Efficient proximity labeling in living cells and organisms with TurboID

Tess C Branon1–4, Justin A Bosch5, Ariana D Sanchez4, Namrata D Udeshi6, Tanya Svinkina6, Steven A Carr6, Jessica L Feldman4, Norbert Perrimon5,7 & Alice Y Ting1–4,8

Protein interaction networks and protein compartmentalization underlie all signaling and regulatory processes in cells. Enzyme-catalyzed proximity labeling (PL) has emerged as a new approach to study the spatial and interaction characteristics of proteins in living cells. However, current PL methods require over 18 h of labeling time or utilize chemicals with limited cell permeability or high toxicity. We used yeast display-based directed evolution to engineer two promiscuous mutants of biotin ligase, TurboID and miniTurbo, which catalyze PL with much greater efficiency than BioID or BioID2, and enable 10-min PL in cells with non-toxic and easily deliverable biotin. Furthermore, TurboID extends biotin-based PL to flies and worms.

Enzyme-catalyzed PL is an alternative to immunoprecipitation and biochemical fractionation for proteomic analysis of macromolecular complexes, organelles, and protein interaction networks1. In PL, a promiscuous labeling enzyme is targeted by genetic fusion to a specific protein or subcellular compartment. Addition of a small-molecule substrate, such as biotin, initiates covalent tagging of endogenous proteins within a few nanometers of the promiscuous enzyme (Fig. 1a). Subsequently, the biotinylated proteins are harvested using streptavidin-coated beads, and identified by mass spectrometry (MS).

Two enzymes are commonly used for PL: APEX2, an engineered soybean ascorbate peroxidase2–4, and BioID, a promiscuous mutant of Escherichia coli biotin ligase5,6. The advantages of APEX2 are its speed—proximal proteins can be tagged in 1 min or less6,7—and its versatility, as APEX2 also captures endogenous RNAs8 and generates contrast for electron microscopy9. However, APEX labeling requires the activity of our starting template (R118S) and input library were too stringency (Supplementary Note 1) and hence we selected this mutant rather than BioID as our starting template for evolution.

As in previous work3,31, we combined yeast surface display of our enzyme library with fluorescence-activated cell sorting (FACS) to perform the evolution. We used error-prone PCR to mutagenize BirA-R118S, generating a library of ~107 mutants, each with an average of about two amino acid mutations relative to template. This library was then displayed on the yeast surface as a fusion to the Aga2p mating protein (Fig. 1b). We added biotin and ATP to the yeast pool to initiate promiscuous biotinylation, followed by streptavidin-fluorophore to stain biotinylation sites on the surface of each yeast cell. FACS was used to enrich cells displaying a high degree of self-biotinylation over cells displaying low or moderate self-biotinylation (Fig. 1b). We gradually reduced the biotinylation time window from 18 h to 10 min over 29 rounds of selection, in order to progressively increase selection stringency (Supplementary Note 1 and Supplementary Fig. 3).

We encountered some technical hurdles during the evolution. First, the activity of our starting template (R118S) and input library were too low to be detected on the yeast surface. Thus, we used tyramide signal amplification (TSA)12 on the yeast surface to boost the biotin signal for proteomic analysis. This precludes the use of BioID for studying dynamic processes that occur on the timescale of minutes or even a few hours. Furthermore, the low catalytic activity makes BioID difficult or impossible to apply in some contexts, such as in worms, flies, or the endoplasmic reticulum lumen of cultured mammalian cells. Recently, new promiscuous biotin ligase variants, BioID2 (ref. 26) and BASU27, have been reported, but the former still requires >16 h of labeling26,28–30, while BASU enriched a proteome of only two proteins27. Further characterization (see below) shows that the activities of BioID, BioID2, and BASU are all comparable.

A new PL enzyme that combines the simplicity and non-toxicity of BioID with the catalytic efficiency of APEX2 would greatly enhance PL applications. To achieve this, we undertook the directed evolution of E. coli biotin ligase (BirA) to generate new promiscuous mutants. To begin, we compared BioID (BirA-R118G) to seven other mutations at the R118 position. We found that R118S is about twofold more active than R118G under identical conditions (Supplementary Figs. 1 and 2 for all full blot images and Supplementary Note 1), and hence we selected this mutant rather than BioID as our starting template for evolution.

Received 23 September 2017; accepted 2 July 2018; published online 20 August 2018; doi:10.1038/nbt.4201
until the activity of the pool was high enough to no longer require it. Second, to avoid enriching mutants that strongly tagged their own lysine residues but failed to biotinylate neighboring proteins, we treated yeast with the reducing agent TCEP in some rounds of selection, to cleave off the ligase after the biotinylation reaction (Supplementary Fig. 3c). Finally, we introduced negative selection to deplete mutants that exhibited strong biotinylation activity before exogenous biotin addition, because the ability to utilize the low levels of biotin naturally present in yeast media would result in loss of user control of the labeling window (Supplementary Fig. 3f).

Our engineering efforts yielded two promiscuous ligases: 35 kD TurboID, with 15 mutations relative to wild-type BirA; and 28 kD miniTurbo, with the N-terminal domain deleted and 13 mutations relative to wild-type BirA (Fig. 1d, Supplementary Table 1 and Supplementary Note 2). Fig. 1e and Supplementary Fig. 4 show the activity of these ligases on the yeast surface in a side by side comparison to BioID, BirA-R118S, and various intermediate clones from our evolution (G1-G3A).

To test TurboID and miniTurbo in mammalian cells, we expressed them in the cytosol of HEK 293T cells. Labeling was initiated with the addition of 50 or 500 µM C6 exogenous biotin and terminated by cooling cells to 4 °C and washing away excess biotin (Supplementary Fig. 5). Streptavidin blot analysis of whole cell lysates showed that TurboID and miniTurbo biotinylated endogenous proteins much more rapidly than BioID, giving ~3- to 6-fold difference in signal at early time points, and ~15- to 23-fold difference in signal at later time points (Fig. 1f,g, Fig. 2a,b, Supplementary Fig. 4b, and Supplementary Fig. 6).

We included the newer promiscuous ligases BioID2 (ref. 26) and BASU27 in the comparison, and after normalization to account for differences in ligase expression levels, their activities were similar to that of BioID (Fig. 2a,b). Notably, TurboID gave nearly as much biotinylated product in 10 min as BioID/BioID2/BASU gave in 18 h (Fig. 2a,b). Overall, miniTurbo was 1.5- to twofold less active than TurboID, but exhibited less labeling before addition of exogenous biotin; this feature makes miniTurbo potentially superior for precise temporal control of the labeling window. The resulting trends were the same when we ran the same experiment with 50 µM biotin instead of 500 µM biotin for labeling (Supplementary Fig. 6c–e).

To compare ligases by a different modality, we also fixed ligase-expressing HEK 293T cells after biotinylation, stained them with neutravidin-fluorophore, and performed confocal microscopy. TurboID and miniTurbo gave clearly detectable biotinylation in most transfected cells after 10 min of biotin labeling (Supplementary Fig. 7). By contrast, BioID, BioID2, and BASU-catalyzed biotinylation was undetectable even at 1 h, and only dimly detectable at 6 h in a small fraction of transfected cells.

Different organelles have distinct pH, redox environments, and endogenous nucleophile concentrations, which may influence PL activity. We therefore compared TurboID, miniTurbo, and BioID in the nucleus, mitochondrial matrix, ER lumen, and ER membrane of HEK 293T cells (Fig. 2c). We found that the absolute activities of each ligase, as well as the relative activities between ligases, varied across compartments (Supplementary Note 3). However, TurboID signal was clearly detectable after 10 min in each compartment, and even stronger than BioID 18-h labeling in the mitochondrial matrix and ER lumen. TurboID was superior to miniTurbo in each of these four organelles. Given our observations, we recommend that users test both TurboID and miniTurbo for PL applications, given the context-dependent variations in their activities. We next evaluated TurboID and miniTurbo in full-scale proteomic experiments. We asked whether 10-min labeling with these ligases would produce proteomic data sets of similar quality to BioID labeling for 18 h, in terms of specificity, coverage, and labeling radius (Supplementary Note 4 and Supplementary Fig. 8). We selected three mammalian organelles for the analysis: the mitochondrial matrix, nucleus, and ER membrane (ERM)-facing cytosol (Fig. 2d–h and Supplementary Figs. 9–11). Because the ERM is continuous with the cytosol, it is valuable for assessing labeling radius; a good PL enzyme should strongly enrich ERM-localized proteins over immediately adjacent cytosolic proteins.

The HEK293T samples we prepared for proteomic analysis are depicted in Figure 2d and Supplementary Figure 10a. TurboID and miniTurbo labeling were each performed for 10 min, whereas BioID

Figure 1 Directed evolution of TurboID. (a) Proximity-dependent biotinylation catalyzed by promiscuous biotin ligases. Ligases catalyze the formation of biotin-5′-AMP anhydride, which diffuses out of the active site to biotinylate proximal endogenous proteins on nucleophilic residues such as lysine. (b) Yeast-display-based selection scheme. A >10^12 library of ligase variants is displayed on the yeast surface as a fusion to mating protein A. Aga2p. All ligases have a C-terminal myc epitope tag. Biotin and ATP are added to the yeast library for between 10 min and 24 h. Ligase-catalyzed promiscuous biotinylation is detected by staining with streptavidin-phycocerythrin and ligase expression is detected by staining with anti-myc antibody. Two-dimensional FACs sorting enables enrichment of cells displaying a high ratio of streptavidin to myc staining. (c) Tyramide signal amplification (TSA) improves ligase detection sensitivity on the yeast surface. In the top row, the three indicated yeast samples (G1 is the winning ligase mutant from the first generation of evolution) were labeled with exogenous biotin for 18 h then stained for FACS as in b. The y-axis quantifies biotinylation extent, and the x-axis quantifies ligase expression level. In the second row, after 18 h of biotin incubation, yeast were stained with streptavidin-HRP, reacted with biotin-phenol 2,3′ to create additional biotinylation sites (TSA protocol), then stained with streptavidin-phycocerythrin and anti-myc antibody before FACS. The third row omits biotin. Percentage of cells in upper right quadrant (Q2/Q4) is shown in top right of each graph. This experiment was performed once, but each yeast sample was analyzed under identical conditions at least twice in separate experiments with similar results. (d) E. coli ligase gradient screen (PDB: 2EWN) with sites mutated in TurboID (left) and miniTurbo (right) shown in red. The N-terminal domain (aa1–63) is also removed from miniTurbo. A non-hydrizable analog of biotin-5′-AMP, biotinol-5′-AMP, is highlighted in yellow. (e) FACs plots summarizing progress of directed evolution. G1–G3 are the winning clones from generations 1–3 of directed evolution. G3A has its N-terminal domain (aa1–63) deleted. ‘Omit biotin’ samples were grown in biotin-deficient media for the entire induction period (~18–24 h). This experiment was performed twice with similar results, except G3A ‘omit biotin’, which was performed once. All results shown here were performed side by side in a single experiment. (f) Comparison of ligase variants in the HEK cytosol showing that TurboID and miniTurbo are much more active than BioID, as well as the starting template BirA-R118S and various intermediate clones from the evolution. Indicated ligases were expressed as NES (nuclear export signal) fusions in the HEK cytosol. 50 µM exogenous biotin was added for 3 h, then whole cell lysates were analyzed by streptavidin blotting. Ligase expression detected by anti-V5 blotting, U, untransfected. S, BirA-R118S. Asterisks indicate ligase self-biotinylation. BioID labeling for 18 h (50 µM biotin) shown for comparison in the last lane. This experiment was performed twice with similar results. All results shown here were performed side by side in a single experiment. (g) Quantification of streptavidin blot data in f and from a 30-min labeling experiment shown in Supplementary Figure 4b. Quantitation excludes self-biotinylation band. Sum intensity of each lane is divided by the sum intensity of the ligase expression band; ratios are normalized to that of BioID/18 h, which is set to 1.0. Gray dots indicate quantitation of signal intensity from each replicate, colored bars indicate mean signal intensity calculated from the two replicates.
labeling was carried out for 18 h. Cells were lysed and biotinylated proteins enriched with streptavidin beads. After on-bead digestion of proteins to peptides, we chemically labeled the peptides with isotopically distinct TMT (tandem mass tag) labels. This enabled us to quantify the relative abundance of each protein across samples. After liquid chromatography (LC)-MS/MS analysis of pooled peptides, we filtered the data via receiver operating characteristic (ROC) analysis (Supplementary Fig. 9c and Supplementary Fig. 10f, g), using true-positive and false-positive protein lists for each organelle (Supplementary Tables 2–4), to obtain BioID-, TurboID-, and miniTurbo-derived proteomes for the ERM (Supplementary Table 5), nucleus (Supplementary Table 6), and mitochondrial matrix (Supplementary Table 7).

TurboID- and miniTurbo-derived 10-min proteomes had similar size and specificity as BioID-derived 18-h proteomes in all three compartments (Figure 2e–h). In particular, we note that TurboID was just as effective as BioID in enriching secretory proteins over cytosolic proteins when localized to the ERM (Fig. 2e–g), suggesting a similar labeling radius despite much faster labeling kinetics. Depth of coverage was similar in the mitochondrial matrix and ERM for the three

---

**Diagram Descriptions**

**Diagram a**
- Promiscuous biotin ligase
- ATP + B
- AMP
- Biotin
- Biotinylated proteins
- Distal endogenous proteins
- Biotinylated proximal endogenous protein

**Diagram b**
- Aga1p
- Aga2p Ligase
- Yeast cell
- myc
- Biotin, ATP
- 10 min–24 h
- Aga1p
- Aga2p Ligase
- Library
- Yeast cell
- myc

**Diagram c**
- R118S
- R118S library
- G1
- 18 h
- 30 min biotin
- Streptavidin-phycoerythrin
- Anti-myc (AlexaFluor647, log10)

**Diagram d**
- TurboID (35 kD)
- N-terminal domain aa 1–63
- Biotinol-5′-AMP
- Q65P
- I87V
- R118S
- E140K
- Q141R
- S150G
- L151P
- V160A
- T192A
- K194I
- M209V
- M241T
- S263P
- I305V

**Diagram e**
- R118S
- G1
- G2
- G3
- G3a
- miniTurbo
- TurboID
- 6 h biotin
- Omit biotin

**Diagram f**
- 3 h biotin
- Streptavidin-HRP
- Anti-V5

**Diagram g**
- Relative biotinylation activity, normalized to ligase expression level
- Zoom
- 30 min biotin
- 3 h biotin
ligases, but slightly lower for TurboID and miniTurbo in the nucleus (Supplementary Figs. 9f and 10j).

Given the extremely high activity of TurboID, we wondered whether increasing the labeling time would produce a bigger and better proteome. For the ER, we found that 1-h labeling with TurboID did increase proteome size by 46% compared to 10-min labeling, but at the expense of specificity (Fig. 2e). With increased labeling time, proximal nucleophile may become saturated with biotin, enabling TurboID-generated biotin-AMP to travel farther and biotinylate distal, non-specific proteins.

