recombinant protein expression, as well as E. Fessler and J. P. Medema for generation of WNT3A/R-spondin-conditioned medium. This work was supported by the Dutch Cancer Society (NKI 2015-7609), the Cancer Genomics Center, an Ammodo KNAW Award 2015 for Biomedical Sciences to T.R.B., by the Netherlands Organization for Scientific Research (NWO) as part of the National Roadmap Large-scale Research Facilities of the Netherlands, Proteins@Work (project number 184.032.201) to O.B.B. and A.F.M.A., and by a Vidi grant (723.012.102) to A.F.M.A.

Author Contributions M.B., V.A.B., J.N., M.R. L.T.J., and T.R.B. were responsible for the overall design of the study. V.A.B. and E.S. performed the bioinformatics. E.S. developed the Phenosaurus platform. O.B.B. and A.F.M.A. designed, performed, and analysed the proteomics experiments. M.B., V.A.B., L.T.J., and T.R.B. wrote the manuscript; all authors commented on it.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to L.T.J. (jae@genzentrum.lmu.de) or T.R.B. (t.brummelkamp@nki.nl).

Reviewer Information *Nature* thanks J. Moffat and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Cell lines and culture. HAP1 cells[17] were cultured in IMDM-medium (ThermoFisher Scientific) supplemented with 10% heat-inactivated fetal calf serum (FCS; ThermoFisher Scientific or Sigma) and penicillin–streptomycin–glutamine solution (ThermoFisher Scientific). HEK293T-, SKBR3-, A549-, and U2OS cells were obtained from authentic stocks (American Type Culture Collection) and maintained in DMEM (ThermoFisher Scientific) containing the aforementioned supplements. HAP1 wild-type and knockout cell lines were monitored for ploidy, authenticated by genotyping, and all lines were tested for mycoplasma and verified to be negative.

Generation of knockout cell lines. Knockout cell lines (Supplementary Table 17) were generated by using the CRISPR/Cas9 system. HAP1 cells were transfected with the plasmid px330 (Addgene 42230) containing the guide RNA (gRNA) for a gene of interest (Supplementary Table 18) and a vector containing a gRNA to the zebrafish *TIA* gene (5′-GGTATGTCGGGAACCTCTCC-3′) as well as a P2A-blasticidin resistance cassette flanked by two *TIA* target sites. This allowed incorporation of the blasticidin resistance gene into the locus, resulting in a stable knockout similarly as described[31]. After blasticidin selection (10 μg ml^{-1}), resistant clones were expanded. To create double knockouts, *KCTD5*-knockout cells were transfected with two gRNAs targeting RIC8A.

In contrast to HAP1 cells, HEK293T cells were co-transfected with two different gRNAs (px330) and pMX-ires-Blast, followed by blasticidin selection (80 μg ml^{-1}).

gRNAs against KCTD5 were also cloned into a lentiviral red fluorescent protein (RFP)–CRISPR backbone (Addgene 75162). SKBR3-, A549-, and U2OS cell lines were infected with lentiviral supernatant and RFP-positive cells were sorted after 4 days.

Generating libraries of mutagenized HAP1 cells. To produce gene-trap retrovirus required for mutagenesis of HAP1 cells, HEK293T cells were seeded in T175 flasks at 40% confluence. Medium was replaced with DMEM supplemented with 30% FCS before transfection with 6.6 μg gene-trap vector and either green fluorescent protein[32] or blue fluorescent protein. In addition, the packaging plasmids Gag-pol, VSVg, and pAdv[16] were co-transfected. Viral supernatant was harvested 48 h after transfection and subsequently concentrated by ultracentrifugation at 22,000 r.p.m. for 2 h at 4 °C. Pellets were resuspended in 200 μl phosphate buffered saline (PBS, Life Technologies) and stored overnight at 4 °C. This procedure was repeated twice a day for 3 consecutive days.

To mutagenize HAP1 cells, 40 million cells were seeded and transduced with concentrated virus in the presence of 8 μg ml^{-1} protamine sulphate (Sigma). After two additional rounds of transduction (>90% fluorescent cells), the mutant library was expanded before analysis of an intracellular phenotype by FACS staining.

FACS-based phenotypic screens. Mutagenized HAP1 cells (3×10^9) were either treated to induce the phenotype of interest (Supplementary Table 16) or directly harvested by dissociation (trypsin-EDTA, Life Technologies). To fix the state of a particular phenotype, dissociated cells were resuspended in a volume of BD fix buffer I (BD Biosciences) corresponding to pellet volume at 37 °C for 10 min. Cells were washed with PBS containing 10% FCS and subsequently permeabilized for 30 min on ice using cold BD permeabilization buffer III (BD Biosciences). After washing twice in PBS/10% FCS, cells were filtered through a 40 μm strainer (BD Falcon). Primary antibody staining was performed in 100 μl per 10^7 cells (Supplementary Table 18) for 1 h at room temperature. After two washing steps (PBS/10% FCS), cells were incubated with secondary antibodies (coupled to Alexa Fluor 488, 568, or 647, Life Technologies) at 1:1,000 dilution for 1 h, protected from light. DNA was counterstained using either 1 μg ml^{-1} 4′,6-diamidino-2-phenylindole (DAPI) or 10 μg ml^{-1} propidium iodide (Life Technologies) solution. If propidium iodide staining was performed, cells were also treated with 100 μg ml^{-1} RNase A (Qiagen).

Depending on the gene-trap fluophore/secondary antibody fluophore combination, cells were sorted on an Astrios Moflo or a Biorad S3 cell sorter, using appropriate laser-filter combinations. For every phenotype of interest, the specificity of the staining was determined with a negative (no primary antibody) control. To remove diploid cells containing heterozygous mutations, cells were first gated on the basis of DNA content (1*n*). Subsequently, gates were set on the basis of the log(signal intensity) in the appropriate channel, to include the lowest (dim) and highest (bright) 1–5% of the whole cell population. Cells (8×10^6 to 20×10^6) were sorted into PBS/FCS for low and high populations each.

