SCIENTIFIC REPORTS

OPEN

Comprehensive list of SUMO targets in *Caenorhabditis elegans* and its implication for evolutionary conservation of SUMO signaling

Krzysztof Drabikowski[1,2], Jacqueline Ferralli[2], Michal Kistowski[1], Jacek Oledzki[1], Michal Dadlez[1,3] & Ruth Chiquet-Ehrismann[2,4]

Post-translational modification by small ubiquitin-related modifier (SUMO) is a key regulator of cell physiology, modulating protein-protein and protein-DNA interactions. Recently, SUMO modifications were postulated to be involved in response to various stress stimuli. We aimed to identify the near complete set of proteins modified by SUMO and the dynamics of the modification in stress conditions in the higher eukaryote, *Caenorhabditis elegans*. We identified 874 proteins modified by SUMO in the worm. We have analyzed the SUMO modification in stress conditions including heat shock, DNA damage, arsenite induced cellular stress, ER and osmotic stress. In all these conditions the global levels of SUMOylation was significantly increased. These results show the evolutionary conservation of SUMO modifications in reaction to stress. Our analysis showed that SUMO targets are highly conserved throughout species. By comparing the SUMO targets among species, we approximated the total number of proteins modified in a given proteome to be at least 15–20%. We developed a web server designed for convenient prediction of potential SUMO modification based on experimental evidences in other species.

Small ubiquitin-related modifier (SUMO) proteins are covalently and reversibly coupled to intracellular protein targets, modulating protein-protein and protein-DNA interactions[1]. The post-translational protein modification by SUMO is achieved by similar mechanism as ubiquitination by a set of homologous but distinct enzymes[2]. The SUMO activation enzyme E1 (Aos1/Uba2 dimer) forms, in an ATP dependent manner, a thioester bond between its catalytical cysteine and the C-terminal carboxyl group of SUMO. Next, SUMO is transferred to the cysteine of the SUMO E2 conjugating enzyme, Ubc9. Finally, SUMO is transferred to the ε-group of lysine of the target protein, forming an isopeptide bond. This step is often achieved by the help of E3 SUMO ligase such as PIAS and RanBP2[3]. The SUMO moiety is removed from the target proteins by SUMO isopeptidases from the SENP family[4]. Modification by SUMO regulates various aspects of cell physiology including cell cycle, DNA replication and repair, transcription, epigenetic gene silencing, intracellular trafficking and cell signaling[5]. Despite the vast implications of SUMO in various cellular processes, only few enzymes mediating SUMOylation are known. In comparison, the ubiquitin conjugation system involves almost thousand proteins. Recent work suggested that SUMOylation specificity comes from spatiotemporal determinants and thus, such a big number of proteins, as in case of ubiquitination, is not necessary[6].

The variety and number of cellular processes regulated by the post-translational modification by SUMO is comparable to processes regulated by phosphorylation and ubiquitination. Several SUMO targets belong to pathways involved in cancer and neurodegenerative diseases[7]. For example, increased levels of UBC-9, the SUMO conjugating enzyme, result in tumor formation[8]. In several tumors the levels of the SUMO E3 ligase PIAS are elevated[9]. SUMO modification regulates the protein levels and activity of tumor suppressor proteins, for example p53, p63, p73[10] and MDM2[11]. Moreover, SUMO regulates the protein levels of huntingtin, amyloid precursor

[1]Institute of Biochemistry and Biophysics, Polish Academy of Science, Warszawa, Poland. [2]Friedrich Miescher Institute for Biomedical Research, Department of Mechanisms of Cancer, Maulbeerstrasse 66, CH-4058, Basel, Switzerland. [3]Institute of Genetics and Biotechnology, Biology Department, Warsaw University, Warszawa, Poland. [4]University of Basel, Faculty of Sciences, CH-4056, Basel, Switzerland. Correspondence and requests for materials should be addressed to K.D. (email: drabikowski@ibb.waw.pl)

protein, DJ-1, ataxin-1, tau and superoxide dismutase[12]. Thus, SUMO attachment to disease related proteins is discussed as a potential therapeutic target[13].

Post-translation modification by SUMO was shown to be influenced under stress conditions. SUMOylation increases in response to heat shock in human cells[14], *Drosophila* embryos[15] and *Arabidopsis*[16]. Ischemia, the cold shock, also increases global levels of SUMOylation in neurons[17]. Furthermore, SUMOylation is increased in the re-oxidation phase of oxygen/glucose deprivation, a model for ischemia[18]. In yeast SUMOylation increases in response to DNA damage[6]. Protein modification by SUMO conjugation machinery was shown to be directly regulated by the redox state of the cell[19,20]. Along the same lines, hydrogen peroxide treatment increased SUMOylation by SUMO2/3 but not via SUMO1[21]. On the other hand, SUMOylation decreased upon ER stress in *C. elegans*[22]. Furthermore, SUMO isopeptidases have been shown to be inactivated by heat[23].