Despite the widespread application of BioID, there have been only two in vivo demonstrations to date, which may be is related to BioID's low catalytic activity. We wondered whether TurboID and miniTurbo's increased activity might enable biotin-based PL in new settings. We first tested these ligases in bacteria (E. coli) and yeast (Saccharomyces cerevisiae). As in mammalian cells, TurboID and miniTurbo were considerably more active than BioID (Fig. 3a,b). In particular, we and others observe that BioID activity was nearly undetectable in yeast, perhaps in part because yeast is cultured at 30 °C whereas BioID functions optimally at 37 °C;26. Because we carried out our directed evolution in yeast at 30 °C, TurboID and miniTurbo exhibit activity at 30 °C.

BioID has not previously been reported in flies (Drosophila melanogaster) or worms (Caenorhabditis elegans), despite their appeal as highly genetically tractable model organisms. To test biotin-based PL in flies, we expressed BioID, TurboID, or miniTurbo selectively in the larval wing disc, which gives rise to the adult wing, and raised animals on biotin-containing food for 5 d from early embryo stages (Fig. 3c). TurboID- and miniTurbo-catalyzed biotinylation were 22-fold and ten-fold higher, respectively, than BioID-catalyzed biotinylation, as shown by staining of dissected wing discs with streptavidin-fluorophore (Fig. 3d,e). Consistent with our observations in HEK 293T cells, TurboID also gives some low biotinylation signal in flies fed regular, non-biotin supplemented, food.

We also generated flies expressing BioID, TurboID, or miniTurbo in all tissues (Act-Gal4 driver, Fig. 3f,g), in muscle (Mef2-Gal4 driver, Supplementary Fig. 12a,b), and in all tissues at non-permissive temperature (tub-Gal4/tub-Gal80° driver, Supplementary Fig. 12c,d). Animals were raised on either biotin-containing food from early-embryo stages to adulthood (13 d) (Fig. 3f,g), or regular food to adulthood (13 d), and then switched to biotin-supplemented food for 4 or 16 h (Supplementary Fig. 12). Streptavidin blotting of whole fly lysate showed extensive biotinylation in TurboID and miniTurbo flies, as early as 4 h post-biotin addition (Supplementary Fig. 12), but no signal was detectable in BioID flies, even after 13 d of biotin exposure (Fig. 3g). The absence of detectable BioID signal here, compared to the wing experiment (Fig. 3d), may have resulted from endogenous biotinylated proteins drowning out specific signal in the streptavidin blot.

To test for possible toxicity of TurboID, miniTurbo, and BioID expression in flies, we performed morphological and survival assays. We observed no evidence of toxicity when any of the ligases were expressed tissue-specifically. However, we did find a decrease in fly viability and size when TurboID was expressed ubiquitously and constitutively, and exogenous biotin was withheld (Supplementary Fig. 13, Supplementary Note 5, and Supplementary Fig. 14). Under these conditions, TurboID may consume all the biotin, effectively starving cells of biotin.

We also tested BioID, TurboID, and miniTurbo in C. elegans. We expressed the ligases early in the intestinal lineage (~150 min after the first cleavage) and assessed biotinylation activity ~4 h later, at the embryonic bean stage (stage 1), ~5.5 h later at the embryonic comma stage (stage 2), or 3 d later in the adult worm (Fig. 3h). TurboID and miniTurbo were significantly more active than BioID by both imaging and streptavidin blotting at all observed developmental stages (Fig. 3i,j and Supplementary Fig. 15). We also observed that TurboID was expressed at higher levels than miniTurbo expression in adult worms, resulting in much stronger labeling (Fig. 3i and Supplementary Fig. 15g). TurboID and miniTurbo labeling yield could be further increased by raising worms at higher temperatures (25 °C vs. 20 °C; Supplementary Fig. 15g).

While we observed background labeling by TurboID in adult biotin-depleted worms (Fig. 3i), similar to our observations in

Figure 2 Characterization of TurboID and miniTurbo in mammalian cells. (a) Comparison of TurboID and miniTurbo to three other promiscuous ligases (BioID25, BiolD226, and BASU27) in the cytosol of HEK 293T cells. Here, 500 μM exogenous biotin was used for labeling, whereas 50 μM was used in Supplementary Figure 6c-e. Streptavidin-HRP blotting detects promiscuously biotinylated proteins, and anti-V5 blotting detects ligase expression. U, untransfected. Asterisks denote ligase self-biotinylation bands. This experiment was performed twice with similar results. All results shown here were performed side by side in a single experiment. (b) Quantitation of experiment in a. For shorter time points (~biotin and 10 min), we used a longer-exposure image of the same blot, shown in Supplementary Figure 6a; for longer time points (1 h, 6 h, 18 h), we used a shorter-exposure image of the blot in a, shown in Supplementary Figure 6b. Quantitation performed as in Figure 1g. Gray dots indicate quantitation of signal intensity from each replicate, colored bars indicate mean signal tyramide signal amplification (TSA) intensity calculated from the two replicates. (c) Comparison of promiscuous ligases in multiple HEK organelles. Each ligase was fused to a peptide targeting sequence (Supplementary Table 8) directing them to the locations indicated in the scheme at right. BioID samples were treated with 50 μM biotin for 18 h. TurboID and miniTurbo samples were labeled for 10 min with 50 (+) or 500 (++) μM biotin. U, untransfected. Asterisks denote ligase self-biotinylation. This experiment was performed five times for nuclear constructs, three times for mitochondrial constructs, four times for ER membrane constructs, and twice for ER lumen constructs with similar results. (d) MS-based proteome experiment comparing TurboID and BiolD on the ER membrane (ERM), facing cytosol. Experimental design and labeling conditions. Ligase fusion constructs were stably expressed in HEK 293T. BiolD samples were treated with 50 μM biotin for 18 h, while TurboID samples were treated with 500 μM biotin for 10 min or 1 h. After labeling, cells were lysed and biotylated proteins were enriched with streptavidin beads, digested to peptides, and conjugated to TMT (tandem mass tag) labels. All 11 samples were then combined and analyzed by LC-MS/MS. Proteomic lists in Supplementary Table 5; further analysis of proteome data in Supplementary Figure 9. This experiment was performed once with two replicates per condition. All results shown here are from two replicates performed within one experiment. (e) Specificity analysis for proteomic data sets derived from experiment in d. Size of each ERM proteome at top. Bars show percentage of each proteome with prior secretory pathway annotation, according to GOCC, Phobius, human protein atlas, human plasma proteome database, and literature (Supplementary Table 2). (f) Same as e, except for each ERM proteome, we analyze the subset with ER, Golgi, or plasma membrane annotation. Annotations from GOCC were assigned in the priority order: ER > Golgi > plasma membrane (Supplementary Table 2). (g) Breakdown of ER proteins enriched by TurboID and BiolD, by transmembrane or soluble. Soluble proteins were further divided into luminal or cytosol-facing. Annotations obtained from GOCC, UniProt, TMHMM, and literature (Supplementary Table 2). (h) Characterization of nuclear and mitochondrial matrix proteomes obtained via BiolD (18 h), TurboID (10 min), and miniTurbo (10 min)-catalyzed labeling. Proteome sizes across top. Bars show fraction of each nuclear (left) or mitochondrial (right) proteome with prior nuclear or mitochondrial annotation, according to GOCC, MitoCarta, or literature (Supplementary Tables 3 and 4). Design of proteomic experiment shown in Supplementary Figure 10a, proteomic lists in Supplementary Tables 6 and 7; further analysis of proteome data in Supplementary Figure 10.
flies and mammalian cell culture, we found that miniTurbo, but not TurboID, gave some background labeling in biotin-depleted worm embryos at stage 2 (Fig. 3k and Supplementary Fig. 15a). We also assessed viability and developmental timing, and did not observe decreased survival in worms expressing any of the three ligases in intestinal cells; however, developmental delay was evident in worms expressing TurboID (Supplementary Fig. 16 and Supplementary Note 5).

In summary, we have performed yeast-display-based directed evolution, incorporating TSA signal amplification, reductive removal of
ligases, and negative selections, to generate two new ligases for PL applications: TurboID and miniTurbo. TurboID is the most active, and should be used when the priority is to maximize biotinylation yield and sensitivity and/or recovery. However, in many contexts, we observe a small degree of labeling before exogenous biotin is supplied, indicating that TurboID can utilize the low levels of biotin present in cells and/or organisms grown in typical biotin-containing media or food. Hence, if the priority is to precisely define the labeling time window, miniTurbo may be preferable to TurboID. Though 1.5- to twofold less active than TurboID, miniTurbo gives much less background in the biotin-omitted condition, and it is also 20% smaller (28 vs. 35 kD), which may reduce interference with fusion protein trafficking and function. Yet another factor to consider when choosing a ligase for PL is ligase stability. Our results indicate that miniTurbo is less stable than TurboID (likely due to removal of its N-terminal domain), resulting in lower expression levels in the adult worm intestine and adult fly, for example. miniTurbo also exhibits biotin-dependent stability, similar to BioID (e.g., see anti-V5 western blots in Fig. 2a).

Up to now, in vivo applications of PL have required very long labeling times34,35 or extensive genetic or manual manipulation36-37 to deliver chemical substrates to relevant cells. TurboID and miniTurbo offer facile substrate delivery and rapid labeling in vivo. In addition to increased catalytic efficiency, we believe that the temperature-activity profiles of TurboID and miniTurbo help to explain their superior performance to BioID in vivo. Whereas BioID is derived from E. coli (37 °C), TurboID and miniTurbo were evolved in yeast (30 °C). Flies grow at 25 °C, while worms are typically grown at 20 °C.

Our toxicity analyses in flies, worms, and mammalian cell culture (Supplementary Figs. 13, 14 and 16) do suggest some necessary precautions when using TurboID and miniTurbo in vivo. First, if TurboID is expressed ubiquitously, it can sequester endogenous biotin and cause toxicity; the solution is to supplement animals with exogenous biotin. Second, users should empirically optimize the in vivo labeling time window, and use the shortest labeling time that produces sufficient biotinylated material for analysis. Longer-than-necessary labeling can cause toxicity via chronic biotinylation of endogenous proteomes, and/or degrade spatial specificity due to saturation of proximal labeling sites (Fig. 2c).

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

FACS was performed at the Koch Institute Flow Cytometry Core (MIT) and Stanford Shared FACS Facility. S. Han (Stanford) synthesized neutravidin-AlexaFluor647. S. Ax (Stanford) cloned the cell surface TurboID and miniTurbo constructs. We are grateful to I. Drojniene (Harvard) for advice on biotin labeling in D. melanogaster. Biotin auxotrophic E. coli MG1655bioKkan was kindly donated by J. Cronan (University of Illinois). This work was supported by NIH R01-GM086588 (to A.Y.T.), Howard Hughes Medical Institute Collaborative Award (to A.Y.T., S.C., and N.P.), and NIH New Innovator Award DP2GM119136 (to J.L.F.). T.C.B. was supported by Dow Graduate Research and Lester Wolfe Fellowships. J.A.B. was supported by a Damon Runyon Post-Doctoral Fellowship. A.D.S. was supported by NIH Training Grant 2T32GM007276.

AUTHOR CONTRIBUTIONS

T.C.B. and A.Y.T. designed the research and analyzed all the data except those noted. T.C.B. performed all experiments except those noted. T.C.B., A.Y.T., N.D.U., and S.A.C. designed the proteomics experiments. T.C.B. prepared the proteomic samples. N.D.U. and T.S. processed the proteomic samples and performed mass spectrometry analyses.

Figure 3 TurboID and miniTurbo in flies, worms, and other species. (a) Comparison of ligases in yeast. EBY100 S. cerevisiae expressing BioID, TurboID, or miniTurbo in the cytosol were treated with 50 µM biotin for 18 h. Whole cell lysates were blotted with streptavidin-HRP to visualize biotinylated proteins, and anti-V5 antibody to visualize ligase expression. U, untransfected. Asterisks denote ligase self-biotinylation. Bands in untransfected lane are endogenous, naturally biotinylated proteins. This experiment was performed twice with similar results. (b) Comparative results in E. coli. Ligases, fused at their N-terminal ends to His6-maltose binding protein, were expressed in the cytosol of BL21 E. coli and 50 µM exogenous biotin was added for 18 h. Whole cell lysates were analyzed as in a. This experiment was performed twice with similar results. (c-g) Comparison of ligases in flies. (c) Scheme for tissue-specific expression of ligases in the wing disc of D. melanogaster. ptc-Gal4 induces ligase expression in a strip of cells within the wing imaginal disc that borders the anterior/posterior compartments. (d) Imaging of larval wing discs after 5 d of growth on biotin-containing food. Biotinylated proteins are detected by staining with streptavidin-AlexaFluor555, and ligase expression is detected by anti-V5 staining. Panels show the ptc region of the wing disc, indicated by the dashed line in c. Scale bar, 40 µm. Each experimental condition has at least three technical replicates; one representative image is shown. This experiment was independently repeated twice with similar results. (e) Quantitation of streptavidin-AlexaFluor555 signal intensities in d. Error bars, s.e.m. Average fold-change shown as text above bars. Sample size values (n) from left column to right: 5, 6, 3. (f) Scheme for ubiquitous expression of ligases in flies, at all developmental time points, via the act-Gal4 driver. (g) Western blotting of fly lysates prepared as in f. Biotinylated proteins detected by blotting with streptavidin-HRP, ligase expression detected by anti-V5 blotting. In control sample, act-Gal4 drives expression of UAS-luciferase. Bands in control lanes are endogenous, naturally biotinylated proteins. This experiment was performed twice with similar results. (h-k) Comparison of ligases in worms. (h) Scheme for tissue-specific expression of ligases in C. elegans intestine via ges-1p promoter. Transgenic strains are fed either biotin-producing E. coli OP50 (biotin+), or biotin-auxotrophic E. coli MG1655bioB:kan (biotin–). Promoter ges-1p drives ligase expression ~150 min after the first cell cleavage. (i) Adult worms prepared as in h were shifted to 25 °C for one generation, then lysed and analyzed by western blotting. Control worms (NZ) do not express ligase. Anti-HA antibody detects ligase expression. Streptavidin-IRDye detects biotinylated proteins. This experiment was performed five times (n = 5). In biotin+ conditions, BioID biotinylation activity was undetectable and TurboID gave robust biotinylation signal (n = 5/5). Despite high activity detected in embryos (see j,k), we only detected low levels of biotinylation by miniTurbo in adults (n = 2/5), likely due to low ligase expression. (j) Representative images of comma-stage worm embryos (stage 2) worm embryos from h. See Supplementary Figure 15a for representative images of comma-stage worm embryos (stage 2). Embryos were fixed and stained with streptavidin-AF488 to detect biotinylated proteins, and anti-HA antibody to detect ligase expression. Intestine is outlined by a white dotted line. Scale bar, 10 µm. (k) Quantitation of streptavidin-AF488 signal acquired from staining of embryonic stages 1 and 2 shown in j and Supplementary Figure 15a. Mean streptavidin pixel intensities for each embryo assessed are plotted for BioID (B), TurboID (T), and miniTurbo (mT). Two independent transgenic lines for BioID and TurboID and one for miniTurbo were assessed. Number of embryos imaged (n) from left to right: 26, 18, 11, 16, 25, 8, 19, 23, 14, 14, 23, 9. Statistical significance via Mann–Whitney U test (two-sided). **P < 0.001, *P < 0.01, P < 0.05. Pink asterisks indicate significance of pairwise comparisons between biotin- and corresponding biotin-treated embryos. Mean (reported in Supplementary Fig. 15b) is shown as a black horizontal line for each condition, and error bars indicate s.e.m. Note that the streptavidin-AF488 pixel intensities for miniTurbo are an underrepresentation of the signal as camera exposure settings were lowered to avoid pixel saturation (see Methods). See Supplementary Figure 15 for more details.
spectrometry. J.A.B. performed D. melanogaster experiments. J.A.B. and N.P. analyzed D. melanogaster data. T.C.B., A.Y.T., A.D.S., and J.L.F. designed the C. elegans experiments. A.D.S. performed C. elegans experiments. A.D.S. and J.L.F. analyzed C. elegans data.