Amplification of insertion sites. After sorting, insertion sites were retrieved similarly as described[13]. Briefly, cells were pelleted by centrifugation and resuspended in a 1:1 mix of PBS and lysis buffer AL (Qiagen) with the addition of proteinase K. De-crosslinking was performed overnight at 56 °C with agitation, and genomic DNA was purified with a QIAamp DNA Mini Kit (Qiagen) and used as input for linear amplification-mediated polymerase chain reactions (LAM-PCR). These were performed in a total volume of 50 μl with 0.5–2 μg genomic DNA, 1 mM $MgSO_4$, 0.75 pmol double-biotinylated primer (5′/double biotin/GGTCTCCA AATCTCGGTGGAAC-3′), Accuprime Taq HiFi (0.4 μl), and the supplied buffer II (Life Technologies) and amplified for 120 cycles.

PCR reactions were pooled, and biotinylated single-stranded DNA (ssDNA) was isolated by addition of 2× binding buffer (6 M LiCl, 10 mM Tris, 1 mM EDTA, pH 7.5) and M270 streptavidin-coated Dynabeads (Life Technologies) for 2 h at room temperature.

Bead-coupled DNA was washed with PBS containing 0.05% Triton X-100 (Sigma). Before introducing Illumina sequencing adaptors, ssDNA linkers were ligated to the non-biotinylated (3′) end of the PCR products using one of the following two protocols. Depending on the number of LAM-PCR reactions (*N*), the ligation was performed in a total volume of $N \times 10$ μl, using $N \times 12.5$ pmol ssDNA linker (5′/phospho/ATCGTATGCCGTCTTCTGCTTGACTCAGTAGTTGTGCG ATGGATTGATG/dideoxycytidine/3′), 2.5 mM $MnCl_2$, 1 M betaine, $N \times 0.5$ μl of Circligase II (Illumina), and supplied reaction buffer. Alternatively, we used $N \times 12.5$ pmol of a pre-adenylated linker (5′/adenyl/ATCGTATGCCGTCTTCT GCTTGACTCAGTAGTTGTGCGATGGATTGATG/dideoxycytidine/3′) and 2 μg *Escherichia coli*-purified TS2126 thermostable RNA ligase 1 from *Thermus scotoductus* bacteriophage[33] in an $N \times 10$ μl reaction containing 18.75% PEG6000, $N \times$ (2.5 μg BSA, 2.5 mM $MnCl_2$, 1 μl buffer (500 mM MOPS, 100 mM KCl, 50 mM $MgCl_2$, 10 mM dithiothreitol)). Ligation was performed at 60 °C for 2 h in non-stick 1.5 ml tubes (Life Technologies), followed by three wash steps with PBS/0.05% Triton X-100 (Sigma). To introduce the adaptor sequences required for Illumina sequencing (P5 and P7), 50 μl reactions were set up (one for two LAM-PCR reactions), containing 25 pmol of each primer, 5 μl buffer II, and 0.6 μl Accuprime Taq HiFi.

Eighteen cycles of PCR were performed at an annealing temperature of 55 °C for 30 s followed by an extension (at 68 °C) for 105 s, using primers 5′-AATGA TACGGCGACCACCGAGATCTGATGGTTCTCTAGCTTGCC-3′ and 5′-CAAG CAGAAGACGGCATACGA-3′. Libraries were purified with PCR purification columns (Qiagen) and sequencing was performed with a concentration of 18 pM per lane (51-bp or 65-bp single reads) on either HiSeq2000 or HiSeq2500 instruments (Illumina) using sequencing primer 5′-CTAGCTTGCCAAACCTAC AGGTGGGGTCTTTCA-3′.

Insertion site mapping and analysis. After deep sequencing of the low- and high-sorted populations, gene-trap insertion sites were determined as unique reads aligning unambiguously to the human genome (hg19) using Bowtie[34], allowing for a single mismatch. Aligned reads were mapped using hg19 protein-coding gene coordinates (Refseq) to identify intragenic insertion sites and their orientation with respect to the gene using intersectBED[35]. For this analysis, insertion sites integrated in sense within a gene were considered disruptive. To prevent potential confounding, insertion sites in genomic regions assigned to overlapping genes were discarded, as well as integrations in the 3′ untranslated region (UTR) of genes as the gene-trap cassette might have been be less effective there in ablating gene function. To identify genes enriched for disruptive gene-trap integrations in either the high- or low-query populations, the number of unique disruptive mutations in each gene and in the total of one population (for example, signal high) was counted and compared with those values in the other population (for example, signal low) using a two-sided Fisher's exact test. Resulting *P* values were adjusted for multiple testing using the Benjamini–Hochberg false discovery rate correction. For each gene, a mutation index (MI) was calculated corresponding to the ratio of the number of disruptive integrations per gene in both populations normalized by the number of total integrations in each channel:

$$\mathrm{MI} = \frac{\text{Number of sense insertions in gene in high population}}{\text{(Total number of sense insertions in high population)} - \text{(Number of sense insertions in gene in high population)}} \Big/ \frac{\text{Number of sense insertions in gene in low population}}{\text{(Total number of sense insertions in low population)} - \text{(Number of sense insertions in gene in low population)}}$$

For genes without a single insertion site in only one of the channels, a value of 1 was assigned so as not be omitted from the plots. Reactome pathway[36] enrichment analysis for gene groups was performed using ConsensusPathDB[37], and network analysis was done in Cytoscape 3.