Initially, research was focused on the functions of SUMO modification in the nucleus, for example chromatin remodeling, genome stability and RNA. Recently, involvement of SUMO in non-nuclear processes is emerging such as in protein folding, lipogenesis, autophagy[24], membrane transport and proteostasis[25]. Research in *C. elegans* revealed functions of SUMO in intermediate filament assembly[26], ER stress[22], chromosome assembly[27] and assembly of adherens junctions to actin cytoskeleton[28].

Recently, several whole proteome studies increased our knowledge about SUMO targets in human cells[29–33]. For recent compilation of SUMO targets in human cells please refer to a review by Henriks *et al.*[33]. However, SUMOylation is a very dynamic process which depends on the cell type, cell cycle stage and interactions of an organism with its environment. Thus, we investigated the targets of SUMO modification at the level of an entire organism, the nematode *Caenorhabditis elegans. C. elegans* is a complex higher eukaryote, but in contrast to vertebrates, contains only one SUMO gene. This gene is essential and needs to confer all functions of SUMO modifications in a higher organism. Our study identified a comprehensive list of SUMO targets. Interestingly, we show that those targets are remarkably evolutionary conserved. Thus, we propose a novel bioinformatics approach to predict SUMO modification based on experimental data of other species.

Results

SUMO is expressed throughout life of *C. elegans*.   In *C. elegans* only one gene is coding for SUMO, *smo-1*. Deletion of *smo-1* is lethal. To analyze the posttranslational modification by SUMO in *C. elegans*, we generated a worm strain carrying a transgene composed of a N-terminally fused 8His-GFP to the worm SUMO gene (*smo-1*), called herein SUMO-GFP (Fig. 1A). The SUMO-GFP transgene rescues the lethal *smo-1* gene deletion in the VC186 strain. The SUMO-GFP fusion was expressed in all cells, including germ cells where the transgenes are often silenced, and throughout all developmental stages (Fig. 1B). The SUMO-GFP protein is localized predominantly in the nucleus but the cytoplasmic expression is also detectable, especially in the gut cells (Fig. 2). In the nucleus the protein is localized into speckles (Fig. 2A, arrows) or label dividing chromosomes (Fig. 2A, star). In the cytoplasm of gut cells SUMO-GFP can be also detected in the apical membrane (arrow heads) and some vesicles (arrows) (Fig. 2D).

Modification by SUMO is increased under stress conditions.   We analyzed the levels of SUMO modification applying various stress conditions to *C. elegans*. We studied the stress response throughout time after applying short, non-lethal stress and followed it until the level of SUMOylation reverted to steady state levels (Fig. 3, Fig. S1). SUMOylation was dramatically increased in response to heat shock (Fig. 3A), arsenite induced cellular damage (Fig. 3B), and UV induced DNA damage (Fig. 3C). Short osmotic stress (10 min 500 mM NaCl) induced also transient increase in SUMOylated proteins, albeit to a lower extent (Fig. 3F). Oxidative stress induced in the mitochondria by applying paraquat did not induce global increase in SUMOylation (Fig. 3D). Endoplasmic reticulum (ER) stress induced by tunicamycin initially increased SUMOylation to a small extend, followed by decreased SUMOylation (Fig. 3E). The period of elevated SUMOylation upon transient stress condition varied depending on the stressor, from less than an hour in case of osmotic stress to over 6 hours in case of UV response. We confirmed the increase in SUMOylation in response to heat-shock in wild-type N2 worms using SUMO specific antibodies (Fig. S2).

