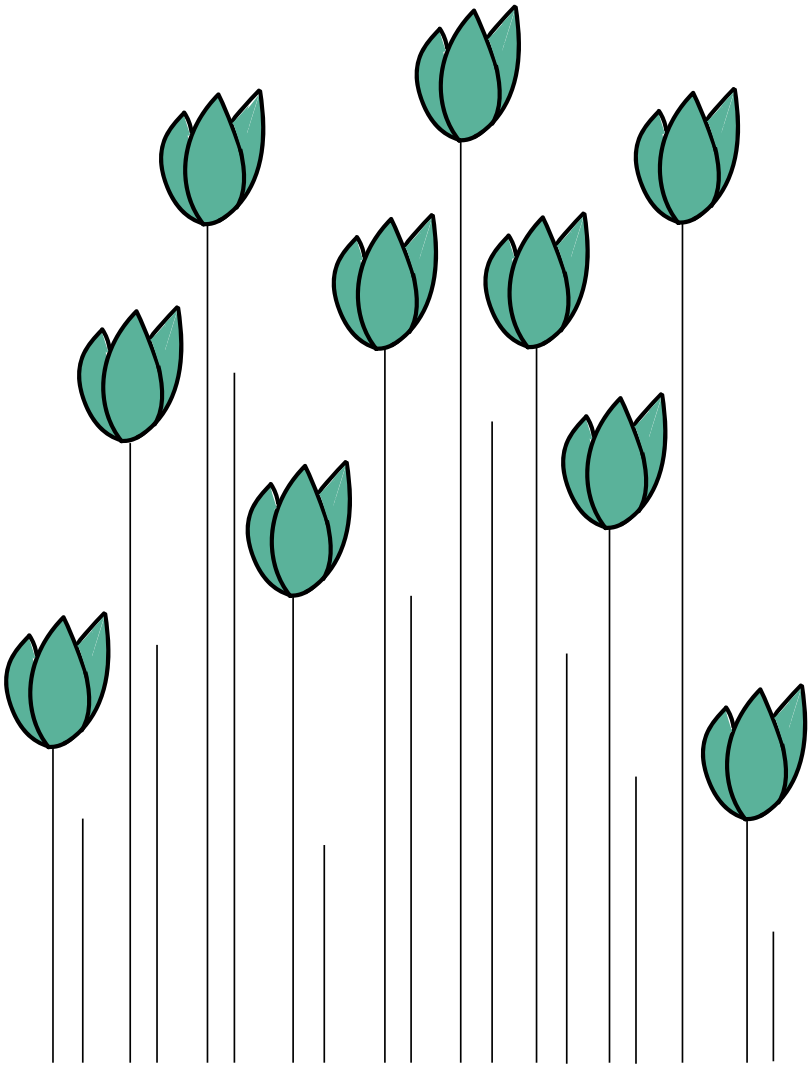
Proteomics and Functional Investigation of SUMO and Ubiquitin E3 ligases

Salas Lloret, D.



4

SUMO Activated Target Traps (SATTs) enable the identification of a comprehensive E3-specific SUMO proteome

Daniel Salas-Lloret[1] , Nicolette S. Jansen[1,8], Easa Nagamalleswari[2,8], Ekaterina Gracheva[1], Coen van der Meulen[1], Arnoud H. de Ru[3] , H. Anne Marie Otte[3] , Peter A. van Veelen[3] , Andrea Pichler[2,4], Joachim Goedhart[5] , Alfred C.O. Vertegaal[1] , Román González-Prieto[1,6,7,9]

[1] Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands.

[2] Max Plank Institute for Immunobiology and Epigenetics, Freiburg, Germany.

[3] Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands.

[4] Institute of Biochemistry, ETH Zürich, Zürich, Switzerland.

[5] University of Amsterdam, Amsterdam, The Netherlands

[6] Genome Proteomics laboratory, Department of Genome biology, Andalusian Center for Molecular Biology and regenerative Medicine (CABIMER), University of Seville, Seville, Spain.

[7] Department of Cell Biology, University of Seville, Seville, Spain.

[8] Equal contribution

[9] Corresponding author

This chapter has been published in Science Advances. 9, eadh2073 (2023).

Abstract

Ubiquitin and ubiquitin-like conjugation cascades consist of dedicated E1, E2 and E3 enzymes with E3s providing substrate specificity. Mass spectrometry-based approaches have enabled the identification of more than 6,500 SUMO2/3 target proteins. The limited number of SUMO E3s provides the unique opportunity to systematically study E3-substrate wiring. We developed SUMO Activated Target Traps (SATTs) and systematically identified substrates for eight different SUMO E3s, PIAS1, PIAS2, PIAS3, PIAS4, NSMCE2, ZNF451, LAZSUL(ZNF451-3) and ZMIZ2. SATTs enabled us to identify 427 SUMO1 and 961 SUMO2/3 targets in an E3-specific manner. We found pronounced E3 substrate preference, even at the substrate isoform level. Quantitative proteomics enabled us to measure substrate specificity of E3s, quantified using the SATT index. Furthermore, we developed the Polar SATTs web-based tool (https://amsterdamstudygroup.shinyapps.io/polarVolcaNoseR_revised/) to browse the dataset in an interactive manner, increasing the accessibility of this resource for the community. Overall, we uncover E3-to-target wiring of 1681 SUMO substrates, highlighting unique and overlapping sets of substrates for eight different SUMO E3 ligases.

Keywords: Mass-Spectrometry; SUMO; E3s; SATTs

INTRODUCTION

Protein fate and function is controlled by numerous Post-Translational Modifications (PTMs). Among them, ubiquitination is the second most important PTM after phosphorylation (1) and controls virtually every process in eukaryotic cells in a dynamic manner. Ubiquitination consists of the covalent attachment of the small 76 amino acids ubiquitin protein to acceptor proteins. It is performed by an enzymatic cascade in which ubiquitin-activating enzymes (E1) activate ubiquitin and transfer it to ubiquitin-conjugating enzyme (E2) which conjugates ubiquitin to the substrate assisted by a ubiquitin-ligase enzyme (E3). E3s are responsible for determining substrate specificity. The human genome encodes for two ubiquitin E1s, 30-40 E2s and more than 600 E3s (2).

Similar to ubiquitin, other ubiquitin-like (Ubl) modifiers exist, which have dedicated E1-E2-E3 enzymatic cascades. Among these Ubls, Small Ubiquitin-like Modifiers (SUMOs) are the most abundant after ubiquitin. In vertebrates, there are three different types of active SUMOs: SUMO1, SUMO2 and SUMO3. Mature SUMO2 and SUMO3 differ only in a couple of amino acids and are commonly referred to as SUMO2/3. In contrast to ubiquitin, vertebrates express a single E1, a single E2 and less than a dozen *bona fide* E3s for SUMOs (2).

Recent advances in mass spectrometry technologies and the optimization of sample preparation methodologies (3) have enabled the identification of several tens of thousands of acceptor sites on thousands of proteins in human cells both for ubiquitin and SUMOs (4-10). However, our knowledge on E3-substrate wiring is still very limited. Determining which E3 modifies which substrate is a major challenge.

For ubiquitin, given the high number of E3s, solving the E3-to-target wiring in a proteome-wide manner is virtually impossible. However, for SUMOs, the E3 complexity is limited, simplifying this task. A proposed approach has been the quantification of changes on the SUMO proteome after SUMO E3 overexpression(11), which in principle, is an indirect measure. Another applied approach has been the performance of SUMOylation assays on protein array-based screens (12), which is an *ex vivo* system that misses out on the restricted subcellular localization of proteins and lacks protein-protein complexes that are abundant in cells.

Here, we took advantage of our previous experience in the systematic identification of ubiquitination substrates using Ubiquitin Activated Interaction Traps (UbAITs) (13) in the Targets of Ubiquitin Ligase Identified by Proteomics (TULIP) methodologies (14-16) and applied it for the identification of SUMO E3-specific substrates in a systematic manner for SUMO E3s in a proteome-wide approach.

RESULTS

SUMO E3 overexpression causes SUMO2/3 depletion in an RNF4-dependent manner.

Aiming to identify putative E3-specific SUMOylation substrates, we employed a similar approach as previously done with PIAS1(11). We made GFP-tagged constructs for different SUMO E3s, including NSMCE2, PIAS1, PIAS2, PIAS3, PIAS4, ZNF451, the LAP2α isoform of the ZNF451 SUMO Ligase (LAZSUL) and, additionally, another PIAS-like enzyme, ZMIZ2(17-20) (**Figure 1A**), which , we previously tested for *in vitro* SUMO E3 activity both for SUMO1 and SUMO2 ligase activity (**Figure 1B**). As a result, we observed that ZMIZ2 has E3 enzymatic activity for SUMO2 but not for SUMO1. Next, we transfected the GFP-tagged constructs of the E3s indicated in Figure 1A in U2OS cells. To evaluate the transfection efficiency of our constructs, we analyzed our cells by fluorescence microscopy after immunostaining for SUMO2/3 (**Figure 1C, Supplementary Figure 1A**). GFP-positive cells could be observed for every construct at different efficiencies, except for GFP-PIAS2, which transfection did not lead to the appearance of GFP-positive cells. Unexpectedly, the immunofluorescence SUMO2/3 signal was highly reduced in GFP-positive cells for NSMCE2, PIAS1, ZNF451 and LAZSUL (**Figure 1C, Supplementary Figure 1A**). Therefore, we quantified the SUMO2/3 signal by immunofluorescence for GFP-positive and -negative cells from three independent experiments (**Figure 1D**). While GFP-NSMCE2, -PIAS1, -PIAS3, -ZNF451 and -LAZSUL reduced the average SUMO2/3 nuclear signal by 41%, 63%, 19%, 77% and 82% respectively, GFP-PIAS4 positive cells presented a slight increase of 4% in SUMO2/3 signal. Interestingly, GFP-ZMIZ2 positive cells had a remarkable 46% increase in SUMO2/3 signal.

In a previous screen for targets of the SUMO Targeted Ubiquitin Ligase (STUbL) RNF4, we had observed that SUMO E3s were targets of RNF4 for ubiquitination and subsequent degradation by the proteasome, with ZNF451 and PIAS1 being the strongest RNF4 ubiquitination targets, and PIAS4 the weakest. ZMIZ2 was not a substrate for RNF4 (15, 16). We hypothesized that overexpression of these E3s was promoting their hyperactivation leading to their auto-SUMOylation and increased SUMOylation of their substrates and subsequent degradation in an RNF4-dependent manner. Thus, proteasome inhibition should rescue the effect on SUMO2/3 levels in response to the different E3s overexpression. Therefore, we compared the effect on SUMO2/3 levels of the different SUMO E3s overexpression in the presence or absence of the proteasome inhibitor MG132 for 5 hours (**Supplementary Figure 1B**). The 5 hours MG132 treatment could rescue the effect on SUMO2/3 levels of the E3s which overexpression had milder phenotypes, namely NSMCE2, PIAS3 and PIAS4, but was not sufficient to rescue strong effects of PIAS1, ZNF451, LAZSUL overexpression. Furthermore, proteasome inhibition by MG132 has many pleiotropic effects.

To further test our hypothesis, we made stable inducible U2OS cells for GFP-LAZSUL, which was the E3 with the strongest phenotype (**Figure 1C-D, Supplementary Figure 1 A-B**). Cells were treated with a control or an RNF4-targetting siRNA, GFP-LAZSUL was dox induced and analyzed by immunostaining (**Figure 1E-F**) and immunoblotting (**Figure 1G**). RNF4 knock down caused an increase in the fraction of GFP-LAZSUL positive cells and rescued the SUMO2/3 depletion phenotype. Consistently, RNF4 knockdown increased the levels of both modified and non-modified GFP-LAZSUL.

RNF4 knockdown increases cellular SUMO2/3 levels but does not affect SUMO1 levels (21). Thus, we also decided to investigate the effect of NSMCE2, PIAS1, ZNF451, and LAZSUL transient overexpression on SUMO1 levels by immunofluorescence (Supplementary Figure 1C-D). Accordingly, the overexpression of these E3s did not cause SUMO1 depletion as previously observed for SUMO2/3 (Figure 1C-F).

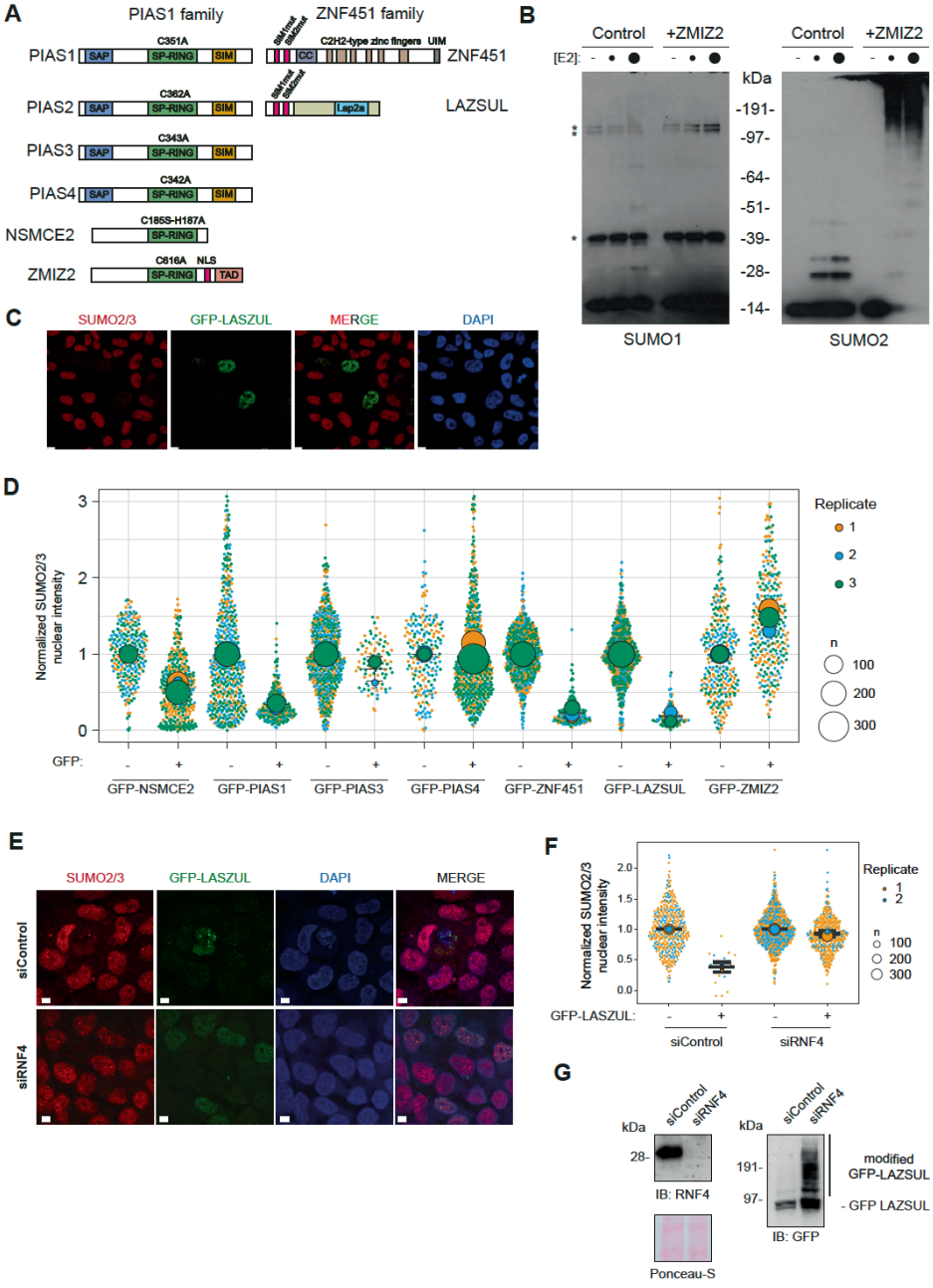


Figure 1. (A) E3s studied in this article. The mutations performed on each E3 to construct the catalytic-dead mutant controls are indicated **(B))** *In vitro* sumoylation assays including or ZMIZ2 SUMO E3 enzyme and different concentrations of the SUMO E2 (0, 100 and 200 ng). Assays were carried out using either SUMO1 or SUMO2. **(C)** Representative immunofluorescence image of U2OS cells transiently transfected with GFP-LAZSUL immunostained for SUMO2/3. **(D)** Superplot depicting relative SUMO2/3 nuclear intensities after immunostaining of individual U2OS cells transiently transfected with GFP-tagged constructs of different E3s. Values were normalized to the average SUMO2/3 nuclear intensity of GFP-negatives from each individual experiment. Values from 3 independent experiments are depicted. **(E)** Stable inducible GFP-LAZSUL expressing U2OS cells were treated with control or RNF4-targeting siRNAs. 36h after siRNA transfection, GFP-LAZSUL expression was induced with 20 µg/mL doxycycline. Cells were fixe 48h after siRNA transfection and analyzed by immunostaining. **(F)** Quantification of the normalized nuclear SUMO2/3 intensities from the cells in (E). Independent values from two independent experiments are depicted. **(G)** Analysis by immunoblotting of the cells in (C). Size bars in fluorescence microscopy images represent 10 µm.