RNA-seq data of ten independent wild-type HAP1 cells were aligned using Tophat[38], assigned to Ensembl genes, and expression levels determined by

HTSeq-count[39]. Read counts in protein-coding genes were normalized to 10 million reads followed by $\log_2$ transformation. To avoid negative normalized values, 1 was added to each gene expression value. To map mutations to the genome, a file based on Refseq coordinates containing all unique genomic protein-coding regions was used. Genes occurring in both the Ensembl-based expression data as well as the Refseq-based insertion site data were considered for further expression analysis (16,800 genes). For the binned expression analysis, 17 bins of approximate equal size (~988 genes per bin) were created.

KCTD5 cloning and expression. The KCTD5 gene was cloned into the pETNKI-strepII-3C-LIC-kan vector[40] for expression of strepII-3C-KCTD5. The construct was transformed into *E. coli* BL21(DE3) and cells were grown in Lysogeny broth (LB) medium supplemented with 30 μg ml^{-1} kanamycin at 37 °C until absorbance at 600 nm was 0.6. The temperature was decreased to 20 °C and protein expression was induced by addition of 0.4 mM IPTG. After 16 h of expression at 20 °C, cells were harvested by centrifugation. The cell pellet was resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 1 mM TCEP, and protease inhibitor tablet) and cells were lysed by sonication. The lysate was cleared by centrifugation and the soluble fraction was applied to Streptactin beads (IBA Lifesciences). Beads were washed with 50 mM Tris-HCl pH 8.0, 250 mM NaCl, 1 mM TCEP, and protein was eluted by 2.5 mM desthiobiotin in the same buffer. The strepII tag was removed upon incubation of the protein with GST-3C protease overnight at 4 °C. KCTD5 was further purified by size-exclusion chromatography on an S200 16/60 column (GE Healthcare) in 50 mM Tris pH 8.0, 250 mM NaCl, 1 mM TCEP. KCTD5 eluted from the column in a single peak. Fractions containing the protein were pooled and glycerol was added (20% final concentration) before the protein was concentrated and stored in aliquots at −80 °C.

Conditioned medium. For the genetic screen for active β-catenin, before sorting HAP1 cells were stimulated for 4 h with 10% WNT3A-conditioned medium and 4% R-spondin1-conditioned medium[41].

Immunoprecipitations and immunoblotting. To detect protein–protein interactions, cells were lysed with ice-cold CHAPS buffer (30 mM Tris-Cl pH 7.5, 150 mM NaCl, 1% CHAPS) containing protease and phosphatase inhibitor cocktail (Roche). After sonication, cell debris was pelleted and the cleared supernatant was incubated with anti-Flag M2 agarose beads (Sigma-Aldrich) at 4 °C for 2 h. An aliquot of the lysate was kept and served as input control. After five wash steps with CHAPS buffer, SDS sample buffer was added to the beads. After boiling the samples at 95 °C for 5 min, precipitated proteins were separated by gel electrophoresis. Proteins were transferred onto polyvinylidene fluoride (PVDF) membranes and detected with the indicated antibodies (Supplementary Table 19). For gel source data, see Supplementary Fig. 1. For western blots, a representative example is presented obtained from at least two independent experiments.

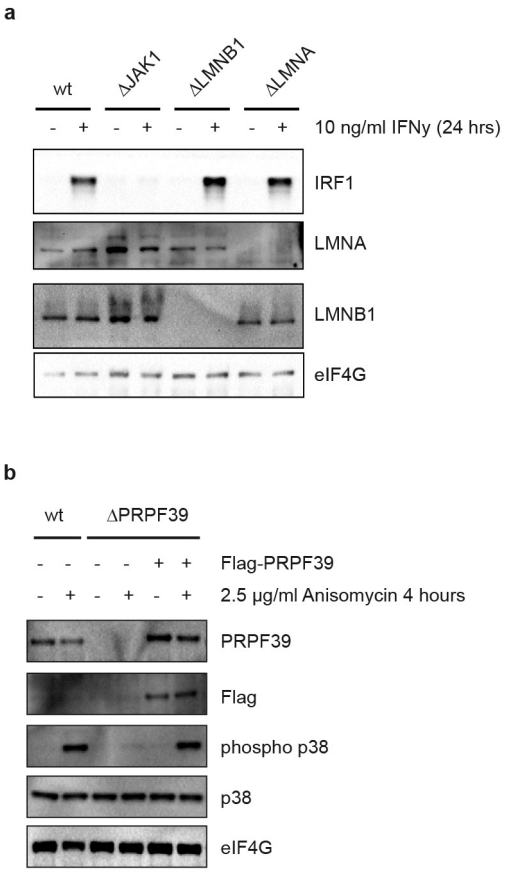
Proteomics. For ubiquitination site profiling, 15 mg amounts of protein of KCTD5-knockout and wild-type HAP1 cells were alkylated with 10 mM chloroacetamide and digested overnight with trypsin (1:50 at 37 °C); for proteome profiling, 50 μg amounts were additionally pre-digested with LysC (1:75, 4 h at 37 °C). Ubiquitinated peptides were enriched by immunoaffinity purification using a PTMScan Ubiquitin Remnant Motif (K-ε-GG) Kit (Cell Signaling Technology)[42,43]. (Ubiquitinated) peptide mixtures were analysed by nanoLC-MS/MS on an Orbitrap Fusion Tribrid mass spectrometer equipped with a Proxeon nLC1000 system (Thermo Scientific) using a nonlinear 210 min gradient, as described previously[44]. Raw data files were processed with MaxQuant version 1.5.0.30, searching against the human-reviewed Uniprot database (October 2014, 20,195 entries). A false discovery rate was set to 1% for both protein and peptide levels, and GG(K) was set as an additional variable modification for analysis of ubiproteome samples. (Ubiquitinated) peptides and proteins were quantified with label-free quantitation using default settings[45].