COMPETING INTERESTS
A.Y.T. and T.C.B. have filed a patent application covering some aspects of this work.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Kim, D.I. & Roux, K.J. Filling the void: proximity-based labeling of proteins in living cells. Trends Cell Biol. 26, 804–817 (2016).
2. Rhee, H.-W. et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 339, 1328–1331 (2013).
20. Kehrer, J., Frischknecht, F. & Mair, G.R. Proteomic analysis of the Escherichia coli biotin protein ligase. Protein Sci. 13, 3043–3050 (2004).

21. Gaji, R.Y. et al. Phosphorylation of a myosin motor by TgCDPK3 facilitates rapid initiation of motility during Toxoplasma gondii egress. PLoS Pathog. 11, e1005268 (2015).

22. Batsios, P., Ren, X., Baumann, O., Larochelle, D.A. & Gräf, R. Src1 is a protein of the inner nuclear membrane interacting with the Dictyostelium Lamin NE81. Cells 5, 13 (2016).

23. Meyer, I. et al. CP39, CP75 and CP91 are major structural components of the Dictyostelium centrosome’s core structure. Eur. J. Cell Biol. 96, 119–130 (2017).

24. Uezu, A. et al. Identification of an elaborate complex mediating postsynaptic inhibition. Science 353, 1123–1129 (2016).

25. Opitz, N. et al. Capturing the Asci1p receptor for activated C kinase I (RACK1) microenvironment at the head region of the 4OS ribosome with quantitative BioID in yeast. Mol. Cell. Proteomics 16, 2199–2218 (2017).

26. Kim, D.I. et al. An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell 27, 1188–1196 (2016).

27. Ramanathan, M. et al. RNA-protein interaction detection in living cells. Nat. Methods 15, 207–212 (2018).

28. Birendra, K.C. et al. VRK2A is an A-type lamin-dependent nuclear envelope kinase that phosphorylates BAF. Mol. Biol. Cell 28, 2241–2250 (2017).

29. Redwine, W.B. et al. The human cytoplasmic dynein interactome reveals novel activators of motility. eLife 6, e28257 (2017).

30. Jung, E.M. et al. Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. Nat. Neurosci. 20, 1694–1707 (2017).

31. Martell, J.D. et al. A split horseradish peroxidase for the detection of intercellular protein-protein interactions and sensitive visualization of synapses. Nat. Biotechnol. 34, 774–780 (2016).

32. Bobrow, M.N., Harris, T.D., Shaughnessy, K.J. & Litt, G.J. Catalyzed reporter deposition, a novel method of signal amplification. Application to immunoassays. J. Immunol. Methods 125, 279–285 (1989).

33. Hung, V. et al. Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. eLife 6, e24463 (2017).

34. Dingar, D. et al. BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. J. Proteomics 118, 95–111 (2015).

35. Reineke, A.W., Balla, K.M., Bennett, E.J. & Troemel, E.R. Identification of microsporidia host-exposed proteins reveals a repertoire of rapidly evolving proteins. Nat. Commun. 8, 14023 (2017).

36. Reineke, A.W., Mak, R., Troemel, E.R. & Bennett, E.J. In vivo mapping of tissue- and subcellular-specific proteomes in Caenorhabditis elegans. Sci. Adv. 3, e1602426 (2017).

37. Chen, C.-L. Proteomic mapping in live Drosophila tissues using an engineered ascorbate peroxidase. Proc. Natl. Acad. Sci. USA 112, 12093–12098 (2015).
ONLINE METHODS

Cloning. See Supplementary Table 8 for a list of genetic constructs used in this study, with detailed description of construct designs, linker orientations, epitope tags, and signal sequence identities. All ligase variants were derived from E. coli biotin protein ligase, have the residue A146 deleted to suppress dimerization38, and are codon optimized for expression in mammalian cells. For cloning, PCR fragments were amplified using Q5 polymerase (New England BioLabs (NEB)). The vectors were double-digested using standard enzymatic restriction digest and ligated to gel purified PCR products by T4 DNA ligation or Gibson assembly. Ligated plasmid products were introduced by heat shock transformation into competent XL1-Blue bacteria. Ligase mutants were either generated using QuikChange mutagenesis (Stratagene) or isolated from individual yeast clones and transferred to mammalian expression vectors using standard cloning techniques.

Yeast cell culture. For yeast-display (Fig. 1c,e and Supplementary Figs. 3 and 4a), S. cerevisiae strain EBY100 was cultured according to previously published protocols39. Cells were propagated at 30 °C in synthetic dextrose plus casein amino acid (SDCAA, ‘regular’) medium supplemented with tryptophan (20 µg/mL). Yeast cells were transformed with the yeast-display plasmid pCTCON239 using the Frozen E-Z Yeast Transformation II kit (Zymoprep) according to manufacturer protocols. Transformed cells containing the TRP1 gene were selected on SDCAA plates and propagated in SDCAA medium at 30 °C. Protein expression was induced by inoculating saturated yeast culture into 10% SD/GCAA (SDCAA medium with 90% of dextrose replaced with galactose), or into “biotin-depleted” medium40 (1.7 g/L YNB-Biotin (Sunrise Science Products), 5 g/L ammonium sulfate, 2 g/L dextrose, 18 g/L galactose, complete amino acids, 0.125 ng/mL d-biotin), at a 1:100-1000 dilution and incubating at 30 °C for 18 – 24 h.

Generation of ligase libraries for yeast display. Libraries of ligase mutants were generated by error-prone PCR according to published protocols41. 150 ng of the template ligase in vector pCTCON2 (ref. 41) was amplified for 10 – 20 rounds with 0.4 µM forward and reverse primers: F: 5'-CTAAGGCAAGCCGCTGCTGCTGCGTGCTGTCG-3' R: 5'-TATCGACGCGTCCGACGCTGCTGCTGCTGCTGTCG-3'. The PCR products were then gel purified and reamplified for another 30 cycles under normal PCR conditions using: F: 5'-CAAGGCAAGCCGCTGCTGCTGCGTGCTGTCG-3' R: 5'-CTCACGCTGCTGCTGCTGCTGCGTGCTGTCG-3'.

The inserts were then electroporated into electrocompetent S. cerevisiae EBY100 (ref. 41) with the BamHI-Nhel linearized pCTCON2 vector (10 µg insert/1 µg vector) backbone. The electroporated cultures were rescued in 2 mL yeast extract peptone dextrose (YPD) complete medium42 for 1 h at 30 °C with no shaking. Cells were vortexed briefly, and 1.99 µL of the rescued cell suspension was transferred to 100 mL of SDCAA medium supplemented with 50 units/mL penicillin and 50 µg/mL streptomycin and grown for 2 days at 30 °C. The remaining 10 µL of the rescued cell suspension was diluted 100×, 1000×, 10,000×, and 100,000×; 20 µL of each dilution was plated on SDCAA plates and incubated at 30 °C for 3 days. After 3 days, each colony observed in the 100×, 1000×, 10,000×, or 100,000× segments of plates will correspond to 10^4, 10^5, 10^6, or 10^7 transformants in the library, respectively.

General methods for yeast display–based directed evolution. For each round of evolution (Supplementary Fig. 3), we input 10-fold more yeast cells than the estimated library size. For the first round, library size was estimated by the transformation efficiency of the initial ligase library. For subsequent rounds, library size was taken to be the number of yeast cells collected during the previous sort. Ligase protein expression was induced by inoculating saturated yeast culture into 10% SD/GCAA or biotin-depleted medium at a 1:1000 dilution and incubating at 30 °C for 18 – 24 h. However, when inducing yeast populations with high diversity, for example the initial libraries, saturated yeast cultures were diluted at as high as 1:100 dilution into the appropriate induction media in order to conserve media.

For samples biotin labeled for 18 h, yeast were induced in 10% SD/GCAA or biotin-depleted medium supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 at 30 °C for 20-24 h (to allow time for protein expression and ensure labeling occurs for at least 18 h). For samples labeled for shorter time periods, yeast were induced in 10% SD/GCAA or biotin-depleted medium for 18 h at 30 °C (no biotin), then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for the labeling times indicated. For ‘omit biotin’ samples, yeast were induced in 10% SD/GCAA or biotin-depleted medium at 30 °C for 18-24 h. After labeling, approximately 5 million cells (assuming 1 OD600 = 3 × 10^6 cells^4^) were pelleted at 5000×g for 30 s at 4 °C and washed five times with 1 mL PBS (phosphate buffered saline) + 0.5% bovine serum albumin (BSA; 1 mg/mL) (PBS-B).

For tyramide signal amplification42 (TSA, Supplementary Fig. 3b), yeast cells were incubated in 50 µL PBS-B + 1:100 streptavidin-horseradish peroxidase (HRP) for 1 h at 4 °C, then washed three times with 1 mL PBS-B. HRP labeling was performed by incubating yeast in 750 µL PBS-B with 50 µM biotin-phenol and 1 mM H2O2 for 1 min at room temperature. The reaction was quenched by adding 750 µL PBS-B + 20 mM sodium ascorbate and 10 mM Trollox followed by rapid mixing via inversion. Cells were then washed two times with 1 mL PBS-B + 10 mM sodium ascorbate and 5 mM Trollox, and once with 1 mL PBS-B. For removal of ligase proteins via TCEP reduction (Supplementary Fig. 3c), yeast were incubated in 500 µL PBS-B + 2 mM TCEP at 30 °C for 90 min, then washed four times with 1 mL PBS-B.

For biotin and epitope tag detection, yeast cells were then incubated in 50 µL PBS-B + 1:400 chicken anti-myc and 1:50 rabbit anti-biotin (when detecting biotinylated proteins with anti-biotin antibody) for 1 h at 4 °C, then washed three times with 1 mL PBS-B. Yeast cells were then incubated in 50 µL PBS-B + 1:200 Alexa Fluor 647 (AF647) goat-anti-chicken IgG and 1:50 phycoerythrin (PE) goat-anti-rabbit IgG (when detecting biotinylated proteins with anti-biotin antibody) or streptavidin-PE (when detecting biotinylated proteins with streptavidin) for 1 h at 4 °C, then washed three times with 1 mL PBS-B for FACS analysis. Refer to Supplementary Table 9 for a list of antibodies used in this study.

For two-dimensional FACS sorting, samples were resuspended in PBS-B at a maximal concentration of 100 million cells/mL and sorted on a BD FACS Aria II cell sorter (BD Biosciences) with the appropriate lasers and emission filters (561 nm excitation laser and 582/15 emission filter for PE, 640 nm excitation laser and 670/30 emission filter for AF647). To analyze and sort single yeast cells, cells were plotted by a forward-scatter area (FSC-A) and side-scatter area (SSC-A) and a gate was drawn around cells clustered between 10^4 – 10^5 FSC-A and 10^3 – 10^4 SSC-A. The population was then sorted using BD FACSDIVA software. For removal of ligase proteins via TCEP reduction, yeast cultures were diluted at as high as 1:100 dilution into the appropriate induction media in order to conserve media.
all data from FACS sorting and analysis. Summaries of all yeast-display directed evolution and resulting mutants are shown in Figure 1e, Supplementary Figures 3 and 4, and Supplementary Table 1, and are described in detail in the ‘Directed evolution of TurboID and miniTurbo’ sections below.

Directed evolution of TurboID and miniTurbo: generation 1. For the first round of evolution (Supplementary Fig. 3b), three libraries were generated using BirA-R118S (Supplementary Table 8) as the starting template. The three libraries were generated using error prone PCR as described above, using the following conditions to produce varying levels of mutagenesis:

Library 1: 1 µM 8-oxo-dGTP, 2 µM dPTP, 10 PCR cycles
Library 2: 2 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles
Library 3: 20 µM 8-oxo-dGTP, 20 µM dPTP, 10 PCR cycles

The library sizes (approximated by number of transformants as described above), were 1.4 × 10^6 for Library 1, 1.7 × 10^6 for Library 2, and 8 × 10^6 for Library 3. FACS analysis of the three libraries showed robust expression and wide range of activities for Library 1 and Library 2, however Library 3 showed poor expression and no activity. Sequencing of 24 clones in Library 1 revealed an average of 1.5 amino acid changes per ligase gene. Sequencing of 24 clones in Library 2 revealed an average of 2.4 amino acid changes per ligase gene.

Library 1 and Library 2 were combined and used as the initial population for the first round of selections. This combined library was induced as described above, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 5 × 10⁶ cells were prepared for sorting (assuming 1 OD₆₀₀ ≈ 3 × 10⁶ cells¹²) as described above with TSA treatment (Supplementary Fig. 3b). 6.24 × 10⁶ cells were sorted by FACS. A square gate that collected cells positive for both anti-myc and streptavidin (conjugated to fluorophores, see Supplementary Table 9) was drawn, and approximately 2.5 × 10⁶ cells were collected (4%) to give population E1-R1.

Population E1-R1 was passaged twice, and analyzed by FACS side-by-side with the original combined library and BirA-R118S to ensure the sort was successful (resulting population still had expression and had higher or equivalent activity). Sequencing of 24 clones from E1-R1 revealed an average of 1.5 mutations per ligase gene. Population E1-R1 was induced, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 10-fold excess cells were prepared for sorting (assuming 1 OD₆₀₀ ≈ 3 × 10⁶ cells¹²) as described above with TSA treatment. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 3.8% of cells were collected to give population E1-R2.

Population E1-R2 was passaged twice, and analyzed by FACS side-by-side with previous rounds and BirA-R118S. Sequencing of 24 clones from E1-R2 revealed an average of 1.5 mutations per ligase gene. Population E1-R2 was induced for ~18 h then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 5 × 10⁶ cells were prepared for sorting (assuming 1 OD₆₀₀ ≈ 3 × 10⁶ cells¹²) as described above with TSA treatment. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 0.7% of cells were collected to give population E1-R3.

Population E1-R3 was passaged twice, and analyzed by FACS side-by-side with previous rounds and BirA-R118S. Population E1-R3 was induced for ~18 h then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 6 hours. From this culture, approximately 10-fold excess cells were prepared for sorting with TSA treatment. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 2.4% of cells were collected to give population E1-R4.

Population E1-R4 was passaged twice, and analyzed by FACS side-by-side with previous rounds and BirA-R118S. Population E1-R4 was induced for ~18 h then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 3 hours. From this culture, approximately 10-fold excess cells were prepared for sorting. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 2.6% of cells were collected to give population E1-R5.

Population E1-R5 was passaged twice, and analyzed by FACS side-by-side with previous rounds and BirA-R118S. Population E1-R5 was induced for ~18 h then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 0.9% of cells were collected to give population E1-R6.