4

We conclude that SUMO E3 overexpression-based screens to identify SUMOylation substrates could potentially be misleading due to a negative control loop mediated by RNF4. This loop is activated upon SUMO E3 overexpression and leads to SUMO2/3 depletion in cells. Therefore, SUMO E3 overexpression screens must be carefully evaluated.

SUMO Activated Target Traps (SATTs) to identify the E3-specific SUMO proteome.

Previously, in an effort to identify E3-specific ubiquitin substrates, Ubiquitin Activated Interaction Traps (UbAITs) were engineered (13), which we later adopted and optimized for systematic screening in the Targets for Ubiquitin Ligases Identified by Proteomics (TULIP) methodology (15, 16). However, due to the high number of Ubiquitin E3 enzymes in the human proteome, performing the TULIP methodology on each E3 is an incredibly challenging task.

In contrast to Ubiquitin, the number of *bona fide* SUMO E3 enzymes is more limited, comprising the Siz/Pias Really Interesting New Gene (S-P RING) family, the ZNF451 family and RANBP2 (2). Therefore, addressing the E3-substrate wiring for SUMO E3s is a more manageable challenge.

Thus, similar to the TULIP2 methodology (15), we designed the SUMO Activated Target Traps (SATTs) approach, in which lentiviral doxycycline-inducible plasmids consisting of 10xHIS tag and a Gateway cloning sequence, followed by 10xHIS and either mature SUMO1 or mature $SUMO2_{Q87R}$ were constructed. **(Figure 2A**) The Gateway sequence enables the straightforward shuttling of any SUMO E3 of interest. The $SUMO2_{Q87R}$ mutation facilitates the identification of SUMO acceptor sites by mass spectrometry-based proteomics (7, 8). Consistently, the rationale behind this approach is that if we generate a linear fusion between an E3 and activated SUMO, the E3 will be prone to use the attached SUMO moiety to modify its substrate, enabling the co-purification of the E3 together with its substrate and subsequent identification by mass spectrometry-based proteomics. In line with TULIP2 methodology (15), we included two different negative controls in our screens. The first control is a ΔGG construct where the SUMO moiety lacks the C-terminal di-Gly motif and thus cannot be conjugated to a substrate. The second control is a catalytic dead mutant where the interaction with the SUMO E2 enzyme is abolished, thus the transfer of the SUMO moiety from the E2 to the substrate cannot be catalyzed **(Figure 2B**).

Accordingly, we built SATTs for the E3s indicated in Figure 1A. SUMO1 SATTs for the S-P RING SUMO E3 enzymes, PIAS1, PIAS2, PIAS3, PIAS4 and NSMCE2. Additionally, for SUMO2 SATTs, in addition to the S-P RING SUMO E3 enzymes, we also included ZNF451, LAZSUL and ZMIZ2 as they are exclusive for SUMO2/3 (22, 23) (**Supplementary Figure 2A**). Other characterized SUMO E3s were not included for different reasons. RANBP2 was left out from our screen due to the size of the protein (3,224 amino acids). Also, the ZNF451 family E3 KIAA1586 was left out because it is exclusively found in primates, and not in other vertebrates (23). To generate the catalytic-dead mutant controls, we introduced specific mutations in each E3 (**Figure 1A**). For the S-P RING family E3s, we mutated cysteines in the S-P RING domain, for the ZNF451 family E3s, we mutated the SUMO Interaction Motifs (SIMs) by substituting the long hydrophobic amino acids into alanines (23).

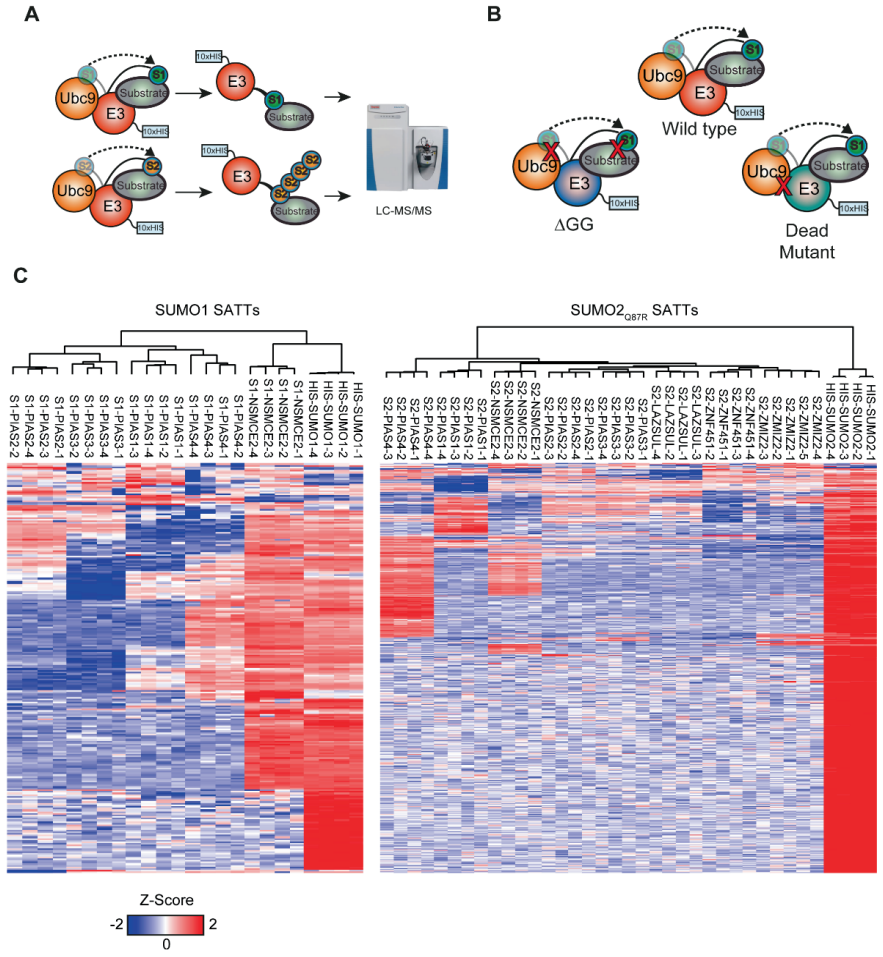


Figure 2. (A) SATTs screen rationale. SUMO moieties fused to the C term of an E3 of interest will be attached to E3 substrates, enabling the co-purification of the E3 together with the SUMOylation target, which will be later identified by mass spectrometry. **(B)** SATTs negative controls rationale. While ΔGG SATTs lack the C-terminal SUMO diGly motif, unable to conjugation to the substrate, catalytic-dead mutants prevent

interaction with the SUMO E2. **(C**) Heatmap of Z-scores for different SATT targets. Only HIS-SUMO1 and HIS-SUMO2$_{Q87R}$ targets are included.

Next, we constructed U2OS cells expressing HIS-SUMO1 or HIS-SUMO2$_{Q87R}$ in a constitutive manner or stably expressing the inducible E3 SATTs constructs indicated in Supplementary Figure 2A, including the ΔGG and catalytic-dead mutant negative controls. We induced the expression of the constructs for 24h, lysed the cells in denaturing conditions and purified the SATT conjugates from four independent biological repeats (five in the case of ZMIZ2) yielding a total of 171 samples. In order to avoid RNF4-mediated degradation of the SATTs due to auto-SUMOylation (**Figure 1**), and to increase the number of SUMOylation conjugates (8), the proteasome was inhibited for 5h with MG132.

Sample analysis by immunoblotting (**Supplementary Figure 2B-C**) showed that the expression levels of the SATTs were below or close to endogenous counterparts for every construct. Moreover, signal could be observed in a higher molecular weight smear for the wild type and catalytic-dead mutant construct for every SATT corresponding to E3-SUMO-target conjugates. This smear was absent in the ΔGG constructs. Consistently, the catalytic-dead SATT smears had different profiles than their wild type counterparts, indicating that the SUMO moieties in the mutant SATTs could still be used for conjugation by other endogenous E3s.

For the cell lines expressing either HIS-SUMO1 or HIS-SUMO2$_{Q87R}$ in a constitutive manner, mass spectrometry analysis of the samples enabled the identification of 244 SUMO1 targets and 1509 SUMO2 targets (**Supplementary Datasets 1-2**) after 5h of proteasome inhibition with MG132. Among the 244 SUMO1 targets, 171 could be considered a SUMO1 SATT conjugate for at least one SUMO E3. In the case of SUMO2 SATT conjugates, the numbers were 570 out of 1509 which were preferential or specific for the different E3s (**Figure 2C, Supplementary Datasets 3-4**). When compared with the biggest SUMO2 target study to date in U2OS cells (7) the number of SUMO2 SATT targets that had been identified as SUMO2 targets increased to 656 (**Supplementary Dataset 5**).

However, the SUMO proteome is highly diverse depending on the cell type and experimental conditions of study (7). Therefore, we considered the possibility that the SATTs could enable the modification and enrichment of SUMOylation substrates which are not constitutively modified in the cell line and condition of study or not detectable when a total SUMO proteome purification is performed. Accordingly, we also analyzed the mass spectrometry data from the SATTs in an unbiased manner. Namely, proteins not considered as His-SUMO targets were included when identified as SATT substrates (**Supplementary Figure 3, Supplementary Datasets 6-7**). This way, we could identify 302 extra putative substrates for SUMO1 SATTs, mainly for NSMCE2 and PIAS2 and 459 new putative substrates for SUMO2 SATTs, mainly for PIAS2 and LAZSUL.

The SATT Index.

Although the substrates we identified for each tested SUMO E3 were relatively specific for every E3 when comparing a wild type SATT with its ΔGG counterpart, all the substrates did not remain equally significant compared to their catalytic dead mutant counterpart (**Supplementary datasets 3,4,6,7**), indicating that, as previously shown for UbAITs (13, 15, 16), the SUMO moiety attached to the mutant SATT can also be conjugated to a substrate by another endogenous SUMO E3.

Therefore, we used the relation between the differences of the enrichment of a substrate for a specific E3 comparing both the ΔGG and the mutant counterpart to wildtype, which we termed SATT index:

$$SATT_i = \frac{[\log_2 SATT^{WT}] - [\log_2 SATT^{Mut}]}{[\log_2 SATT^{WT}] - [\log_2 SATT^{\Delta GG}]}$$

Values close to 1 and higher are considered very specific and values close to 0 and lower are considered not specific.

Different E3s have different preferences towards SUMO1 or SUMO2/3.

It could be argued that making a SUMO1 SATT with an E3 which normally catalyzes SUMO2/3 conjugation might force SUMO1 conjugation on SUMO2/3 substrate. Thus, we decided to investigate if SUMO E3s could discriminate substrate specificity depending on the SUMO type they were conjugating. First, we looked at the overlap between SUMO1 and SUMO2 substrates (**Figure 3A**). Overall, and similar to previous studies, where SUMO2 is the most abundant and important SUMO(10, 24) , while most of the identified SUMO1 targets (87%) can be also modified by SUMO2/3, only 14% of the SUMO2 substrates can be also modified by SUMO1. Next, we looked at the substrate preference and overlap for the different S-P RING E3s that had been investigated both for SUMO1 and SUMO2 SATTs (**Figure 3B**). On one side, in contrast to SUMO proteome data (**Figure 3A**), NSMCE2, PIAS1and PIAS2 data indicated a preference for SUMO1 modification. For PIAS1, only 9% of the SUMO1 substrates were also SUMO2/3 substrates (87% in SUMO proteome), and 45% of the SUMO2/3 were also substrates for SUMO1 (14% in SUMO proteome). This preference was milder for NSMCE2, where the numbers were 63% and 57%, respectively, and PIAS2, 69% for SUMO1 and 23% for SUMO2. For PIAS3 and PIAS4 the preference for SUMO2 modification was more acute than in total SUMO proteome analysis. For PIAS3 there was only one protein identified for SUMO1 conjugation. This protein was also found for SUMO2 conjugation (100%). For PIAS4, 93% of SUMO1 conjugates were also modified by SUMO2. These values indicate that PIAS3 and PIAS4 are mainly a SUMO2/3 E3 enzyme which is consistent with previous studies on SUMO specificity, and, consistently, SUMO1 ligase activity has also been reported to be higher for PIAS1 and PIAS2 (12).

Gene Ontology analysis.

Gene Ontology analysis for biological processes of the SUMOylation substrates for the different E3s indicated that different E3s are involved in different processes (**Figure 3C, Supplementary Dataset 8**). As expected, PIAS1, PIAS4, NSCME2 and ZNF451 substrates are enriched in Gene Ontology terms relative to genome biology (25-31), and PIAS3 substrates are enriched for maintenance of proteins at the nucleus (32-34). LAZSUL highest enrichment term was protein SUMOylation. ZMIZ2 substrates were not enriched for specific cellular processes. Interestingly, PIAS2 substrates are enriched for membrane translocation and ADP/ATP mitochondrial transport.