Properties of genes with broader or narrower phenotypic range. Genes that scored significantly in at least one of the phenotypes analysed in this study (see Fig. 2b) were intersected with public protein–protein interaction data from BioGRID, release 3.4.144 (https://thebiogrid.org)[46]. BioGRID data were filtered for human proteins and physical interactions. Intersection of the data in this manner allowed 1,988 of 2,085 genes scoring in at least one genetic screen to be mapped to the protein–protein interaction data. For statistical analysis, these 1,988 genes were grouped into having a narrower phenotypic range (that is, scoring in one or two screens; 1,478 genes) or having a broader phenotypic range (that is, scoring in three or more screens; 510 genes). For these two groups, the difference in the number of protein–protein interactions was analysed using an unpaired two-sided *t*-test. To test the representation of genes required for fitness in either group, the 2,085 genes contributing to at least one phenotype were intersected with fitness-affecting genes in HAP1 cells[13]. The proportion of fitness-related genes in either the narrower (that is, scoring in one or two screens) or broader (that is, scoring in three or more screens) phenotypic range group was tested using a χ^2 test.

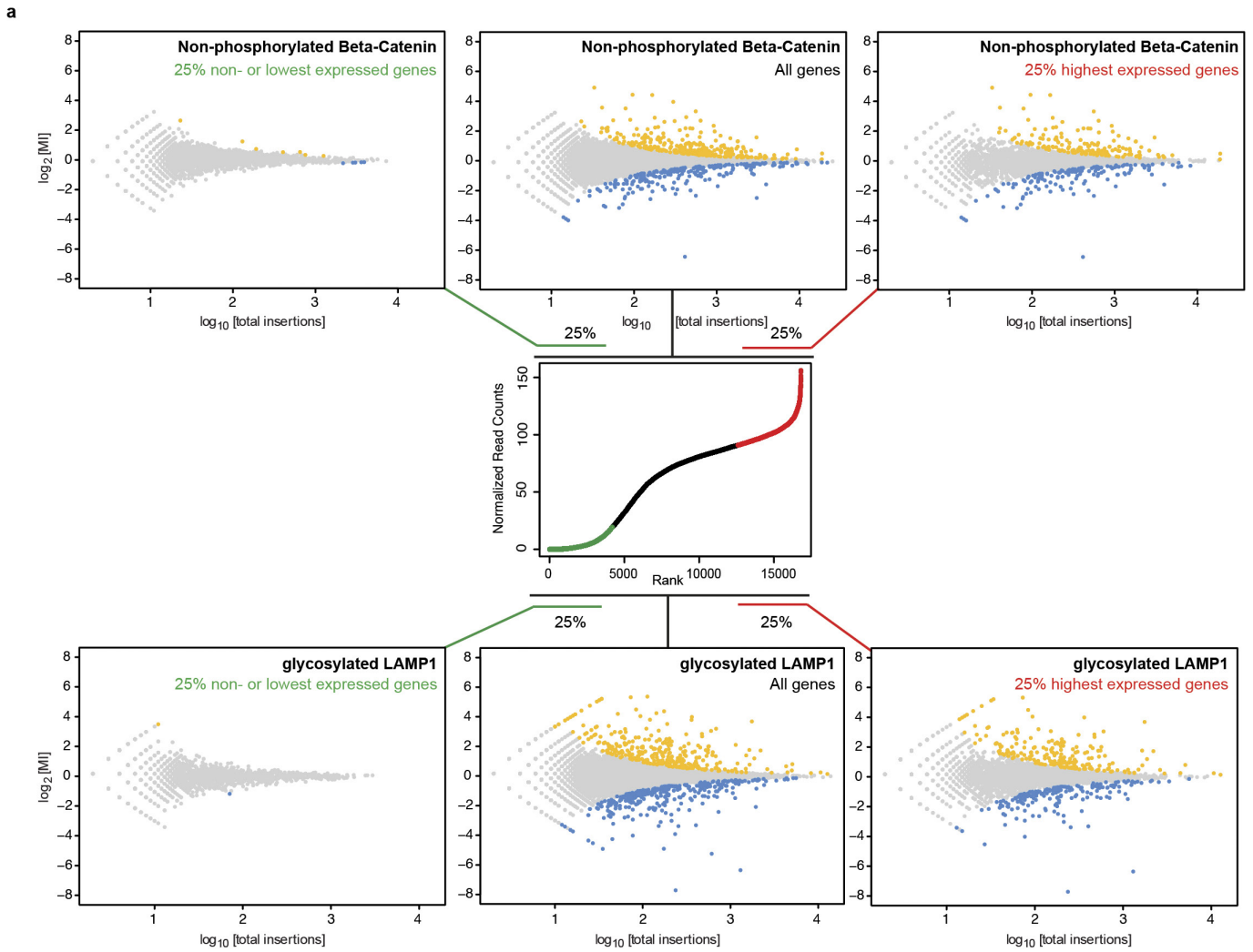
Data availability. All sequencing datasets have been deposited in the NCBI Sequence Read Archive under accession number SRP099134. In addition, all processed screen results are accessible in an interactive database (https://phenosaurus.nki.nl/). Source Data for the main and Extended Data figures are provided in the online version of the paper.

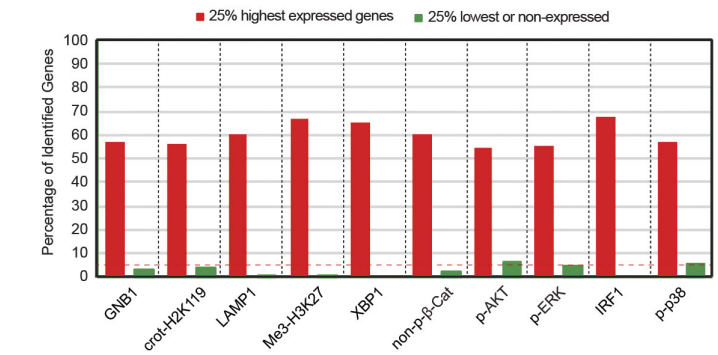
Code availability. Code used for data analysis or other data are available from the corresponding authors upon reasonable request.