Population E1-R6 was passaged twice, and analyzed by FACS side-by-side with previous rounds and BirA-R118S. Sequencing of E1-R6 revealed several mutants with the mutation E313K. Several mutants with and without this mutation were assayed as single clones on the yeast surface, and the most promising mutants, including two with the E313K mutation, were assayed in the mammalian cell cytosol. While neither of the E313K mutants showed significant difference in activity to R118S over 24 h, they both showed very strong self-labeling at shorter time points, e.g. 1 h. The crystal structure of BirA¹³ shows that this residue points directly into the active site, where a lysine mutation could easily react with the phosphate group of biotin-5'-AMP. We removed this mutation from the two promising clones bearing it and assayed again in the mammalian cell cytosol. One of the mutants, denoted in this study as G1 (Supplementary Table 1), displayed significantly higher promiscuous activity than R118S after 24 hours of labeling. Another mutant from the mammalian cell screen, denoted in this study as R6-1 (Supplementary Table 1), also displayed significantly higher promiscuous activity than R118S after 24 h of labeling. Both of these mutants, with 4 mutations each, had each of their mutations removed individually and in different combinations. Analysis of the resulting mutants in mammalian cells showed that each mutation was contributing to increased activity relative to R118S observed for R6-1 and G1.

Directed evolution of TurboID and miniTurbo: generation 2. For the second round of evolution (Supplementary Fig. 3c), six libraries were generated. Three libraries were recombined using R6-1 (Supplementary Tables 1 and 8) as the starting template, and the three libraries were made using G1 (Supplementary Tables 1 and 8) as the starting template, both using error prone PCR with the following conditions:

Library 1: R6-1, 2 µM 8-oxo-dGTP, 2 µM dPTP, 10 PCR cycles
Library 2: R6-1, 2 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles
Library 3: R6-1, 20 µM 8-oxo-dGTP, 20 µM dPTP, 10 PCR cycles
Library 4: G1, 2 µM 8-oxo-dGTP, 2 µM dPTP, 10 PCR cycles
Library 5: G1, 2 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles
Library 6: G1, 20 µM 8-oxo-dGTP, 20 µM dPTP, 10 PCR cycles

The library sizes, as calculated by transformation efficiency, were 3.8 × 10⁷ for Library 1, 1.9 × 10⁷ for Library 2, 1.6 × 10⁷ for Library 3, 3.8 × 10⁷ for Library 4, 3.9 × 10⁷ for Library 5, and 3.9 × 10⁷ for Library 6. FACS analysis of the three libraries showed robust expression and wide range of activities for Libraries 1, 2, 4, and 5, however Libraries 3 and 6 showed poor expression and no activity.

Libraries 1, 2, 4, and 5 were combined and used as the initial population for the first round of selections. This combined library was induced, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 10-fold excess cells were prepared for sorting with TSA treatment. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 8.4% of cells were collected to give population E2-R1.

Population E2-R1 was passaged twice, and analyzed by FACS side-by-side with the combined library template. Population E2-R1 was induced, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 10-fold excess cells were prepared for sorting with TCEP treatment (Supplementary Fig. 3c) followed by TSA treatment. A square gate that collected cells positive for streptavidin but negative for anti-myc was drawn, and approximately 1.2% of cells were collected to give population E2-R2.

Population E2-R2 was passaged twice, and analyzed by FACS side-by-side with the combined library template and previous rounds. Population E2-R2 induced, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 10-fold excess cells were prepared for sorting with TSA treatment. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 19% of cells were collected to give population E2-R3.

Population E2-R3 was passaged twice, and analyzed by FACS side-by-side with previous rounds. Population E2-R3 was induced, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂, for 24 h. From this culture, approximately 10-fold excess cells were prepared for sorting. A square gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 1.4% of cells were collected to give population E2-R6.
were collected to give population E2-R4. From here on, only trapezoidal gates as described here were used for double-negative selections.

Population E2-R4 was passaged twice, and analyzed by FACS side-by-side with previous rounds. Population E2-R4 was induced for ~18 h, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and approximately 1.1% of cells were collected to give population E2-R5.

Population E2-R5 was passaged twice, and analyzed by FACS side-by-side with the combined library template and previous rounds. Population E2-R5 was induced for ~18 h, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 6 h. From this culture, approximately 10-fold excess cells were prepared for sorting with TCEP treatment followed by TSa. A square gate that collected cells positive for streptavidin and negative for anti-myc was drawn, and approximately 1.5% of cells were collected to give population E2-R6.

Population E2-R6 was passaged twice, and analyzed by FACS side-by-side with previous rounds and the combined library template. Sequencing of E2-R6 revealed several mutations that appeared in multiple clones. Several of these mutations were assayed as single clones on the yeast surface, however it was found after re-sequencing that many of the most promising mutants had mutated stop codons. After mutating back the stop codons, the mutants were re-assayed on the yeast surface, and the mutants that remained promising were assayed in the mammalian cell cytosol. One of the mutants, denoted in this study as G2, displayed significantly higher promiscuous activity than R118S, G1 (its template), or any other mutant tested after 1 h of labeling. G1, with 2 additional mutations relative to G1, had each or both of its mutations removed. Analysis of the resulting mutants in mammalian cells showed that each mutation was contributing to activity boost observed for G2.

Directed evolution of TurboID and miniTurbo: generation 3. For the third round of evolution (Supplementary Fig. 3d), three libraries were made using G2 as the starting template (Supplementary Tables 1 and 8) using error prone PCR with the following conditions:

- Library 1: 2 µM 8-oxo-dGTP, 2 µM dPTP, 10 PCR cycles
- Library 2: 2 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles
- Library 3: 10 µM 8-oxo-dGTP, 20 µM dPTP, 10 PCR cycles

The library sizes, as calculated by transformation efficiency, were 3.5 × 10⁸ for Library 1, 3.6 × 10⁸ for Library 2, and 6.8 × 10⁸ for Library 3. FACS analysis of the three libraries showed robust expression and wide range of activities for Library 1 and Library 2, however Library 3 showed weak expression and no activity.

Libraries 1 and 2 were combined and used as the initial population for the first round of selections. This combined library was induced for ~18 h, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and less than 0.1% of cells were collected to give population E3-R1.

Population E3-R1 was passaged twice, and analyzed by FACS side-by-side with G2 and the combined library template. Population E3-R1 was induced for ~18 h, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.15% of cells were collected to give population E3-R2.

Population E3-R2 was passaged twice, and analyzed by FACS side-by-side with G2, the combined library template, and previous rounds. Population E3-R2 was induced for ~18 h, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 10 min. From this culture, approximately 10-fold excess cells were prepared for sorting. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and less than 0.1% of cells were collected to give population E3-R3.

At E3-R3, it was noted that the population had strong streptavidin signal in the absence of exogenous biotin addition. Sequencing of population E3-R3 revealed that the majority of clones had a large insertion at the 5' of the ligase gene. Removal of this insertion restored biotin dependence, but also resulted in decreased activity (5-fold less than E3-R3). The library was 'cleaned' by removing this insertion via PCR with primers that restored the wild-type N-terminal sequence, and subjected to one additional round of double-positive selection with 10 minute labeling and 0.1% cells collected. The resulting population was E3-R4.

Population E3-R4 was passaged twice, and analyzed by FACS side-by-side with previous rounds. Sequencing of E3-R4 revealed several mutations that appeared in multiple clones. Several of these mutants were assayed as single clones on the yeast surface, the most promising mutants were assayed in the mammalian cell cytosol. Two mutants had significantly higher activity than the template G2 or any other mutants. The mutations from these mutants were combined in various combinations, resulting in the highest activity mutant, denoted in this study as G3 (Supplementary Table 1).

Directed evolution of TurboID and miniTurbo: Generations 4 and 5. G3 was the highest activity mutant found to date, but it also appeared to give labeling even without the addition of exogenous biotin. This was observed in yeast, where this signal proved to be biotin-dependent (Supplementary Fig. 3e), and also in the mammalian cytosol (Fig. 1f, Supplementary Fig. 4b). From this point, we continued with two evolutionary paths as follows:

In one path, we truncated the N-terminal domain (aa 1-63) of G3 to give G3∆ (Supplementary Table 1). Consistent with literature⁴⁴,⁴⁵, this truncation resulted in reduced streptavidin signal when exogenous biotin was omitted (Supplementary Fig. 3e). Using G3∆ as the starting template (Supplementary Tables 1 and 8) for another round of evolution (Supplementary Fig. 3g), we generated three libraries using error prone PCR with the following conditions:

- Library 1: 2 µM 8-oxo-dGTP, 2 µM dPTP, 10 PCR cycles
- Library 2: 2 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles
- Library 3: 4 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles

The library sizes, as calculated by transformation efficiency, were 4.9 × 10⁸ for Library 1, 4.6 × 10⁸ for Library 2, and 3.7 × 10⁸ for Library 3. FACS analysis of the three libraries showed robust expression and wide range of activities for all libraries, therefore all were combined and used for the first round of selections.

This combined library was induced in biotin-depleted media, supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 18 h. From this culture, approximately 10-fold excess cells were prepared for sorting with streptavidin. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.1% of cells were collected to give population E4-R1.

Population E4-R1 was passaged twice, and analyzed by FACS side-by-side with G3∆ and the combined library template. Population E4-R1 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP and 5 mM MgCl₂ for 3.5 h. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A trapezoidal gate that collected cells positive for both anti-myc and anti-biotin was drawn, and 1% of cells were collected to give population E4-R2.

Population E4-R2 was passaged twice, and analyzed by FACS side-by-side with G3∆, the combined library template, and previous rounds. Population E4-R2 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting with streptavidin. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.2% of cells were collected to give population E4-R3.

Population E4-R3 was passaged twice, and analyzed by FACS side-by-side with G3∆, the combined library template, and previous rounds. Population E4-R3 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl₂ for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A trapezoidal gate that collected cells positive for both anti-myc and anti-biotin was drawn, and 0.1% of cells were collected to give population E4-R4.

Population E4-R4 was passaged twice, and analyzed by FACS side-by-side with G3∆, the combined library template, and previous rounds. Population E4-R4 was induced for ~18 h in biotin-depleted media, labeling was omitted for negative selection (Supplementary Fig. 3f). From this culture, approximately 10-fold excess cells were prepared for sorting with streptavidin. A square gate that collected cells positive for anti-myc and negative for streptavidin was drawn, and 50% of cells were collected to give population E4-R5.
Population E4-R5 was passaged twice, and analyzed by FACS side-by-side with G33, the combined library template, and previous rounds. Two selections were performed on E4-R5. In the first selection, population E4-R5 was induced for ~18 h in biotin-depleted media, labeling was omitted for negative selection. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A square gate that collected cells positive for both anti-myc and anti-biotin was drawn, and 45% of cells were collected to give population E4-R6.1.

In the second selection, population E4-R5 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for 20 min. From this culture, approximately 10-fold excess cells were prepared for sorting with streptavidin. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.1% of cells were collected to give population E4-R6.2.

One more round of selections was performed on E4-R6.1, which was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for 1 h. From this culture, approximately 10-fold excess cells were prepared for sorting with streptavidin. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.2% of cells were collected to give population E4-R7.

Population E4-R7 was passaged twice, and analyzed by FACS side-by-side with previous rounds. Sequencing of E4-R7 revealed several mutations that appeared in multiple clones. Several of these mutations were assayed as single mutations and in various combinations in the mammalian cytosol. One mutation, K194I, was found to significantly increase activity while not increasing signal exogenous when biotin is omitted. Introducing K194I into G33 resulted in miniTurbo (Supplementary Table 1).

In a second evolutionary path, we continued with evolving G3 (Supplementary Fig. 3h). Two libraries were made using G3 as the starting template (Supplementary Tables 1 and 8) using error prone PCR with the following conditions:

- Library 1: 2 µM 8-oxo-dGTP, 2 µM dPTP, 10 PCR cycles
- Library 2: 2 µM 8-oxo-dGTP, 2 µM dPTP, 20 PCR cycles

The library sizes, as calculated by transformation efficiency, were 2 × 10^6 for Library 1 and 1.1 × 10^6 for Library 2. FACS analysis of the libraries showed robust expression and wide range of activities for Library 1 and Library 2. Libraries 1 and 2 were combined and used as the initial population for the first round of selections. This combined library was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for 10 min. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody (Supplementary Table 9) in place of streptavidin. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.1% of cells were collected to give population E5-R1.

Population E5-R1 was passaged twice, and analyzed by FACS side-by-side with G3 and the combined library template. Population E5-R1 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for 10 min. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A trapezoidal gate that collected cells positive for both anti-myc and anti-biotin was drawn, and 0.1% of cells were collected to give population E5-R2.

Population E5-R2 was passaged twice, and analyzed by FACS side-by-side with G3, the combined library template, and previous rounds. Population E5-R2 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for 10 min. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A trapezoidal gate that collected cells positive for both anti-myc and anti-biotin was drawn, and 1.7% of cells were collected to give population E5-R3.

Population E5-R3 was passaged twice, and analyzed by FACS side-by-side with G3, the combined library template, and previous rounds. Population E5-R3 was induced for ~18 h in regular media, labeling was omitted for negative selection. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A square gate that collected cells positive for anti-myc and negative for anti-biotin was drawn, and 34% of cells were collected to give population E5-R4.

Population E5-R4 was passaged twice. FACS analysis side-by-side with G3, the combined library template, and previous rounds showed that the negative selection that resulted E5-R4 reduced overall activity of the population. Population E5-R4 was induced for ~18 h in biotin-depleted media, then supplemented with 50 µM biotin, 1 mM ATP, and 5 mM MgCl2 for 10 min. From this culture, approximately 10-fold excess cells were prepared for sorting with streptavidin. A trapezoidal gate that collected cells positive for both anti-myc and streptavidin was drawn, and 0.8% of cells were collected to give population E5-R5.

Population E5-R5 was passaged twice, and analyzed by FACS side-by-side with G3, the combined library template, and previous rounds. Population E5-R5 was induced for ~18 h in regular media, labeling was omitted for negative selection. From this culture, approximately 10-fold excess cells were prepared for sorting with anti-biotin antibody. A square gate that collected cells positive for both anti-myc and negative for anti-biotin was drawn, and 11.6% of cells were collected to give population E5-R6.

Population E5-R6 was passaged twice, and analyzed by FACS side-by-side with previous rounds. Sequencing of E5-R6 revealed several mutations that appeared in multiple clones. Several of these mutations were assayed as single mutations and in various combinations in the mammalian cytosol. None of the mutations gave dramatic increases in activity, but one mutation M241T, appeared to impart benefits to activity.

Screening of mutations present in E4-R6.2 in the mammalian cell cytosol revealed one mutation, S263P, which boosted activity, but also increased signal when biotin was omitted. This mutation, along with K194I from E4-R7 and M241T from E5-R6, were introduced into G33 to give TurboID (Supplementary Table 1). We also tested M241T in miniTurbo, however it was not added because it increased background signal when biotin was omitted.