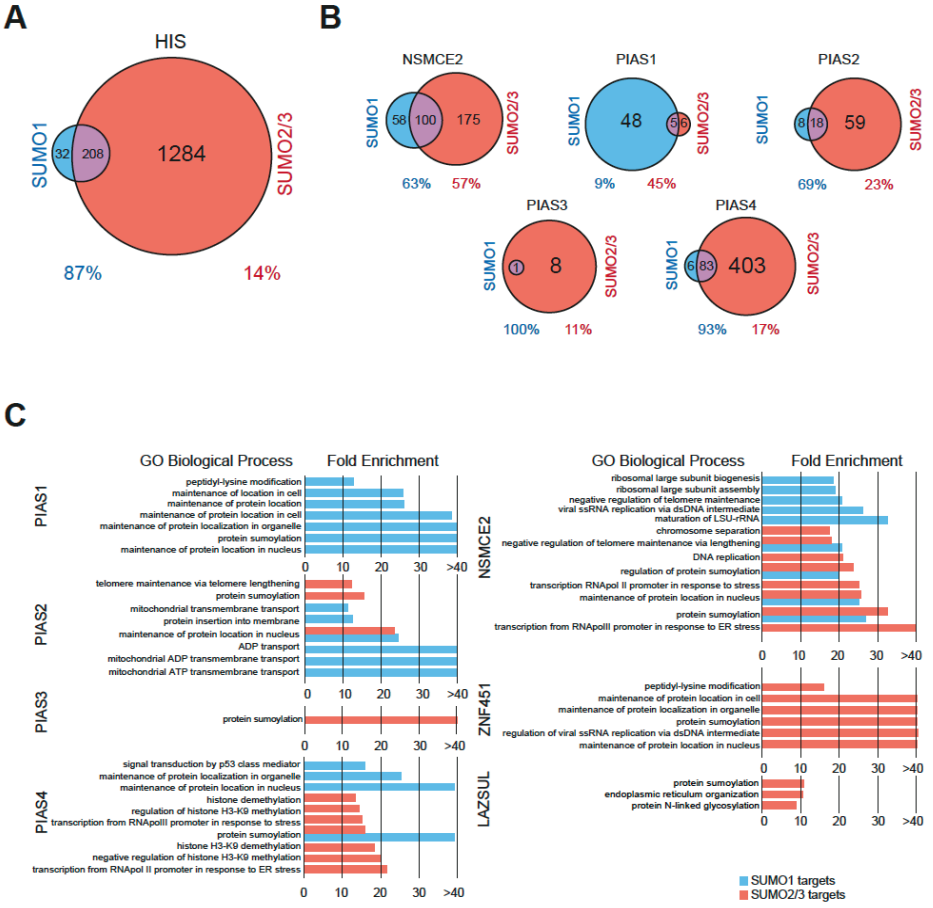


Figure 3. (A) Overlap between HIS-SUMO1 and HIS- SUMO2$_{Q87R}$ targets. **(B)** Overlap between SUMO1- and SUMO2-SATT substrates for the indicated E3s. **(C)** Gene ontology analysis for the SUMOylation substrates of the different E3 SATTs analyzed in this study.

PIAS4 and NSMCE2 make hybrid SUMO1-SUMO2/3 chains.

$SUMO2_{Q87R}$ SATTs leave a QQTGG remnant after tryptic digestion on acceptor lysines which can be identified by mass spectrometry-based proteomics (8). Although K11 is known to be the canonical site to make SUMO2/3 chains (35), several other SUMO2/3 sites at the endogenous level have been identified (6). Therefore, in addition to K11- SUMO2/3 chains, other chain types exist.

Mass spectrometry analysis of our samples enabled us to obtain MS/MS spectra in which the QQTGG remnant could be localized on SUMOs in an unambiguous manner (**Supplementary Figure 4**) and the intensity of these SUMOylation sites could be quantified (**Figure 4**). SUMO2/3 K11 chains, were found with every E3. As expected, no QQTGG-modified peptides were found in ΔGG SATTs samples. In contrast, signal for the modification with SUMO2/3 on K11 either on SUMO2 or SUMO3 could be detected in every SATT and, at less intense level, on K5. This included both wild type and catalytic-dead mutant SATTs. Importantly, only NSCME2 and PIAS4 wild type SATTs were able to modify SUMO1 with SUMO2/3 at K7, being completely dependent on the catalytic activity of the SATT. Similarly, SUMO3 K7 chains were also formed with NSCME2 and PIAS4, but only depending on the catalytic activity of NSMCE2. Finally, SUMO2/3/4 K32/33/33 chains were only detected for NSCME2 and were completely depending on NSMCE2 catalytic activity.

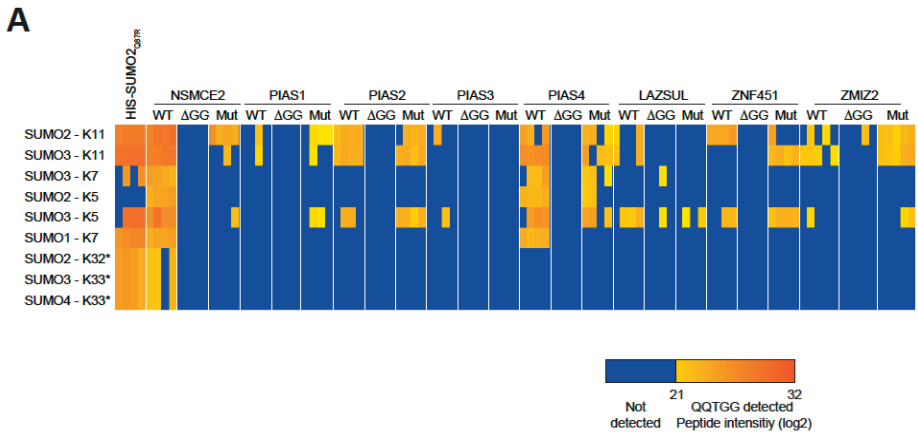


Figure 4 (A). Total peptide intensities of the different peptides corresponding to different acceptor lysines in SUMO1, SUMO2, SUMO3 and SUMO4 depicted in Supplementary Figure 4 in the different SUMO2Q87R SATT samples. * Indicates that corresponds to a peptide present in the three indicated SUMO types, thus, not enabling distinction.

SATTs complement analysis

Finally, we made statistical comparisons of proteins that were enriched or depleted in the wild type SATTs samples using all the other wild type SATTs as control. This was done both for considering and not considering exclusively the SUMO1 or SUMO2-3 substrates identified in Supplementary Datasets 1 and 2 (**Supplementary Datasets 9-12, Figure 5 and Supplementary Figure 5**).

Altogether, these analyses enable the identification of very high confidence E3-specific SUMOylation substrates in which wild type SATTs target proteins were statistically enriched when

compared to their ΔGG and catalytic dead mutant counterparts and compared to all the other wild type SATTs (**Supplementary Dataset 13**).

Polar SATTs – A user-friendly site to browse the dataset.

Most proteomic screens, including this one, usually consist of large spreadsheet datasets full of gene/protein names and values and comparisons. The interpretation of these datasets can be daunting for researchers from other disciplines. To overcome these potential outreach hurdles, we developed an online web app tool to browse the dataset, which is freely accessible (https://amsterdamstudygroup.shinyapps.io/polarVolcaNoseR_revised/). This tool enables users to select a protein of interest, and, if present in this study, will pop up in a polar plot in the sectors corresponding to the relevant E3s, indicating enrichment in terms of p-value and difference between wild type and either ΔGG SATTs, or all the other wild type SATTs as complement control. Additionally, for the ΔGG SATTs, how relatively specific the substrate is for the E3 in terms of SATT index is depicted with a color scale.

Moreover, the app can be used to customize the data visualization by enabling adjustment of the p-value and differences, choosing to hide the values that exceed the limits. The size of the datapoints can also be adjusted to facilitate visualization, and the resulting figure highlighting the desired substrates can be exported in .pdf or .png. A “Dark Theme” is also available.

Finally, if users prefer browsing in independent Volcano Plots instead of the default Polar Plot, that is also enabled. An example of visualization of different SUMOylation substrates for different E3s using the app is shown in Figure 6.

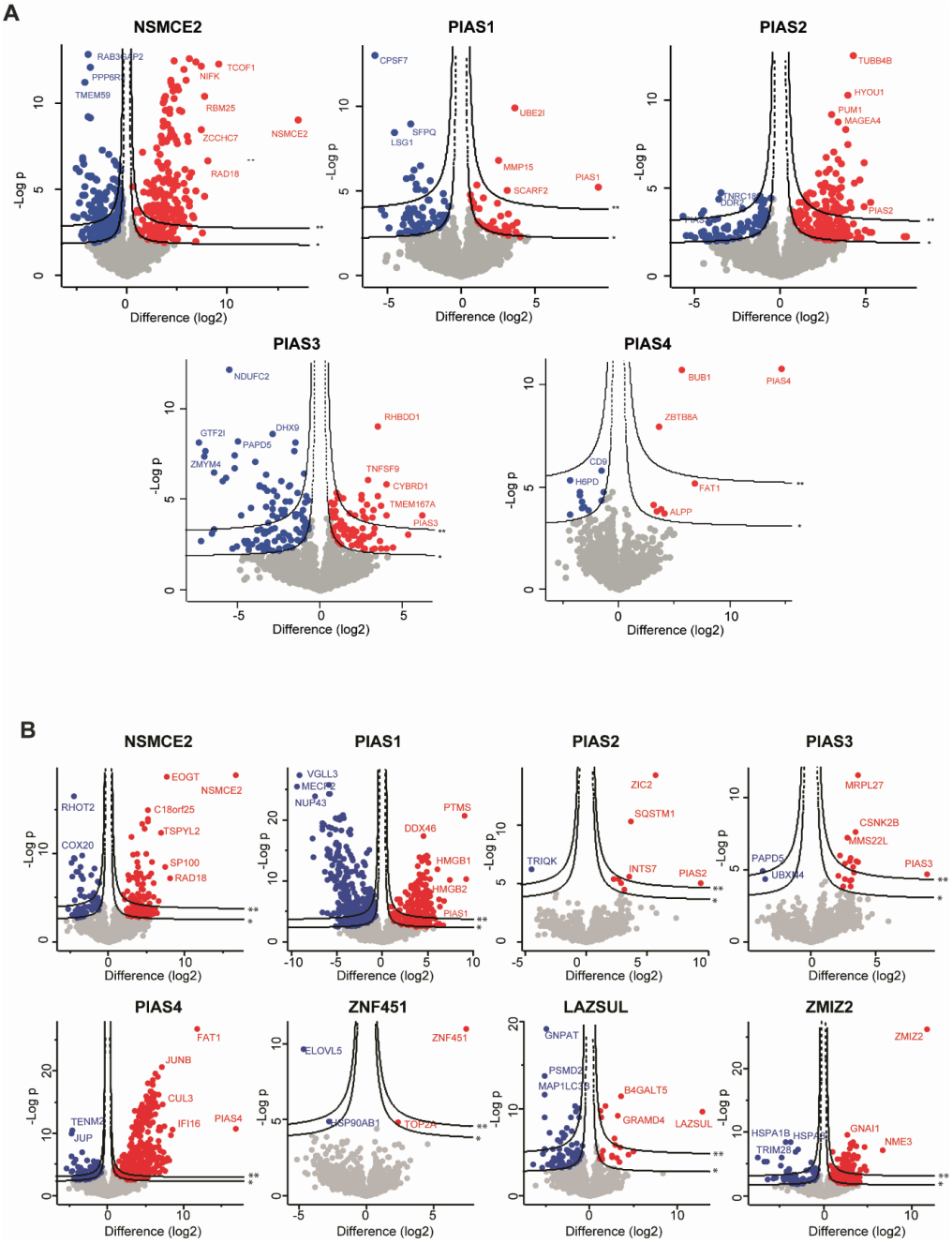


Figure 5. Multiple volcano plots (Hawaii plot) depicting statistical differences of different unbiased SUMO1- **(A)** or SUMO2- **(B)** SATTs targets using the values from the rest of wild type SATTs as complement negative control. Cutoffs represent a Pearson of 100 and an FDR=0.05 (*) or 0.01 (**) and an S0=0.1.

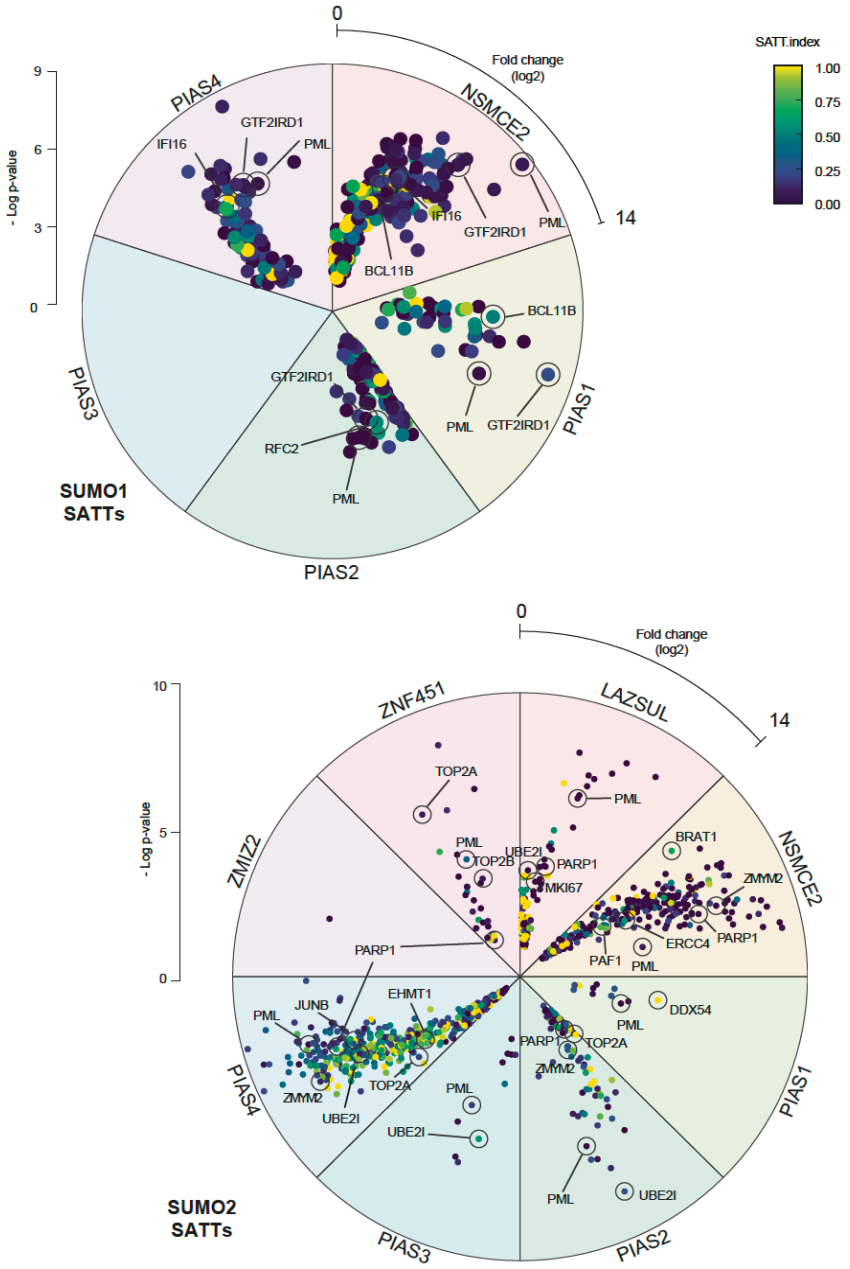


Figure 6. SATT Polar plots extracted from the polarVolcaNoseR web app. Some of the most prominent or specific substrates from different E3s are indicated. Depicted plots use ΔGG as control, includes the SATT index and only consider HIS-SUMO targets.

Orthogonal Validation

Finally, we performed orthogonal validation of the newly identified substrates for some E3 enzymes, including NSMCE2, LAZSUL and PIAS4.

