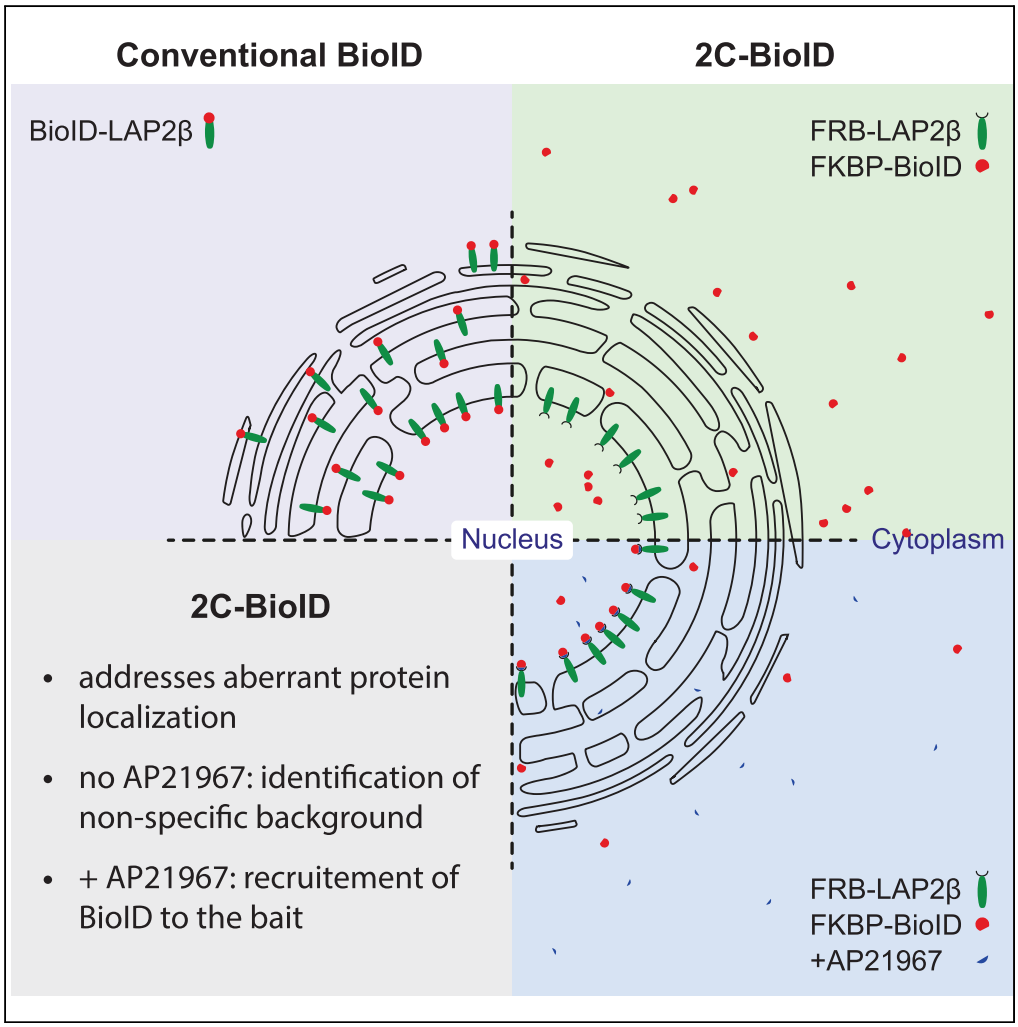
iScience

CellPress

Article

2C-BioID: An Advanced Two Component BioID System for Precision Mapping of Protein Interactomes



Alexandre Chojnowski, Radoslaw M. Sobota, Peh Fern Ong, ..., Oliver Dreesen, Brian Burke, Colin L. Stewart

alexandre.chojnowski@imb.a-star.edu.sg (A.C.)
brian.burke@imb.a-star.edu.sg (B.B.)

HIGHLIGHTS

Addresses aberrant protein localization that can be observed in conventional BioID

Recruitment of the biotin-protein ligase only after addition of the dimerizing agent

Facilitates the elimination of non-specific interactors and false-positives *in silico*

Chojnowski et al., iScience *10*, 40–52
December 21, 2018
https://doi.org/10.1016/j.isci.2018.11.023

iScience

CellPress

Article

2C-BioID: An Advanced Two Component BioID System for Precision Mapping of Protein Interactomes

Alexandre Chojnowski,[1,8,*] Radoslaw M. Sobota,[2,3,7] Peh Fern Ong,[4,7] Wei Xie,[5] Xianrong Wong,[1] Oliver Dreesen,[4] Brian Burke,[5,*] and Colin L. Stewart[1,6]

SUMMARY

The modulation of protein-protein interactions (PPIs) is an essential regulatory activity defining diverse cell functions in development and disease. BioID is an unbiased proximity-dependent biotinylation method making use of a biotin-protein ligase fused to a protein of interest and has become an important tool for mapping of PPIs within cellular contexts. We devised an advanced method, 2C-BioID, in which the biotin-protein ligase is kept separate from the protein of interest, until the two are induced to associate by the addition of a dimerizing agent. As proof of principle, we compared the interactomes of lamina-associated polypeptide 2β (LAP2β) with those of lamins A and C, using 2C- and conventional BioID. 2C-BioID greatly enhanced data robustness by facilitating the *in silico* elimination of non-specific interactors as well as overcoming the problems associated with aberrant protein localization. 2C-BioID therefore significantly strengthens the specificity and reliability of BioID-based interactome analysis, by the more stringent exclusion of false-positives and more efficient intracellular targeting.

INTRODUCTION

BioID is a proximity-based assay developed by Roux and colleagues to define and map potential protein-protein interactions (PPIs) (Roux et al., 2012). The method relies upon the construction of a bait protein (or protein of interest [PoI]) fused to a promiscuous biotin-protein ligase (pBPL). In the original and most widely used version of BioID, the pBPL is a mutant form (R118G) of the *E. coli* BirA protein (BioID1). BioID1 was engineered to non-specifically biotinylate neighboring proteins containing exposed primary amines. The assay depends upon the expression of the BioID1-PoI fusion protein in an appropriate cell type, with the cellular location of the fusion protein being determined by the PoI. Subsequent addition of biotin to the culture medium promotes biotinylation of proteins within a 10- to 20-nm radius of the BioID1 moiety (Kim et al., 2014). The biotinylated proteins are recovered on streptavidin-coated magnetic beads and identified by mass spectrometry, or in some situations, by western blot. After *in silico* sifting, the remaining proteins represent either direct or indirect candidate interactors of the PoI.

As a tool for PPI analysis, BioID provides a powerful complement to other established and widely used methods such as the yeast two-hybrid (Y2H) system, co-immunoprecipitation (IP), or pull-down approaches (Moosavi et al., 2017). The BioID method has several significant advantages. First, in contrast to Y2H, potential interactions are monitored in a more physiological cellular context employing, mostly, the full-length PoI. Moreover, BioID efficiently detects weak or transient interactions that may otherwise be missed. Importantly, BioID is unaffected by PoI or target protein solubility because the biotinylation step that flags potential interactors takes place before cell lysis. This allows for biotinylated proteins to be recovered under denaturing conditions, thereby reducing non-specific background interactions (Roux et al., 2012). On the other hand, because of their simplicity combined with the overall speed of implementation, pull-down or IP should represent methods of choice when first embarking on interactome analyses. The drawback associated with these approaches, however, is that they can be exquisitely sensitive to cell lysis conditions and rely on experimental designs that tread a fine line between the solubilization of protein complexes and retention of meaningful interactions. Although these issues can be overcome in part by employing cross-linking strategies, the use of chemical cross-linkers introduces an additional layer of complexity demanding additional controls, and with its own attendant artifacts. Cross-linking notwithstanding, the problem of solubility has, for instance, bedevilled investigations into the mammalian nuclear envelope (NE), as well as

[1]Developmental and Regenerative Biology, Institute of Medical Biology, Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, #06-06 Immunos, Singapore 138648, Singapore

[2]Functional Proteomics Laboratory, Institute of Molecular & Cell Biology, A*STAR, Singapore, Singapore

[3]Institute of Medical Biology, A*STAR, Singapore, Singapore

[4]Cell Ageing, Skin Research Institute, A*STAR, Singapore, Singapore

[5]Nuclear Dynamics and Architecture, Institute of Medical Biology, A*STAR, Singapore, Singapore

[6]School of Biological Sciences, Nanyang Technical University, Singapore, Singapore

[7]These authors contributed equally

[8]Lead Contact

*Correspondence: alexandre.chojnowski@imb.a-star.edu.sg (A.C.), brian.burke@imb.a-star.edu.sg (B.B.)

https://doi.org/10.1016/j.isci.2018.11.023

focal adhesions at cell surface-substrate contacts. The former contains a number of poorly soluble components such as the nuclear lamina, whereas the latter are highly labile. In both systems, BioID has provided valuable new insights into the interactomes of these cellular structures (Dong et al., 2016; Kim and Roux, 2016).

There are, nonetheless, some significant shortcomings in the standard BioID system. BioID cannot provide a "snapshot" of changing PoI interactions with the original pBPL repertoire, because biotinylation occurs during an incubation period of several hours. Instead, the technique provides a complete history of possible functional associations for the PoI, from its site of synthesis to its final cellular location. For very-short-duration time-resolved analyses, the use of recently developed high-activity pBPLs may complement available techniques such as APEX (Branon et al., 2018; Hung et al., 2014) and extend the range of tools at our disposal. It should also be clear that in addition to more or less stable PPIs, BioID also detects stochastic interactions, thus increasing the non-specific background over time. This necessitates the recognition and subsequent filtering of such background as an essential aspect of the BioID workflow.

As with all tagging methods, the construction and expression of a fusion protein to mimic the behavior and the physiological interactions of its endogenous counterpart, both qualitatively and quantitatively, has its own limitations. For instance, with a mass of 35.1 kDa (83 residues larger than GFP), BioID1 could potentially affect the localization and/or function of the PoI. As a particular example, the presence of BioID1 may restrict passage through nuclear pore complexes (NPCs) of integral inner nuclear membrane (INM) proteins, in this way interfering with their correct targeting. Worman and colleagues indeed demonstrated that the nucleoplasmic domain of INM proteins hinders their entry into the nucleus when extended beyond 60–70 kDa (Soullam and Worman, 1995). A case in point is the lamina-associated polypeptide 2β (LAP2β), a resident protein of the INM, for which the addition of BioID1 to its N-terminus would exceed the NPC exclusion limit. This size problem can be circumvented to some extent by the recent development of a smaller 26.6-kDa second-generation pBPL (BioID2) from *Aquifex aeolicus* (Kim et al., 2016). However, for many INM proteins, BioID2 will still compromise intracellular sorting.

To minimize issues associated with both the size of the BioID-PoI fusion protein (BioID1-PoI and to a lesser extent BioID2-PoI) and the ever-present challenge of non-specific background subtraction, we devised a modified BioID system. This new, two-component BioID (2C-BioID) takes advantage of the well-established FKBP:FRB (FK506-binding protein:FKBP-rapamycin-binding domain of mammalian target of rapamycin [mTOR]) oligomerization system (Putyrski and Schultz, 2012). In 2C-BioID, the PoI and pBPL are fused to FRB and FKBP. The two components, pBPL and PoI, are only brought together following FKBP:FRB oligomerization induced by the biologically inactive rapamycin analog AP21967. To assess the efficacy of 2C-BioID, we compared the interactomes of LAP2β generated with either conventional BioID or 2C-BioID. In addition, we employed both techniques to interrogate the interactomes of lamins A and C, major components of the nuclear lamina and the subjects of previous BioID studies. The two methods yielded overlapping but non-identical sets of potential interactors. Consistent with this, whereas the BioID1 moiety interfered with efficient targeting of LAP2β to the INM, no such impairment was observed when LAP2β was fused to the much smaller (~11.2kDa) FRB domain. Furthermore, the FKBP-BioID activity, in the absence of the AP21967 dimerizer, provides an internal baseline of the non-specific levels of biotinylation. Accordingly, 2C-BioID affords a more stringent appraisal of potential false-positives, thereby improving the overall robustness of the assay.

RESULTS

2C-BioID Results in the Appropriate Localization of LAP2β to the Inner Nuclear Membrane

Human LAP2β (~51kDa) is the largest of four resident INM proteins encoded by the *TMPO* gene. LAP2β is typical of tail-anchored transmembrane proteins, with the bulk of the molecule (~46 kDa) exposed on the nucleoplasmic side of the INM. Its amino terminal region harbors a LEM (LAP2, emerin, MAN1) domain, mediating its interaction with a small chromatin-associated protein, barrier-to-autointegration factor (BAF). The LEM domain is also found in the other transmembrane isoforms of LAP2 (ε,δ,γ), in its two soluble isoforms (α, ζ), and in several other INM components, including emerin, MAN1 (LEMD3), and LEM2. Although the role of LAP2β in normal cell physiology has yet to be fully elucidated, it has been implicated in transcriptional repression through its association with histone deacetylase 3 (Nili et al., 2001). Moreover, the interaction of LEM domain proteins, including LAP2β, with BAF, appears to have a role in NE reassembly following mitosis (Gant et al., 1999). Functional divergence among LAP2 isoforms has focused to a large

extent on LAP2α, which has roles in aspects of cell cycle progression via interactions with A-type lamins and pRb (Dorner et al., 2006; Naetar et al., 2008). LAP2α has also been implicated in telomere maintenance, a function that is impaired in the premature aging syndrome Hutchinson-Gilford progeria (Chojnowski et al., 2015; Dechat et al., 2004; Vidak et al., 2015). Comparing the interactomes of LAP2α and LAP2β would therefore be of particular interest to further define their functional specializations. However, this is less straightforward than might first appear.