Mammalian cell culture, transfection, and stable cell line generation. HEK 293T cells from ATCC (passage number ~25) were cultured as a monolayer in growth media (either MEM (Cellgro) or a 1:1 DMEM:MEM mixture (Cellgro) supplemented with 10% (w/v) fetal bovine serum (VWR)), 50 units/mL penicillin, and 50 µg/mL streptomycin at 37 °C under 5% CO2. Mycoplasma testing was not performed before experiments. For confocal imaging experiments, cells were grown on 7 × 7 mm glass coverslips in 48-well plates with 250 µL growth medium. To improve adherence of HEK 293T cells, glass coverslips were pretreated with 50 µg/mL fibronectin (Millipore) in MEM for at least 20 min at 37 °C before cell plating. For Western blotting, cells were grown on polystyrene 6-well plates (Greiner) with 2.5 mL growth medium.

For transient expression (Fig. 1f, 2c and Supplementary Figs. 1, 4b, 5, 10b, and 14c,d), cells were typically transfected at approximately 60% confluence using 3.2 µL/mL Lipofectamine2000 (Life Technologies) and 800 ng/mL plasmid in serum-free media (250 µL total volume for 48-wells, 2.5 mL total volume for 6-wells) for 3–4 h, after which time Lipofectamine-containing media was replaced with fresh serum-containing media.

In an attempt to achieve similar expression levels of ligase in the experiment presented in Figure 2a, b, Supplementary Figure 6a, b, and Supplementary Figure 7, cells were transfected at approximately 60% confluence using 1.6 µL/mL Lipofectamine2000 (Life Technologies) in serum-free media with the following amounts of each plasmid (250 µL total volume for 48-wells, 2.5 mL total volume for 6-wells): 160 ng/mL V5-BioID-NES, 80 ng/mL V5-TurboID-NES, 200 ng/mL V5-miniTurbo-NES, 30 ng/mL V5-BioID2-NES, and 1000 ng/mL V5-BASU-NES (Supplementary Table 8). After 3–4 h, the Lipofectamine-containing media was replaced with fresh serum-containing media.

In an attempt to achieve similar expression levels of ligase in the experiment presented in Supplementary Figure 6c–e, cells were transfected at approximately 60% confluence using 1.6 µL/mL Lipofectamine2000 (Life Technologies) in serum-free media with the following amounts of each plasmid (250 µL total volume for 48-wells, 2.5 mL total volume for 6-wells): 320 ng/mL V5-BioID-NES, 160 ng/mL V5-TurboID-NES, 400 ng/mL V5-miniTurbo-NES, 60 ng/mL V5-BioID2-NES, and 1000 ng/mL V5-BASU-NES (Supplementary Table 8). After 3–4 h, the Lipofectamine-containing media was replaced with fresh serum-containing media.

For preparation of lentiviruses, HEK 293T cells in T25 flasks (BioBasic) were transfected at ~60–70% confluence with the lentiviral vector pxL304 containing the gene of interest (2500 ng; Supplementary Table 8), and the lentiviral packaging plasmids pVSVG (250 ng; Supplementary Table 8) and Δ8.9 (2250 ng; Supplementary Table 8) with 30 µL Lipofectamine2000 in

doi:10.1038/nbt.4201

NATURE BIOTECHNOLOGY
serum-free media for 3 h, after which time the Lipofectamine-containing media was replaced with fresh serum-containing media. Approximately 60 h after transfection, the cell medium containing the lentivirus was harvested and filtered through a 0.45-µm filter.

To generate stable cell lines, HEK cells were then infected at ~50% confluence, followed by selection with 8 µg/mL blasticidin in growth medium for at least 7 days before further analysis (Fig. 2c and Supplementary Figs. 8, 9a, b, 10c and 14a, b).

Biotin labeling with TurboID and miniTurbo in live mammalian cells.

For labeling of transiently transfected cells, we initiated biotin labeling 18 to 36 h following transfection. From a 100 mM biotin stock in dimethyl sulfoxide (DMSO), we diluted biotin directly into serum-containing cell culture medium, to the desired final concentration. For BioID, we used a final concentration of 50 µM biotin. For TurboID and miniTurbo, we typically used a final concentration of 500 µM biotin (unless indicated otherwise). Labeling was stopped after the desired time period by transferring the cells to ice and washing five times with ice-cold PBS. Supplementary Figure 5 shows that cooling samples to 4°C terminates the biotinylation reaction. For negative controls, we omitted exogenous biotin, or omitted the ligase.

Gels and Western blots. For gels and Western blots shown in Figures 1f and 2a,c and Supplementary Figures 1, 2a, 2b, 5b, 6b, 8b-d, 9a and 10b,c, HEK 293T cells expressing the indicated constructs were plated, transfected, and labeled with biotin as described above, and subsequently detached from the flask by gently pipetting a stream of ice-cold PBS directly onto the cells. Pellets were collected by centrifuging the resulting cell suspension at 1,500 rpm for 3 min. The supernatant was removed, and the pellet was lysed by resuspending in RIPA lysis buffer (50 mM Tris pH 8, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 1x protease inhibitor cocktail (Sigma-Aldrich), and 1 mM PMSE) by gentle pipetting and incubating for 5 min at 4°C. Lysates were clarified by centrifugation at 10,000 rpm for 10 min at 4°C. Protein concentration in clarified lysate was estimated with Pierce BCA Protein Assay Kit (ThermoFisher) prior to separation on a 9% SDS-PAGE gel. Silver-stained gels (Supplementary Figures 9a, 10b–c) were generated using Pierce Silver Stain Kit (ThermoFisher).

For the Western blot experiment in Figure 3a comparing ligase variants in the yeast cytosol, S. cerevisiae strain BY4741 cells were propagated at 30°C in supplemented minimal medium (SMM; 6.7 g/L Difco nitrogen base without amino acids, 20 g/L dextrose, 0.54 g/L CSM – Ade – His – Leu – Lys – Try – Ura (Sunrise Science Products), 20 mg/L adenine, 20 mg/L uracil, 20 mg/L histidine, 30 mg/L lysine) supplemented with leucine (100 mg/L). Yeast cells were transformed with pBS415 plasmids using the Frozen-E-Z Yeast Transformation II kit (Zymoprep) according to manufacturer protocols. Transformed cells containing the LEU2 gene were selected on SMM plates (SMM with 20 g/L agar) and propagated in SMM at 30°C. Ligase expression was induced by inoculating saturated yeast culture into 10% D/G SMM (SMM medium with 90% of dextrose replaced with galactose) supplemented with 50 µM biotin at a 1:100 dilution and incubating at 30°C. After ~12 h, the saturated induced culture was diluted 1:30 in fresh induction media supplemented with 50 µM biotin and allowed to grow for approximately 6 h more until reaching OD600 ~1. Three milliliters of this culture was pelleted (normalized across samples so that the same approximate amount of cells are collected for each sample), and resuspended in 15 µL 6x protein loading buffer (0.33 M Tris-HCl pH 8, 34% glycerol, 94 mg/mL SDS, 88 mg/mL DTT, 113 µg/mL bromophenol blue). The protein was boiled for 5 min, diluted to 1x, and then separated on a 9% SDS-PAGE gel.

For all Western blots in Figures 1f, 2a,c and 3a,b and Supplementary Figures 1, 4b, 5b, 6b–d, 9a and 10b,c, proteins separated on SDS-PAGE gels were transferred to nitrocellulose membrane, and then stained by Ponceau S (5 min in 0.1% (w/v) Ponceau S in 5% acetic acid/water). The blots were then blocked in 5% (w/v) milk (LabScientific) in TBS-T (Tris-buffered saline, 0.1% Tween 20) for at least 30 min at room temperature, or as long as overnight at 4°C. Blots were then washed with primary antibodies (Supplementary Table 9) in 3% BSA (w/v) in TBS-T for 1-1.5 h at 4°C, washed four times with TBS-T for 5 min each, then washed with secondary antibodies or 0.3 µg/mL streptavidin–HRP (Supplementary Table 9) in 3% BSA (w/v) in TBS-T for 1-1.5 h at 4°C. The blots were washed four times with TBS-T for 5 min each time before development with Clarity Western ECL Blotting Substrates (Bio-Rad) and imaging on a UVP BioSpectrum Imaging System, except for blots in Figures 1f, 2c and 3a,b, and Supplementary Figures 5a (miniTurbo) and 10b,c, which were imaged on a Bio-Rad Gel Doc 2000. Quantitation of Western blots was performed using ImageJ on raw images under non-saturating conditions.

Confocal fluorescence imaging of cultured cells. For fluorescence imaging experiments in Supplementary Figures 7, 8e, 9b, 10d,e and 14e,f, HEK 293T cells expressing the indicated constructs were plated, transfected, and labeled with biotin as described above, and subsequently fixed with 4% (v/v) paraformaldehyde in PBS at 4°C for 45 min. Cells were then washed three times with PBS and permeabilized with cold methanol at -20°C for 5 min. Cells were then washed three times with PBS, and then incubated with primary antibody (Supplementary Table 9) in PBS supplemented with 3% (w/v) BSA for 1 h at 4°C. After washing three times with PBS, cells were then incubated with DAPI/secondary antibody, and neutravidin–Alexa Fluor647 (Supplementary Table 9) in PBS supplemented with 3% (w/v) BSA for 1 h at 4°C. Cells were then washed three times with PBS and imaged by confocal fluorescence microscopy.

Confocal imaging was performed using a Zeiss AxioObserver.Z1 microscope, outfitted with a Yokogawa spinning disk confocal head, a Cascade II:512 camera, a Quad-band notch dichroic mirror (405/488/568/647), 405 (diode), 491 (DPSS), 561 (DPSS), and 640 (diode) nm lasers (all 50 mW). DAPI (405 laser excitation, 445/40 emission), Alexa Fluor568 (561 laser excitation, 617/73 emission), and Alexa Fluor647 (640 laser excitation, 700/75 emission), and differential interference contrast (DIC) images were acquired through a 63x oil-immersion objective; Acquisition times ranged from 50 to 100 ms. All images were collected and processed using SlideBook 6.0 software (Intelligent Imaging Innovations). The data in Supplementary Figures 7, 8e, 9b, 10d,e and 14e,f are representative of at least 10 fields of view.

Conjugation of AlexaFlour647 to neutravidin. A reaction mixture was assembled in a 1.5 mL Eppendorf tube with the following components (added in this order): 200 µL of 5 mg/mL Neutravidin (Life Technologies) in PBS, 20 µL of 1 M sodium bicarbonate in water, and 10 µL of 10 mg/mL AlexaFluor647–NHS Ester (Life Technologies) in anhydrous DMSO. The tube was incubated at room temperature with rotation in the dark for 3 h. The neutravidin–AlexaFluor647 conjugate was purified from unreacted dye using a NAP-5 size-exclusion column (GE Healthcare Life Sciences) according to the manufacturer’s instructions. The conjugate was typically eluted from the column in 500 µL cold PBS. Absorbance values, determined using a Nanodrop 2000C UV-vis spectrophotometer (Thermo Scientific), were typically as follows: A280 = -0.284 and A647 = -1.625. The conjugate was stable at 4°C in the dark for at least 4 months and was flash frozen and stored at 80°C for longer term storage. For mammalian cell labeling experiments, the conjugate was diluted 1,000-fold in PBS containing 1% BSA.

Sample preparation for proteomics and for the Western blot experiment in Supplementary Figure 8. For each sample, HEK 293T cells were grown as a monolayer in 1:1 DMEM/MEM mixture (Cellgro) supplemented with 10% (v/v) fetal bovine serum (VWR), 50 µg/mL penicillin, and 50 µg/mL streptomycin in T150 flasks at 37°C under 5% CO2. Nuclear proteomic samples were generated by transfecting cells at approximately 60% confluency with...
30 µg DNA using 150 µL Lipofectamine 2000 for 4 h. Mitochondrial matrix, ER membrane, and outer mitochondrial membrane samples were generated using HEK 293T cell lines that stably express the respective ligase. BioID samples were labeled using 50 µM biotin in 18 h; TurboID and miniTurbo samples were labeled using 500 µM biotin for 10 min. Labeling was stopped by placing cells on ice and washing five times with ice-cold PBS (Supplementary Fig. 5). Cells were detached from the flask by gently pipetting a stream of PBS directly onto the cells, then pellets were collected by centrifuging the resulting cell suspension at 1,500 r.p.m. for 3 min. The supernatant was removed, and the pellet was lysed in −1.5 mL RIPA lysis buffer by gentle pipetting and incubating for 5 min at 4 °C. Lysates were clarified by centrifugation at 10,000 r.p.m. for 10 min at 4 °C.

To enrich biotinylated material from proteomic samples, 350 µL streptavidin-coated magnetic beads (Pierce) were washed twice with RIPA buffer, incubated with clarified lysates containing ~3 mg protein for each sample with rotation for 1 h at room temperature, then moved to 4 °C and incubated overnight with rotation. The beads were subsequently washed twice with 1 mL of RIPA lysis buffer, once with 1 mL of 1 M KCl, once with 1 mL of 0.1 M Na2CO3, once with 1 mL of 2 M urea in 10 mM Tris-HCl (pH 8.0), and twice with 1 mL RIPA lysis buffer. The beads were then resuspended in 1 mL fresh RIPA lysis buffer, transferred to a new Eppendorf tube, and shipped to Steve Carr’s laboratory (Broad Institute) on ice for further processing and preparation for LC-MS/MS analysis.

For proteomic samples, 2.5% of the lysate was removed prior to enrichment used to estimate the protein concentration in clarified lysates using Pierce BCA Protein Assay Kit (ThermoFisher), as well as to verify ligase expression and confirm successful biotinylation by Western blotting as shown in Supplementary Figures 9a and 10b,c. After enrichment, 5% of beads were removed and biotinylated proteins were eluted by boiling the beads in 75 µL of 3x protein loading buffer supplemented with 20 mM DTT and 2 mM biotin. The eluted proteins were separated on an SDS-PAGE gel and stained using Pierce Silver Stain Kit to ensure that enrichment of biotinylated material was successful (Supplementary Figures 9a and 10b,c).

To enrich biotinylated material from samples prepared for Supplementary Figure 8, 240 µL streptavidin-coated magnetic beads (Pierce) were washed twice with RIPA buffer, incubated with clarified lysates containing approximately 3 mg protein for each sample with rotation for 1 h at room temperature. The beads were subsequently washed twice with 1 mL of RIPA lysis buffer, once with 1 mL of 1 M KCl, once with 1 mL of 0.1 M Na2CO3, once with 1 mL of 2 M urea in 10 mM Tris-HCl (pH 8.0), and twice with 1 mL RIPA lysis buffer. Biotinylated proteins were then eluted from the beads by boiling the beads in 300 µL of 3x protein loading buffer supplemented with 20 mM DTT and 2 mM biotin. For each sample, 25 µL of this eluate was then separated on SDS-PAGE gel alongside 35 µg protein from the corresponding clarified lysate prior to enrichment, and then transferred to nitrocellulose for Western blotting as described above with antibodies against the endogenous proteins indicated in Supplementary Figure 8 (also see Supplementary Table 9).