For orthogonal validation of NSMCE2 substrates, we stably expressed or not HIS-SUMO2Q87R in parental and NSMCE2 knockout (NSMCE2-KO) cells (36) and rescued NSMCE2-KO with either wild type NSMCE2 or with a catalytic dead mutant $NSMCE2_{C185S/H187A}$ (37). Next, cells were cultured, treated with proteasome inhibitor MG132 as previously done for the SATTs and lysed, for subsequent Ni-NTA purification of the HIS-SUMO2Q87R proteome. Mass spectrometry based proteomics analysis was performed and the obtained data processed together with the NSCME2-SUMO2-SATT data (**Figure 7A-B, Supplementary dataset 14**). Among the statistically significant NSMCE2-SUMO2-SATT substrates compared to both their ΔGG and catalytic mutant counterparts and both in this analysis and analysis from Supplementary Dataset 4, 31 proteins were also considered as putative SUMOylation substrates for both U2OS and NSCME2-KO parental cells. Of those 31, 10 proteins were statistically reduced in their SUMOylation levels depending on NSMCE2 catalytic activity. Among the 21 proteins which SUMOylation status was not affected by the lack of NSMCE2 catalytic activity, 17 of them were also strong substrates for other E3 SATTs. Two proteins, namely, ERCC4 and PAF1, were strong substrates for NSMCE2-SUMO2-SATT and had their SUMOylation levels affected by lack of NSMCE2 catalytic activity and were not strong E3 substrates for other SATTs. We decided to analyze these two candidates by immunoblotting (**Figure 7C**). Consistently, a big reduction in the SUMOylation signal could be observed when the catalytic activity of NSMCE2 was not present.

Additionally, we generated LAZSUL knockout U2OS cells (**Supplementary Figure 6**), introduced HIS-SUMO2 and validated the proliferation marker protein Ki-67 as a target which SUMOylation was considerably reduced when LAZSUL was not present (**Figure 7D**).

In an alternative approach, we performed an shRNA-mediated knockdown of PIAS4 in parental, or HIS-SUMO2Q87R U2OS cells. Analysis by immunoblotting of the HIS-SUMOQ87R proteome validated strong PIAS4-S2-SATT targets with high SATT index value, which had not been previously described in the literature as the Histone-lysine N-methyltransferase EHMT1, or the transcription factor jun-B (**Figure 7E**).

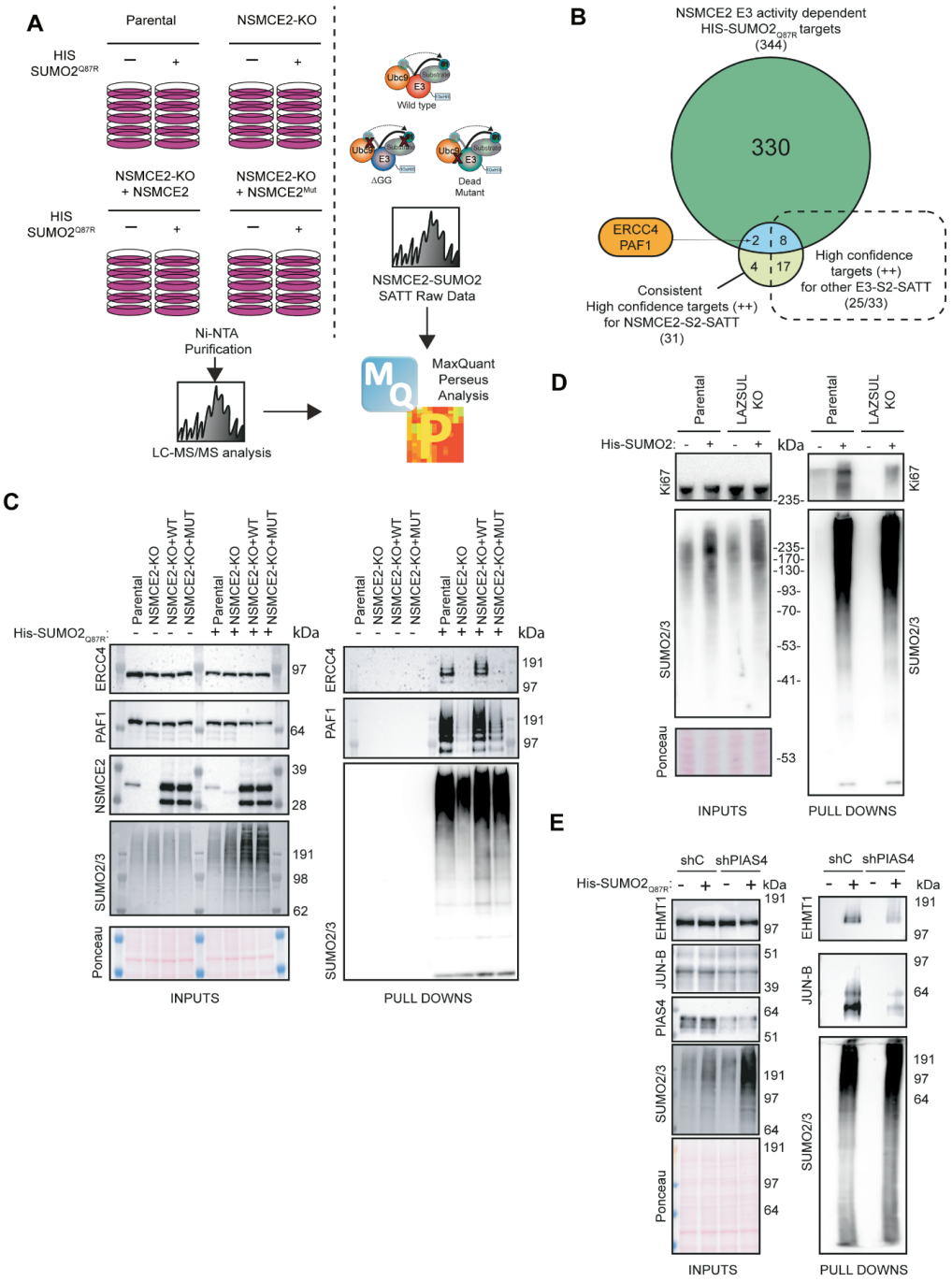


Figure 7. Orthogonal Validation. **(A)** Experimental setup. Parental and NSMCE2-KO U2OS cells rescued or not with either wild type NSMCE2 or a catalytic dead mutant and expressing or not HIS-SUMO2Q87R were cultured, incubated for 5h with MG132, lysed and analyzed by LC-MS/MS. Next, data was processed together with the NSMCE2-SUMO2-SATT data by MaxQuant and Perseus. **(B)** Venn Diagram depicting

overlap between NSMCE2-SUMO2-SATT substrates, HIS-SUMO2Q87R substrates affected by lack of NSMCE2 catalytic activity, and NSMCE2-SUMO2-SATT substrates shared with other E3-SATTs. (C-E) Immunoblot analysis of HIS-SUMO2 substrates decreasing upon NSMCE2 catalytic activity **(C)**, LAZSUL knockout **(D)** or PIAS4 knockdown **(E)**.

DISCUSSION

This resource identifies 427 SUMO1 and 961 SUMO E3 targets in an E3-specific manner. Nevertheless, while the number of SUMOylation substrates that have been identified for SUMO1 and SUMO2 in U2OS cells is higher (7, 9, 10) **(Supplementary Datasets 1-2**), proteins that have not been previously described as SUMOylation substrates upon proteasome inhibition with MG132, were identified as E3 substrates for SUMO conjugation in this work. It is noteworthy, that the screens performed in this project only comprised a single condition of 5h of proteasome inhibition with MG132 in a single cell line, U2OS, for 8 different SUMO ligases.

The results obtained from this screen also corroborate previous observations on protein arrays regarding E3 preferences for SUMO1 or SUMO2/3(12). PIAS3 and PIAS4 had been proposed to have a big preference for SUMO2. Consistently, for PIAS3, enrichments are much higher in SUMO2 SATTs than in SUMO1 SATTs (**Figure 3B**), and for PIAS4, 93% of SUMO1 substrates are shared with SUMO2. PIAS1 was observed to have a more balanced preference, in line with our SATT screen. For PIAS2 and NSMCE2 there is no data available from protein arrays to make this comparison, but they show a more balanced preference towards SUMO1.

Additionally, substrates that are highly SUMOylated and shared by every E3, such as PML, RANGAP, RNF216 or ZBTB33 have a very low SATT index for every E3, which indicates that either are redundant substrates, can be modified by different E3s or require no E3 at all for SUMOylation (**Supplementary Datasets 1-2**). Moreover, some substrates with high SATTi values for specific E3s have been previously independently identified substrates in other laboratories for the relevant E3s. Namely, PARP1 (38), ZMYM2 (39), and TOP2a (27, 30) for PIAS4 and SUMO2/3. (**Figure 6, Supplementary Datasets 4,7**). SMC5 for NSMCE2 and SUMO1 (26) (Supplementary Datasets 3,6)

SUMO2/3 is not recycled at the proteasome

Whereas the classical model suggests that SUMO2/3 is deconjugated and recycled at the proteasome (40), we have observed that when overexpressing PIAS1, ZNF451 or LAZSUL, SUMO2/3 is depleted from the nucleus in an RNF4-dependent manner (**Figure 1**). This indicates that SUMO2/3 moieties attached to STUbL targets are also degraded. Whether the proteasome discriminates between mixed SUMO2/3-Ubiquitin chains on substrates and substrates that are co-modified independently by SUMO2/3 and Ubiquitin chains requires further investigation. We previously showed that the oncogene c-Myc, which SUMOylated form highly accumulates upon proteasome inhibition and is an RNF4 target (8, 15, 41), is SUMOylated and ubiquitinated on different residues in an RNF4-dependent manner. Based on these results, we favor the hypothesis that independent ubiquitination of the SUMOylated substrate is sufficient for the degradation of SUMO moieties attached to the substrate without the need for mixed chains. Nevertheless, we also found that mixed SUMO/ubiquitin polymers are efficiently stabilized by proteasome inhibitors, indicating that these mixed polymers also constitute an efficient proteasomal degradation signal (5).

In contrast to GFP-PIAS1, -ZNF451, -LAZSUL and, partially, -PIAS3, which overexpression caused nuclear SUMO2/3 depletion, GFP-PIAS4 overexpression was linked to a slight increase in nuclear SUMO2/3 levels (**Figure 1**). Interestingly, PIAS4 and NSMCE2 were the only E3s capable of assembling SUMO2/3-SUMO1 mixed chains (**Figure 4**). We therefore hypothesize that SUMO2/3-SUMO1 mixed chains are poor substrates for RNF4. Similarly, SUMO1-capped SUMO2/3 chains are poor substrates for RNF4 but efficient substrates for RNF111/Arkadia, another STUbL (42).

Perspective and implications

SUMOylation of proteins occur in response to many different types of cellular stresses, such as DNA damage or replication stress and heat shock among others (31, 43). The E3 that modifies a specific target may vary depending on the scenario, and SATTs should be screened for specific targets at specific conditions. One E3 can be specific for the SUMOylation of a protein in a certain context, and another in a different context. In this regard, PIAS4 and NSCME2 shared many SUMOylation substrates from the DNA damage response. We speculate that, while NSMCE2 SUMOylates these substrates in response to DNA damage in the context of DNA replication as part of the SMC5/6 complex (44), PIAS4 SUMOylates these substrates in a replication-independent manner.

Moreover, altering the expression levels of different components of the SUMO pathway affects the SUMOylation landscape (7, 11, 45-47). Although here we expressed the SATTs at close-to-endogenous or below levels, this might affect both the SUMOylation status of target proteins, E3 specificity, or SUMO type modification, as proposed in Figure 3A.

Previously, it was shown that both SUMO1 and SUMO2 were recruited to sites of DNA damage, with SUMO1 recruitment depending on PIAS4 (25). Accordingly, DNA damage repair pathways are significantly enriched among PIAS4 substrates for SUMO1 **(Supplementary Dataset 8**). However, we did not detect affinity of DNA Damage Response-related proteins for SUMO1 moieties (10). Which might indicate that SUMO1 moieties incorporated at DNA damage sites in a PIAS4-dependent manner correspond to hybrid SUMO1-SUMO2/3 chains. Our results indicate that different types of mixed SUMO polymers and mixed SUMO/ubiquitin polymers constitute differential signals.

Nonetheless, results from this screen might open new lines of investigation. Gene ontology analysis revealed that PIAS2 substrates are significantly enriched for serine import into mitochondrion (**Figure 3B**) as biological function, which deregulation has been very recently described as causative for Parkinson disease (48). Interestingly, PIAS2 has also been recently linked to Parkinsonism (49). This suggest a potential role for the identified PIAS2 substrates in the development of this neurological disorder. Future mining of the resource presented here could improve our understanding of the biological functions of the different SUMO E3 ligases.

MATERIALS AND METHODS

Antibodies

Antibodies are listed in Supplementary Table 1 with working dilutions. For the generation of the Rabbit anti-LAZSUL antibody, recombinant ZNF451-3 (LAZSUL) was produced as in (23) and supplied to Cambridge Research Biochemicals (United Kingdom) for antibody production. Two rabbits were injected and the antibody affinity-purified for LAZSUL.

Generation of SATT toolbox

For the generation of the SATTs plasmids AgeI-10HIS-SUMO1-XmaI, AgeI-10HIS-SUMO1ΔGG-XmaI, AgeI-10HIS-SUMO2Q87R-XmaI, AgeI-10HIS-SUMO1ΔGGQ87R-XmaI restriction fragments from PCR products amplified with primers FW-AgeI-10HIS-SUMO1, FW-AgeI-10HIS-SUMO2, RV-XmaI-SUMO1, RV-XmaI-SUMO1noGlyGly, RV-XmaI-SUMO2-Q87R and RV-XmaI-SUMO2-Q87R-noGlyGly, were cloned into AgeI-SpeI sites of pCW57.1-nonStop plasmid (16). The N-terminal 10xHIS tag was cloned as previously done for the H-TULIP2 plasmids (15). Primer sequences are listed in Table T2.

Generation of lentiviral plasmids

SUMO1 and SUMO2$_{Q87R}$ SATT plasmids containing a SUMO E3 ligase of interest were generated using Gateway® cloning LR reaction (Thermo Fisher Scientific). LR reactions were performed using a donor plasmid containing an E3 enzyme cDNA without stop codon and a SATT plasmid as destination vector. We used different donor plasmids: pENTR223-PIAS1, pDNOR221-PIAS2, pENTR223-PIAS3, pDNOR221-PIAS4, pENTR223- NSMCE2, and pENTR233-ZMIZ2 were obtained from DNASU repository with the following IDs: HsCD00505402, HsCD00042133, HsCD00514170, HsCD00041383, HsCD00287670 and HsCD00505806 respectively. pDNOR207-ZNF451-1 (isoform 1), pDNOR207-ZNF451-3 (LAZSUL) and pDNOR207-GPF-LAZSUL were generated using the Gateway® cloning BP reaction (Thermo Fisher Scientific) upon cDNA amplification using BP-tailed primers and pDNOR207 as donor vector. Catalytic dead mutants of each SUMO E3 ligase were generated by site-directed mutagenesis on donor plasmids and subsequent LR reaction into SATT plasmids. Primer sequences are listed in Table T2.

Stable-inducible GFP-LAZSUL construct was generated by LR Gateway cloning between pDNOR207-GFP-ZNF451-3 and pCW57.1 (Addgene#41393). NSMCE2 rescued constructs were generated by LR between either pENTR223- NSMCE2-WT or pENTR223- NSMCE2C185S-H187A and pLX303 (Addgene#25897).

Transfection of GFP-E3 constructs

For the transient transfection experiments of GFP-tagged SUMO E3 ligases in Figure 1 and Supplementary Figure 1, 1x10^4 U2OS cells were transferred to 6-well plates containing 18 mm coverslips and left to attach overnight. The next day, 300 µL of transfection mixture consisting of 150 mM NaCl containing 1 µg of plasmid DNA and 6 µg of Polyethyleneimine were added to the cells. 24h after transfection, culture medium was replaced for fresh medium. Cells were fixed with 1% paraformaldehyde 48h after transfection for immunofluorescence analysis.