Given the size of its nucleoplasmic domain, about 46 kDa, LAP2β presents specific challenges when using conventional BioID approaches. Because of size limitations, linking either BioID1 or BioID2 to the LAP2β N terminus may restrict access of the fusion protein to the INM. To explore this further, we fused the BioID1 cDNA to the 5′ end of the corresponding LAP2β-coding sequence and used a doxycycline (DOX)-inducible lentiviral system to express the BioID-LAP2β fusion protein in human fibroblasts (Chojnowski et al., 2015). Both western blot analysis and immunofluorescence microscopy revealed that BioID-LAP2β was duly expressed upon DOX addition as shown in Figures 1A and 1B. However, although some of the BioID-LAP2β fusion protein did localize to the NE, a significant fraction remained in the peripheral endoplasmic reticulum (ER) (Figures 1B, S1A, and S1B). This suggests that BioID-LAP2β is unable to efficiently localize to the INM, thus complicating the analysis of subsequent BioID results. In contrast, V5-tagged LAP2β localized exclusively to the NE, indicative of correct INM targeting (Figures 1C and S1B). These findings are consistent with a diffusion retention model of INM protein sorting, whereby increasing the size of the nucleoplasmic domain beyond 60–70 kDa restricts passage between the outer nuclear membrane (ONM) and INM at the level of the NPC membrane (Lusk et al., 2007; Soullam and Worman, 1995).

To skirt these targeting issues, and to address some of the limitations of the conventional BioID assay, we designed a two-component strategy, 2C-BioID, based upon the FKBP:FRB rapamycin-induced dimerization (Putyrski and Schultz, 2012). To accomplish this, we fused FKBP to BioID2 and FRB to the N terminus of LAP2β (Figure S1C). Using our lentiviral delivery system, we stably transduced human fibroblasts with the DOX-inducible V5-tagged FRB-LAP2β construct. DOX-dependent expression of the construct was verified by both western blot and immunofluorescence microscopy (Figures 1D and 1E, left hand panels; see also Figure S1E). Fusing FRB to the N-terminal nucleoplasmic domain of LAP2β had no adverse affect on the appropriate targeting of LAP2β to the NE, in striking contrast to the BioID-LAP2β fusion protein whose targeting was clearly compromised (Figure 1B). We subsequently transduced the FRB-LAP2β stable line with the BioID-FKBP construct. We employed a constitutive EF1α promoter to drive the expression of BioID-FKBP. As before, expression levels and localization were monitored by a combination of western blot and immunofluorescence microscopy (Figures 1D and 1E, right hand panels). The latter revealed that the BioID-FKBP fusion protein was distributed between both the cytoplasm and the nucleus (Figure 1E, top right panel; Figure S1D, left panel), a significant portion being clearly nuclear. There was, however, no evidence for any association of BioID-FKBP with the NE. Upon addition of the rapamycin analog AP21967, the BioID-FKBP fusion protein localized to the nuclear periphery and adopted a cellular distribution indistinguishable from that of FRB-LAP2β (Figure 1E, bottom right panel; Figure S1D, rightmost panel). Taken together, these results demonstrate that FRB-LAP2β localizes correctly to the INM, and that BioID-FKBP can then be tethered to it, in an AP21967-dependent manner (Figure 1F).

Comparison of Conventional BioID and 2C-BioID-LAP2β Interactomes

To gauge the potential advantages and disadvantages of 2C-BioID versus conventional BioID, we used each system to generate a LAP2β interactome. For both systems, we followed the basic protocol originally described by Roux et al (Roux et al., 2012). Additional sample pairs were, nevertheless, generated depending upon the presence or absence of the dimerizing agent AP21967. Both BioID protocols specifically identified proteins of the INM and nuclear lamina, in particular MAN1 (LEMD3), emerin (EMD), and lamin A/C (LMNA) (Figure 2). As expected for the 2C-BioID approach, these proteins were only detected upon induced dimerization of BioID-FKBP and FRB-LAP2β. The conventional BioID-LAP2β method identified the same targets after DOX induction of the fusion protein and addition of biotin. In this case, the same spectrum of biotinylated proteins was duly detected in both the presence and absence of the dimerization compound. The BirA ligases used in the conventional BioID (BioID1), in the 2C-BioID (BioID2), and the FRB domain of mTOR used in the 2C-BioID (identified by mass spectrometry as mTOR) were also detected (Figure 2). Significantly, other NE components such as LAP1 (TOR1AIP1) and lamin B receptor (LBR) were differentially identified by the two BioID methods. Conventional BioID identified both TOR1AIP1 and LBR as potential interactors of LAP2β, whereas the 2C-BioID method indicated that their detection was most likely

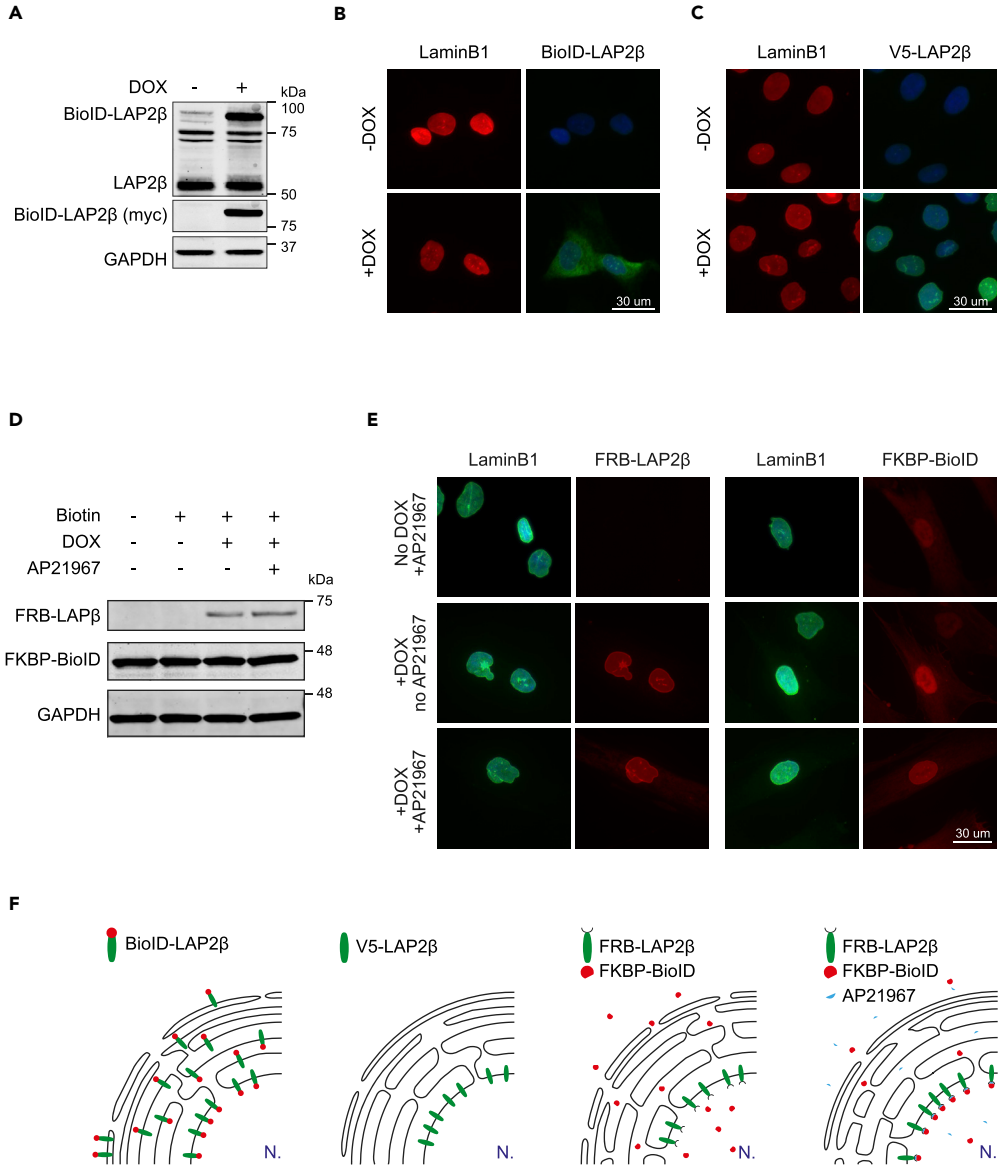


Figure 1. 2C-BioID Allows Proper Localization of FRB-LAP2β and FKBP-BioID to the Nuclear Lamina/Envelope

(A) Western blot showing doxycycline-dependent expression of myc-tagged BioID-LAP2β in primary human fibroblasts using LAP2 antibody (top panel) and myc-tagged antibody (middle panel) as indicated. LAP2, myc, and GAPDH are shown.

(B) Immunofluorescence microscopy using lamin B1 antibody (left panels, red) and myc-tagged antibody with DAPI counterstaining (right panels), showing doxycycline-dependent expression of myc-tagged BioID-LAP2β in human fibroblasts and subcellular localization at the nuclear periphery/ER. Scale bar, 30 μm.

(C) Immunofluorescence microscopy using lamin B1 antibody (left panels, red) and V5-tagged antibody with DAPI counterstaining (right panels) showing doxycycline-dependent expression and appropriate subcellular localization at the nuclear periphery of V5-tagged LAP2β in human fibroblasts. Scale bar, 30 μm.

(D) Western blot showing inducible expression of V5-tagged FRB-LAP2β and constitutive expression of myc-tagged FKBP-BioID in primary human fibroblasts, using V5 antibody (top panel) and myc-tagged antibody (middle panel). Cell culture condition (+/-biotin, +/− doxycycline, +/− AP21967) is indicated; V5-FRB-LAP2β, myc-FKBP-BioID, and GAPDH are shown.

(E) Immunofluorescence microscopy showing (left panels) lamin B1 (green) with DAPI counterstaining and V5-tagged antibody (red) and (right panels) lamin B1 (green) with DAPI counterstaining and myc-tagged antibody (red).

Figure 1. *Continued*
Relocalization of myc-tagged FKBP-BioID to V5-tagged FRB-LAP2β occurs only upon addition of AP21967 and doxycycline-dependent induction (bottom right panel). Scale bar, 30 μm.
(F) Illustration of BioID-LAP2β and 2C-BioID assays and respective localizations of the resulting fusions BioID/protein of interest/FKBP.
See also Figure S1.

non-specific, as revealed by their lower overall BioID score (Figure 2A). The BioID score used here represents a composite index based on a combination of prey absolute abundance, abundance ratio between samples and controls, and scores available from the CRAPome database (see Transparent Methods) (Mellacheruvu et al., 2013). The likelihood that both proteins were non-specifically detected is reflected by the fact that for both TOR1AIP1 and LBR, significant levels were detected in the absence of dimerizer. Several other proteins show similar differences between the two methods. The 2C-BioID score classifies them as low ranked, whereas in the absence of a priori functional input the conventional BioID assay suggests otherwise. Figures S2A and S2B shows a gene ontology (GO) terms enrichment analysis of proteins belonging to the two largest clusters of proteins identified in the LAP2β interactome using ClueGO (Bindea et al., 2009). These are mostly ribosome-related proteins for cluster 1 and actin-related proteins for cluster 2. As LAP2β is unlikely to have any extensive role related to either of these two clusters, this would indicate that most of the cluster constituents should be classified as background signal. Presumably many of these protein preys were biotinylated either owing to stochastic encounters with the bait or by transient proximity-based biotinylation occurring during the translation or trafficking of LAP2β to its final destination, i.e., the INM/lamina (Liu et al., 2018). It is also likely that some proteins could be endogenously biotinylated, or be non-specifically adsorbed to the streptavidin-coated beads in a biotin-independent fashion. Conventional BioID highlighted 18 such proteins as being of possible interest based solely on their score (BioID scores above 1), whereas only 3 would be classified as such using the 2C-BioID scoring, a decrease of ∼83% (PLS3, TES, and transforming growth factor β1-induced transcript 1 [TGFβ1I1]; see also Figures S3A and S3B). The complete and unfiltered list of identified proteins, quantification, and BioID/2C-BioID scores is shown in Table S1.

Among the highest ranked candidates in 2C-BioID, we then selected potential but currently unknown LAP2β interactors: TGFβ1I1, NHP2, SWAP70, STRN3, CRBN, AGFG1, and C1Orf198. We then used a yellow fluorescent protein (YFP)-based bimolecular fluorescence complementation assay to assess their association with LAP2β. In this system, PPI is revealed by reconstitution of a fluorescent YFP protein from two separate YFP segments fused to two separate putative partners (Bischof et al., 2018; Kerppola, 2008; Ohashi et al., 2012). As shown Figure S4A, the interaction between free FRB and FKBP was specifically detected after induction of their dimerization with AP21967: no YFP fluorescence could be observed without the addition of the dimerizer, despite both FRB and FKBP sharing similar intracellular localization. To assess the interaction of the selected candidates with LAP2β, each protein was fused at its N or C terminus to the N- or C-terminal segments of the split YFP and co-transfected with the corresponding LAP2β constructs (Figure 3A). Two of the candidates, TGFβ1I1/ARA55 (androgen receptor-associated protein of 55 kDa) and NHP2, were found to support LAP2β-dependent YFP fluorescence complementation and were thus chosen for further analysis (Figures 3C and 3D). Significantly, TGFβ1I1 and NHP2 were specifically identified as potential interactors by the 2C-BioID LAP2β and were not identified as such in the subsequent 2C-BioID lamin A or C experiments (Table S4).