On-bead trypsin digestion of biotinylated proteins. To prepare samples for mass spectrometry analysis, proteins bound to streptavidin beads (~300 µL of slurry) were washed twice with 200 µL of 50 mM Tris HCl buffer (pH 7.5) followed by two washes with 2 M urea/50 mM Tris (pH 7.5) buffer. The final volume of 2 M urea/50 mM Tris buffer (pH 7.5) was removed and beads were incubated with 80 µL of 2 M urea/50 mM Tris containing 1 mM DTT and 0.4 µg trypsin for 1 h at 25 °C with shaking. After 1 h, the supernatant was removed and transferred to a fresh tube. The streptavidin beads were washed twice with 60 µL of 2 M urea/50 mM Tris buffer (pH 7.5) and the washes were combined with the on-bead digest supernatant. The eluate was reduced with 4 mM DTT for 30 min at 25 °C with shaking. The samples were alkylated with 10 mM iodoacetamide for 45 min in the dark at 25 °C with shaking. An additional 0.5 µg of trypsin was added to the sample and the digestion was completed overnight at 25 °C with shaking. After overnight digestion, samples were acidified (to pH < 3) by adding formic acid (FA) such that the sample contained ~1% FA. Samples were desalted on C18 StageTips and evaporated to dryness in a vacuum concentrator, as previously described46.

TMT labeling and fractionation of peptides. Desalted peptides were labeled with TMT (6-plex or 11-plex) reagents. Peptides were reconstituted in 100 µL of 50 mM HEPES. Each 0.8 mg vial of TMT reagent was reconstituted in 41 µL of anhydrous acetone and added to the corresponding peptide sample for 1 h at room temperature. Labeling of samples with TMT reagents was completed with the design shown in Fig. 2d and Supplementary Figure 10a. TMT labeling reactions were quenched with 8 µL of 5% hydroxyamine at room temperature for 15 min with shaking, evaporated to dryness in a vacuum concentrator, and desalted on C18 StageTips. For each TMT 6-plex cassette and the TMT 11-plex cassette, 50% of the sample was fractionated by basic pH reversed phase using StageTips while the other 50% of each sample was reserved for LC-MS analysis by a single-shot, long gradient. One StageTip was prepared per sample using 2 plugs of Styrene Divinylbenzene (SDB) (3M) material. The StageTips were conditioned twice with 50 µL of 100% methanol, followed by 50 µL of 50%MeCN/0.1% FA, and two times with 75 µL of 0.1% FA. Sample, resuspended in 100 µL of 0.1% FA, was loaded onto the stage tips and washed with 100 µL of 0.1% FA. Following this, sample was washed with 60 µL of 20mM NH4HCO3/2% MeCN, this wash was saved and added to fraction 1. Next, sample was eluted from StageTip using the following concentration gradients of MeCN in 20 mM NH4HCO3: 10%, 15%, 20%, 25%, 30%, 40%, and 50%. For a total of 6 fractions, 10 and 40% (fractions 2 and 7) elutions were combined, as well as 15 and 50% elutions (fractions 3 and 8). The six fractions were dried by vacuum centrifugation.

Liquid chromatography and mass spectrometry. Desalted peptides were resuspended in 9 µL of 3% MeCN/0.1% FA and analyzed by online nanoflow liquid chromatography tandem mass spectrometry (LC-MS/MS) using an Orbitrap Fusion Lumos Tribrid MS (ThermoFisher Scientific) coupled online to a Proxence Easy-nLC 1200 (ThermoFisher Scientific). Four microliters of each sample was loaded onto a microcapillary column (360 µm outer diameter × 75 µm inner diameter) containing an integrated electrospray emitter tip (10 µm), packed to approximately 24 cm with ReproSil-Pur C18-AQ 1.9 µm beads (Dr. Maisch GmbH) and heated to 50 °C. The HPLC solvent A was 3% MeCN, 0.1% FA, and the solvent B was 90% MeCN, 0.1% FA. The SDF fractions were measured using a 110 min MS method, which used the following gradient profile: (min:%B) 0:2; 1:6; 85:30; 94:60; 95:90; 100:90; 101:50; 110:50 (the last two steps at 500 nL/min flow rate). Non-fractionated samples were analyzed using a 260 min MS method with the following gradient profile: (min:%B) 0:2; 1:6; 235:30; 246:60; 245:90; 250:90; 251:50; 260:50 (the last two steps at 500 nL/min flow rate).

The Orbitrap Fusion Lumos Tribrid was operated in the data-dependent mode acquiring HCD MS/MS scans (resolution = 15,000 for TMT6-plex, or resolution = 50,000 for TMT11-plex) after each MS1 scan (resolution = 60,000) on the most abundant ions within a 2 s cycle time using an MS1 target of 3 × 106 and an MS2 target of 5 × 104. The maximum ion time utilized for MS/MS scans was 50 ms for TMT6-plex experiments and 105 ms for the TMT11-plex experiment; the HCD normalized collision energy was set to 34 for TMT6 and 38 for TMT11; the dynamic exclusion time was set to 45 s, and the peptide match and isotope exclusion functions were enabled. Charge exclusion was enabled for charge states that were unassigned, 1 and >6.

Analysis of mass spectrometry data. Collected data were analyzed using Spectrum Mill software package v6.1pre-release (Agilent Technologies). Nearby MS scans with the similar precursor m/z were merged if they were within ±60 s retention time and ±1.4 m/z tolerance. MS/MS spectra were excluded from searching if they failed the quality filter by not having a sequence tag length 0 or did not have a precursor MH+ in the range of 750–4000. All extracted spectra were searched against a UniProtdatabase containing human reference proteome sequences. Search parameters included: parent and fragment mass tolerance of 20 p.p.m., 30% minimum matched peak intensity, and ‘calculate reversed database scores’ enabled. The digestion enzyme search parameter used was Trypsin Allow P, which allows K-P and R-P cleavages. The missed cleavage allowance was set to 4. Fixed modifications were carbamidomethylation at cysteine. TMT labeling was required at lysine, but peptide N termini were allowed to be either labeled or unlabeled. Allowed variable modifications were protein N-terminal acetylation and oxidized methionine. Individual spectra were automatically assigned a confidence score using the
Supplementary Table 5

To calculate optimal cut-offs, we then calculated the true positive rate (TPR) and false positive rate (FPR) we would obtain if we retained only proteins above the TMT ratio in question. We selected TMT ratios that maximize the difference between TPR and FPR as our cutoffs (Supplementary Fig. 9c).

To select cutoffs for proteins biotinylated by the indicated ERM-ligase over non-specific bead binders, we classified the detected proteins into three groups:

1. ERM proteins (Supplementary Table 2; true positive list of 90 well-established ERM proteins).
2. Non-secretory proteins (Supplementary Table 2; false positive list of 7421 human proteins that are not predicted to be secretory by Phobius48 or are not annotated with the following Gene Ontology49,50 terms: GO:0005783, GO:0005798, GO:0005802, GO:0006887, GO:0006890, GO:0005801, GO:0012510, GO:0005635, GO:0007009, GO:0012505).
3. All other proteins

We then normalized the TMT ratios in order to account for differences in total protein quantity between samples within the TMT 11-plex experiment. To do this, the Log2(TMT ratios) corresponding to ERM-ligase/untransfected (Log2[(127N/126C), Log2(128N/126C), Log2(129N/126C), Log2(128C/126C)], Log2(130N/126C), Log2(131N/126C), Log2(131C/126C)]) were normalized to the median for class (2) proteins, which was set to 0 (i.e. TMT ratios set to 1). To calculate optimal cut-offs, we then calculated the true positive rate (TPR) and false positive rate (FPR) we would obtain if we retained only proteins above that TMT ratio. We defined TPR as the fraction of class (1) proteins above the TMT ratio in question, and FPR as the fraction of class (2) above the TMT ratio in question. We selected TMT ratios that maximize the difference between TPR and FPR as our cutoffs (Supplementary Table 5).

Generation of proteomic lists for the ER membrane. Complete mass spectrometry data for the ER membrane (ERM) proteomic experiment are shown in Supplementary Table 5. To select cutoffs for proteins biotinylated by the indicated ligase over non-specific bead binders, we classified the detected proteins into three groups:

1. ERM proteins (Supplementary Table 2; true positive list of 90 well-established ERM proteins).
2. Soluble matrix proteins (Supplementary Table 2; false positive list of 173 soluble mitochondrial matrix proteins).
3. All other proteins

To assess the specificity of our proteomes (Fig. 2e), we report the percentage of proteins present in Supplementary Table 2, a list of 11,838 human proteins with secretory annotation according to Phobius48, the Human Protein Atlas51 (protein localized to endoplasmic reticulum, Golgi apparatus, plasma membrane, vesicles, nuclear membrane, cell junctions; or predicted membrane proteins and predicted secreted proteins), the Plasma Proteome Database52, literature (referenced in cited table, or are annotated with the following Gene Ontology49,50 terms: GO:0005783, GO:0005798, GO:0007029, GO:0038087, GO:0048237, GO:0061163, GO:0016320, GO:0030868, GO:006983, GO:0001319, GO:0051645, GO:0031985, GO:0005796, GO:0005795, GO:0005794, GO:0007030, GO:0032587, GO:0016021, GO:0005886, GO:0007009, GO:1903561, GO:0070062, GO:0035426, GO:1903053, GO:1903551, GO:0005578, GO:0005783, GO:0005798, GO:0005793, GO:0006887, GO:0006890, GO:0005801, GO:0012510, GO:0005637, GO:0071765, GO:1905719, GO:0031985, GO:0035646, GO:0090341, GO:0090340, GO:0005635, GO:0007077, GO:006989, GO:0051081, GO:0065641, GO:0031965, GO:0005637, GO:0071765, GO:0048471, GO:0030968, GO:1902236, GO:0048278, GO:0032587, GO:0016021, GO:0050887, GO:00071816, GO:0031526, GO:0091593, GO:0072546, GO:1990440, GO:0039068, GO:0005913, GO:0072546, GO:1902236, GO:1994041, GO:0034976, GO:0057788, GO:0005790, GO:1902237, GO:0007009, GO:0005786, GO:0005793, GO:0044322, GO:0098554, GO:0005791, GO:0021060, GO:0043001, GO:0005802, GO:006989, GO:0005801, GO:0012510, GO:0008924, GO:0041247, GO:0034999, GO:0032588, GO:0003104, GO:0005689, GO:0005689, GO:0000042, GO:0032580, GO:0003173, GO:0008924, GO:0030198, GO:0031668, GO:0010715, GO:0035426, GO:1903055, GO:0001560, GO:0022167, GO:0006887, GO:00012505).

To assess the recall/sensitivity of our ERM proteomes, we utilized a list of true positive ERM proteins (Supplementary Fig. 9f, Supplementary Table 2). In the scatter plots shown in Supplementary Figure 9e, true positive ERM proteins (Supplementary Table 2) are shown in green, cytosolic proteins (Supplementary Table 2; human proteins with Gene Ontology49,50 term GO:0005783 that lack annotated or predicted transmembrane domains according to UniProt47 or TMHMM53) are shown in red, and all other proteins are shown in black.

Generation of mitochondrial matrix and nuclear proteomic lists. Complete mass spectrometry data for both the nucleus and mitochondrial matrix are shown in Supplementary Tables 6 and 7 respectively. Each of the two replicates for each proteomic experiment (mitochondrial matrix and nucleus) were analyzed separately. To select cutoffs for proteins biotinylated by the indicated ligase over non-specific bead binders, we classified the detected proteins into five groups:

1. Nuclear annotated proteins (Supplementary Table 3; true positive list of 6710 human proteins annotated with the following Gene Ontology49,50 terms: GO:0005783, GO:0005798, GO:0007029, GO:0038087, GO:0048237, GO:0061163, GO:0016320, GO:0030868, GO:006983, GO:0001319, GO:0051645, GO:0031985, GO:0005796, GO:0005795, GO:0005794, GO:0007030, GO:0032587, GO:0016021, GO:0050887, GO:00071816, GO:0031526, GO:0091593, GO:0072546, GO:1990440, GO:0039068, GO:0005913, GO:0072546, GO:1902236, GO:1994041, GO:0034976, GO:0057788, GO:0005790, GO:1902237, GO:0007009, GO:0005786, GO:0005793, GO:0044322, GO:0098554, GO:0005791, GO:0021060, GO:0043001, GO:0005802, GO:006989, GO:0005801, GO:0012510, GO:0008924, GO:0041247, GO:0034999, GO:0032588, GO:0003104, GO:0005689, GO:0005689, GO:0000042, GO:0032580, GO:0003173, GO:0008924, GO:0030198, GO:0031668, GO:0010715, GO:0035426, GO:1903055, GO:0001560, GO:0022167, GO:0006887, GO:00012505).
Ontology terms: GO:0015629, GO:0016235, GO:0030054, GO:0005813, positive list of 6815 human proteins annotated with the following Gene centers, nucleoplasm, or nucleus) and hyperLOPIT data (hyperLOPIT loci were assembled using Cell Atlas data (proteins detected by validated antibodies).

Figure 3d. Western blotting of Drosophila wing discs. For Figure 3d, wandering 3rd instar larvae were bisection and inverted to express the imaging discs. Inverted carasses were fixed for 20 min in 4% paraformaldehyde in 1× PBS. Fixed carasses with attached wing discs were permeabized with PBS + 0.1% Triton-X100 (PBST) for 20 min and blocked with PBST + 5% normal goat serum (PBST-NGS) for 1 hour at room temperature. Blots were washed extensively with PBST and exposed using Pico Chemiluminescent Substrate (Thermo Fisher 34577). To detect expressed V5-tagged ligases, blots were incubated with 1:10,000 mouse anti-V5 (Invitrogen R960-25) with PBST-BSA overnight at 4 °C, washed with PBST, incubated with 1:5000 anti-mouse Alexa 800 (Thermo Fisher A32730), washed with PBST, and imaged on an Aerius Fluorescent imager (LI-COR 9250).

Immunohistochemistry of Drosophila wing discs. For Figure 3d, wandering 3rd instar larvae were bisection and inverted to express the imaging discs. Inverted carasses were fixed for 20 min in 4% paraformaldehyde in 1× PBS. Fixed carasses with attached wing discs were permeabized with PBS + 0.1% Triton-X100 (PBST) for 20 min and blocked with PBST + 5% normal goat serum (PBST-NGS) for 1 hour. Blocked carasses were incubated overnight at 4 °C in PBST-NGS with 1:500 mouse anti-V5 (Invitrogen R960-25) and with PBST-BSA overnight at 4 °C, washed with PBST, incubated with 1:5000 anti-mouse Alexa 800 (Thermo Fisher A32730), washed with PBST, and imaged on an Aerius Fluorescent imager (LI-COR 9250).

Quantitation of fluorescence signal intensity from Drosophila wing discs inFigure 3e. Average signal intensity of fluorescence of streptavidin-555 in wing discs was measured using raw images obtained under identical confocal settings and under non-saturating exposure settings. Using ImageJ software, the polygon tool was used to select a rectangular region of the ptc-Gal4 expressing wild type or transgenic strains of Drosophila melanogaster. The age and sex of animals involved in experiments are indicated in figure legends and methods below. The Harvard Medical School Standing Committee on Animals (through the Office of the Institutional Animal Care and Use Committee (IACUC)) deems flies as invertebrates with limited sentence and therefore not subject to formal review and approval by the committee.
domain in the wing pouch. The average signal intensity in this selected region was determined separately for the streptavidin-555 channel and the anti-V5 channel. The average signal intensity in control samples (very low background staining) was subtracted from signal intensity of experimental conditions (BioID, turboID, miniturboID). For each wing disc, the signal intensity of streptavidin-555 was normalized to the signal intensity of anti-V5 (streptavidin-555/anti-V5). Fold change was determined by normalizing streptavidin-555/anti-V5 values from TurboID and miniturboID to values from BioID. Measurements were taken from at least three wing discs for each condition.