31. Lackner, D. H. *et al.* A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nat. Commun.* **6,** 10237 (2015).
32. Jae, L. T. *et al.* Deciphering the glycosylome of dystroglycanopathies using haploid screens for Lassa virus entry. *Science* **340,** 479–483 (2013).
33. Blondal, T. *et al.* Isolation and characterization of a thermostable RNA ligase 1 from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties. *Nucleic Acids Res.* **33,** 135–142 (2005).
34. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
35. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
36. Joshi-Tope, G. *et al.* Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.* **33,** D428–D432 (2005).
37. Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB–a database for integrating human functional interaction networks. *Nucleic Acids Res.* **37,** D623–D628 (2009).
38. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25,** 1105–1111 (2009).
39. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
40. Luna-Vargas, M. P. A. *et al.* Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases. *J. Struct. Biol.* **175,** 113–119 (2011).
41. Fessler, E. *et al.* TGFβ signaling directs serrated adenomas to the mesenchymal colorectal cancer subtype. *EMBO Mol. Med.* **8,** 745–760 (2016).
42. Xu, G., Paige, J. S. & Jaffrey, S. R. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* **28,** 868–873 (2010).
43. Udeshi, N. D., Mertins, P., Svinkina, T. & Carr, S. A. Large-scale identification of ubiquitination sites by mass spectrometry. *Nat. Protocols* **8,** 1950–1960 (2013).
44. Ameziane, N. *et al.* A novel Fanconi anaemia subtype associated with a dominant-negative mutation in *RAD51*. *Nat. Commun.* **6,** 8829 (2015).
45. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* **13,** 2513–2526 (2014).
46. Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34,** 535–539 (2006).

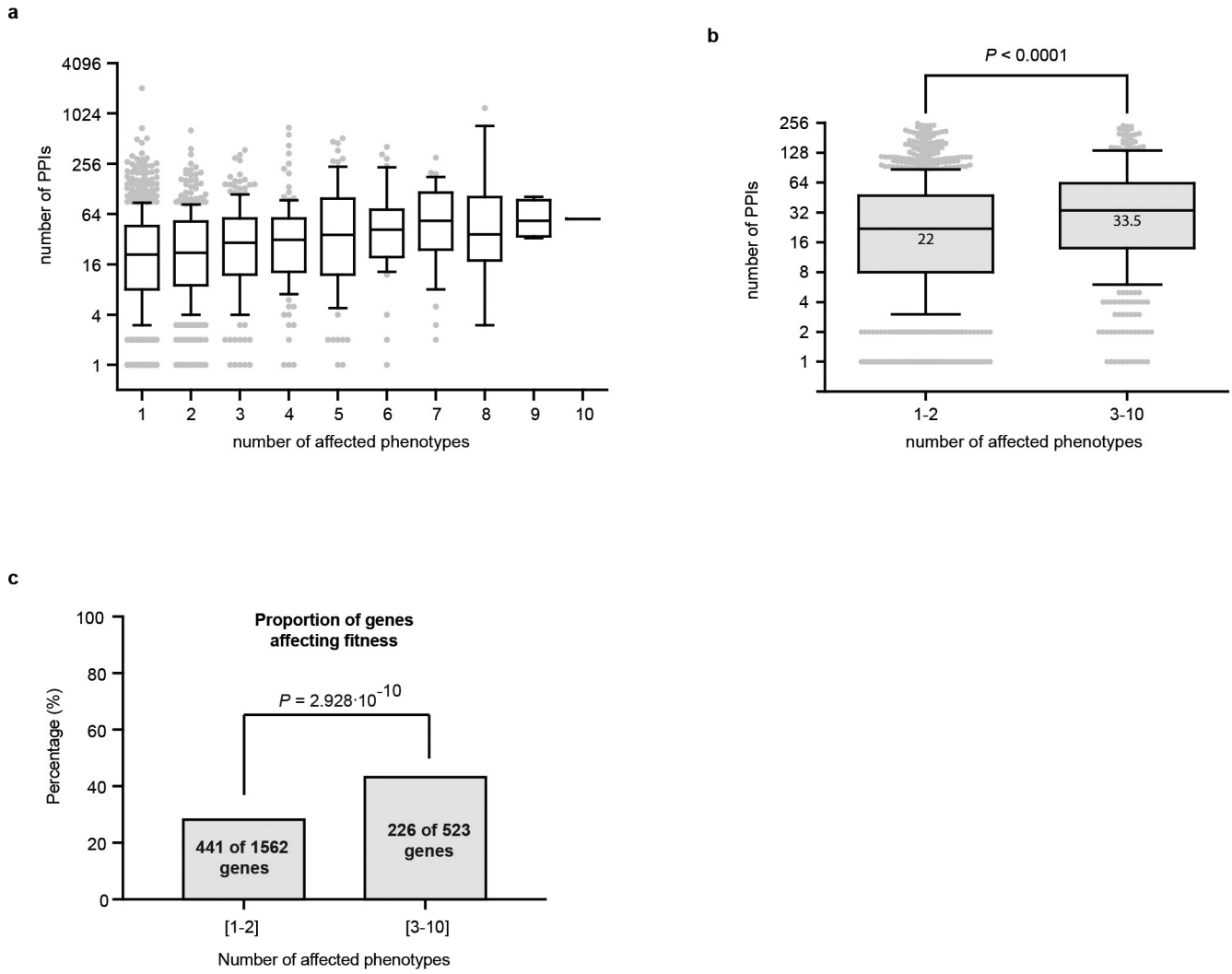


Extended Data Figure 1 | Validation of selected identified regulators. **a**, Wild-type, JAK1-, LMNB1-, and LMNA1-deficient HAP1 cells were treated with IFN-γ for the indicated amount of time, lysates were prepared and analysed by immunoblotting. **b**, Wild-type, PRPF39-deficient, and PRPF39-deficient HAP1 cells reconstituted with Flag-tagged PRPF39 were treated with the protein synthesis inhibitor anisomycin for 4 h; lysates were prepared and analysed by immunoblotting.