To further assess their potential for association with LAP2β, we co-expressed either TGFβ1I1 or NHP2 in cells transduced with DOX-inducible V5-tagged FRB-LAP2β (Figures 3E, 3F, S5A, and S5B). Although largely cytoplasmic, TGFβ1I1 has a number of documented interactions with nuclear proteins, including the SMAD family members (Shola et al., 2012). In addition, it has a role as a co-activator of the androgen receptor and shuttles into the nucleus upon androgen stimulation (Leach et al., 2014). Structured illumination microscopy confirms that in cells transiently transfected with a TGFβ1I1 cDNA, the protein is restricted mainly to the cytoplasm (Figure 3E). However, a fraction of TGFβ1I1 is clearly recruited to the NE upon induction of LAP2β expression, which would be consistent with an association between TGFβ1I1 and LAP2β. The functional significance of this association, direct or indirect, and whether TGFβ1I1 might act as a regulator of some aspect of LAP2β function, or vice versa, will require further study. Nevertheless, the inference is that LAP2β and TGFβ signaling may be more entwined than previously suspected. Such a notion would be consonant with a previous report of cross-talk between TGFβ signaling pathways and LAP2α, the major soluble LAP2 isoform, in lamin A-deficient mice (Cohen et al., 2013).

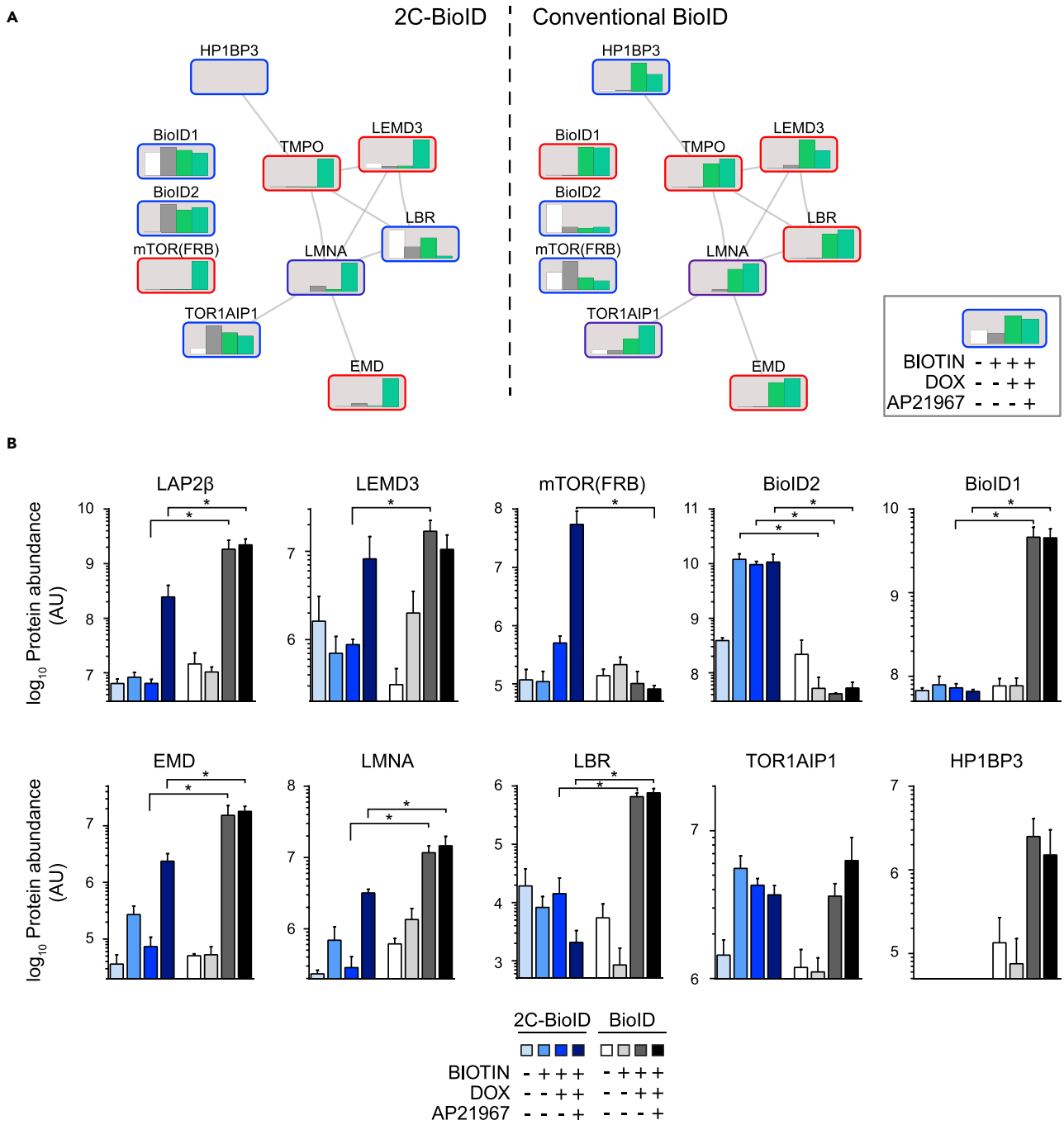


Figure 2. Comparison between 2C-BioID and BioID Analysis for the Nuclear Lamina and Selected Proteins

(A) Graphic representation of the network of selected proteins identified by 2C- and conventional BioID. The edge color of the nodes is mapped to the BioID score, from high score (red) to low score (blue, see Transparent Methods). The bar graphs of each node represent the relative quantity of its stated protein compared to its maximum value, across each condition (+/-biotin, +/− doxycycline, +/− AP21967, see inset on the right) for either 2C- or conventional BioID. (B) Quantification of protein abundance (a.u.) from mass spectrometric analysis of the BioID assay for LAP2β, LEMD3, mTOR (FRB domain), BioID2 (FKBP-BioID), BioID1 (BioID-LAP2β), EMD, LMNA, LBR, TOR1AIP1, and HP1BP3. Values are represented as mean ± SEM, *p < 0.05, two-way ANOVA with Sidak's post-hoc test.

See also Figures S2 and S3 and Table S1.

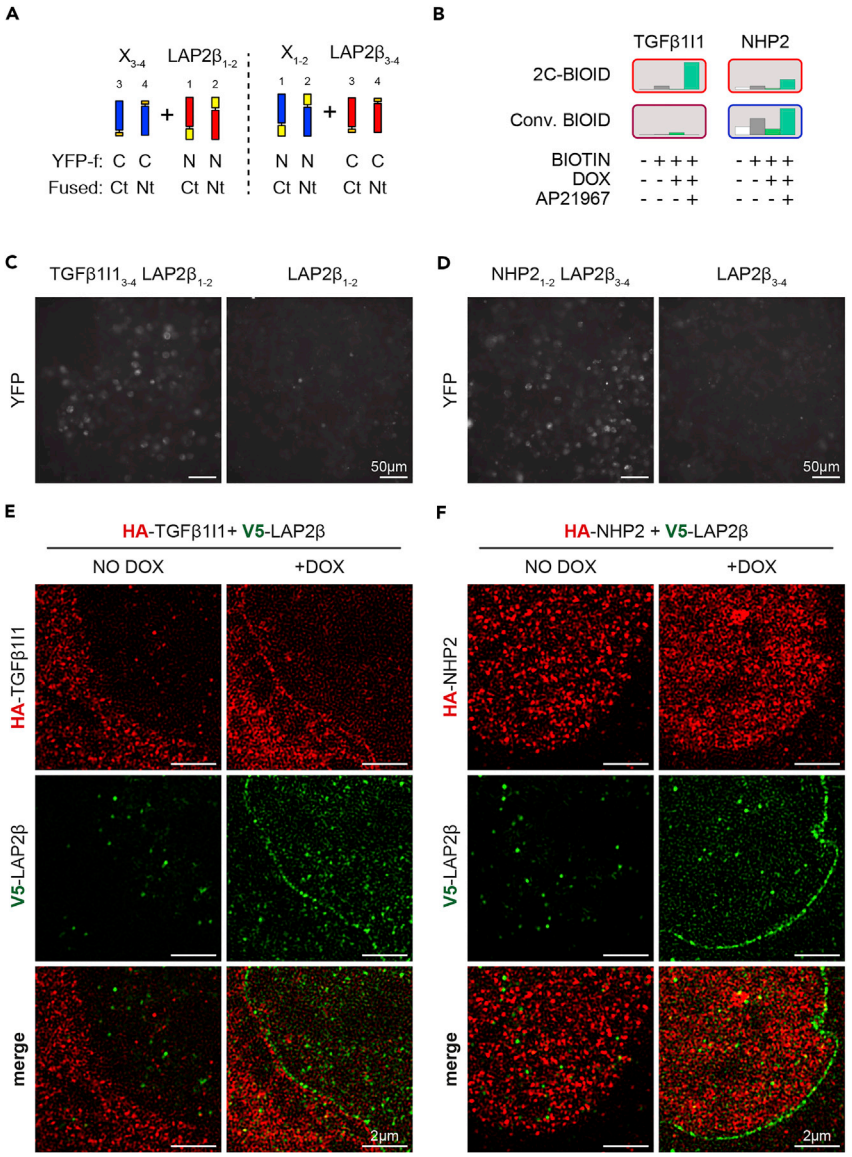


Figure 3. Interaction of LAP2β with TGFβ1I1 and NHP2

(A) Fusion constructs and transfection strategy for BiFC using N- or C-terminal fusions of the N- (yellow, numbered 1 and 2) and C-terminal (orange, numbered 3 and 4) fragments of split-YFP (YFP-f) to the bait (X, blue), or to LAP2β (red).

(B) Graphic representation of TGFβ1I1 and NHP2 as identified by 2C- and conventional BioID. The bar graphs of each node represent the min-max quantity of its stated protein identified across both 2C- and conventional BioID experiments for each condition (+/-biotin, +/− doxycycline, +/− AP21967).

(C and D) YFP fluorescence in live HEK293T transfected with selected combinations of YFP fusions for (C) TGFβ1I1 and (D) NHP2, and corresponding LAP2β-only constructs fused to split-YFP N- or C- terminal fragments without associated YFP-fused bait. Scale bar, 50 μm.

(E and F) 3D structured illumination microscopy immunofluorescence microscopy using hemagglutinin (HA)- (top panels, red) and V5-tagged (middle panels, green) antibodies showing HA-tagged TGFβ1I1 (E) and HA-tagged NHP2 (F) localization in human dermal fibroblasts expressing doxycycline-dependent V5-tagged FRB-LAP2β. Scale bar, 2 μm.

See also Figures S4 and S5.

In contrast to TGFβ1I1, NHP2 is restricted largely to the nuclear compartment (Figure 3F). It has been described as a component of both the telomerase complex (H/ACA ribonucleoprotein complex subunit 2) and small nucleolar ribonucleoprotein complexes (Leach et al., 2014; Vulliamy et al., 2008). The fact that we could find no convincing recruitment to the NE upon induction of LAP2β expression could simply

be a reflection of a weak or transient interaction. It is also possible that the association is indirect and that the expression of LAP2β alone may be insufficient to recruit NHP2 to the nuclear periphery. In any event, further experiments will be necessary to confirm the association, if any, between NHP2 and LAP2β.

Collectively, these results indicate that the 2C-BioID approach efficiently identifies potential bait interactors based on their spatial proximity, similar to the original BioID method. Furthermore, of the two novel candidate interactors for LAP2β, TGFβ1I1 and NHP2, only the latter could have been identified using the conventional BioID (Figure 3B). Finally, the data suggest a significant advantage of 2C-BioID. In the conventional BioID method the biotin ligase is always located with the PoI. This limits the possibility of detecting non-specific biotinylation without construction of a separate control cell line that expresses the pBPL alone. In contrast, in 2C-BioID, expression of the autonomous BioID-FKBP fusion protein, in the absence of a dimerizer, provides a simple internal control for non-specific biotinylation.

2C-BioID Reduces the Level of False-Positives in the BioID-LAP2β Interactome

The BioID score provides a quantitative ranking of proteins identified by BioID, by taking into account their abundance and CRAPome representation (Mellacheruvu et al., 2013). As shown in Figure S6A, scores from 2C-BioID and conventional BioID- LAP2β analyses share similar profiles across their respective protein ranks. Furthermore, in both datasets a majority of the top 10% of putative interactors was classified as "nuclear," if only transiently, by either GO or the Human Protein Atlas database (Uhlén et al., 2015). This corresponded to 34/54 and 27/54 nuclear proteins in the respective 2C- and conventional BioID datasets (Table S2). Considering the physiological localization of LAP2β at the INM, such an outcome is not unexpected. However, a comparison of the same top 10% of putative interactors for each BioID strategy revealed that only 10 of 54 proteins were common to both lists (Figure S6B, Table S2). Thus, although the two strategies share most of their respective bottom 90% hits, they differ significantly in their ranking of the putative top 10% hits. The average abundance ratio of the top 10% ranked proteins for 2C- or conventional BioID reveals a major difference in the outcome using the two methods (Figures 4A and 4B). As shown Figure 4A, the top 10% of proteins identified and scored using the 2C-BioID method (blue bars) shows the expected significant abundance ratio differences between control and sample conditions. The same proteins also show a significant difference in abundance between control and sample conditions using detection and scoring by the conventional BioID method (Figure 4A, gray bars). This indicates that the top hits identified by the 2C-BioID would not readily be discarded if the same experiment were to be performed using BioID. In contrast, although the top 10% of proteins identified and scored using conventional BioID display the expected abundance profile (Figure 4B, gray bars), the same proteins identified and scored using 2C-BioID do not show the expected difference between sample and control conditions (Figure 4B, blue bars). This suggests that a proportion of the putative interactors identified using conventional BioID may be false-positives. To confirm this, we submitted to GO terms enrichment analysis the list of proteins that were in the top 10% for one BioID method but were ranked low for the other BioID approach (ranked by their scores in the bottom 70%, Table S3). As shown in Figures S6C and S6D, a subset of proteins that ranked high in BioID but low in the 2C-BioID belonged to the vesicle organization/Golgi transport cluster. Conversely, five proteins ranked high in 2C-BioID were found clustered around cell adhesion/molecule-binding GO terms. This suggests that at least part of the difference between conventional and 2C-BioID within the top hits arises from the very different ways that the BirA ligase reaches its final cellular destination: either fused to the PoI from the start of its translation or separated from the PoI until the dimerization is induced in 2C-BioID. Obviously, the inefficient targeting of BioID-LAP2β to the INM also contributes to this discrepancy. Nevertheless, the two methods show similar profiles of abundance ratios for the lowest ranked proteins (Figures 4C and 4D).