Generation of LAZSUL-KO cell lines

Three different gRNAs targeting LAZSUL exon 4 start and LAP2α domain (**table T2**) were cloned, independently, into a Cas9-GFP-containing pX458 backbone plasmid (AddGene #48138). U2OS cells were seeded in 15 cm diameter plates at 10% confluency and left to attach overnight. The next day, 2 mL of transfection mixture consisting of 150 mM NaCl containing 5 µg DNA of each plasmid and 100 µg of Polyethylenimine were added to the cells. Transfection medium was replaced by fresh culture medium after 24 h. 48 h after transfection, cells were GFP FACS-sorted by FACSAria III (BD Biosciences) and seeded for monoclonal expansion.

Selection of positive clones was performed by genomic PCR and immunoblotting (Supplementary Figure 6). Two primers targeting the LAP2α coding sequence were employed for clones validation by genomic PCR (**table T2**). 6 out of 30 clones were found positive with a 972 bp deletion in the LAP2α coding sequence (**Supplementary Figure 6A**). Parental and clone 27 LAZSUL-KO cell lines were subjected to immunoblotting against an in-house produced LAZSUL antibody (**table T1**) (**Supplementary Figure 6B**).

siRNA transfection for RNF4 knockdown

siRNA-mediated knockdowns were performed as previously described (16). DharmaFect 1 Transfection Reagent (GE Lifesciences) was used, according to the manufacturer's instructions, using on-target plus RNF4 siRNAs (J-006557-08) and the non-targeted control siGENOME non-Targeting siRNA #1 (GE Lifesciences).

Lentivirus production

293T HEK cells were seeded at 30% confluency in a T175 flask containing 16 mL of DMEM + 10% FBS. The following day, a 2 mL transfection mixture containing lentiviral packaging plasmids 7.5 µg pMD2.G (#12259, Addgene), 11.4 µg pMDLg-RRE (#12251, Addgene), 5.4 µg pRSV-REV (#12253, Addgene) and 13.7 µg SATT plasmid with 114 µL of 1 mg/mL Polyethyleneimine (PEI) was prepared in 150 mM NaCl. After vortexing, solutions were incubated 10 min at room temperature before adding to the HEK cells. The day after transfection, culture medium was changed by fresh DMEM/FBS/Pen/Strep. three days after transfection, lentiviral suspension was harvested by filtering through a 0.45 µm syringe filter (PN4184, Pall Corporation). Lentiviral particle concentration was determined using the HIV Type 1 p24 antigen ELISA Kit (ZeptoMetrix Corporation).

Generation of SATT, GFP-LAZSUL and NSMCE2 rescued cell lines

U2OS cells were seeded in a 15 cm diameter plates at 10% confluency with DMEM + 10% FBS. The next day, cell culture medium was replaced with either lentiviral SATT, GFP-LAZSUL or NSMCE2 rescued constructs containing medium and polybrene 8 µg/mL. Lentiviral containing medium was replaced by fresh DMEM/FBS/Pen/Strep culture medium after 24h. After two days in fresh medium, 3 µg/mL puromycin was added to the medium for selection of SATT positive clones.

shPIAS4 transduction for PIAS4 knockdown

Transductions were performed in DMEM containing 8 µg/mL polybrene either with PIAS4 or a control non-targeting shRNA. Cells were infected with a multiplicity of infection (MOI) of 3 with third generation lentiviruses encoding shRNA. DMEM containing virus was replaced after 24 hours of infection. Cells were harvested and lysed 3 days post infection. Plasmids used for shRNA mediated knockdown were derived from the MISSION ® shRNA library (Sigma) with number TRCN0000004115 (shPIAS4), and SHC-002 (shControl).

Cell culture

293T HEK, U2OS and RPE1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 % Fetal Bovine Serum (FBS) and 100 U/mL penicillin 100 µg/mL

streptomycin at 37°C and 5% CO2 unless specified. Cells were regularly tested for mycoplasma contamination.

Anti-SUMO1 and SUMO2/3 immunostaining

Cells were grown on 9 mm coverslips and fixed with 1% paraformaldehyde (PFA), 0.3% Triton X-100 for 20 min at RT. A second round was performed with 1% PFA, 0.3% Triton X-100 and 0.5% methanol for 20 minutes at Room Temperature (RT). Next, cells were washed three times with PBS and then blocked for 30 minutes with 0.5% blocking reagent (Roche) in 0.1 M Tris, pH 7.5 and 0.15 M NaCl (TNB). Cell were then incubated with either anti-SUMO1 or anti-SUMO2/3 antibody in TNB for one hour. Coverslips were washed five times with PBS and incubated with the secondary antibody (Goat anti-mouse coupled to Alexa-594) in TNB for one hour. Next, coverslips were washed five times with PBS- and mounted onto a microscopy slide using citifluor/DAPI solution (500 ng/mL). Immunofluorescence image analysis was performed using the FiJi – ImageJ distribution(50).

Generation of His-SUMO1 and His-SUMO2 Q87R U2OS cell lines

U2OS cells were infected using a bicistronic lentivirus encoding either a 10xHis-SUMO1-IRES-GFP or a 10xHis-SUMO2Q87R-IRES-GFP separated by an IRES, which was modified from previously described 10xHis-SUMO2-WT (51). Following infection, U2OS cells were sorted in an FACSAria III (BD Biosciences) for low GFP levels.

Purification of 10xHis-SUMO1, 10xHis-SUMO2 Q87R and SATT conjugates

Following the TULIP2 methodology (15), five 15 cm diameter plates of U2OS control cells or expressing either 10xHis-SUMO1, 10xHis-SUMO2Q87R or a particular SATT, were grown up to 60% to 80% confluence. Expression of SATTs constructs was induced with 1 µg/mL doxycycline once 60-80% confluency was reached. 24 h after doxycycline induction, cells were treated with proteasome inhibitor MG132 (Sigma Aldrich) at 10 µM for 5 h. After proteasome inhibition, cells were washed twice with ice-cold PBS and scraped. Cells were spun down and collected in 2 mL ice-cold PBS, 100 µL of sample was taken as input and lysed in 200 µL SNTBS buffer (2% SDS, 1% NP-40, 50mM TRIS pH 7.5, 150 mM NaCl). After additional centrifugation, cells were lysed in 10 mL Guanidinium buffer (6M guanidine-HCl, 0.1M Sodium Phosphate, 10mM TRIS, pH 7.8) and snap frozen in liquid nitrogen.

After thawing, lysates were homogenized at room temperature by sonication at 80% amplitude during 5 s using a tip sonicator (Q125 Sonicator, QSonica, Newtown, USA). Sonication was performed twice. Subsequently, protein concentration was determined by BiCinchoninic Acid (BCA) Protein Assay Reagent (Thermo Scientific). Total protein in cell lysates was equalized accordingly. After equalization, cell lysates were supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. 100 µL of dry nickel-nitrilotriacetic acid-agarose (Ni-NTA) beads (QIAGEN), were equilibrated with Guanidinium buffer supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. Equilibrated Ni-NTA beads were added to the cell lysates and incubated overnight at 4°C under rotation.

After lysate-bead incubation, Ni-NTA beads were spun down and transferred with Wash Buffer 1 (6 M Guanidine-HCl, 0.1M Sodium Phosphate, 10 mM Tris, 10 mM Imidazole, 5 mM β-mercaptoethanol, 0.2 % Triton X-100, pH 7.8) to an Eppendorf LoBind tube (Eppendorf). After

mixing and spinning down again, the beads were moved to a new LoBind tube with Wash buffer 2 (8 M Urea, 0.1M Sodium Phosphate, 10 mM Tris, 10 mM imidazole, 5mM β-mercaptoethanol, pH 8). This procedure was repeated with Wash buffer 3 (8 M urea, 0.1M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 6.3). Ultimately, beads were washed twice with Wash buffer 4 (8 M urea, 0.1M Sodium Phosphate, 10 mM TRIS, 5 mM β-mercaptoethanol, pH 6.3). When washing with wash buffer 3 and 4, beads were allowed to equilibrate with the buffer for 15 min under rotation.

For the analysis by immunoblotting the beads were boiled at 99 oC for 10 min with 2x LDS loading buffer.

Trypsin digestion of SATT-purified conjugates

After the final wash with Wash buffer 4, Ni-NTA beads were resuspended in 7 M urea, 0.1 M NaH2PO4/Na2HPO4, 0.01 M Tris/HCl, pH 7 and digested with 500 ng recombinant Lys-C (Promega) at RT while shaking at 1,400 rpm. After 5h with Lys-C, urea buffer was diluted to <2M by adding 50 mM ABC. A second digestion was performed o/n at 37°C while shaking at 1,400 rpm using 500 ng of sequencing grade modified trypsin (Promega). Trypsin digested peptides were separated from Ni-NTA beads by filtering through a 0.45 µm filter Ultrafree-MC-HV spin column (Merck-Millipore).

Mass Spectrometry sample preparation

Digested peptides were acidified by adding 2% TriFlourAcetic (TFA) acid. Subsequently, peptides were desalted and concentrated on triple-disc C18 Stage-tips as previously described (52). Stage-tips were in-house assembled using 200 µL micro pipet tips and C18 matrix. Stage-tips were activated by passing through 100 µL of methanol. Subsequently 100 µL of Buffer B (80% acetonitrile, 0.1% formic acid), 100 µL of Buffer A (0.1% formic acid), the acidified peptide sample, and two times 100 µL Buffer A were passed through the Stage-tip. Elution was performed in 50 µL of 32.5% acetonitrile, 0.1% formic acid.

Samples were vacuum dried using a SpeedVac RC10.10 (Jouan, France) and stored at -20°C. Prior to mass spectrometry analysis, samples were reconstituted in 10 µL 0.1% formic acid and transferred to autoload vials.

LC-MS/MS analysis

All the experiments were analyzed by on-line C18 nanoHPLC MS/MS with a system consisting of an Ultimate3000nano gradient HPLC system (Thermo, Bremen, Germany), and an Exploris480 mass spectrometer (Thermo, Bremen, Germany). Samples were injected onto a cartridge precolumn (300 µm × 5 mm, C18 PepMap, 5 µm, 100 A) with a flow of 10 µl/min for 3 minutes (Thermo, Bremen, Germany) and eluted via a homemade analytical nano-HPLC column (50 cm × 75 µm; Reprosil-Pur C18-AQ 1.9 µm, 120 A) (Dr. Maisch, Ammerbuch, Germany). The gradient was run from 2% to 38% solvent B (80% acetonitrile, 0.1% formic acid) in 120 min. The nano-HPLC column was drawn to a tip of ~10 µm and acted as the electrospray needle of the MS source. The temperature of the nano-HPLC column was set to 50°C (Sonation GmbH, Biberach, Germany). The mass spectrometer was operated in data-dependent MS/MS mode for a cycle time of 3 seconds, with a HCD collision energy at 28 V and recording of the MS2 spectrum in the orbitrap, with a quadrupole isolation width of 1.2 Da. In the master scan (MS1) the resolution was 120,000, the

scan range 350-1600, at a standard AGC target with maximum fill time of 50 ms. A lock mass correction on the background ion m/z=445.12 was used. Precursors were dynamically excluded after n=1 with an exclusion duration of 45 s, and with a precursor range of 10 ppm. Charge states 2-5 were included. For MS2 the scan range mode was set to automated, and the MS2 scan resolution was 30,000 at a normalized AGC target of 100% with a maximum fill time of 60 ms.

Mass Spectrometry data analysis

All raw data was analyzed using MaxQuant (version 1.6.14) as previously described (53). We performed two searches, first one analyzing the U2OS samples, HIS-SUMO1 and SUMO1 SATTs, and the other one analyzing U2OS, HIS-SUMO2Q87R and SUMO2Q87R SATTs. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens including canonical and isoform sequences (5th July 2021). Database searches were performed according to standard settings with the following modifications: Digestion with Trypsin/P was used, allowing 4 missed cleavages. Oxidation (M), Acetyl (Protein N-term), Phospho (S, T) and, in the SUMO2Q87R SATTs analysis, also QQTGG (K) (for SUMOylation sites) were allowed as variable modifications with a maximum number of 3. Carbamidomethyl (C) was disabled for SATTs analysis as a fixed modification. Label-Free Quantification was enabled, not allowing Fast LFQ. At least 2 peptides needed to be identified to calculate LFQ for a protein. All peptides were used for protein quantification.

Output from the analysis in MaxQuant was further processed in the Perseus computational platform version 1.6.14 (54) for statistical analysis. LFQ values were log2 transformed and contaminants, proteins identified by site and reverse peptides were excluded from the analysis. Next, samples were separated in two packages. The first package consisted of U2OS control samples and HIS-SUMO samples whereas the second package consisted of the SATT samples. The first package was analyzed to determine which proteins could be statistically considered a HIS-SUMO target. Proteins that were not identified in at least four replicates of at least one condition were removed. Then, missing values were imputed from a normal distribution with a width of 0.3 and a downshift of 2.5. Which resulted in a percentage of imputed valid values of 29.7% (Supplementary Dataset 3), 39,7 (Supplementary Dataset 4), 7.6% (Supplementary Dataset 6) and 11.8% (Supplementary Dataset 7). We performed t-test corrected with an FDR=0.05 and S0=0.1. Statistically enriched proteins were considered HIS-SUMO substrates (Supplementary Datasets 1-2). Next, we proceeded only considering proteins that were HIS-SUMO substrates and merge the table with package 2 samples (left sided), so only proteins which were HIS-SUMO substrates remained in the datasets. For the unbiased analysis, this step was omitted. Each SATT set was independently analyzed as performed with U2OS and HIS-SUMO samples. Finally, all the analyses were merged and exported and further processed in Microsoft Excel 365 for comprehensive data visualization and calculation of SATT indexes. For the heatmaps, the ΔGG and Mut samples were removed and Z-score was calculated for heatmap visualization.

Similarly, for the NSMCE2-KO orthogonal validation, LFQ values were log2 transformed and contaminants, proteins identified by site and reverse peptides were excluded from the analysis. Proteins that were not identified in at least four replicates of at least one condition were removed. Then, missing values were imputed from a normal distribution with a width of 0.3 and a downshift

of 2.5. Groups comparisons were performed by t-test with an FDR=0.05 and S0=0.1 and data exported into MS Excel 365 for comprehensive data browsing.

Gene Ontology Analysis

Gene Ontology analyses from the SATT substrates were performed using the PANTHER overrepresentation test (released 20221013) from the Gene Ontology Consortium (55). The Gene Ontology Database used was released on 2022-07-11 and the test was a Fisher test with Bonferroni correction. The whole human proteome was used as reference for comparison.

Electrophoresis and immunoblotting

Samples were separated on Novex 4-12% gradient gels (Thermo Fisher Scientific) using NuPAGE® MOPS SDS running buffer (50mM MOPS, 50mM Tris-base, 0.1% SDS, 1mM EDTA pH 7.7) and transferred onto Amersham Protran Premium 0.45 NC Nitrocellulose blotting membranes (GE Healthcare) using a Bolt Mini-Gel system (Thermo Fisher Scientific), which was used for both the gel electrophoresis and the protein transfer to the membrane according to vendor instructions. Membranes were stained with Ponceau-S (Sigma Aldrich) to determine total amount of protein loaded. Next, membranes were blocked with blocking solution (8% Elk milk, 0.1% Tween-20 in PBS) for 1h prior to primary antibody incubation. Chemiluminescence reaction was initiated with Western Bright Quantum Western blotting detection kit (Advansta – Isogen) and measured in a ChemiDocTM imaging system (BIO-RAD, Hercules, CA, USA).