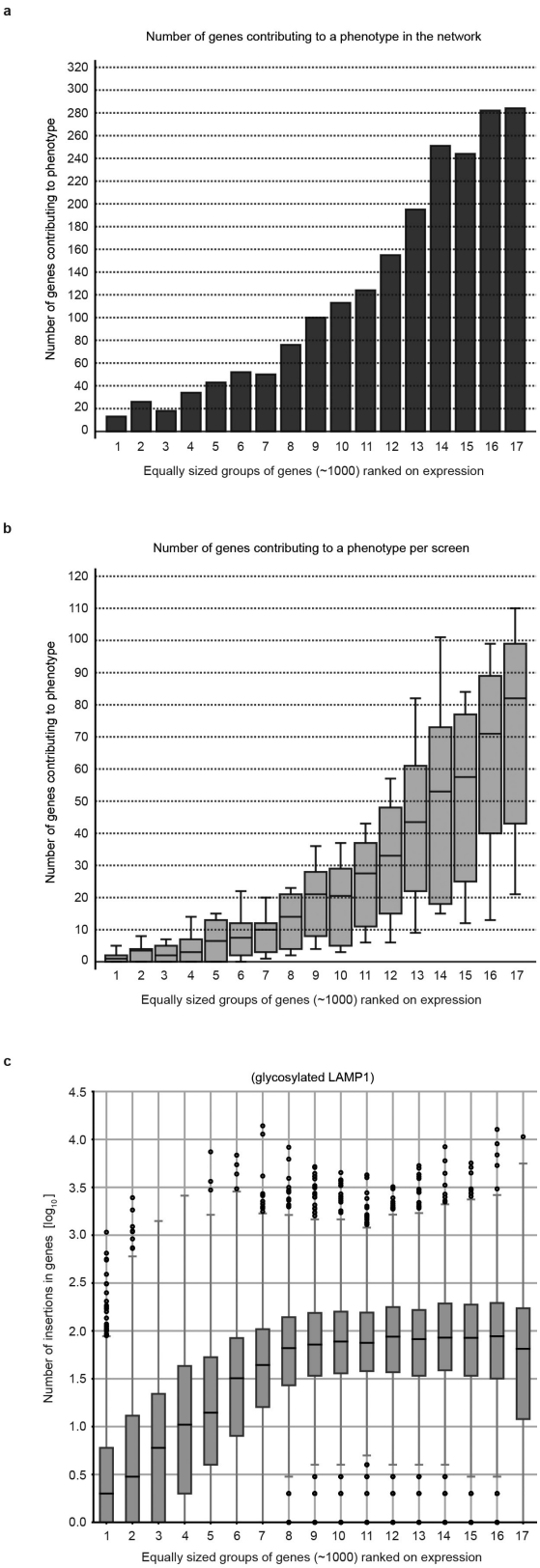




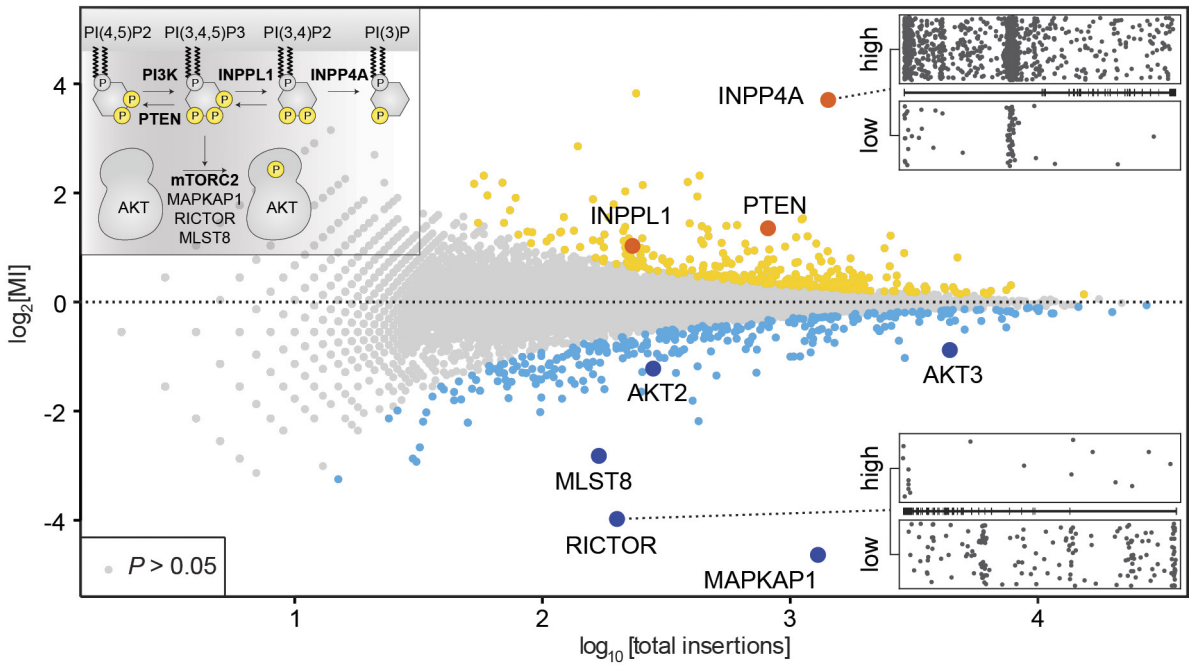
Extended Data Figure 2 | Gene expression is a requirement for phenotypic contribution. a, The datasets for the two screens were filtered to display only the genes falling within the top 25% (4,200 genes) highest and non- or lowest-expressed genes in HAP1 cells. **b**, Bar plot representing the quantification of all screens (analysed as in **a**).