To further explore the advantages of 2C-BioID, particularly in the elimination of false-positives, we extended this assay to two additional proteins, lamins A and C (Table S4). This enabled us to compare 2C-BioID analyses with our previously published conventional lamins A and C interactomes (Xie et al., 2016). To accomplish this, we functionally grouped the proteins identified in the lamin A, lamin C, conventional, and 2C-BioID assays. After filtering out the proteins non-ambiguously belonging to the background (conservatively only keeping BioID scores above 0.1, see Figure S7C), the results show that the proportion of proteins belonging to each of the identified functional groups common to all four experiments is extremely similar between 2C- and conventional BioID (Figure S7A). Together with our previous conventional and 2C-BioID LAP2β comparison, this confirms that 2C-BioID is, at minimum, performing similarly to the conventional BioID. This also demonstrates that the recruitment of FKBP-BioID to the bait results

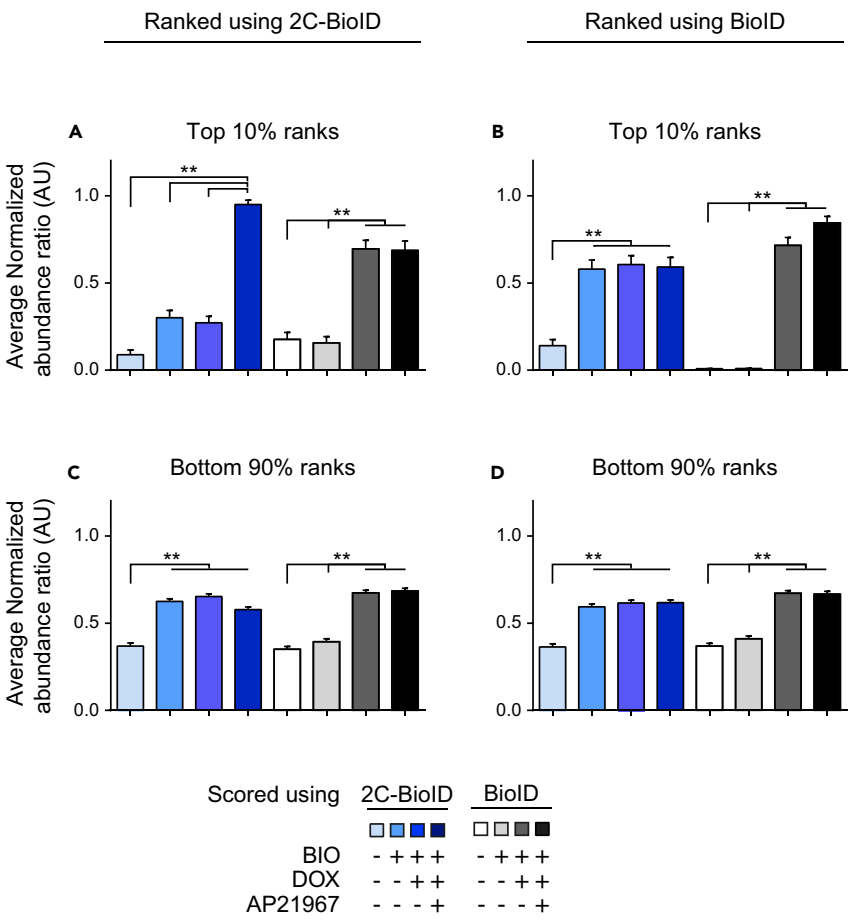


Figure 4. 2C-BioID Facilitates the Identification of False-Positives and Non-specific Interactors

(A) Average min-max abundance ratio of the top 10% ranks based on the 2C-BioID scores, from 2C-BioID quantifications (blue bars), or from conventional BioID quantifications (gray bars).

(B) Average min-max abundance ratio of the top 10% ranks based on the conventional BioID scores, from 2C-BioID quantifications (blue bars), or from conventional BioID quantifications (gray bars).

(C) Average min-max abundance ratio of the bottom 90% ranks based on the 2C-BioID scores, from 2C-BioID quantifications (blue bars), or from conventional BioID quantifications (gray bars).

(D) Average min-max abundance ratio of the bottom 90% ranks based on the conventional BioID scores, from 2C-BioID quantifications (blue bars), or from conventional BioID quantifications (gray bars). Cell culture conditions (+/-biotin, +/− doxycycline, +/− AP21967) are indicated below. Values are represented as mean ± SEM, **$p < 0.01$, one-way ANOVA with Tukey's post-hoc test.

See also Figures S6 and S7 and Tables S2, S3, and S4.

in a biotinylation profile as specific as that of the conventional BioID, in which the BioID is permanently fused to the PoI.

The distribution of the scores from the 2C- and conventional BioID is similar for both lamin A and lamin C BioID assays (Figure S7C), echoing our LAP2β results. Of note is the reduction of the BioID scores for the preys in the intermediate ranks, which suggests that 2C-BioID should be able to more efficiently discard potential preys compared with the conventional system. As shown in Figure S7B, functional inspection of the first 100 preys identified in both lamin A and lamin C assays indeed shows a reduction in groups related to nuclear import, including passage through NPCs, as well as vesicular trafficking. Taken together, both the lamin A and lamin C data, as well as the LAP2β results, lend support to the notion that 2C-BioID provides a simple strategy to screen out stochastic encounters of the PoI during its translation and shuttling to its final destination. Such encounters are not easily recognizable in the conventional BioID system, where the biotin-protein ligase is permanently fused to the bait PoI.

In summary, our results demonstrate that 2C-BioID is able to identify the majority of potential interactors that would have been detected using the original BioID method, but with the added benefit of providing an internal control that eliminates many false-positives.

DISCUSSION

The use of BioID as an entry point into interactome analysis has been steadily increasing since its first description by Roux and colleagues in 2012. The principal attributes of BioID are that it identifies potential PPIs within a physiologically relevant cellular environment while at the same time being insensitive to the solubility properties of both the bait and prey molecules. The method, nevertheless, has certain quirks. First, and the most important, it requires the addition of a sizable polypeptide, consisting of a pBPL, to the PoI. This additional polypeptide, of 26–35 kDa, could interfere with PoI localization or function. Second, because the pBPL is constitutively active, biotinylation of vicinal proteins may commence as soon as the pBPL-*PoI* fusion protein is produced by the cellular translation machinery. As described in the current study, for N-terminal fusion proteins, this may occur even before synthesis is complete and would account for the almost unvarying representation of ribosomal proteins in BioID experiments. The situation is further complicated by the fact that the PoI may, of necessity, associate with the cellular machinery involved in delivery to its final destination. For instance, lamins A and C must engage with soluble nuclear import receptors to traverse the NE via NPCs. Similarly, LAP2β and other INM proteins, which are initially inserted either post- or co-translationally into the ER or ONM, must transit the NPC membrane domain to reach their final destination. It should come as no surprise then, that assorted components of the nuclear import machinery should be detected in these experiments.

As the NPC imposes a size limit on the movement of membrane proteins between the ONM and INM (Lusk et al., 2007; Soullam and Worman, 1995), the most recent BioID studies on the NE and nuclear lamina interactome have made use of the smaller (26 kDa) second-generation pBPL (BioID2) (Birendra et al., 2017). Our 2C-BioID system, in which FRB (~10kDa) replaces the pBPL, was designed in part to further reduce interference with the bait protein's localization and physiological interactions. Only after the FRB-PoI fusion protein reaches its proper cellular localization does the FKBP counterpart of the oligomerization system couple the BioID2 moiety to the PoI. This provides additional flexibility in the use of BioID to map interactomes, because many proteins can be distributed between two or more cellular compartments. By controlling FKBP-BioID cellular localization, it would therefore be possible to map interactomes within specific cell compartments, using the very same FRB-PoI construct. As an example, to monitor the interactomes of nuclear versus cytoplasmic pools of a certain PoI, this could be performed in cell lines harboring FKBP-BioID bearing either a nuclear localization sequence or a nuclear export sequence. In addition, the FKBP:FRB oligomerization occurs in the presence of dimerizer (Putyrski and Schultz, 2012) over a timescale far shorter than the biotin labeling period, and therefore places on 2C-BioID few, if any, limits on time-constrained PPI studies.

When generating the interactome of LAP2β, the 2C-BioID approach clearly sidestepped problems associated with inefficient targeting of conventional BioID-LAP2β fusion proteins. The localization of FRB:LAP2β was equivalent to that observed with V5-LAP2β, the widely used V5-tag being only 14 amino acids long. In the particular case of LAP2β, the replacement of BioID1 by the smaller and newer variant BioID2 should improve the overall localization of the PoI, if only partially. For bulkier proteins the use of FRB in 2C-BioID would clearly represent the most favorable strategy. In the case where even minimal tagging of a PoI would interfere with its physiological activities, more global approaches such as protein correlation profiling or cellular thermal stability shift assay would be recommended, because both avoid protein fusions altogether (Dai et al., 2018; Larance et al., 2016).

Size notwithstanding, the principal advantage of 2C-BioID is its capacity to filter out background noise and associated false-positives, a major complication of the conventional system. Here we have shown that this can be achieved without compromising or skewing the detection of candidate interactors. Indeed, interactors that most likely belong to the background noise inherent to this type of experiment (the lowest ranked identifications) were quite similar between 2C- and conventional BioID. The main difference was seen in the highest ranked candidate interactors: in 2C-BioID such interactors were also highly ranked in the conventional BioID. Strikingly, the converse was not true: the highest ranked interactors identified by the conventional BioID were not consistently highly ranked in the 2C-BioID strategy. In this case, the 2C-BioID results indicated that their identification was most probably due to non-specific detection rather

than to actual interactions with the PoI. This is consistent with the fact that even in the lowest ranked potential interactors, conventional BioID still indicated a trend toward increased detection of false-positives when compared with 2C-BioID. This highlights the principal advantage of 2C-BioID: the constitutively active biotin ligase in the absence of dimerizer can be used as an effective internal control to identify non-specific interactors and background chatter. Furthermore, this is done in the same cellular context, in the presence of exactly the same PoI and pBPL fusion proteins. There is no requirement for the preparation of separate cell lines or controls, the only variable being the presence or absence of dimerizer.

For these reasons, the 2C strategy should be easily applicable to similar proximity labeling methods such as TurboID or APEX (Branon et al., 2018; Kim and Roux, 2016; Lam et al., 2015). In contrast to BioID and TurboID, APEX promotes biotinylation of vicinal proteins by means of an ascorbate peroxidase enzyme employing hydrogen peroxide and biotin phenol as substrates. APEX's extremely high activity makes it particularly suitable for time-resolved studies and provides an effective labeling method to analyze the entire proteome of membrane-limited compartments, mitochondria being a case in point (Rhee et al., 2013). TurboID and miniTurbo evolutions of the BioID enzymes were similarly developed with improved proximity labeling efficiency in mind. However, targeting or localization of bait proteins in TurboID or APEX assays may equally be restricted by the bulk of the TurboID (35 kDa), miniTurbo (28 kDa), or peroxidase (28kDa) enzymes. The former is akin to BioID1 in size, whereas the latter two more closely resemble BioID2. It goes without saying that our 2C-BioID system can easily accommodate the new-generation high-activity TurboID enzymes as well as the APEX ascorbate peroxidase tag. In each of these situations, this system should reduce problems associated with bait protein localization and functionality, while at the same time providing a built-in control for non-specific labeling. Taken as a whole, our 2C approach should significantly reduce the number of false-positives that bedevil these related proximity labeling techniques.

In another recent development, Schopp et al. (2017) introduced a split-BioID system in which the BirA enzyme itself is divided into a pair of inactive N- and C-terminal fragments. The biotin ligase activity is restored only when the two fragments are brought together. This system provides an exemplary method to validate binary associations by fusing N- and C-terminal BirA domains to known interacting partners and can identify additional interactors within multiprotein complexes. Clearly, this may not be feasible in all molecular contexts because it requires the initial constructive modification of two discrete interacting partners. Notwithstanding this proviso, it has been employed successfully to probe new associations within microRNA processing complexes. We would therefore argue that 2C-BioID in all of its guises and split--BioID provide highly complementary tools in the identification and validation of PPIs *in vivo*. 2C-BioID is ideally suited to the efficient detection of new protein pairs while eliminating spurious interactions, whereas split-BioID will highlight the mutual associations of known partners. In this way, 2C-BioID will underpin subsequent split-BioID analyses.

Taken together, our results show that 2C-BioID brings significant improvements to conventional BioID. In addition, generating stable lines or animal models constitutively expressing FKBP-BioID should be straightforward. This will facilitate subsequent 2C-BioID analyses that will have no more experimental complexity than conventional BioID. Together with other novel refinements designed for whole-proteome analysis or conversely for specific purposes, 2C-BioID should further extend and improve our current capabilities in refining both global and subcellular protein network analyses.