In vitro SUMOylation assay

Recombinant SUMOylation machinery (56), including 300 ng E1, 0/100/200 ng E2, 4 µg SUMO1/SUMO2 and 0/300 ng E3 was incubated for 3h at 37ºC in in vitro SUMOylation buffer (50mM Tris pH 7.5, 2 mM ZnCl2 with Protease Inhibitors cocktail without EDTA (11836170001, Roche) and Energy Regeneration Solution 1:25 (B-10, Boston Biochem) in a total volume of 25 µL. SUMO chain formation was analyzed by immunoblotting.

Data Representation

Super plots were constructed using SuperPlotsOfData (57).

For the SATT polar plots, a web app for the display of multiple volcano plots side-by-side, named polarVolcaNoseR, was made with R/Shiny.

The code was written using R (https://www.r-project.org) and Rstudio (https://www.rstudio.com). To run the app, several freely available packages are required: shiny, ggplot2, magrittr, dplyr, ggrepel, htmlwidgets, ggiraph, glue, scales. The web app is freely accessible at: https://amsterdamstudygroup.shinyapps.io/PolaRVolcaNoseR/ and the code is available at github: https://github.com/ScienceParkStudyGroup/polarVolcaNoseR. In the default ‘polar’ representation, the volcano plots are plotted in a circle, where the radius depicts the -Log10(p-value) and the circumference reflects the positive Log2(Fold-change). Labels of proteins can be added by point-and-click and the data of individual dots are displayed when the cursor hovers over a datapoint. A customized, interactive plot can be exported as a HTML file.

ACKNOWLEDGEMENTS

Reagents and Funding statement

Authors would like to thank Dr. Geert Hamer for sharing Parental and NSMCE2-KO U2OS cell lines This work was supported by a Young Investigator Grant from the Dutch Cancer Society (KWF-KIG 11367/2017-2) the EMERGIA program from the Andalusian Regional Government, Spain (Junta de Andalucia – EMERGIA20_00276), the Spanish Ministry of Science and Innovation, the Spanish Research Agency, and the European Regional Development Fund (Proyecto PID2021-122361NA-I00 financiado por MCIN /AEI /10.13039/501100011033 / FEDER, UE) to RG-P. Work in the laboratory of A.C.O.V. has been supported by the European Research Council (ERC; grant 310913) and the Dutch Research Council (NWO; grant 724.016.003).

Competing interest

The authors declare that they have no competing interests.

Author contributions

RG-P and ACOV conceived the project. DSL, CvdM and RG-P performed the GFP-tagged E3 overexpression experiments. DS-L prepared all the resource reagents and samples. AdR, AO and PvV injected mass spectrometry samples. DS-L and RG-P analyzed mass spectrometry data. EG made recombinant ZMIZ2. RG-P and EN performed in vitro SUMOylation assays. DSL, NSJ and RG-P preformed orthogonal validation. EN raised the anti-LAZSUL antibody, JG programmed the polarVolcanoseR web App. AP supervised EN. RG-P supervised DS-L and CvdM. ACOV supervised EG. PvV supervised AdR and AO. RG-P lead the project and wrote the manuscript together with DS-L and input from all the authors.

Data availability statement

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (58) with the dataset identifiers PXD038326.

SUPLEMENTARY FIGURES AND DATASET LEGENDS

Supplementary Dataset 1. His-SUMO1 targets dataset. Spreadsheet including total identified SUMO1 targets, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 2. His-SUMO2 targets dataset. Spreadsheet including total identified SUMO2 targets, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 3. SUMO1 SATTs among HIS-SUMO1 targets. Spreadsheet including identified SUMO1 targets for the different SATTs among SUMO1 targets identified in Supplementary Dataset 1, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 4. SUMO2 SATTs among HIS-SUMO2 targets. Spreadsheet including identified SUMO2 targets for the different SATTs among SUMO2 targets identified in Supplementary Dataset 2, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 5. SUMO2 SATTs targets found in Hendriks et al, 2017. Spreadsheet including identified SUMO2 targets for the different SATTs among SUMO2 targets identified in Hendriks et al, 2017, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 6. Total SUMO1 SATTs. Spreadsheet including identified SUMO1 targets for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 7. Total SUMO2 SATTs. Spreadsheet including identified SUMO2 targets for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 8. Gene Ontology spreadsheet.

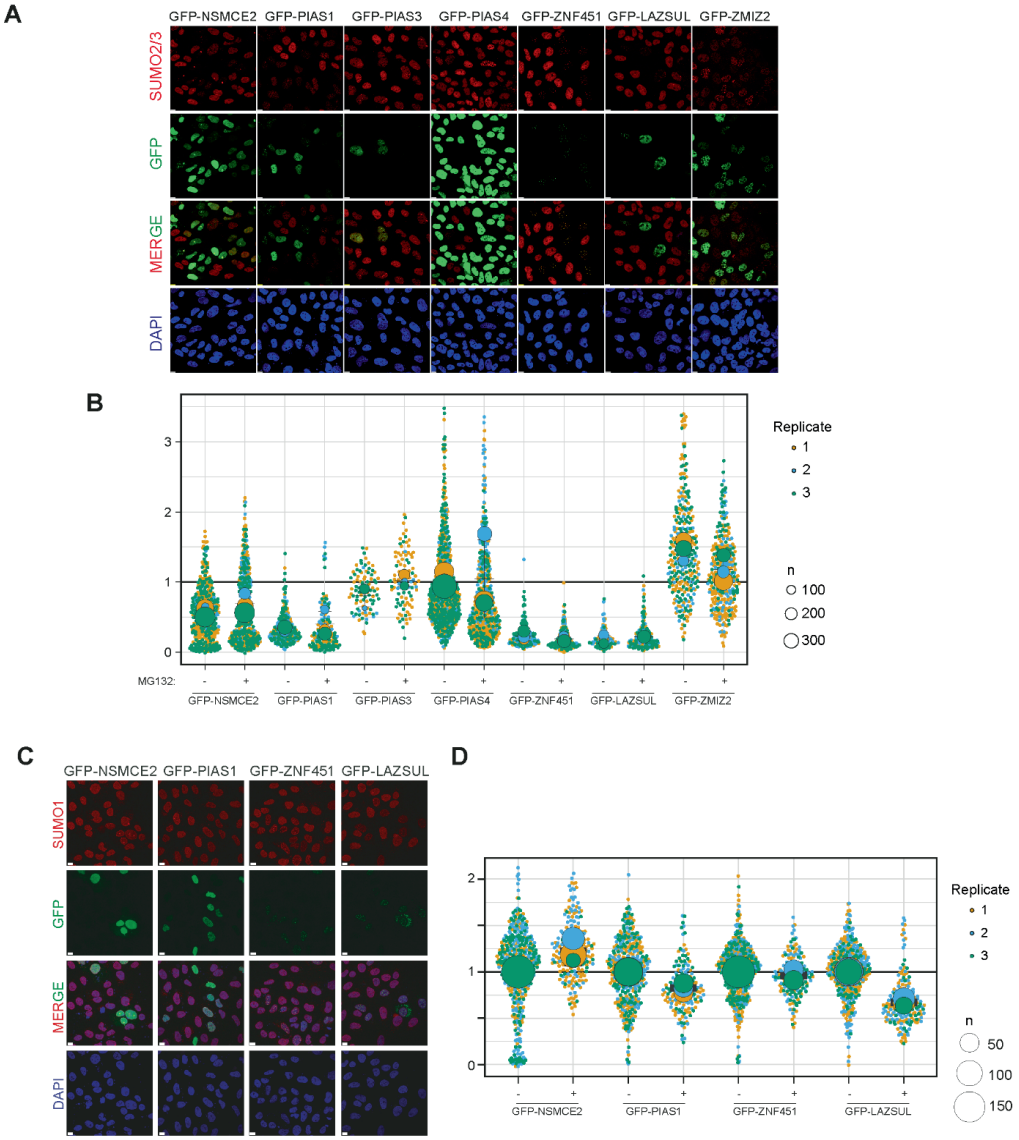
Supplementary Dataset 9. SATTs complement analysis for SUMO1 targets among HIS-SUMO1 targets. Spreadsheet including identified SUMO1 targets identified in Supplementary Dataset 1 after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 10. SATTs complement analysis for SUMO2 targets among HIS-SUMO2 targets. Spreadsheet including identified SUMO2 targets identified in Supplementary Dataset 2 after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

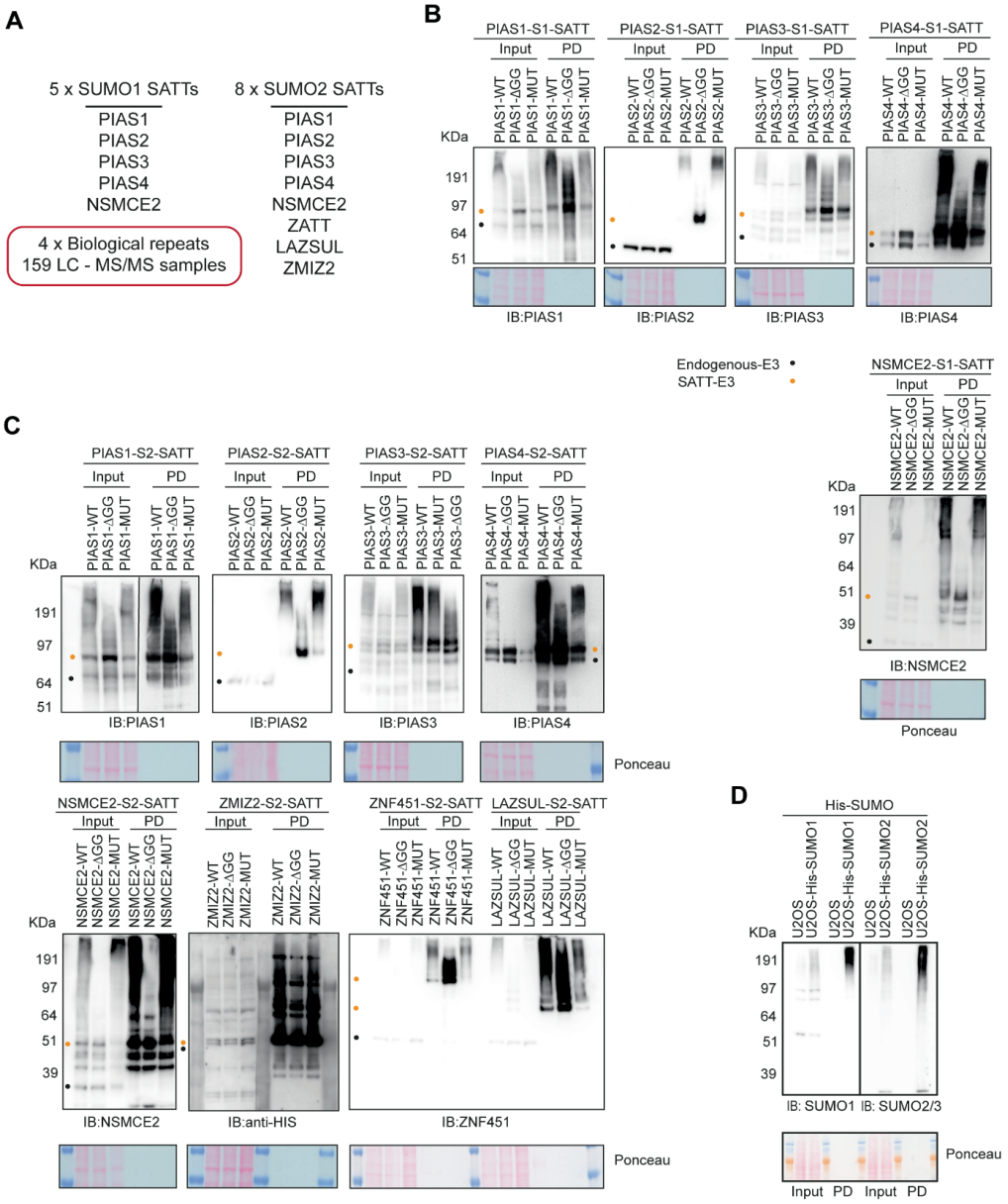
Supplementary Dataset 11. SATTs complement analysis for SUMO1 targets. Spreadsheet including identified SUMO1 targets after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 12. SATTs complement analysis for SUMO2 targets. Spreadsheet including identified SUMO2 targets after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

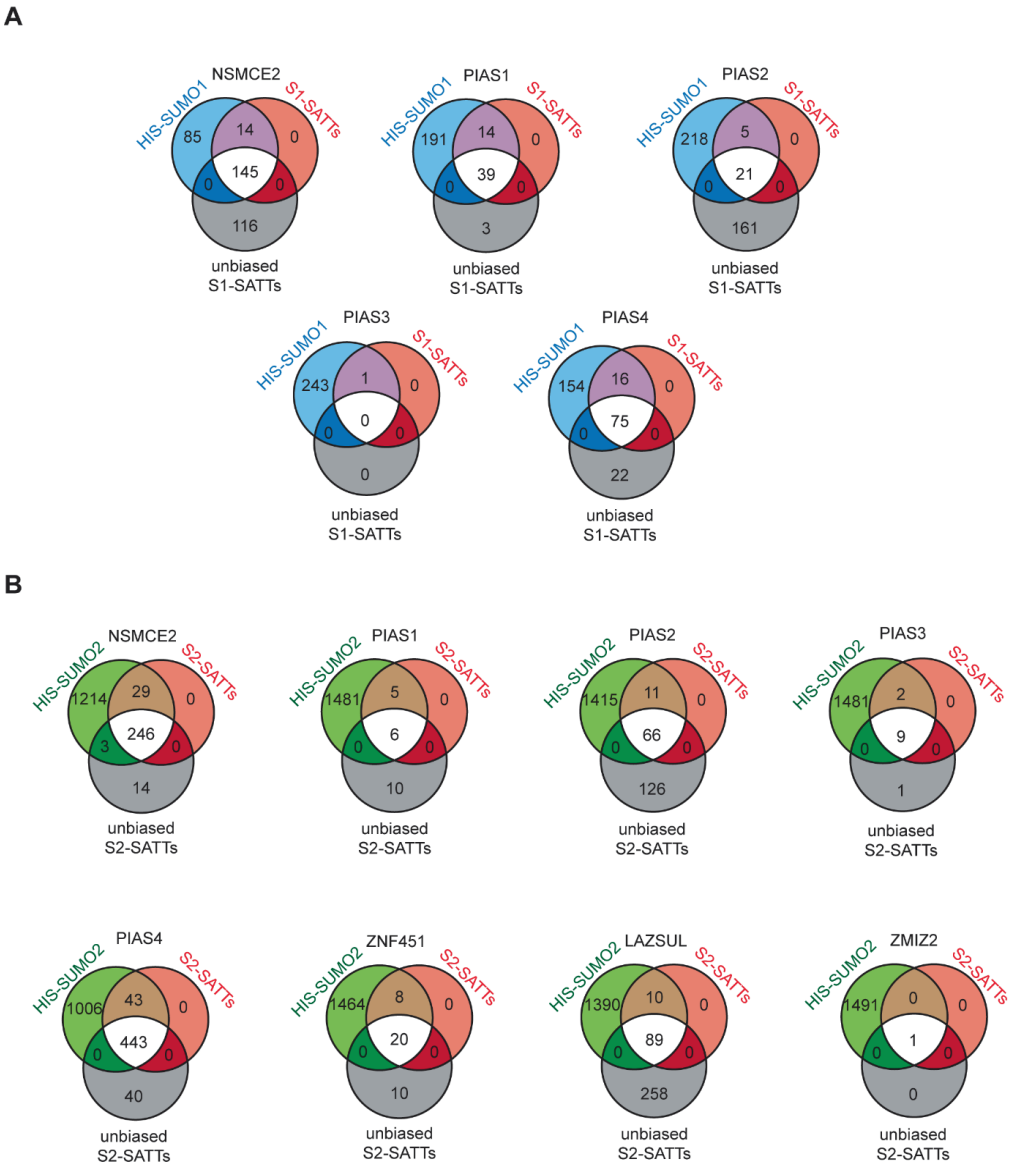
Supplementary Dataset 13. Very high confidence preferential targets. Spreadsheet including high confidence identified SUMO1 and SUMO2 targets for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.