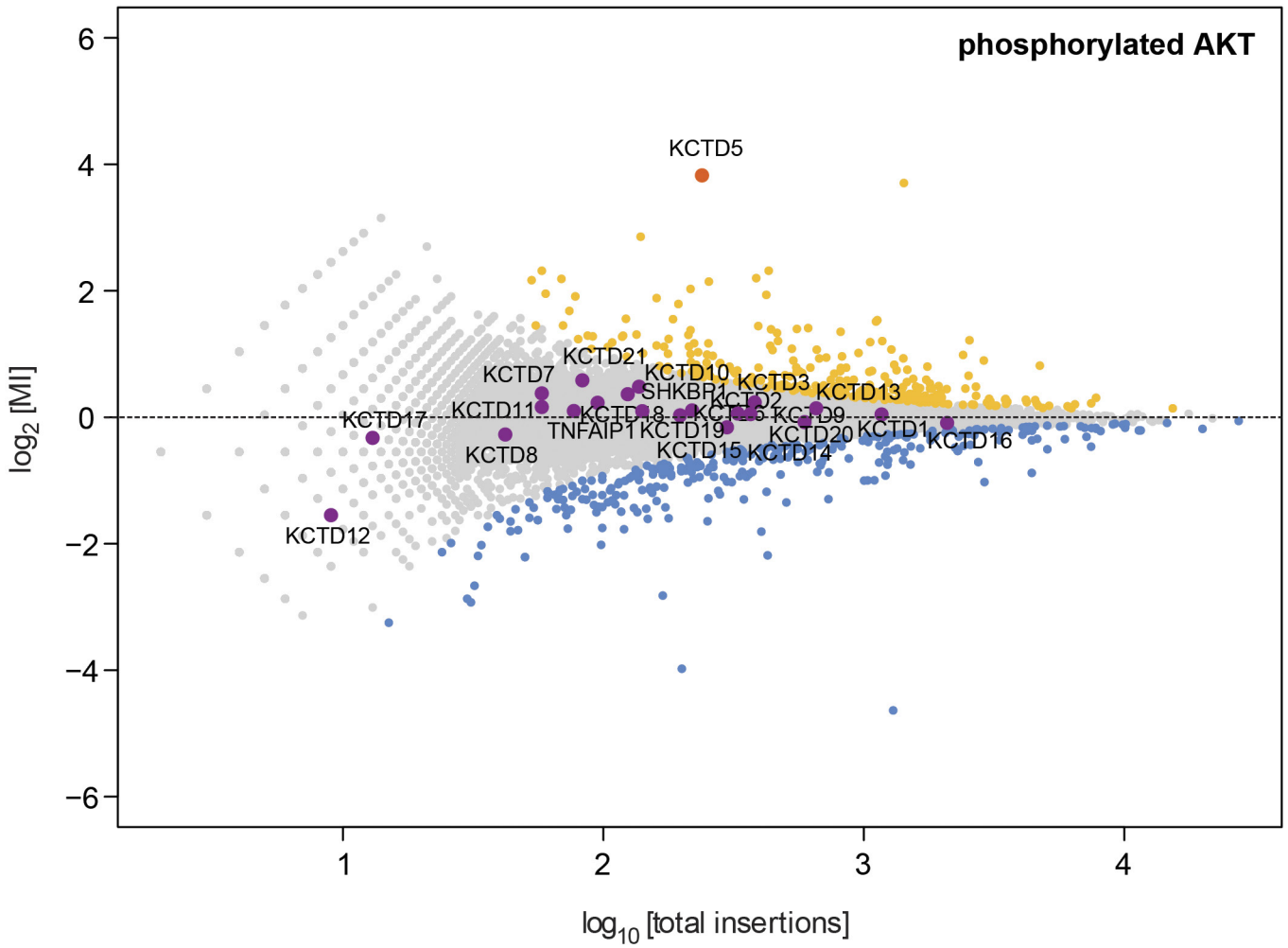
Extended Data Figure 3 | Analysis of genes linked to few or many phenotypes. a, Number of reported physical protein–protein interactions as a function of the number of phenotypes analysed in this study affected by a gene. **b**, As in **a** but with genes categorized as affecting either few (one or two; 1,478 genes) or many (three or more; 510 genes) traits. Two-sided unpaired *t*-test shows a modest but significant difference in the average number of protein–protein interactions between both groups. The *y* axis is cropped at 256 protein–protein interactions for better visibility and the median number of protein–protein interactions in each group is indicated. Box plots and error bars drawn according to Tukey's representation. **c**, Comparison of fitness contribution for genes affecting few (one or two) versus many (three to ten) phenotypes. Genes specifically required for fitness in HAP1 cells[13] were intersected with the genes contributing to phenotype-affecting genes and the proportion occurring in either group was tested using a two-sided χ^2 test.