Limitations of the Study

BioID is a proximity labeling method that can be used to map PPIs. A significant advantage over conventional methods is that it can be applied to proteins in their normal physiological contexts. BioID is also insensitive to the solubility properties of both the PoI and its potential interactors. There are, however, limitations to the technique. First, attachment of a relatively large biotin-protein ligase to the PoI may interfere with its correct targeting. Second, BioID will detect all associations, stochastic and otherwise, that the PoI experiences during its entire lifetime, thereby generating significant background identifications. 2C-BioID was originally devised to circumvent the protein targeting problem, but its most critical feature is its ability to computationally reduce the non-specific background. This still leaves an important limitation of the 2C-BioID system: as it relies on proximity labeling, hence detecting near neighbors, the method cannot be used to formally demonstrate interactions. It is essential, therefore, that all candidate PPIs identified by 2C-BioID be validated using other independent approaches. To this end, we employed bimolecular

fluorescence complementation and super-resolution microscopy, but the alternative scheme could involve pull-downs from cell lysates. However, this may miss weak interactions as well as interactions involving poorly soluble proteins. Ultimately multi-faceted strategies may be required to confirm potential PPIs.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods, seven figures, and four tables can be found with this article online at https://doi.org/10.1016/j.isci.2018.11.023.

ACKNOWLEDGMENTS

This work was supported by the Singapore Biomedical Research Council and the Singapore Agency for Science, Technology and Research (A*STAR), the Progeria Research Foundation, IMCB, and IMB (SPF) A*STAR, YIG 2015 (BMRC, A*STAR), and NMRC MS-CETSA platform grant MOHIAFCAT2/004/2015 to R.M.S.

AUTHOR CONTRIBUTIONS

Study conception and design, A.C., B.B., and C.L.S.; Acquisition of data, A.C, R.M.S, P.F.O., W.X., and X.W.; Analysis and interpretation of data, A.C., R.M.S., P.F.O., W.X., X.W., O.D., B.B., and C.L.S.; Drafting of manuscript A.C., R.M.S, W.X, O.D., B.B., and C.L.S.

DECLARATION OF INTERESTS

The authors declare that they have no conflict of interest.

Received: June 19, 2018
Revised: September 26, 2018
Accepted: November 12, 2018
Published: December 21, 2018

REFERENCES

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–1093.

Birendra, K.C., May, D.G., Benson, B.V., Kim, D.I., Shivega, W.G., Ali, M.H., Faustino, R.S., Campos, A.R., and Roux, K.J. (2017). VRK2A is an A-type lamin–dependent nuclear envelope kinase that phosphorylates BAF. Mol. Biol. Cell *28*, 2241–2250.

Bischof, J., Duffraisse, M., Furger, E., Ajuria, L., Giraud, G., Vanderperre, S., Paul, R., Björklund, M., Ahr, D., Ahmed, A.W., et al. (2018). Generation of a versatile BiFC ORFeome library for analyzing protein-protein interactions in live Drosophila. Elife *7*, 343483.

Branon, T.C., Bosch, J.A., Sanchez, A.D., Udeshi, N.D., Svinkina, T., Carr, S.A., Feldman, J.L., Perrimon, N., and Ting, A.Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. Nat. Biotechnol. *36*, 880–887.

Chojnowski, A., Ong, P.F., Wong, E.S.M., Lim, J.S.Y., Mutalif, R.A., Navasankari, R., Dutta, B., Yang, H., Liow, Y.Y., Sze, S.K., et al. (2015). Progerin reduces LAP2α-telomere association in Hutchinson-Gilford progeria. Elife *4*, 1–21.

Cohen, T.V., Gnocchi, V.F., Cohen, J.E., Aditi, P., Liu, H., Ellis, J.A., Foisner, R., Stewart, C.L., Zammit, P.S., and Partridge, T.A. (2013). Defective skeletal muscle growth in lamin A/C-deficient mice is rescued by loss of lap2α. Hum. Mol. Genet. *22*, 2852–2869.

Dai, L., Zhao, T., Bisteau, X., Sun, W., Prabhu, N., Lim, Y.T., Sobota, R.M., Kaldis, P., and Nordlund, P. (2018). Modulation of protein-interaction states through the cell cycle. Cell *173*, 1481–1494.e13.

Dechat, T., Gajewski, A., Korbei, B., Gerlich, D., Daigle, N., Haraguchi, T., Furukawa, K., Ellenberg, J., and Foisner, R. (2004). LAP2alpha and BAF transiently localize to telomeres and specific regions on chromatin during nuclear assembly. J. Cell Sci. *117*, 6117–6128.

Dong, J.M., Tay, F.P., Swa, H.L., Gunaratne, J., Leung, T., Burke, B., and Manser, E. (2016). Proximity biotinylation provides insight into the molecular composition of focal adhesions at the nanometer scale. Sci.Signal. *9*, rs4.

Dorner, D., Vlcek, S., Foeger, N., Gajewski, A., Makolm, C., Gotzmann, J., Hutchison, C.J., and Foisner, R. (2006). Lamina-associated polypeptide 2α regulates cell cycle progression and differentiation via the retinoblastoma–E2F pathway. J. Cell Biol. *173*, 83–93.

Gant, T.M., Harris, C.A., and Wilson, K.L. (1999). Roles of LAP2 proteins in nuclear assembly and DNA replication: truncated LAP2beta proteins alter lamina assembly, envelope formation, nuclear size, and DNA replication efficiency in *Xenopus laevis* extracts. J. Cell Biol. *144*, 1083–1096.

Hung, V., Zou, P., Rhee, H.W., Udeshi, N.D., Cracan, V., Svinkina, T., Carr, S.A., Mootha, V.K., and Ting, A.Y. (2014). Proteomic mapping of the human mitochondrial intermembrane space in live cells via ratiometric APEX tagging. Mol. Cell *55*, 332–341.

Kerppola, T.K. (2008). Bimolecular fluorescence complementation (bifc) analysis as a probe of protein interactions in living cells. Annu. Rev. Biophys. *37*, 465–487.

Kim, D.I., Birendra, K.C., Zhu, W., Motamedchaboki, K., Doye, V., and Roux, K.J. (2014). Probing nuclear pore complex architecture with proximity-dependent biotinylation. Proc. Natl. Acad. Sci. U S A *111*, E2453–E2461.

Kim, D.I., Jensen, S.C., Noble, K.A., Kc, B., Roux, K.H., Motamedchaboki, K., and Roux, K.J. (2016).

An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell *27*, 1188–1196.

Kim, D.I., and Roux, K.J. (2016). Filling the void: proximity-based labeling of proteins in living cells. Trends Cell Biol. *26*, 804–817.

Lam, S.S., Martell, J.D., Kamer, K.J., Deerinck, T.J., Ellisman, M.H., Mootha, V.K., and Ting, A.Y. (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. Nat. Methods *12*, 51–54.

Larance, M., Kirkwood, K.J., Tinti, M., Brenes Murillo, A., Ferguson, M.A.J., and Lamond, A.I. (2016). Global membrane protein interactome analysis using *in vivo* crosslinking and mass spectrometry-based protein correlation profiling. Mol. Cell. Proteomics *15*, 2476–2490.

Leach, D.A., Need, E.F., Trotta, A.P., Grubisha, M.J., DeFranco, D.B., and Buchanan, G. (2014). Hic-5 influences genomic and non-genomic actions of the androgen receptor in prostate myofibroblasts. Mol. Cell. Endocrinol. *384*, 185–199.

Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R.G., Öhman, T., and Varjosalo, M. (2018). An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. Nat. Commun. *9*, 1188.

Lusk, C.P., Blobel, G., and King, M.C. (2007). Highway to the inner nuclear membrane: rules for the road. Nat. Rev. Mol. Cell Biol. *8*, 414–420.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.-P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat. Methods *10*, 730–736.

Moosavi, B., Mousavi, B., Yang, W., and Yang, G. (2017). Yeast-based assays for detecting protein-protein/drug interactions and their inhibitors. Eur. J. Cell Biol. *96*, 529–541.

Naetar, N., Korbei, B., Kozlov, S., Kerenyi, M.A., Dorner, D., Kral, R., Gotic, I., Fuchs, P., Cohen, T.V., Bittner, R., et al. (2008). Loss of nucleoplasmic LAP2alpha-lamin A complexes causes erythroid and epidermal progenitor hyperproliferation. Nat. Cell Biol. *10*, 1341–1348.

Nili, E., Cojocaru, G.S., Kalma, Y., Ginsberg, D., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Berger, R., Shaklai, S., Amariglio, N., et al. (2001). Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). J. Cell Sci. *114*, 3297–3307.

Ohashi, K., Kiuchi, T., Shoji, K., Sampei, K., and Mizuno, K. (2012). Visualization of cofilin-actin and Ras-Raf interactions by bimolecular fluorescence complementation assays using a new pair of split Venus fragments. Biotechniques *52*, 45–50.

Putyrski, M., and Schultz, C. (2012). Protein translocation as a tool: the current rapamycin story. FEBS Lett. *586*, 2097–2105.

Rhee, H., Zou, P., Udeshi, N.D., Martell, J.D., and Mootha, V.K. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science *339*, 1328–1331.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. *196*, 801–810.

Schopp, I.M., Claudia, C., Ramirez, A., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., and Be, J. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. Nat. Commun. *8*, 15690.

Shola, D.T.N., Wang, H., Wahdan-Alaswad, R., and Danielpour, D. (2012). Hic-5 controls BMP4 responses in prostate cancer cells through interacting with Smads 1, 5 and 8. Oncogene *31*, 2480–2490.

Soullam, B., and Worman, H.J. (1995). Signals andstructural features involved in integral membrane proteintargetting to the inner nuclear envelope. J. Cell Biol. *130*, 15–27.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

Vidak, S., Kubben, N., Dechat, T., and Foisner, R. (2015). Proliferation of progeria cells is enhanced by lamina-associated polypeptide 2α (LAP2α) through expression of extracellular matrix proteins. Genes Dev. *29*, 2022–2036.

Vulliamy, T., Beswick, R., Kirwan, M., Marrone, A., Digweed, M., Walne, A., and Dokal, I. (2008). Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. Proc. Natl. Acad. Sci. U S A *105*, 8073–8078.

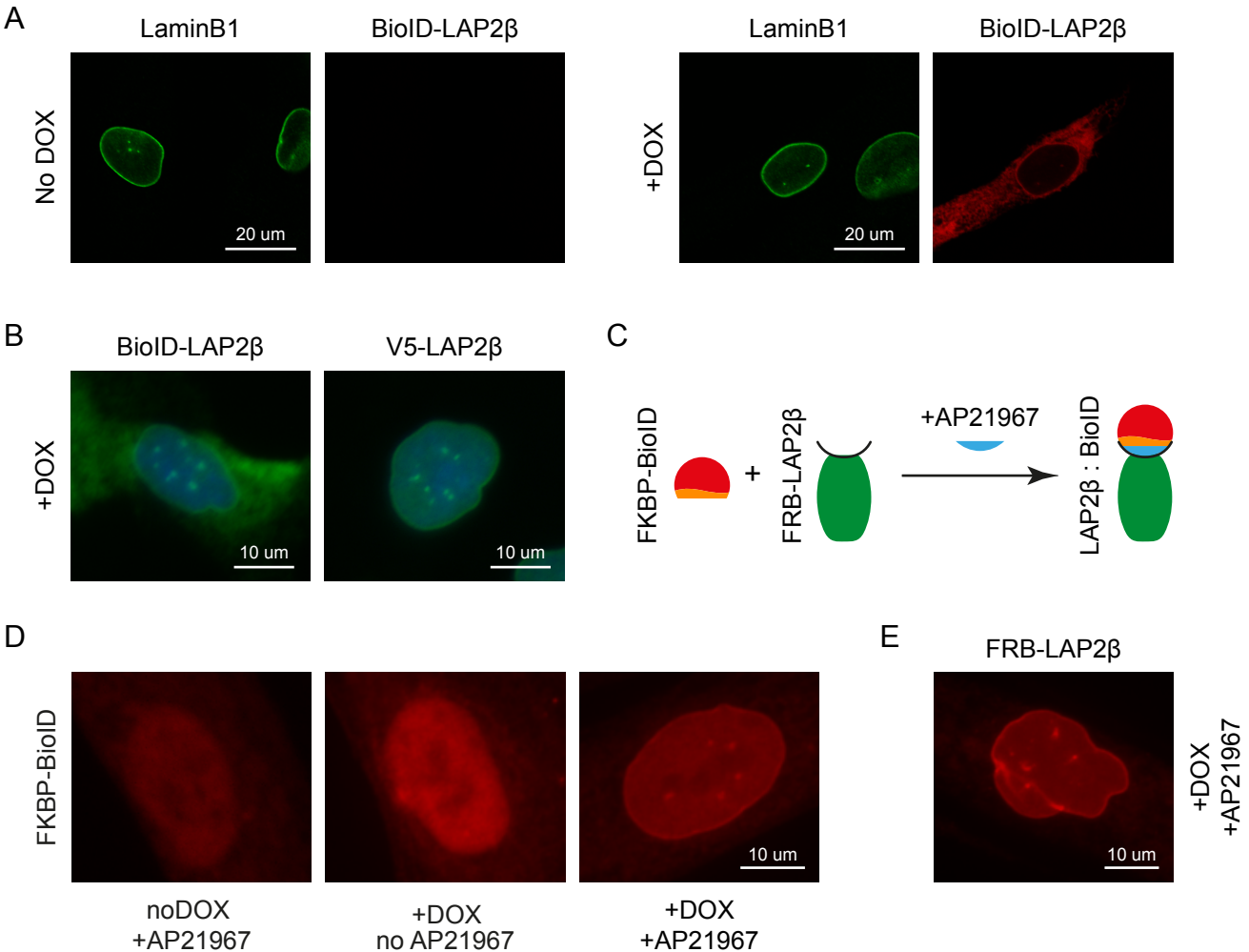
Xie, W., Chojnowski, A., Boudier, T., Lim, J.S.Y., Ahmed, S., Ser, Z., Stewart, C., and Burke, B. (2016). A-type lamins form distinct filamentous networks with differential nuclear pore complex associations. Curr. Biol. *26*, 2651–2658.