Quantification of adult Drosophila wing size and survival after ligase expression during development in Supplementary Figure 13. UAS-V5-ligase transgenes were expressed during development by crossing with different Gal4-expressing lines and their effects on the adult assessed.

To determine if larval wing disc expression of ligases affects adult wing morphology, tub::Gal4 was crossed with UAS-V5-ligase transgenes and the resulting progeny analyzed. tub::Gal4 was crossed with wild-type flies (w1118) as a negative control. Adult flies were aged 3 days after eclosion from pupal cases. Wings were removed from adults, placed in a drop of 50% Permoun/50% Xylenes on a glass slide, and a coverslip added. Mounted wings were imaged using a light microscope with a 10x objective. Wing area was measured using the polygon selection tool in ImageJ. Wings quantified and imaged are from female flies.

To determine if developmental expression of ligases reduces survival to adulthood, we crossed UAS-ligase lines with different Gal4 lines that express in major tissue types (Muscle, Fat, Neurons, Glia, Oocytes, Hemocytes) or ubiquitously (Act5c::Gal4). To quantify toxicity, we counted the number of surviving adult animals after undergoing ~10 days of development (from fertilized egg through pupal stages) expressing UAS-ligase under Gal4 control, and compared to the number of wild-type siblings. UAS-Luciferase was used as a negative control transgene, which is widely considered as non-toxic to cells.

As an example, the following crossing scheme was used for Act5c::Gal4:
P0 Act5c::Gal4/Cyo x UAS-V5-ligase (homozygous)

Segregation of the Act5c::Gal4 chromosome and CyO balancer chromosome results in two possible F1 progeny genotypes:

F1 (genotype 1) Act5c::Gal4/UAS-V5-ligase
F1 (genotype 2) CyO/UAS-V5-ligase

The CyO chromosome has a dominant Cy mutation that causes adult flies to have curly wings. Therefore genotype 1 flies have straight wings and express the ligase transgene, and genotype 2 have curly wings and do not express the transgene. The fraction of surviving flies expressing a given UAS-transgene is calculated as: # genotype 1/(# genotype 1 + # genotype 2) For example, a survival fraction of 0.5 indicates that equal numbers of genotype 1 and genotype 2 were observed in the adult population, and that no reduction in survival from expressing a UAS-transgene during development occurred.

Similar crossing schemes were used for tissue-specific Gal4 lines. Gal4 lines that are normally maintained as a homozygous stock were first outcrossed to an appropriate balancer line to obtain Gal4/Balancer flies, which were then crossed with UAS-Luciferase or UAS-TurboID. Gal4 lines on chromosome II were used with a CyO balancer, and Gal4 lines on chromosome III were used with a TM3, Sh balancer.

Adult flies were aged >3 days after eclosion from pupal cases before being counted. Females and males of the same genotype were counted together.

For imaging whole adults, flies were frozen at -20 °C overnight and images were imaged on LI-COR Odyssey CLx. For imaging whole adults, flies were frozen at -20 °C overnight and images were imaged on LI-COR Odyssey CLx.

C. elegans strains and culture conditions. Experiments on Caenorhabditis elegans were performed with wild type (N2) or transgenic strains carrying extrachromosomal arrays. The age and sex of animals involved in experiments are indicated in figure legends and methods below. The Stanford’s Administrative Panel on Laboratory Animal Care (APLAC) deems C. elegans used in this study as invertebrates and not subject to formal review and approval by the committee.

Unless otherwise noted, C. elegans strains were cultured and maintained at 20 °C on E. coli OP50 bacteria as previously described. To deplete the animals of excess biotin, worms were grown for 2 generations on biotin auxotrophic E. coli (MG1655bioBkan) and washed twice with 1× M9 solution. Biotin auxotrophic E. coli MG1655bioBkan was kindly donated by Dr. John E. Cronan, University of Illinois. Embryos dissected from one-day-old adults of the following genotypes were compared for this study: JLF289 (w00Ex66[ges1p::3×HA::BioID::unc-54, myo-2p::mCherry::unc-54]), JLF290 (w00Ex67[ges1p::3×HA::BioID::unc-54, myo-2p::mCherry::unc-54]), JLF291 (w00Ex68[ges1p::3×HA::TurboID::unc-54, myo-2p::mCherry::unc-54]), JLF292 (w00Ex69[ges1p::3×HA::TurboID::unc-54, myo-2p::mCherry::unc-54]), JLF293 (w00Ex70[ges1p::3×HA::miniTurbo::unc-54, myo-2p::mCherry::unc-54]), JLF294 (w00Ex71[ges1p::3×HA::miniTurbo::unc-54, myo-2p::mCherry::unc-54]).

Transgenic ligase strain construction for C. elegans. C. elegans codon-optimized ligase genes BioID and TurboID (containing the 3 worm introns present in GFP) and miniTurbo (containing 2 worm introns present in GFP) were synthesized (IDT) and inserted into pJF241 to produce plasmids pAS28, pAS31, and pAS32, respectively. Transgenic worms were generated by injecting 30ng/µl ligase gene and 2.5ng/µl of the co-injection marker myo-2p::mCherry into day 1 N2 hermaphrodites.

Western blotting of C. elegans adults. Ligase expression and biotinylation (Fig. 3i, Supplementary Fig. 15g) were assessed by Western blotting one day-old adult worm lysates. For each condition, 50 N2 hermaphrodites (wild-type) or worms expressing a ligase transgene were transferred to Eppendorf tubes containing 1mL of M9 and washed once. Excess M9 was removed until ~50 µL of M9 remained and an additional 50 µL of 4x sample buffer was added. Worms were boiled at 95 °C for 10 min, vortexed 10 seconds, and centrifuged at 13,000xg for 5 min at 4 °C. Equal volume of lysate was loaded onto a 4-20% Mini-PROTEAN TGX PAGE gel (Bio-Rad), transferred to a nitrocellulose membrane (0.4 µm, Bio-Rad), and stained with Ponceau S solution.

Blots were blocked in 5% milk PBST solution, probed with anti-HA (1:5000, rat monoclonal, Roche) and anti-tubulin (1:5000, rat monoclonal, Abcam) primary antibodies, and detected with secondary antibody (1:5000, goat anti-rat IRDye 680RD, Licor) and streptavidin-IRDye (1:5000, 800CW, Licor). Blots were imaged on LI-COR Odyssey CLx.

Immunohistochemistry and microscopy of C. elegans. To visualize ligases and biotinylation (Fig. 3j), embryos were isolated from one-day-old adults, fixed, and stained as previously described. Briefly, embryos were attached to poly-lysine coated microscope slides with Teflon spacers. Slides were frozen on dry ice and embryos were permeabilized by freeze-crack and fixed in 100% MeOH for 5 min at ~20 °C. Embryos were washed in PBS then PBST, and subsequently incubated in anti-HA primary antibody (Abcam, 1:200) overnight at 4 °C to visualize ligase expression. Embryos were washed in PBST and then incubated in CY3-anti-mouse secondary antibody (Jackson Immunoresearch Laboratories, 1:200) Streptavidin Alexa Fluor 488 (Invitrogen, 1:200), and
DAPI (Sigma, 1:10000). Embryos were mounted in Vectashield (Vector Laboratories) and stored at 4°C. Samples were imaged using 405 nm, 488 nm, and 561 nm lasers, a Yokogawa X1 confocal spinning disk head, and a 60x PLAN APO oil objective (NA=1.4) on a Nikon Ti-E inverted microscope (Nikon Instruments) equipped with a 1.5x magnifying lens. Images were captured using NIS Elements software (Nikon) and an Andor Ixon Ultra back thinned EM-CCD camera, at a sampling rate of 0.5μm. All samples were imaged with the same camera and laser settings, with the exception of embryos expressing miniTurbo to avoid pixel saturation, a 25% reduction in camera exposure was used to capture the streptavidin-AF488 signal in miniTurbo expressing embryos. Thus, the miniTurbo images and quantifications in Figure 3 for streptavidin-AF488 are an underrepresentation of the signal resulting from biotinylation. Images were processed and assembled in NIS Elements and Adobe InDesign. In Figure 3j, images shown are maximum intensity projections of two Z-slices with brightness adjusted for visual clarity.

Quantitation of fluorescence signal intensity in C. elegans intestine. Bean and comma stage embryos were chosen for analyses. For each embryo, one slice of the Z-stack was used for analysis. A custom Python script including the modules scikit-image, matplotlib, and NumPy combined with ImageJ was used to analyze C. elegans imaging data. A threshold for the HA:ligase signal was calculated by the Otsu method to create a mask to isolate the intestine region for each embryo. Background streptavidin-AF488 signal for each embryo was determined by drawing a square in the anterior portion of the embryo outside of the intestine and calculating the average pixel intensity within that square. The average background pixel intensity of streptavidin-AF488 was then subtracted from the average streptavidin-AF488 pixel intensity within the isolated intestine region, and the resulting corrected average was plotted for each embryo (Fig. 3k). To measure the ratio of streptavidin-AF488 to HA:ligase pixel intensities, each pixel value for streptavidin-AF488 within the isolated intestine region was corrected for background and then divided by its corresponding HA:ligase pixel value. Then the average of the ratio values for each embryo was plotted (Supplementary Fig. 15). Statistical significance was determined using the Mann-Whitney U test (Fig. 3k). Samples were blinded for statistical analysis.

Quantitation of C. elegans viability in Supplementary Fig. 16. C. elegans strains expressing ligase variants were maintained at 20°C on either biotin+ (OP50) or biotin- (MG1655(bioB:kan)) E. coli. For BioID and TurboID in each bacteria condition, a one-day-old adult worm expressing the ligase transgene from an extrachromosomal array and a sibling one-day-old adult worm lacking the transgene (control) were placed on separate plates containing the appropriate bacteria. For each plate, the adult was removed after laying eggs for 4 hours and the remaining embryos on the plate were counted. Three days later the number of living worms were counted and viability was calculated by dividing the number of hatched worms by the number of eggs that were initially laid. Worms were kept on biotin+ or biotin- bacteria for two generations. Developmentally delayed worms were defined as worms that were larval stage or non-gravid at the time of counting.

Statistics. Figure 3d,e, wing disc imaging results are representative of at least 10 discs present on the microscope slide, and at least 3 of which were imaged. Sample sizes (n) in e from left column to right are 5, 6, 3. Error bars were calculated using s.e.m. This experiment was performed twice with similar results.

Supplementary Figure 13c, sample sizes (n) from left column to right are 17, 14, 17, 15, 19, 18, 19, 18. Error bars were calculated using s.e.m. This experiment was performed twice with similar results. Supplementary Figure 13d, for each food type, a two-sided Chi-square test was used to determine if the difference in proportions (measure of effect) of UAS-ligase transgenes to UAS-Luciferase was statistically significant. Sample size values (n) from left column to right: 512, 586, 466, 563, 286, 524, 513, 459. 95% confidence intervals (CI) for the difference in proportions of BioID, TurboID, and miniTurboID compared to Luciferase were -0.02945 to 0.09046, 0.2757 to 0.384, and -0.02958 to 0.09206 for normal food, and -0.04278 to 0.104, -0.05999 to 0.08728, and -0.07191 to 0.07838 for biotin food. p-values for all datapoints - Columns 1/2: 0.3113, Columns 1/3: <0.0001, Columns 1/4: 0.3027, Columns 5/6: 0.4109, Columns 5/7: 0.718, Columns 5/8: 0.9369. This experiment was performed twice with similar results.

Supplementary Figure 13f, Data was analyzed as in Supplementary Figure 13d. Sample size values (n) from left column to right: 350, 367, 203, 274, 214, 232, 213, 205, 304, 246. 95% confidence intervals (CI) for each Gal4 line from left to right were -0.06874 to 0.08070, 0.04977 to 0.2228, -0.07546 to 0.1081, -0.03849 to 0.148, -0.07737 to 0.1115, 0.1398 to 0.3063, -0.07858 to 0.08951. p-values for all datapoints - Oenocyt: 0.8719, Glia: 0.0017, Muscle: 0.7295, Fat body: 0.2333, Neurons: 0.7143, Gut: <0.0001, Hemocytes: 0.8917. This experiment was performed twice with similar results.

Figure 3i, experiment was repeated 5 times with similar results, with the exception that miniTurbo expression was detectable in only 2 of those 5 replicates. Embryo imaging results shown in Figure 3j and Supplementary Figure 15a are representative images of complete quantitative data shown in Figure 3k. Figure 3k samples sizes (n) from left column to right column are 26, 18, 11, 16, 25, 8, 19, 23, 14, 14, 23, 9. Statistical significance was assessed via Mann-Whitney U test (two-sided). Error bars were calculated using s.e.m. Supplementary Figure 15c-f, statistical significance was assessed via Mann-Whitney U test (two-sided). Error bars were calculated using s.e.m. Supplementary Figure 16a-c, sample sizes (n) are indicated within an in column titled ’Replicates’. Statistical significance was assessed via Mann-Whitney U test (two-sided). Error bars were calculated using s.e.m.

For further detail on the experimental design, reagents, statistics, reproducibility, software, and data collection methods used in this study, please refer to the Life Sciences Reporting Summary.

Code availability. The Python script used for C. elegans imaging analysis is available from the corresponding author upon reasonable request.

Data availability. Source data for Figure 2e and Supplementary Figures 8c-g are provided in the paper in Supplementary Table 5. Source data for Figure 2h and Supplementary Figure 9f-k are provided in the paper in Supplementary Tables 6 and 7. The original mass spectra may be downloaded from MassIVE (http://massive.ucsd.edu) using the identifier: MSV000082304. The data is directly accessible via ftp://massive.ucsd.edu/MSV000082304. Any additional data that support the findings of this study are available from the corresponding author upon reasonable request.