Supplementary Figure 1. (A) Immunofluorescence images of U2OS cells transiently transfected with GFP-tagged constructs of E3 enzymes used in this study. Cells were stained for SUMO2/3 signal. **(B)** Superplot depicting the relative intensity compared to not-treated GFP negative cells of GFP positive cells for the different E3s used in this study after being treated or not with the proteasome inhibitor MG132. **(C)** Immunofluorescence images of U2OS cells transiently transfected with GFP-tagged constructs of indicated E3 enzymes. Cells were stained for SUMO1 signal. **(D)** Superplot depicting the relative SUMO1 intensity of GFP-positive or -negative cells. Intensity was normalized to the average SUMO1 nuclear intensity of GFP-negative cells for each replicate.



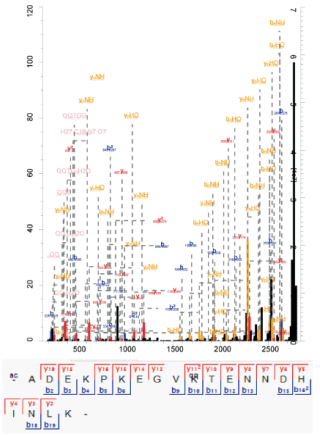
Supplementary Figure 2. (A) SUMO E3 SATTs investigated in this resource. **(B-D)** Immunoblotting analysis of the SUMO1 **(B)** or SUMO2Q87R **(C)** SATTs samples and HIS-SUMO samples **(D)**. Input and pulldown (PD) samples are included. Antibodies used are indicated, Ponceau staining is included as loading control.



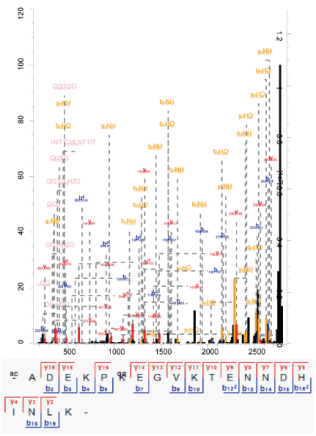
Supplementary Figure 3. Venn diagrams depicting the overlap between the identified **(A)** SUMO1 (blue) or **(B)** SUMO2Q87R (green) identified substrates, the SATTs targets only considering HIS-SUMO targets (red) and the SATTs targets identified in an unbiased manner (grey).

4

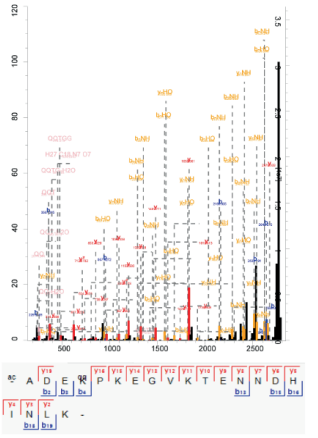
SUMO3 - K11 - SUMO2/3ylation
Score 267.21



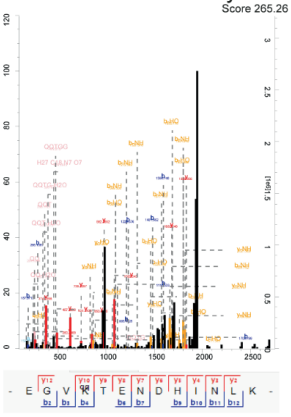
SUMO3 - K7 - SUMO2/3ylation
Score 279.1



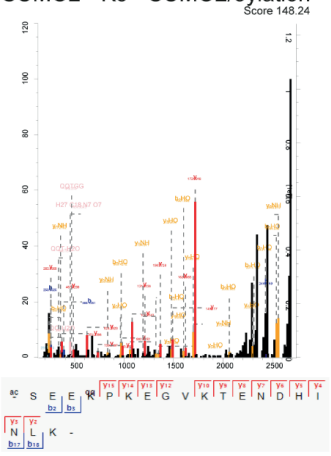
SUMO3 - K5 - SUMO2/3ylation
Score 216.32



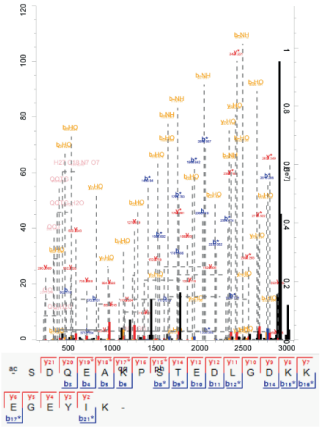
SUMO2 - K11 - SUMO2/3ylation
Score 265.26



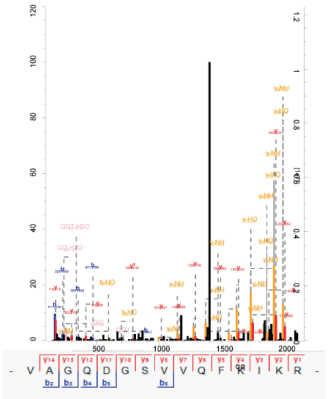
SUMO2 - K5 - SUMO2/3ylation
Score 148.24



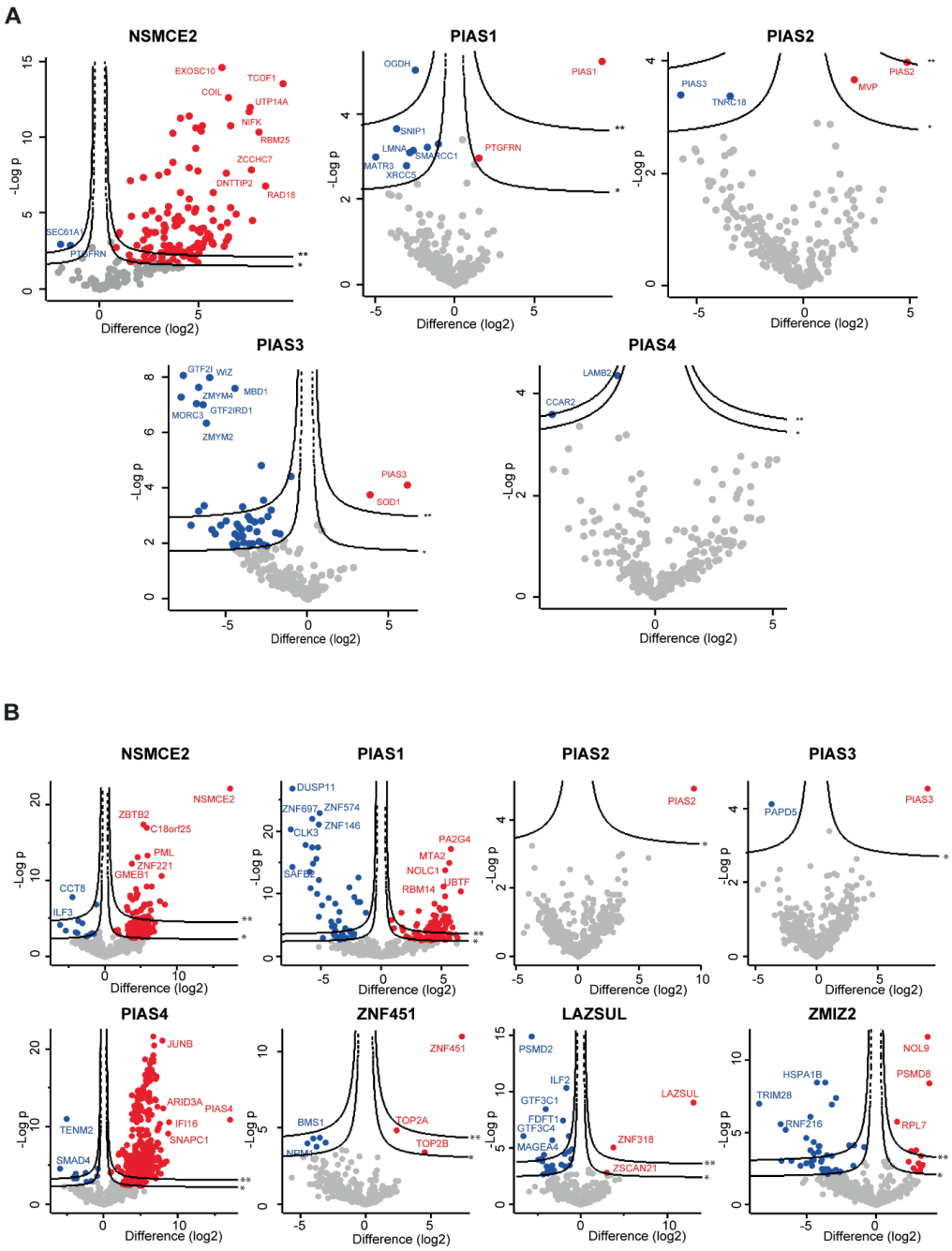
SUMO1 - K7 - SUMO2/3ylation
Score 325.24



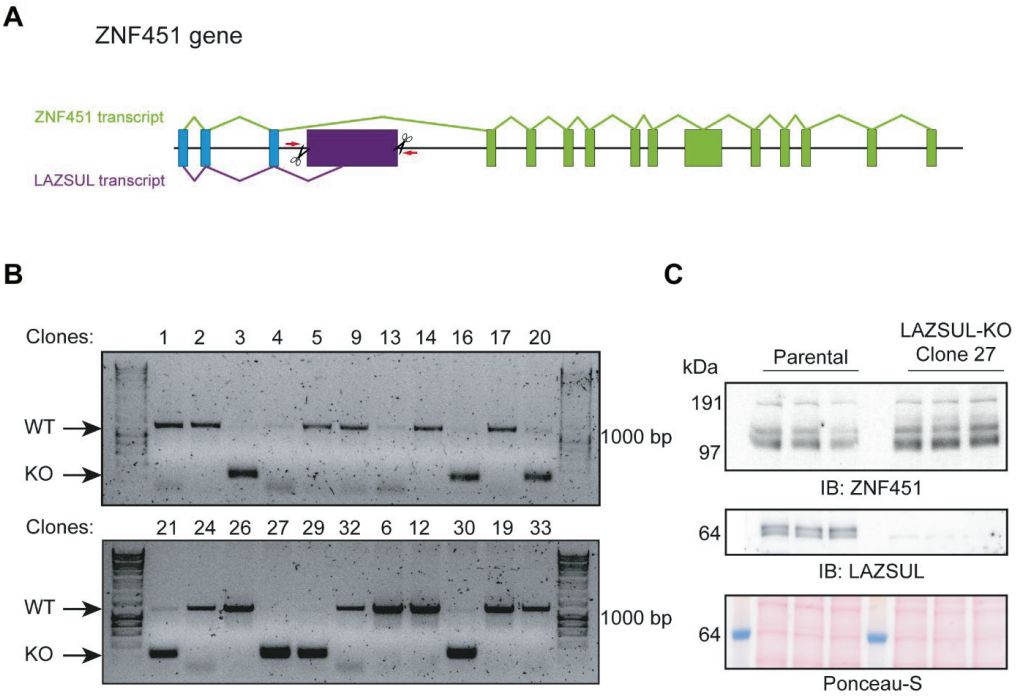
SUMO2/3 - K32/33 - SUMO2/3ylation
Score 192.54



Supplementary Figure 4. Best MS/MS spectra corresponding to the QQTGG sites on SUMOs identified in an unambiguous manner.



Supplementary Figure 5. Multiple volcano plots depicting statistical differences of different SUMO1 **(A)** or SUMO2 **(B)** SATTs targets, only considering HIS-SUMO substrates using the values from the rest of wild type SATTs as complement negative control. Cutoffs represent a Pearson of 100 and an FDR=0.05 (*) or 0.01 (**) and an S0=0.1.



Supplementary Figure 6. U2OS LAZSUL knockout cell line construction. **(A)** Diagram depicting the ZNF451 gene, in blue, common exons for ZNF451 and LAZSUL, in green, ZNF451 exons and in purple the LAP2α exon corresponding to LAZSUL. Scissors represent CRISPR cutting sites employed in the knockout strategy, arrows represent primer sites for clone testing. **(B)** Genomic PCR amplification of LAP2α loci different clones. Size fragments of 1227 bp correspond to LAP2α (WT), while 255 bp fragments correspond to mutated LAP2α (KO). **(C)** Immunoblotting of ZNF451 isoform 1 (ZNF451) and 3 (LAZSUL) of parental and clone 27 LAZSUL-KO cell lines. LAZSUL is detected in parental U2OS cell lines at 63 kDa and is absent in LAZSUL-KO clone 27.

Suplementary Table 1. Antibodies used in this study.

Antibody	Target	Dilution	Company
Primary antibodies			
Rabbit anti-PIAS1 (D33A7)	PIAS1	1:500	Cell Signaling Technology
Rabbit anti-PIAS2 (NBP2-19819)	PIAS2	1:1000	Novus Biologicals
Rabbit anti-PIAS3 (D5F9) 9042S	PIAS3	1:500	Cell Signaling Technology
Rabbit anti-PIAS4 (D2F12) 4392S	PIAS4	1:500	Cell Signaling Technology
Mouse anti-NSMCE2 (13627-1-AP)	NSMCE2	1:5000	SanBio
Rabbit anti-ZNF451	ZNF451	1:1000	Gift from A. Pichler
Mouse anti-His (H1029)	His	1:1000	Sigma-Aldrich
Rabbit anti-histone H1 (PA530055)	histone H1	1:1000	Invitrogen
Mouse anti-SUMO 8A2	SUMO2/3	WB: 1:1000 IF: 1:200	8A2, obtained from Developmental Studies Hybridoma Bank (DSHB), University of Iowa, *in-house* produced
Mouse anti-SUMO1 (4930S)	SUMO1	WB: 1:1000 IF: 1:200	Cell signaling Technology
Rabbit anti-GFP (NB600-308)	GFP	IF	Novus Biologicals
Rabbit anti-JUN-B (210)	JUN-B	1:1000	Santa Cruz
Mouse anti-c-JUN (G-4)	c-JUN	1:1000	Santa Cruz
Rabbit anti-EHMT1/GLP	EHMT1	1:1000	Millipore
Mouse anti-XPF (3F2/3)	ERCC4	1:1000	Santa Cruz
Rabbit anti-PAF1 (A300-172A)	PAF1	1:1000	Bethyl Laboratories
Rabbit anti-LAZSUL	LAZSUL	1:1000	Taylor made A. Pichler (this study)
Secundary antibodies			
HRP-conjugated Donkey anti-Rabbit (31458)	Anti-Rabbit	1:5000	Thermo Fisher Scientific
HPR-conjugated Goat anti-Mouse IgG (H+L) (31432)	Anti-Mouse	1:5000	Thermo Fisher Scientific
HRP-conjugated anti-Rabbit	Anti-Rabbit	1:1000	Jackson ImmunoResearch Laboratories

Supplementary table 2. Primers used in this study.