ISCI, Volume *10*

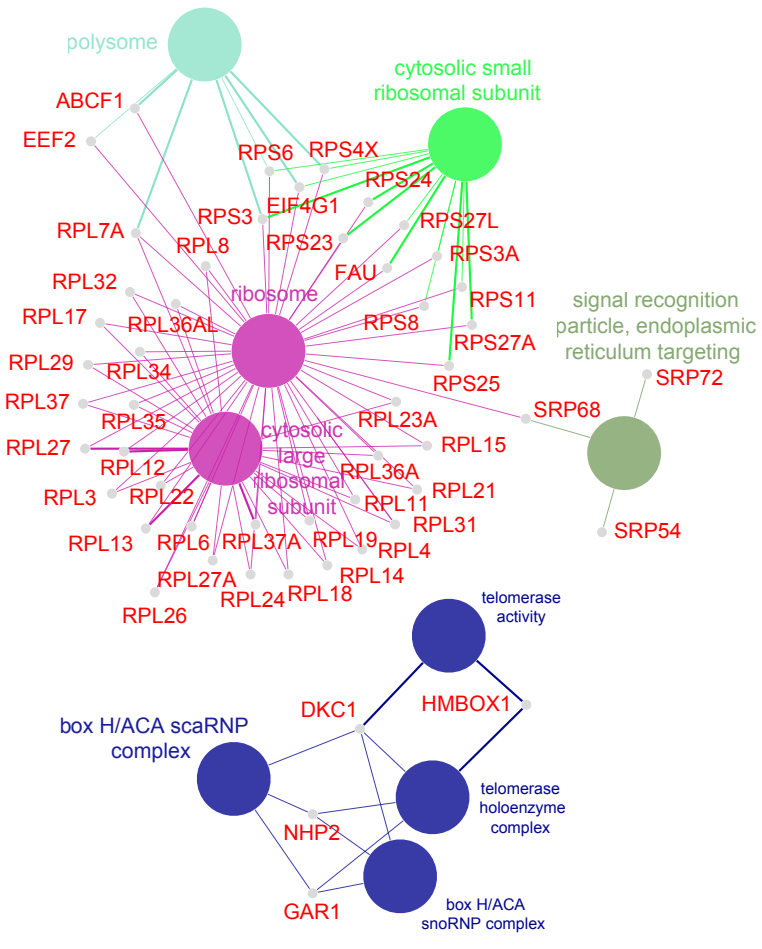
Supplemental Information

2C-BioID: An Advanced Two Component BioID System for Precision Mapping of Protein Interactomes

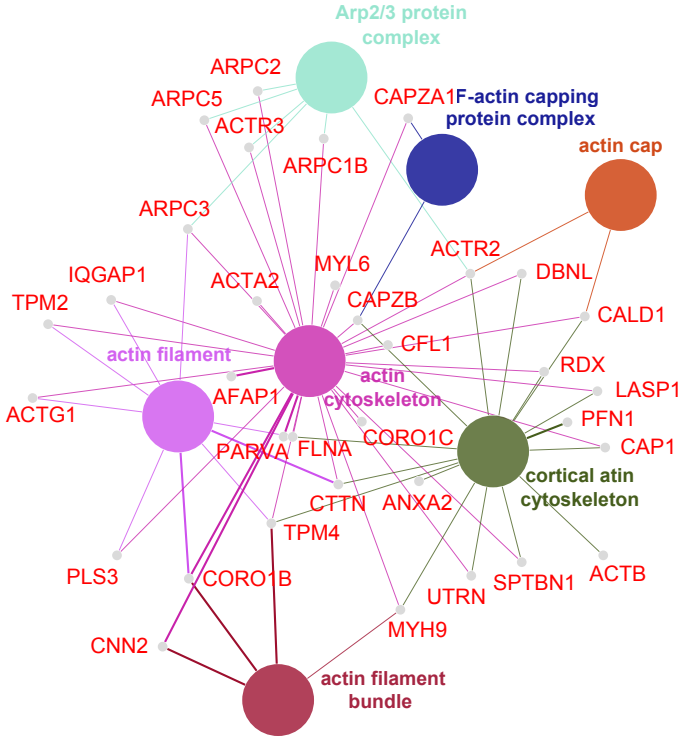
Alexandre Chojnowski, Radoslaw M. Sobota, Peh Fern Ong, Wei Xie, Xianrong Wong, Oliver Dreesen, Brian Burke, and Colin L. Stewart


A
LaminB1
BioID-LAP2β
LaminB1
BioID-LAP2β
No DOX
+DOX
20 um
20 um
B
BioID-LAP2β
V5-LAP2β
+DOX
10 um
10 um
C
FKBP-BioID
+
FRB-LAP2β
+AP21967
LAP2β : BioID
D
FKBP-BioID
10 um
noDOX
+AP21967
+DOX
no AP21967
+DOX
+AP21967
E
FRB-LAP2β
10 um
+DOX
+AP21967

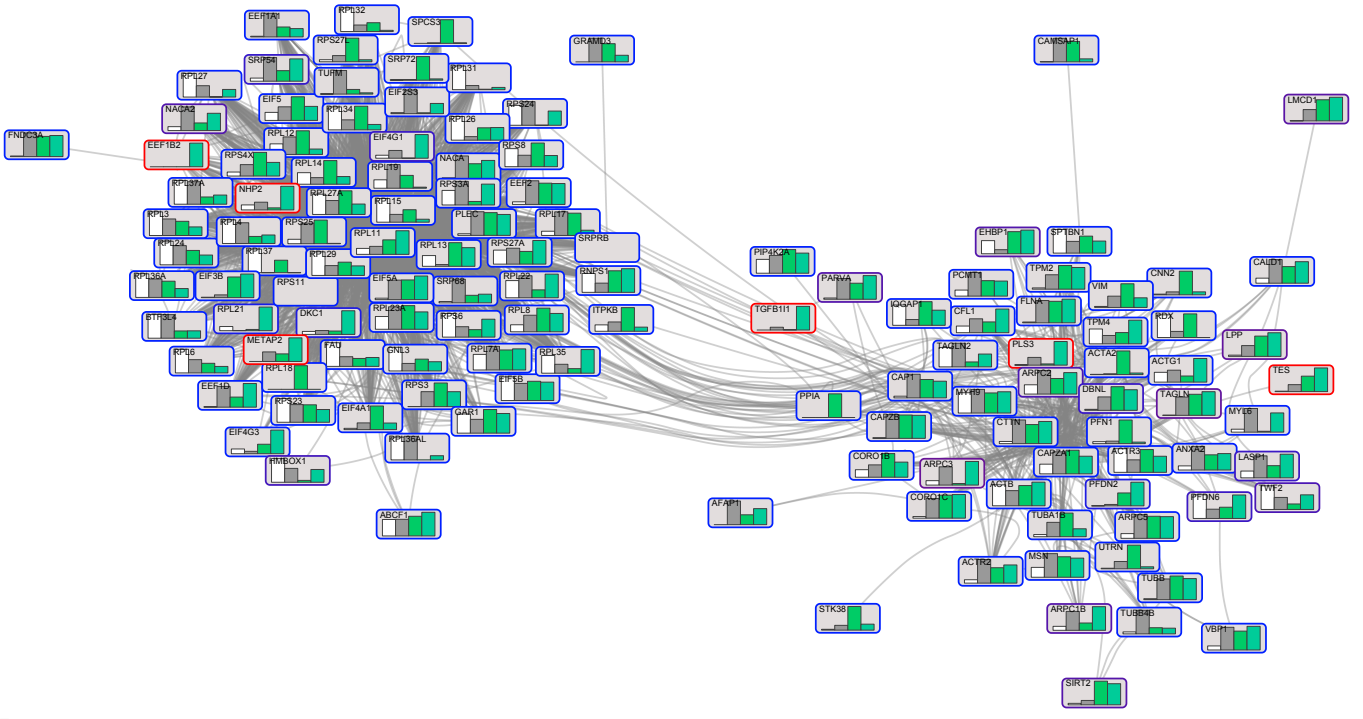
A


polysome
cytosolic small ribosomal subunit
ABCF1
EEF2
RPS6
RPS4X
RPS24
EIF4G1
RPS3
RPS27L
RPL7A
RPL8
RPS23
RPS3A
FAU
RPL32
ribosome
RPS11
RPL17
RPL36AL
RPS8
signal recognition particle, endoplasmic reticulum targeting
RPS27A
RPL29
RPL34
SRP72
RPS25
RPL37
SRP68
RPL35
RPL23A
RPL27
cytosolic large ribosomal subunit
RPL15
RPL12
RPL36A
RPL3
RPL22
RPL21
RPL11
RPL31
SRP54
RPL13
RPL6
RPL37A
RPL19
RPL4
RPL27A
RPL14
RPL24
RPL18
RPL26
telomerase activity
DKC1
HMBOX1
box H/ACA scaRNP complex
telomerase holoenzyme complex
NHP2
GAR1
box H/ACA snoRNP complex

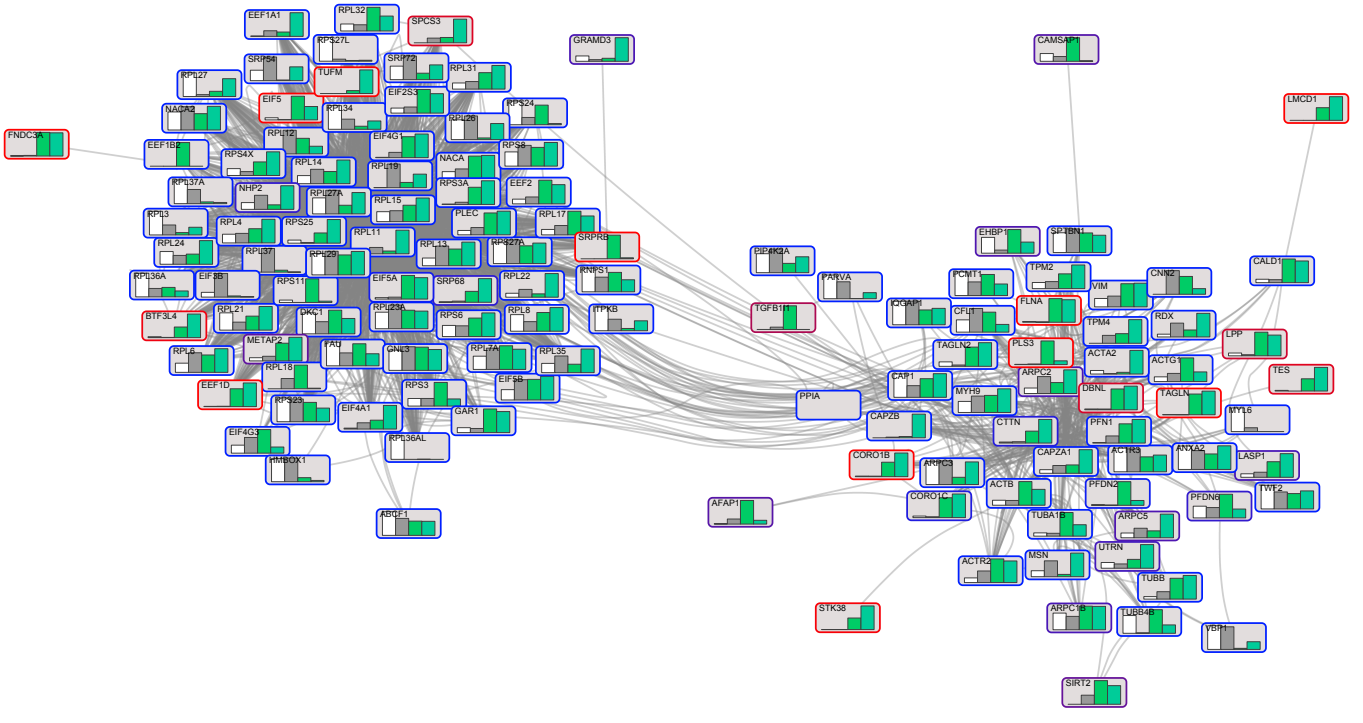
B


Arp2/3 protein complex
ARPC2
ARPC5
CAPZA1
F-actin capping protein complex
ACTR3
actin cap
ARPC1B
ARPC3
ACTR2
IQGAP1
ACTA2
MYL6
DBNL
TPM2
CAPZB
CALD1
CFL1
actin filament
RDX
actin cytoskeleton
LASP1
ACTG1
AFAP1
PFN1
CORO1C
CAP1
PARVA
FLNA
cortical atin cytoskeleton
CTTN
ANXA2
TPM4
PLS3
CORO1B
ACTB
UTRN
SPTBN1
MYH9
CNN2
actin filament bundle