38. Weaver, L.H., Kwon, K., Beckett, D. & Matthews, B.W. Co-repressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator. Proc. Natl. Acad. Sci. USA 98, 6045–6050 (2001).
39. Chao, G. et al. Isolating and engineering human antibodies using yeast surface display. Nat. Protoc. 1, 755–768 (2006).
40. Jan, C.H., Williams, C.C. & Weissman, J.S. LOCAL TRANSLATION. response to comment on "Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling". Science 348, 1217 (2015).
41. Colby, D.W. et al. Engineering affinity by yeast surface display. Methods Enzymol. 388, 348–358 (2004).
42. Aasulie, F.M. et al. Current protocols in molecular biology. Mol. Biol. 1, 13.2.1 (2003).
43. Wood, Z.A., Weaver, L.H., Brown, P.H., Beckett, D. & Matthews, B.W. Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution. J. Mol. Biol. 357, 509–523 (2006).
44. Xu, Y. & Beckett, D. Evidence for interdomain interaction in the Escherichia coli repressor of biotin biosynthesis from studies of an N-terminal domain deletion mutant. Biochemistry 35, 1783–1792 (1996).
45. Egiton, C., Cressman, W.J., Bachas, S., Wade, H. & Beckett, D. Allosteric coupling via distant disorder-to-order transitions. J. Mol. Biol. 427, 1695–1704 (2015).
46. Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. Nat. Protoc. 11, 456–475 (2016).

doi:10.1038/nbt.4201
47. Apweiler, R. et al. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).
48. Käll, L., Krogh, A. & Sonnhammer, E.L.L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).
49. Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res.* **43**, D1049–D1056 (2015).
50. Ashburner, M. et al.; The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
51. Thul, P.J. et al. A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017).
52. Muthusamy, B. et al. Plasma Proteome Database as a resource for proteomics research. *Proteomics* **5**, 3531–3536 (2005).
53. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
54. Calvo, S.E., Clauser, K.R. & Mootha, V.K. MitCoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res.* **44**, D1251–D1257 (2016).
55. Pagliarini, D.J. et al. A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123 (2008).
56. Sultan, M. et al. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956–960 (2008).
57. Perkins, L.A. et al. The transgenic RNAi project at Harvard medical school: resources and validation. *Genetics* **201**, 843–852 (2015).
58. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S.E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.* **40**, 476–483 (2008).
59. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
60. Ortega-Cueilar, D. et al. Biotin starvation with adequate glucose provision causes paradoxical changes in fuel metabolism gene expression similar in rat (*Rattus norvegicus*), nematode (*Caenorhabditis elegans*) and yeast (*Saccharomyces cerevisiae*). *J. Nutrigenet. Nutrigenomics* **3**, 18–30 (2010).
61. Leung, B., Hermann, G.J. & Priess, J.R. Organogenesis of the *Caenorhabditis elegans* intestine. *Dev. Biol.* **216**, 114–134 (1999).
Corresponding author(s): Alice Ting NBT-TR43271B

Reporting Summary

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

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

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

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection
- ImageStudioLite was used to process Western blots imaged with Licor. UVP BioSpectrum Imaging System was used to acquire Western blots imaged with Clarity Western ECL Blotting Substrates (BioRad). Slidebook 6.0 was used to collect mammalian cell imaging data. BD CSampler software was used to acquire the FACS data.

Data analysis
- Slidebook 6.0 was used to analyze mammalian cell imaging data. ImageJ 1.50i was used to quantify Western blot data, Drosophila imaging data, and to measure Drosophila wing area. Prism 7.0 was used to analyze numerical data. BD FACSDIVA v8.0 was used to analyze the FACS data. A custom Python script including the modules scikit-image, matplotlib, and NumPy combined with ImageJ was used to analyze C. elegans imaging data. The script can be made available upon request. R statistical programming language and ggplot2 were used to analyze and plot numerical data for C. elegans data.

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

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source data for Figure 2e and Supplementary Figures 8c-g are provided in the paper in Supplementary Table 25. Source data for Figure 2h and Supplementary Figures 9f-k are provided in the paper in Supplementary Tables 64 and 76. The original mass spectra may be downloaded from MassIVE (http://massive.ucsd.edu) using the identifier: MSV000082304. The data is directly accessible via ftp://massive.ucsd.edu/MSV000082304. Any additional data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size
All mammalian cell imaging results presented were representative of at least 10 independent fields of view. Sample size of mammalian cell image fields of view was chosen in consistency with many other publications (e.g. PMID: 28650461). Drosophila fly wing disc imaging results are representative of at least 10 discs present on the microscope slide, at least 3 of which were imaged. Full information on sample size in figure legend and methods. Sample size of imaged wing discs was chosen as the maximum number of samples processed at the same time without sacrificing consistency of tissue fixation, antibody staining, and slide mounting. The number of adult flies counted for survival assays was performed to maximize accuracy, at least 286 adult flies/sample were counted. Full information on sample size in figure legend and methods. Sample size of counted surviving adults was chosen as the maximum possible number of progeny flies collected from a single cross to control for variables such as food quality. The number of flies counted is consistent with published protocols on quantifying survival in Drosophila (PMID: 24751824). Measurements of wing size from adult flies were performed on at least 14 wings. Full information on sample size in figure legend and methods. Sample size of imaged wing discs was chosen as the maximum number of wings processed at the same time. The number of wings quantified is consistent with many other publications (e.g. PMID: 25107277). For all C. elegans experiments, full information on sample size and replicates are within figures, figure legends, and/ or methods. Images and quantifications shown in Figure 3 and Supplementary figures and qualifications are representative of results. For consistency, two specific stages of worm embryonic development were chosen for IF staining. Sample size of embryos imaged were chosen as the maximum number of samples processed at the same time.

Data exclusions
For consistency, a specific stage of worm development was chosen (embryos). Embryos within that stage were chosen for quantitative analysis.

Replication
The number of times an experimental finding was reproduced is indicated in the associated figure legend for each experiment shown. For Figure 3i, the experiment was performed five times (n = 5). In biotin+ conditions, BioID biotinylation activity was undetectable and TurboID gave robust biotinylation signal in all replicates (n = 5/5). Despite high activity detected by immunofluorescence in embryos, we only detected some low level of biotinylation by miniTurbo in adults (n = 2/5), likely due to its low expression levels. For all other experiments, attempts at replication were successful.

Randomization
No randomization methods were used because this was not applicable for our experiments. Randomization of adult flies or larvae was not used because freely-moving animals are too small to be assigned numbers for subsequent random selection. Fly wing discs and adult wings were selected for imaging randomly, and adult flies were selected for lysis and western blotting randomly. Randomization does not apply to counting survival of adult flies since all flies are counted. Worm embryos were selected for imaging randomly, and adult worms were selected randomly for lysis and subsequent Western blotting. Randomization does not apply to counting survival and developmental delay of worms since all worms in these experiments were counted.

Blinding
Investigator was blinded to group allocation during statistical analyses for C. elegans experiments.

Reporting for specific materials, systems and methods

Nature Biotechnology: doi:10.1038/nbt.4201
Materials & experimental systems

Involved in the study
- Unique biological materials
- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants

Methods

Involved in the study
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

Antibodies

| Antibodies used | vendor, catalog number/lot number: mouse anti-myc, Calbiochem, OP10; goat anti-mouse-HRP, BioRad, 170-6516; chicken anti-myc, Invitrogen, A-21281/1819879; goat anti-chicken-AlexaFluor647, Invitrogen, A-22443/1081817; rabbit anti-biotin, ImmuneChem, ICP0611; goat anti-rabbit-R-phycocerythrin, Life Technologies, P2771MP; mouse anti-V5, Invitrogen, 46-0705/1847319/1865511; mouse anti-calreticulin, Calbiochem, 208912/D00154827; mouse anti-Tom20, Santa Cruz Biotechnology, sc-17764/2117; mouse anti-HXX I, Santa Cruz Biotechnology, sc-46965/F1611; rabbit anti-NDUFS6, Abcam, ab195808; rabbit anti-calnexin, Santa Cruz Biotechnology, sc-11397/1914; rabbit anti-Tom20, Santa Cruz Biotechnology, sc-11415/12711; goat anti-rabbit-AlexaFluor568, Invitrogen, A-11011/1704642; mouse anti-His6, Calbiochem, B005; goat anti-mouse-AlexaFluor647, Invitrogen, A21236; goat anti-mouse-AlexaFluor800, Invitrogen, A32730/R1243148; mouse anti-HA, Abcam, ab130275/GR250145-5; goat anti-mouse-CY3, Jackson Immunoresearch Laboratories, 115-165-166/117091; rat anti-alpha tubulin, Abcam, ab6161; rat anti-HA, Roche, 11867432001; goat anti-rat IRDye 680RD, LI-COR, 92568076/C61115-06.

Validation mouse anti-myc (Calbiochem, OP10): Epitope Tagging - Munro, S., et al. 1986. Cell 46, 291. Immunoblotting - Evan, G.I., et al. 1985. Mol. Cell. Biol. 5, 3610; chicken anti-myc (Invitrogen, A-21281): several published flow cytometry applications cited on website - https://www.thermofisher.com/antibody/product/Myc-Tag-Abène-Polyclonal/A-21281; rabbit anti-biotin (ImmuneChem, ICP0611): Udeshi et al in Nature Methods Vol.14 pages: 1167-1170 (2017); mouse anti-V5 (Invitrogen, 46-0705): several published immunofluorescence and Western blotting applications cited on website - https://www.thermofisher.com/antibody/product/VS-Tag-Abène-Monoclonal/R960-25; mouse anti-calreticulin (Calbiochem, 208912): Nauseef, W.M., et al. 1995. J. Biol. Chem. 270, 4747. Jethmalani, S.M., et al. 1994. J. Biol. Chem. 269, 23603. Smith, M.J., and Koch, G.L.E. 1989. EMBO J. 8, 3581; mouse anti-BCAP31 (Proteintech, 11200-1-AP): validation shown on website - https://www.ptglab.com/products/BCAP31-Abène-11200-1-AP.htm#validation; rabbit anti-Tom20 (Proteintech, 14528-1-AP): validation shown on website - https://www.ptglab.com/products/TOM70-Abène-14528-1-AP.htm#validation; mouse anti-Tom20 (Santa Cruz Biotechnology, sc-17764): several published applications cited on website - https://www.scbt.com/scbt/product/tom20-antibody-f-10; mouse anti-HXX I (Santa Cruz Biotechnology, sc-46695): several published applications cited on website - https://www.scbt.com/scbt/product/hxx-i-antibody-g-17#search; rabbit anti-NDUFS6 (Abcam, ab195808): validation shown on website - https://www.abcam.com/ndufs6-antibody-epr15957-37-ab195808.html; rabbit anti-calnexin (Santa Cruz Biotechnology, sc-11397): several published applications cited on website - https://www.scbt.com/scbt/product/calnexin-antibody-h-707; requestFrom=search; rabbit anti-Tom20 (Santa Cruz Biotechnology, sc-11415): several published applications cited on website - https://www.scbt.com/scbt/product/tom20-antibody-f-145; mouse anti-His6 (Calbiochem, B005); Zentgraf, H., et al. 1995. Nucl. Acids. Res. 23, 3347. Gu, J., et al. 1994. BioTechniques 17, 257. Soreno, E., et al. 1994. J. Biol. Chem., 10991. Sisk, W.P., et al. 1994. J. Virol. 68, 766. Lu, T., et al. 1993. Anal. Biochem. 213, 318. Garner, J., et al. 1992. Cell 69, 833. Hochuli, E., et al. 1987. J. Chromatogr. 411, 177; mouse anti-HA (Abcam, ab130275): several published applications cited on website - https://www.abcam.com/ha-tag-antibody-16b12-ab130275-references.html; rat anti-alpha tubulin (Abcam, ab6161): several published applications cited on website - http://www.abcam.com/tubulin-antibody-yol134-microtubule-marker-ab6161-references.html; rat anti-HA (Roche, 11867432001): several published applications cited on website - https://www.sigmaaldrich.com/catalog/product/roche/roahahafan%en%en%region%US&gclid=CjwKCAjwma3Z8R8wEiwA-Csb1LDeA52Obna8s5aQ7MVoY/K2zRrFHkWtysOCa5cYFlH0W_ZrOrCPycQAvD_BwE.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s) The source of the HEK 293T cell line was ATCC.

Authentication The cell line (HEK293T) was not authenticated.

Mycoplasma contamination The cells were not tested for mycoplasma contamination.

Commonly misidentified lines (See ICLAC register) None. The only cell line used in this paper is HEK293T, which is not listed in the ICLAC database.

Animals and other organisms

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

Experiments on flies were performed with wild type or transgenic strains of Drosophila melanogaster. The age and sex of animals

Nature Biotechnology: doi:10.1038/nbt.4201
Laboratory animals involved in experiments are indicated in figure legends and methods below. Unless otherwise noted, fly stocks were obtained from the Bloomington Drosophila Stock Center and are listed with the corresponding stock number: ptc-Gal4 (2017), Act5c-Gal4/CyO (4414), nub-Gal4 (25754), w1118 (6326), tub-Gal80ts; tub-Gal4/TM6b (Perrimon Lab), UAS-Luciferase (35788), Desat-Gal4 (Oenocyte) (65405), repo-Gal4 (Glialia) (7415), Mef2-Gal4 (Muscle) (27390), Lpp-Gal4 (Fat body) (Perrimon Lab, see transgene in 67043 for information), elav-Gal4 (Neurons) (8760), Myo1a-Gal4 (Gut) (Perrimon Lab, see transgene in 67057 for information), Hml-Gal4 (Hemocytes) (30140). All Details on the flies used are provided in the methods section and figure legends.

Caenorhabditis elegans experiments involved embryos or, hermaphrodite adults. The age of animals involved in experiments are indicated in figure legends or methods. N2 wild-type strain was obtained from the Caenorhabditis Genetics Center (CGC).

Wild animals
None.

Field-collected samples
None.

Flow Cytometry

Plots
Confirm that:
☑️ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
☑️ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
☑️ All plots are contour plots with outliers or pseudocolor plots.
☑️ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation
All FACS experiments in this study used yeast. Sample preparation is described in detail in the following methods sections: Yeast cell culture, Generation of BirA libraries for yeast display, Yeast display selections, Directed evolution of TurboID and miniTurbo.

Instrument
Yeast display FACS sorting was performed on BD FACS Aria II cell sorter, and all FACS analysis was performed on BD Accuri flow cytometer.

Software
BD FACSDIVA software was used to perform FACS sorts and analyze all data. FACS analysis data (performed on BD Accuri flow cytometer) was collected using BD CSampler software.

Cell population abundance
Percentages of pre- and post-sort cell populations analyzed and sorted are described in detail in the following methods sections: Yeast display selections, Directed evolution of TurboID and miniTurbo.

Gating strategy
Descriptions of the gating strategies used are included in the methods sections "Yeast display selections" and "Directed evolution of TurboID and miniTurbo" and are exemplified in Supplementary Figure 2).

☑️ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.
Author Correction: Efficient proximity labeling in living cells and organisms with TurboID

Tess C. Branon, Justin A. Bosch, Ariana D. Sanchez, Namrata D. Udeshi, Tanya Svinkina, Steven A. Carr, Jessica L. Feldman, Norbert Perrimon and Alice Y. Ting

Correction to: Nature Biotechnology https://doi.org/10.1038/nbt.4201, published online 20 August 2018.

In the supplementary information for this article originally posted, there were copy-and-paste errors in Supplementary Tables 2–7. In Supplementary Table 2, binary values in tab 1 indicating the presence of a protein in each proteome were incorrect. In Supplementary Table 3, Uniprot accession IDs in tab 1 and binary values in tab 1 indicating the presence of a protein in each proteome were incorrect. In Supplementary Table 4, Uniprot accession IDs and gene names in tabs 1 and 2 were incorrect. In Supplementary Table 5, binary values in tab 4 indicating the presence of a protein in the miniTurbo proteome were incorrect; some gene names were missing in tab 6; and the following columns were incorrect in tab 6: Log2(127N/127C), Log2(128N/127C), Log2(129N/129C), Log2(128C/129C), Log2(130N/130C), Log2(131N/130C), log2(131C/129C). In Supplementary Table 6, protein names in tabs 1–3 and binary values indicating the presence of a protein in the miniTurbo proteome in tab 4 were incorrect. In Supplementary Table 7, gene names in tab 3 were missing. The errors have been corrected online.

Published online: 20 November 2019
https://doi.org/10.1038/s41587-019-0355-0

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019