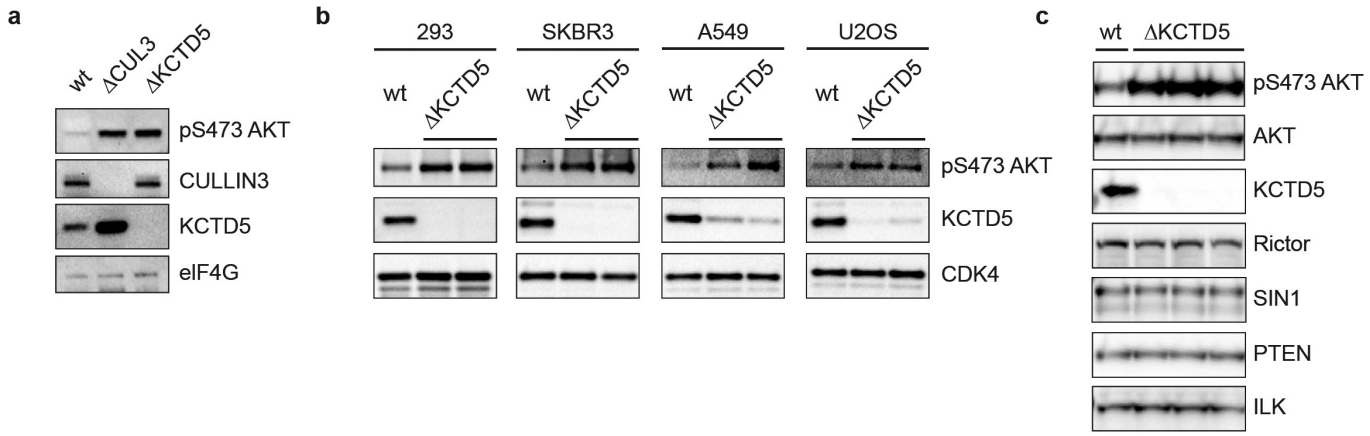
Extended Data Figure 4 | Expression levels relate to phenotypic contribution. a, A total of 16,800 interrogatable genes were ranked on expression levels in HAP1 cells and binned into 17 bins containing approximately 1,000 genes per bin. In each bin, the number of genes identified as a regulator of at least one phenotypic trait was counted. **b**, To account for differences between screens, the same binned approach as in **a** was applied for the number of genes contributing to each individual phenotypic trait additionally. **c**, To analyse the relationship between expression levels and mutation frequencies the number of sense insertions per gene in the glycosylated LAMP1 screen is plotted per bin, demonstrating that the observed increase in phenotypic contribution from bins 8–17 is not due to a higher average mutation frequency. Box plots and error bars drawn according to Tukey's representation.



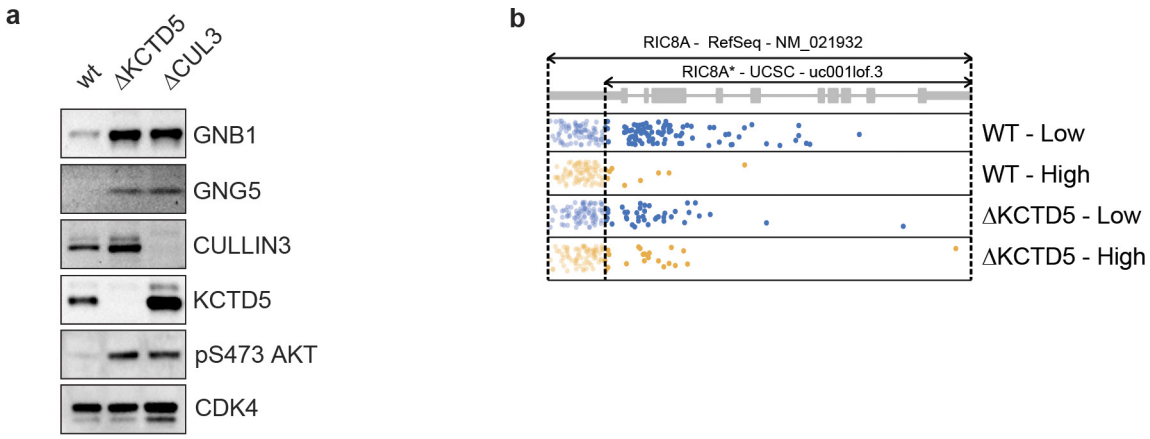
Extended Data Figure 5 | Genetic wiring map for phosphorylation of AKT at S473 identifies known regulators of this process. Outcome of genetic screen for AKT phosphorylation at S473. Data were generated and analysed as in Fig. 1. Selected known factors affecting AKT phosphorylation are labelled and their role in the signalling cascade is indicated in the cartoon. Individual gene-trap insertions (black dots) and their distribution across the gene bodies in the high and low channels (pAKT staining intensity) are shown for *INPP4A* and *RICTOR*.



Extended Data Figure 6 | Effect of different KCTD family members on AKT phosphorylation at S473. KCTD family members are highlighted in the dataset described previously.



Extended Data Figure 7 | KCTD5 regulates phosphorylation of AKT in different human cell lines without affecting levels of common regulators. a, Immunoblot confirming the effect of KCTD5 and CUL3 on AKT phosphorylation (S473) as detected in the genetic screen. Wild-type HAP1 cells and HAP1 cells deficient in KCTD5 or CUL3 were lysed and probed with specific antibodies by immunoblotting. **b**, Indicated wild-type and KCTD5-deficient HEK293 cells (two independent clones) were lysed and probed with specific antibodies by immunoblotting. Three additional cell lines (SKBR3, A549, and U2OS) were infected with a mix of two different lentiviral gRNAs targeting KCTD5 (RFP–CRISPR backbone). RFP-positive cells were sorted after 4 days and immunoblotted with the indicated antibodies. **c**, Wild-type or KCTD5-deficient HAP1 cells (three independent clones) were lysed and analysed with specific antibodies by immunoblotting.



Extended Data Figure 8 | The Gβγ dimer is destabilized in the presence of KCTD5. a, Wild-type HAP1 cells and HAP1 cells deficient in KCTD5 or CUL3 were lysed and probed with specific antibodies by immunoblotting. Increased levels of GNB1 and GNG5, as well as increased phosphorylation of AKT at S473, are comparable in cells deficient for KCTD5 or Cullin3. **b**, For RIC8A*, transcript uc001lof.3 was considered because the longer 5′ UTR in Refseq reduced the observed effect size.