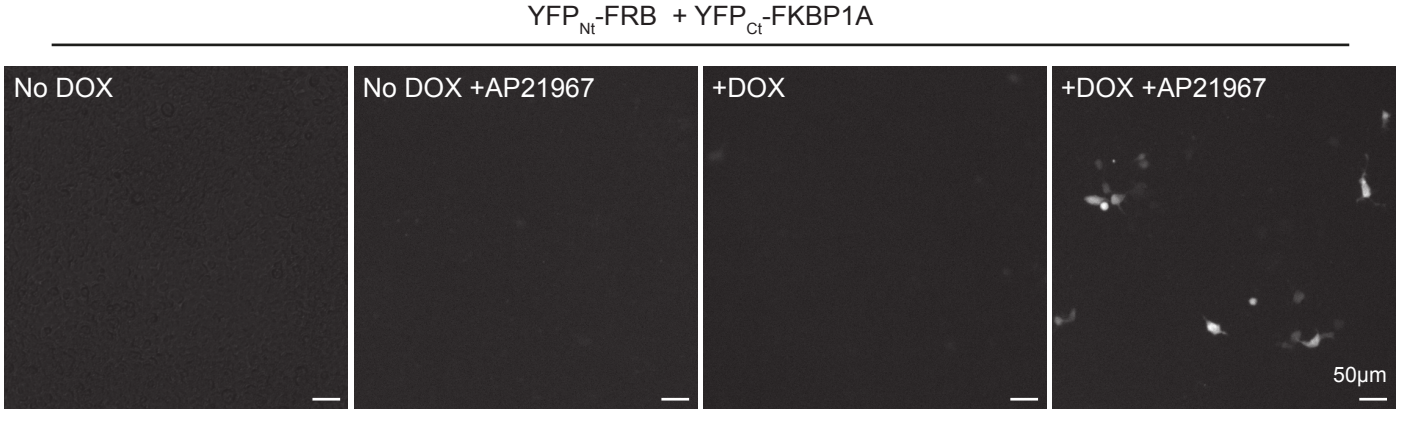
A


EEF1A1
RPL32
SPCS3
RPS27L
GRAMD3
CAMSAP1
SRP54
TUFM
SRP72
RPL31
RPL27
EIF5
EIF2S3
RPS24
LMCD1
NACA2
RPL34
RPL26
FNDC3A
EEF1B2
RPS4X
RPL12
EIF4G1
RPS8
RPL14
RPL19
NACA
RPL37A
NHP2
RPL27A
RPS3A
EEF2
RPL15
RPL13
PLEC
RPL17
SRPRB
RPS25
RPL4
RPL11
RPL24
RPL13
RPS27A
RPL37
RPL29
GNPS1
RPL36A
EIF3B
RPS11
EIF5A
SRP68
RPL22
ITPKB
BTF3L4
RPL21
DKC1
RPL23A
RPS6
RPL8
METAP2
TAU
GNL3
RPL7A
RPL35
RPL6
RPL18
EIF5B
EEF1D
RPS3
RPS23
EIF4A1
GAR1
EIF4G3
RPL36AL
HMBOX1
ABCF1
PIP4K2A
PARVA
TGFB1I1
EHBP1
SPTBN1
PCMT1
TPM2
CNN2
CALD1
IQGAP1
VIM
FLNA
CFL1
TPM4
RDX
LPP
TAGLN2
PLS3
ACTA2
ACTG1
TES
CAP1
ARPC2
MYH9
DBNL
TAGLN
PPIA
CAPZB
CTTN
PFN1
MYL6
CORO1B
CAPZA1
ACTR3
ANXA2
LASP1
ARPC3
ACTB
PFDN2
TWF2
CORO1C
PFDN6
TUBA1B
ARPC5
AFAP1
UTRN
ACTR2
MSN
TUBB
STK38
ARPC1B
TUBB4B
VBP1
SIRT2

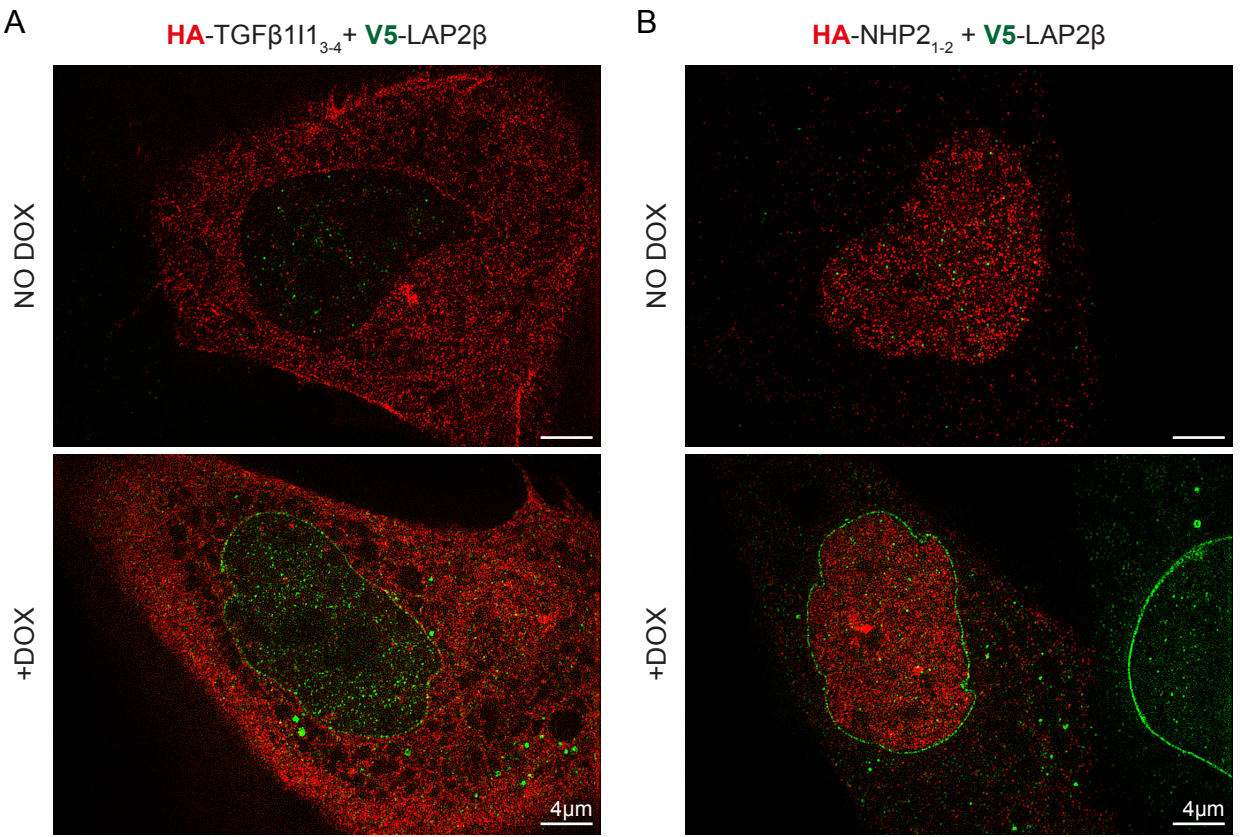
B


EEF1A1
RPL32
SPCS3
RPS27L
GRAMD3
CAMSAP1
SRP54
TUFM
SRP72
RPL31
RPL27
EIF5
EIF2S3
RPS24
LMCD1
NACA2
RPL34
RPL26
FNDC3A
EEF1B2
RPS4X
RPL12
EIF4G1
RPS8
RPL14
RPL19
NACA
RPL37A
NHP2
RPL27A
RPS3A
EEF2
RPL15
RPL13
PLEC
RPL17
SRPRB
RPS25
RPL4
RPL11
RPL24
RPL13
RPS27A
RPL37
RPL29
GNPS1
RPL36A
EIF3B
RPS11
EIF5A
SRP68
RPL22
ITPKB
BTF3L4
RPL21
DKC1
RPL23A
RPS6
RPL8
METAP2
TAU
GNL3
RPL7A
RPL35
RPL6
RPL18
EIF5B
EEF1D
RPS3
RPS23
EIF4A1
GAR1
EIF4G3
RPL36AL
HMBOX1
ABCF1
PIP4K2A
PARVA
TGFB1I1
EHBP1
SPTBN1
PCMT1
TPM2
CNN2
CALD1
IQGAP1
VIM
FLNA
CFL1
TPM4
RDX
LPP
TAGLN2
PLS3
ACTA2
ACTG1
TES
CAP1
ARPC2
MYH9
DBNL
TAGLN
PPIA
CAPZB
CTTN
PFN1
MYL6
CORO1B
CAPZA1
ACTR3
ANXA2
LASP1
ARPC3
ACTB
PFDN2
TWF2
CORO1C
PFDN6
TUBA1B
ARPC5
AFAP1
UTRN
ACTR2
MSN
TUBB
STK38
ARPC1B
TUBB4B
VBP1
SIRT2

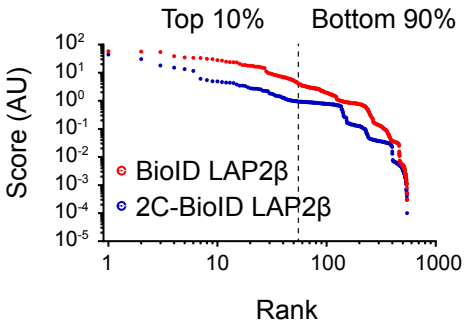
A


YFP$_{Nt}$-FRB + YFP$_{Ct}$-FKBP1A
No DOX
No DOX +AP21967
+DOX
+DOX +AP21967
YFP
50µm

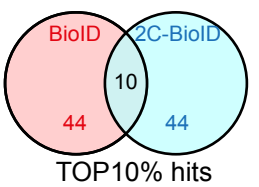
2C-BioID_Figure S5 - Related to figure 3

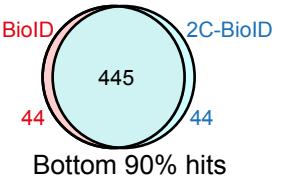


A

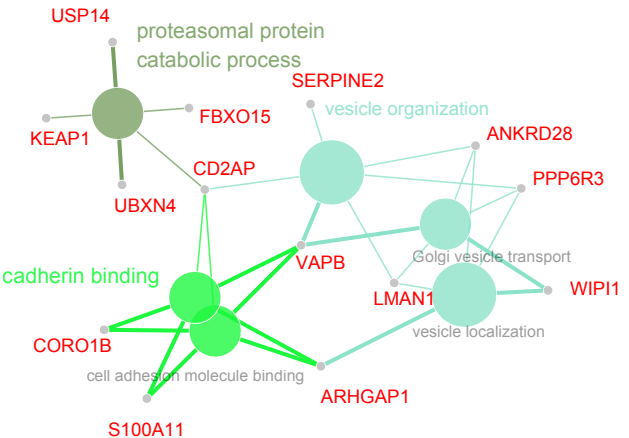

Top 10%
Bottom 90%
Score (AU)
BioID LAP2β
2C-BioID LAP2β
Rank

B

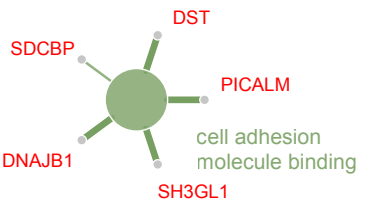

BioID
2C-BioID
10
44
44
TOP10% hits

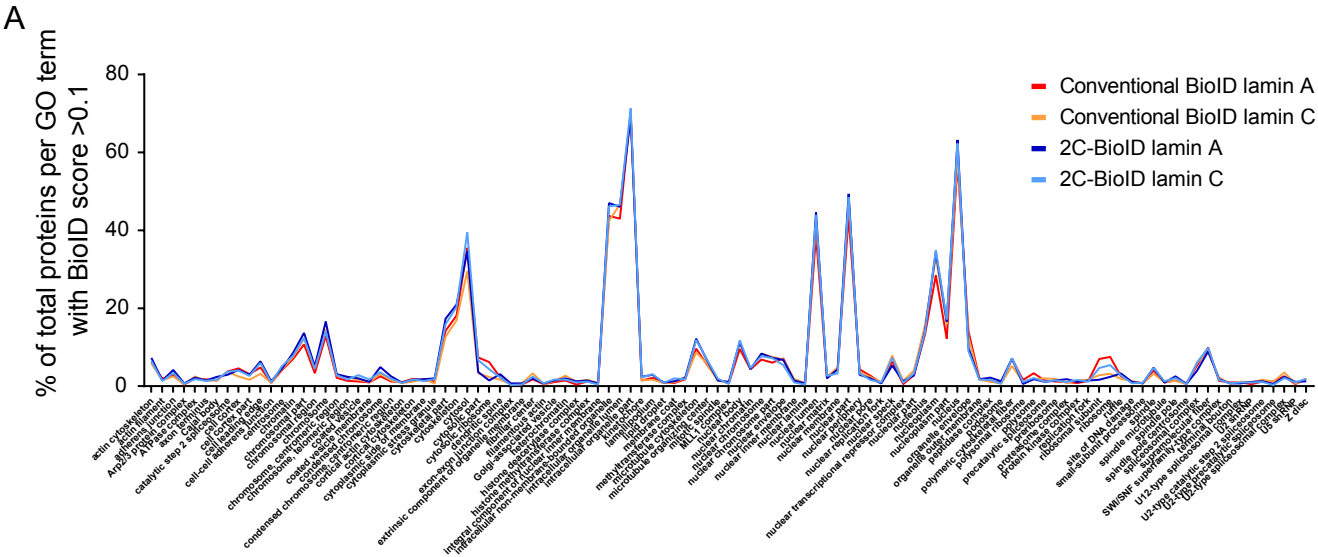

BioID
2C-BioID
445
44
44
Bottom 90% hits

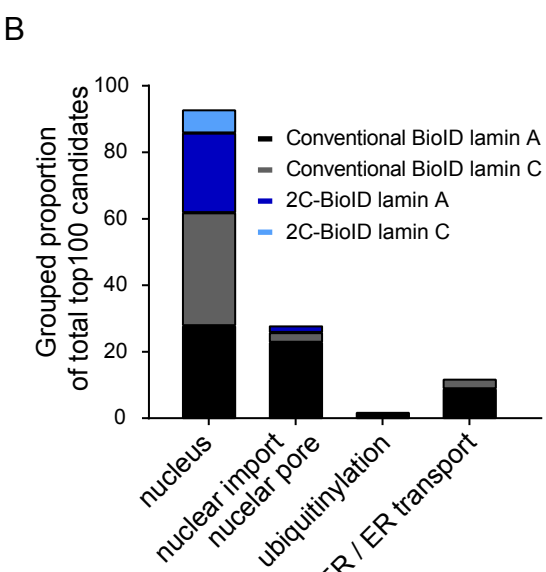
C

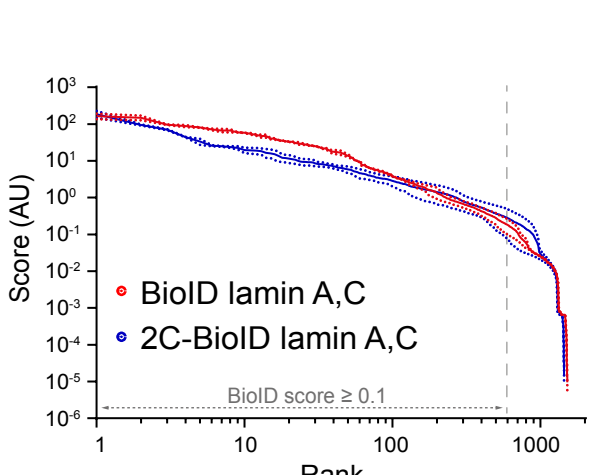

USP14
proteasomal protein catabolic process
SERPINE2
vesicle organization
FBXO15
KEAP1
ANKRD28
CD2AP
PPP6R3
UBXN4
Golgi vesicle transport
cadherin binding
VAPB
LMAN1
WIPI1
vesicle localization
CORO1B
cell adhesion molecule binding
ARHGAP1
S100A11