Primer Name	Sequence	Use
FW-AgeI-10HIS-SUMO1	ACCGGTATGGCTCACCATCACCACCATCATCATCATCATCATGGTGGATCCATGTCTGACCAGGAGGCAAAAC	Generation of SATTs plasmids toolbox SUMO1
FW-AgeI-10HIS-SUMO3	ACCGGTATGGCTCACCATCACCACCATCATCATCATCATCATGGTGGATCC	Generation of SATTs plasmids toolbox SUMO3
RV-XmaI-SUMO1	CCCGGGTCAACCCCCCGTTTGTTCCT	Generation of SATTs plasmids toolbox SUMO1
RV-XmaI-SUMO1noGlyGly	CCCGGGTCACGTTTGTTCCTGATAAACTTCAATCACATC	Generation of SATTs plasmids toolbox SUMO1DGG
RV-XmaI-SUMO3-Q87R	CCCGGGCTAACCTCCCGTCTGCTGC	Generation of SATTs plasmids toolbox SUMO3
RV-XmaI-SUMO3-Q87R-noGlyGly	CCCGGGCTACGTCTGCTGCCGGAACAC	Generation of SATTs plasmids toolbox SUMO3DGG
BP-FW-ZNF451	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGAGACCCGGGGTCG	Generation of pDNOR ZNF451- and pDNOR-ZNF451-3ZNF451
BP-RV-ZNF451-1	GGGGACCACTTTGTACAAGAAAGCTGGGTACATTTCCTCAAGACTTCTTCTAATAGCTTCTTCTA	Generation of pDNOR ZNF451ZNF451
BP-RV-ZNF451-3	GGGGACCACTTTGTACAAGAAAGCTGGGTAATTCTTGTACCCATGGCGTTCTGG	Generation of pDNOR ZNF451-3 and pDNOR-GFP-ZNF451-3
BP-FW-GFP	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGTGAGCAAGGGCGAGGA	Generation of pDNOR-GFP-ZNF451-3
FW-PIAS1-C351A	TTCCGTGTCGGGCCCTTACAgcTTCTCATCTACAATGTTTTGACGCAACTC	Generation of pDNOR PIAS1 Catalytic dead mutant
RV-PIAS1-C351A	GAGTTGCGTCAAAACATTGTAGATGAGAAgcTGTAAGGGCCCGACACGGAA	Generation of pDNOR PIAS1 Catalytic dead mutant
FW-PIAS2-C362A	CCATGCCGTGCAGTGACTgcTACACATCTGCAGTGTTTTGATGCT	Generation of pDNOR PIAS2 Catalytic dead mutant
RV-PIAS2-C362A	AGCATCAAAACACTGCAGATGTGTAgcAGTCACTGCACGGCATGG	Generation of pDNOR PIAS2 Catalytic dead mutant
FW-PIAS3-C343A	TGTCGTGCCCTCACCgcCGCCCACCTGCAGAGCTTC	Generation of pDNOR PIAS3 Catalytic dead mutant
RV-PIAS3-C343A	GAAGCTCTGCAGGTGGGCGgcGGTGAGGGCACGACA	Generation of pDNOR PIAS3 Catalytic dead mutant
FW-PIAS4-C342A	CCTGCCGGGCAGAGACCgcCGCCCACCTGCAGTGCTTC	Generation of pDNOR PIAS4 Catalytic dead mutant
RV-PIAS4-C342A	GAAGCACTGCAGGTGGGCGgcGGTCTCTGCCCGGCAGG	Generation of pDNOR PIAS4 Catalytic dead mutant
FW- NSMCE2-C185S-H187A	GGAAATGAAGAAGCCAGTGAAAAATAAAGTGtctGGCgccACCTATGAAGAGGACGCCATTG	Generation of pDNOR NSMCE2 Catalytic dead mutant
RV- NSMCE2-C185S-H187A	CAATGGCGTCCTCTTCATAGGTggcGCCagaCACTTTATTTTTCACTGGCTTCTTCATTTCC	Generation of pDNOR NSMCE2 Catalytic dead mutant
FW-ZNF451- SIM1	CGGATGAAAATGAAGACGACGCTCAGGCTGCCAGTGAAGGACCATTACGACCTG	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants

RV-ZNF451- SIM1	CAGGTCGTAATGGTCCTTCACTGGCAG CCTGAGCGTCGTCTTCATTTTCATCCG	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants
FW-ZNF451- SIM2	CCATTACGACCTGTTCTTGAATACGCTG ATGCGGCCAGCAGTGATGATGAAGAG CCTAGCACC	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants
RV-ZNF451- SIM2	GGTGCTAGGCTCTTCATCATCACTGCT GGCCGCATCAGCGTATTCAAGAACAG GTCGTAATGG	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants
LAZSUL gRNA #1-FW	caccgGAGGATGGCACCTTAGATTC	Generation of LAZSUL-KO cell line
LAZSUL gRNA #1-RV	aaacGAATCTAAGGTGCCATCCTCc	Generation of LAZSUL-KO cell line
LAZSUL gRNA #2-FW	caccgAGCCCGAGCTGCTAGCTACA	Generation of LAZSUL-KO cell line
LAZSUL gRNA #2-RV	aaacTGTAGCTAGCAGCTCGGGCTc	Generation of LAZSUL-KO cell line
LAZSUL gRNA #3-FW	caccgACGCTGATATTGAGTCATAT	Generation of LAZSUL-KO cell line
LAZSUL gRNA #3-RV	aaacATATGACTCAATATCAGCGTc	Generation of LAZSUL-KO cell line
ZNF451-3 KO seq - FW	ACTTCTCTGATTCCTTTTGCTGACC	Validation of LAZSUL-KO
ZNF451-3 KO seq - RV	CTGGGTCATGAACCAGATCTTCC	Validation of LAZSUL-KO

REFERENCES

1. G. Duan, D. Walther, The roles of post-translational modifications in the context of protein interaction networks. *PLoS Comput Biol* **11**, e1004049 (2015).
2. D. Salas-Lloret, R. Gonzalez-Prieto, Insights in Post-Translational Modifications: Ubiquitin and SUMO. *Int J Mol Sci* **23**, (2022).
3. F. Trulsson, A. C. O. Vertegaal, Site-specific proteomic strategies to identify ubiquitin and SUMO modifications: Challenges and opportunities. *Semin Cell Dev Biol*, (2021).
4. V. Akimov *et al.*, UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nature structural & molecular biology* **25**, 631-640 (2018).
5. F. Trulsson *et al.*, Deubiquitinating enzymes and the proteasome regulate preferential sets of ubiquitin substrates. *Nature communications* **13**, 2736 (2022).
6. I. A. Hendriks *et al.*, Site-specific characterization of endogenous SUMOylation across species and organs. *Nature communications* **9**, 2456 (2018).
7. I. A. Hendriks *et al.*, Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. *Nat Struct Mol Biol* **24**, 325-336 (2017).
8. I. A. Hendriks *et al.*, Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936 (2014).
9. F. Impens, L. Radoshevich, P. Cossart, D. Ribet, Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12432-12437 (2014).
10. R. Gonzalez-Prieto *et al.*, Global non-covalent SUMO interaction networks reveal SUMO-dependent stabilization of the non-homologous end joining complex. *Cell reports* **34**, 108691 (2021).
11. C. Li *et al.*, Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. *Nature communications* **11**, 834 (2020).
12. I. Uzoma *et al.*, Global Identification of Small Ubiquitin-related Modifier (SUMO) Substrates Reveals Crosstalk between SUMOylation and Phosphorylation Promotes Cell Migration. *Molecular & cellular proteomics : MCP* **17**, 871-888 (2018).
13. H. F. O'Connor *et al.*, Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO reports* **16**, 1699-1712 (2015).
14. R. González-Prieto, A. C. O. Vertegaal, in *SUMOylation and Ubiquitination: Current and Emerging Concepts,* V. G. Wilson, Ed. (Caister Academic Press, U.K., 2019), chap. 10, pp. 147-160.
15. D. Salas-Lloret, G. Agabitini, R. Gonzalez-Prieto, TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. *Front Chem* **7**, 802 (2019).
16. R. Kumar, R. Gonzalez-Prieto, Z. Xiao, M. Verlaan-de Vries, A. C. O. Vertegaal, The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. *Nat Commun* **8**, 1809 (2017).
17. R. Moreno-Ayala, D. Schnabel, E. Salas-Vidal, H. Lomeli, PIAS-like protein Zimp7 is required for the restriction of the zebrafish organizer and mesoderm development. *Dev Biol* **403**, 89-100 (2015).
18. Y. Peng, J. Lee, C. Zhu, Z. Sun, A novel role for protein inhibitor of activated STAT (PIAS) proteins in modulating the activity of Zimp7, a novel PIAS-like protein, in androgen receptor-mediated transcription. *The Journal of biological chemistry* **285**, 11465-11475 (2010).
19. J. Beliakoff, Z. Sun, Zimp7 and Zimp10, two novel PIAS-like proteins, function as androgen receptor coregulators. *Nucl Recept Signal* **4**, e017 (2006).
20. H. Lomeli, ZMIZ proteins: partners in transcriptional regulation and risk factors for human disease. *J Mol Med (Berl)*, (2022).
21. M. H. Tatham *et al.*, RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat Cell Biol* **10**, 538-546 (2008).
22. L. Cappadocia, A. Pichler, C. D. Lima, Structural basis for catalytic activation by the human ZNF451 SUMO E3 ligase. *Nature structural & molecular biology*, (2015).
23. N. Eisenhardt *et al.*, A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. *Nature structural & molecular biology*, (2015).
24. L. Wang *et al.*, SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. *EMBO Rep* **15**, 878-885 (2014).
25. Y. Galanty *et al.*, Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939 (2009).

26. N. Varejao *et al.*, DNA activates the Nse2/Mms21 SUMO E3 ligase in the Smc5/6 complex. *The EMBO journal* **37**, (2018).
27. M. Agostinho *et al.*, Conjugation of human topoisomerase 2 alpha with small ubiquitin-like modifiers 2/3 in response to topoisomerase inhibitors: cell cycle stage and chromosome domain specificity. *Cancer research* **68**, 2409-2418 (2008).
28. M. J. Schellenberg *et al.*, ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. *Science* **357**, 1412-1416 (2017).
29. J. Liang *et al.*, SUMO E3 ligase Mms21 prevents spontaneous DNA damage induced genome rearrangements. *PLoS Genet* **14**, e1007250 (2018).
30. Y. Sun *et al.*, A conserved SUMO pathway repairs topoisomerase DNA-protein cross-links by engaging ubiquitin-mediated proteasomal degradation. *Sci Adv* **6**, (2020).
31. P. Sarangi, X. Zhao, SUMO-mediated regulation of DNA damage repair and responses. *Trends in biochemical sciences* **40**, 233-242 (2015).
32. S. Dabir, A. Kluge, A. Dowlati, The association and nuclear translocation of the PIAS3-STAT3 complex is ligand and time dependent. *Mol Cancer Res* **7**, 1854-1860 (2009).
33. M. Sundvall *et al.*, Protein inhibitor of activated STAT3 (PIAS3) protein promotes SUMOylation and nuclear sequestration of the intracellular domain of ErbB4 protein. *The Journal of biological chemistry* **287**, 23216-23226 (2012).
34. J. H. Man *et al.*, PIAS3 induction of PRB sumoylation represses PRB transactivation by destabilizing its retention in the nucleus. *Nucleic acids research* **34**, 5552-5566 (2006).
35. M. H. Tatham *et al.*, Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *The Journal of biological chemistry* **276**, 35368-35374 (2001).
36. D. E. Verver *et al.*, Non-SMC Element 2 (NSMCE2) of the SMC5/6 Complex Helps to Resolve Topological Stress. *Int J Mol Sci* **17**, (2016).
37. A. Jacome *et al.*, NSMCE2 suppresses cancer and aging in mice independently of its SUMO ligase activity. *The EMBO journal* **34**, 2604-2619 (2015).
38. D. B. Krastev *et al.*, The ubiquitin-dependent ATPase p97 removes cytotoxic trapped PARP1 from chromatin. *Nature cell biology* **24**, 62-73 (2022).
39. D. Lee *et al.*, ZMYM2 restricts 53BP1 at DNA double-strand breaks to favor BRCA1 loading and homologous recombination. *Nucleic acids research* **50**, 3922-3943 (2022).
40. J. Schimmel *et al.*, The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Molecular & cellular proteomics : MCP* **7**, 2107-2122 (2008).
41. R. Gonzalez-Prieto, S. A. Cuijpers, R. Kumar, I. A. Hendriks, A. C. Vertegaal, c-Myc is targeted to the proteasome for degradation in a SUMOylation-dependent manner, regulated by PIAS1, SENP7 and RNF4. *Cell Cycle* **14**, 1859-1872 (2015).
42. A. M. Sriramachandran *et al.*, Arkadia/RNF111 is a SUMO-targeted ubiquitin ligase with preference for substrates marked with SUMO1-capped SUMO2/3 chain. *Nature communications* **10**, 3678 (2019).
43. A. Seifert, P. Schofield, G. J. Barton, R. T. Hay, Proteotoxic stress reprograms the chromatin landscape of SUMO modification. *Sci Signal* **8**, rs7 (2015).
44. J. J. Palecek, SMC5/6: Multifunctional Player in Replication. *Genes (Basel)* **10**, (2018).
45. C. Li *et al.*, SUMO Proteomics Analyses Identify Protein Inhibitor of Activated STAT-Mediated Regulatory Networks Involved in Cell Cycle and Cell Proliferation. *Journal of proteome research* **22**, 812-825 (2023).
46. I. A. Hendriks, R. C. D'Souza, J. G. Chang, M. Mann, A. C. Vertegaal, System-wide identification of wild-type SUMO-2 conjugation sites. *Nature communications* **6**, 7289 (2015).
47. J. Schimmel *et al.*, Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Molecular cell* **53**, 1053-1066 (2014).
48. P. Gonzalez-Rodriguez *et al.*, Disruption of mitochondrial complex I induces progressive parkinsonism. *Nature* **599**, 650-656 (2021).
49. J. Magalhaes *et al.*, PIAS2-mediated blockade of IFN-beta signaling: a basis for sporadic Parkinson disease dementia. *Mol Psychiatry* **26**, 6083-6099 (2021).
50. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682 (2012).

51. Z. Xiao *et al.*, System-wide Analysis of SUMOylation Dynamics in Response to Replication Stress Reveals Novel Small Ubiquitin-like Modified Target Proteins and Acceptor Lysines Relevant for Genome Stability. *Molecular & cellular proteomics : MCP* **14**, 1419-1434 (2015).
52. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906 (2007).
53. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016).
54. S. Tyanova *et al.*, The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016).
55. H. Mi *et al.*, PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic acids research* **49**, D394-D403 (2021).
56. K. Eifler *et al.*, SUMO targets the APC/C to regulate transition from metaphase to anaphase. *Nature communications* **9**, 1119 (2018).
57. J. Goedhart, SuperPlotsOfData-a web app for the transparent display and quantitative comparison of continuous data from different conditions. *Molecular biology of the cell* **32**, 470-474 (2021).
58. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic acids research* **47**, D442-D450 (2019).