D


DST
SDCBP
PICALM
cell adhesion molecule binding
DNAJB1
SH3GL1


A
% of total proteins per GO term with BioID score >0.1
80
60
40
20
0
Conventional BioID lamin A
Conventional BioID lamin C
2C-BioID lamin A
2C-BioID lamin C


B
Grouped proportion of total top100 candidates
100
80
60
40
20
0
Conventional BioID lamin A
Conventional BioID lamin C
2C-BioID lamin A
2C-BioID lamin C
nucleus
nuclear import nucelar pore
ubiquitinylation
ER / ER transport


C
Score (AU)
BioID lamin A,C
2C-BioID lamin A,C
BioID score ≥ 0.1
Rank
1
10
100
1000

Supplemental figures and tables legends

Figure S1 - Related to figure 1.

Confocal imaging, high magnification details and representation of dimerization in 2C-BioID. (A) Confocal immunofluorescence microscopy showing lamin B1 antibody (green) and myc-tag antibody (red), showing doxycycline-dependent expression of myc-tagged BioID-LAP2β in human fibroblasts, and subcellular localization at the nuclear periphery/endoplasmic reticulum. Scale bar: 20 µm. (*B*) Details of widefield immunofluorescence microscopy data presented in Figure 1B. (*C*) Representation of the FKBP:FRB dimerization in 2C-BioID applied to LAP2β. (*D-E*) Details of widefield immunofluorescence microscopy data presented in Figure 1E.

Figure S2 - Related to figure 2.

Graphical representation of the ClueGO analysis of the first (A) and second (B) clusters of the network of proteins identified by BioID. Gene names are indicated in red, network nodes and associated GO-terms are indicated in identical colours.

Figure S3 - Related to figure 2.

Detail of the first two clusters from the network of proteins identified by BioID. The edge colour of the nodes is mapped to the BioID score, representing the estimated reliability of each identified protein, from high score (red) to low score (blue, see material and methods). (*A*) The bar graphs on each node represent the min-max relative quantity of its stated protein identified in each condition (+/-biotin, +/-doxycycline, +/- AP21967), from the 2C-BioID experiment. (*B*) Quantification of the same min-max relative quantity of proteins identified from the conventional BioID. See also Table S1.

Table S1 - Related to figure 2.

Full identification list, identifiers, BioID scores and associated quantifications of the average abundance for each identified candidate (Arbitrary Units), for Conventional- and 2C-BioID LAP2β assays, n=3.

Figure S4 - Related to figure 3.

Dimerization of FRB and FKBP after addition of dimerizing agent, by BiFC. (A) Live cell imaging of HEK293T cells transfected with The N- or C-terminus fragment of the split-YFP fused to FRB and FKBP, respectively. The YFP_{Nt}-FRB construct expression is under the control of the doxycycline inducible promoter. Cell culture conditions (+/- doxycycline, +/- AP21967) are indicated. Scale bar: 50µm.

Figure S5 - Related to figure 3.

(A,B) Complete frames of 3D-SIM immunofluorescence microscopy using HA- (red) and V5-tag (green) antibodies shown in Figures 3E and 3F. Scale Bar: 4 µm.

Figure S6 - Related to figure 4.

(*A*) BioID scores ranking for conventional-BioID scores (red) and 2C-BioID scores (blue). (*B*) Number of shared or specific protein between 2C- BioID and BioID, in the top 10% of the interactors (left panel) and bottom 90% identifications (right panel). See also Table S2. (*C*) Graphical representation of the ClueGO analysis of proteins ranked in the TOP 10% by BioID and belonging to the bottom 70% of the 2C-BioID analysis. See also Table S3. (*D*) Graphical representation of the ClueGO analysis of proteins ranked in the TOP 10% by 2C-BioID and belonging to the bottom 70% of the conventional BioID analysis. See also Table S3.

Figure S7 - Related to figure 4.

Comparison between 2C- and conventional-BioID for lamin A and lamin C. (*A*) Graphic representation of the proportion of proteins with BioID score above 0.1 that are associated with the stated GO terms, identified in the conventional-BioID for lamin A (red) and lamin C (orange), or 2C-BioID for lamin A (dark blue) and lamin C (light blue). (*B*) Cumulative proportion of proteins within the top 100 scores that are associated with the stated GO terms (nucleus, nuclear import/nuclear pore, ubiquitinylation and ER / ER transport), as identified in the conventional BioID for lamin A (black) and lamin C (grey), or 2C-BioID for lamin A (dark blue) and lamin C (light blue). (*C*) BioID scores ranking for 2C-BioID lamin A and C (blue), or conventional-BioID scores (red). Values are represented as mean values for lamin A and C ± SEM.

Table S2 - Related to figure 4.

Top 10% and bottom 90% ranks for BioID- and 2C-BioID analysis. BioID1, BioID2 and mTOR(FRB) were excluded from the list. Protein from the TOP 10% with nuclear localization as identified by GO term analysis or Human Protein Atlas are indicated as (GO) and (HPA), respectively.

Table S3 - Related to figure 4.

Complete list of proteins ranked in the Top 10% of interactors in one BioID method (conventional- or 2C-BioID) and ranked in the bottom 70% in the other (2C-or conventional- BioID).

Table S4 - Related to Figure 4.

Complete list of proteins identified in the lamin A and lamin C 2C-BioID assays and associated BioID scores.

Transparent Methods

Cell culture, immunofluorescence, immunoblotting and image acquisition

Human dermal fibroblast and HEK293T maintenance, immunofluorescence, 3D structural illumination microscopy analysis and immunoblotting were performed as previously described (Chojnowski et al., 2015; Xie et al., 2016a). Images were acquired using an Axiovert 200M (Carl Zeiss International), Olympus IX-83 (Olympus), Olympus FV3000 (Olympus) and DeltaVision OMX v4 Blaze (GE Healthcare). Images were processed and exported using Metamorph (Molecular devices) and ImageJ (Schindelin et al., 2012). Images were cropped and figures assembled using Adobe Photoshop CS4 and Adobe Illustrator CS6.

FKBP-FRB dimerization system

FKBP and the FRB_{T2098L} sequences were amplified by PCR from plasmids from the iDimerize system (Clontech). FRB_{T2098L} was then subcloned into the TRIPZ vector (Open Biosystems) together with a N-terminus V5-tag, to obtain the TRIPZ-FRB plasmid. The EZEO backbone was derived from the TRIPZ backbone by replacing the inducible cassette and the antibiotic resistance gene by the human EF1α promoter and Zeocin resistance cassette, respectively. FKBP was then cloned into the EZEO backbone together with a N-terminus myc-tag. Dimerization FKBP-FRB was induced using 500ηM of the rapamycin analog AP21967 (Clontech) 24h before protein extraction.

Constructs and lentiviral production, transduction and selection

For conventional BioID experiments, Full length LAP2β was amplified from TRIPZ-LAP2β and cloned at the C-terminus of TRIPZ-myc-BioID1 (Chojnowski et al., 2015). For the 2C-BioID, LAP2β was cloned in frame into the TRIPZ-FRB plasmid, at the C-terminus end of FRB. The FKBP-BioID was constructed by cloning in frame the BioID2 (Kim et al., 2016), at the C-terminus of FKBP (from the EZEO-FKBP plasmid described above). Lentiviruses were produced according to the manufacturer's protocols and stable cell lines generated as previously described (Chojnowski et al., 2015).

BiFC, Split YFP

cDNA for NHP2, TGFβ1I1, SWAP70, STRN3, CRBN, AGFG1 and C1Orf198 were amplified by PCR using cDNA clones from the ORFeome libraries (Transomic TOH3505) (Lamesch et al., 2007; Wiemann et al., 2016). The split YFP plasmids were a generous gift from Dr Mizuno (Ohashi et al., 2012). For the BiFC assay, the fragments N- (YFP_{1-210}) et C-terminus ($YFP_{210-238}$) of the Venus YFP, with a HA or myc tag respectively, were then cloned in N- and C-terminus of LAP2β, NHP2, TGFβ1I1, SWAP70, STRN3, CRBN, AGFG1 and C1Orf198 cDNA.

Antibodies

Antibodies used were as follow: lamin B1 (Dreesen et al., 2013), lamin A/C (Millipore; MAB3211), LAP2 (all isoforms, Santa Cruz, H-130), V5-tag (Invitrogen; 37-7500), Myc-tag (Cell Signaling Technology, 2276S), GAPDH (Sigma; G9545), HA-tag (Sigma 11 867 423 001.

BioID

Standard BioID was performed as previously described (Chojnowski et al., 2015): BioID-LAP2β expression was induced by 1µg/ml doxycycline (Clontech) prior to analysis, and addition of 50µM biotin (SIGMA) to the medium for 24 hours, after which proteins were extracted. The experiment was conducted in a similar fashion for 2C-BioID, the stable cell line this time expressing FRB-LAP2β under the doxycycline inducible promoter, as well as constitutively expressing the FKBP-BioID construct. The same lentiviral system was used to generate the FRB-LAP2β and FKBP-BioID expressing stable cell line. The dimerizing agent AP21967 was added simultaneously to the biotin supplementation, 24 hours prior protein extraction. Conventional- and 2C-BioID were performed in parallel, using the same media and compound conditions: 1) no DOX, no Biotin, no AP21967 2) no DOX, Biotin 50µM, no AP21967 3) DOX 1µg/ml, Biotin 50µM, no AP21967 4) DOX 1µg/ml, Biotin 50µM, AP21967 500ηM.

Mass spectrometry, BioID interactome

Proteins were extracted, and biotinylated proteins were purified using streptavidin coupled magnetic beads, as previously described (Xie et al., 2016b). Beads were then re-suspended in 50%TFE in 100mM TEAB, reduced in 20mM TCEP (20 min, 55ºC) and alkylated with 55mM CAA (20min, 25ºC). Following dilution with 100mM TEAB

(TFE <5%) two step digestion was performed using LysC enzyme (Wako, Japan) (4h at 37ºC) followed by Trypsin (Promega) for 18h at 37ºC using 1:100 enzyme:protein ratio. The reaction was stopped by acidification with 1% TFA. Peptides were desalted using C-18 SepPak columns (Waters) and vacuum dried. Samples were analyzed using 50cm x 75µm id Easy-Spray column (C-18, 2µm particles, Thermo) connected to Easy nLC1000 (Thermo) chromatography system coupled online with Orbitrap Fusion mass spectrometer (Thermo). Each sample was separated in 120min gradient (0.1% Formic Acid in water and 99.9% Acetonitrile with 0.1% Formic Acid). Data dependent mode was used in a speed mode -3 sec cycle using Orbitrap and iontrap analyzer (OT-MS 4xE5 ions, resolution 60K, IT-MS/MS 8E3 ions, normal scan). Peak lists were generated using Progenesis software (Waters) and searches were done with Mascot 2.5.1 (Matrix Science) and Uniprot Human database combined with BioID. Analysis of mass spectrometry results and network generation was then performed as previously described (Morris et al., 2014; Xie et al., 2016b). Quantification of protein abundance was derived from MS1 ion abundance using the Progenesis software. BioID score is an index built on each prey absolute abundance, abundance ratio between samples and controls, and scores available from CRAPome analysis (Mellacheruvu et al., 2013). Networks were visualized using Cytoscape and functional enrichment analysis was performed with the ClueGO plugin for Cytoscape (Mlecnik et al., 2018; Smoot et al., 2011). Bar graphs of protein quantities depicted in the network representation indicate the relative quantity (from 0-1) of the stated proteins, relative to their maximum abundance value across all experimental conditions. In this representation, the maximum abundance value is therefore set to 1, and all other values expressed relatively to this maximum.

Statistical analysis

Statistical analyses were performed using Excel (Microsoft) and Prism (Graphpad) software. Results are shown as mean ± S.E.M/SD unless otherwise indicated. Data were analyzed using two-way ANOVA and Sidak’s post-hoc test. p-values below 0.05 were considered significant.