Identification of SUMO targets in *C. elegans*.   Our stress assays in worms showed increased amount of SUMO-conjugated proteins. Thus, we reasoned that the analysis of SUMO-conjugated proteins under stress conditions may yield to yet undiscovered SUMO protein targets in worms. Using the strain RU86 expressing HIS-GFP tagged SUMO we set up the purification of proteins modified by SUMO using immobilized metal ion affinity chromatography (IMAC). Control protein purifications and identifications from the wild-type N2 worms revealed several proteins bound to the IMAC column in a non SUMO-GFP specific manner (Table S1). Indeed, *C. elegans* proteome encodes over 500 histidine rich proteins. Thus, we analyzed in parallel proteins purified from SUMO-GFP tagged transgenic worms and N2 wild-type worms (Table S2). We applied single step IMAC purification and compared the identified proteins with the wild-type control (Fig. S3A). Denaturing lysis conditions of boiling 1% SDS followed by SDS precipitation[34] and single step purification using IMAC followed by comparison with proteins isolated from wild-type worms resulted in more specific target identification than purification in native condition, first by IMAC, followed by antibody purification. Lysis in boiling 1% SDS should abolish all protein-protein interactions and thus remove all SUMO or SUMO modified interacting proteins from the purification. Example purification is presented in Fig. S3B. SUMO conjugates were identified by mass spectrometry and data were analyzed using MaxQuant (Fig. S3A). To consider a protein as being SUMOylated, the proteins had to be identified in at least 3 independent experiments, each protein identified with at least two peptides, FDR below 1% and, not present in any of the control purifications from wild-type N2 worms or present with raw intensity greater than 10 times than in control purifications. The danger of this strategy is that some of the genuine targets that were accidently purified from control worms are excluded but we expected to identify very few falls positives.
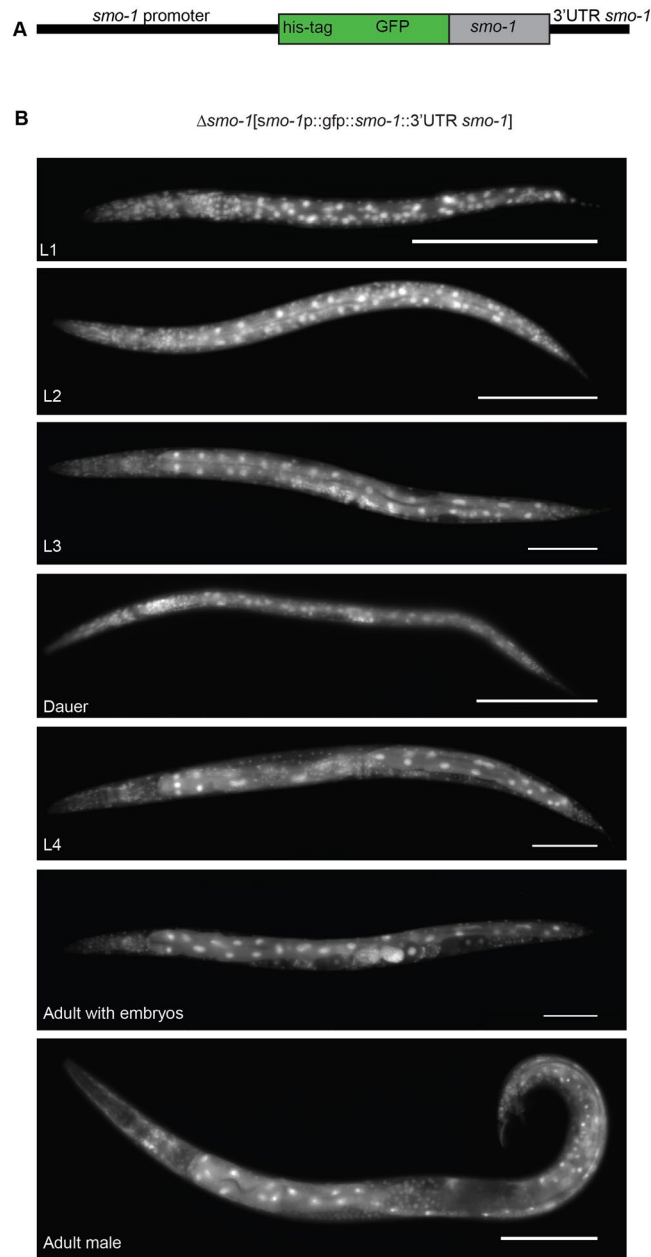
Figure 1. SUMO is expressed throughout development and in all cells in *C. elegans*. (**A**) Schema of translational fusion of HIS-tagged GFP with *smo-1* gene under the control of the endogenous *smo-1* promoter and including the endogenous 3′-UTR. (**B**) Fluorescent images of different developmental stages (L1, L2, L3, dauer, L4, adult hermaphrodite, adult male) of *C. elegans* stably expressing HIS-tagged GFP::*smo-1* construct in RU86. Anterior is to the left, ventral is to the bottom. Scale bar 100 μm.

We applied our purification and identification approach to analyze SUMO conjugates in *C. elegans* under stress conditions. We analyzed the response upon treatment of worms when global changes in SUMOylation were the strongest, namely heat shock, arsenite treatment and UV irradiation (see also Fig. 3). For each condition a non-treated control population was analyzed. To exclude proteins that non-specifically were purified, we analyzed wild-type worms under steady state and heat shock conditions. Each described condition was repeated 3 times independently. Altogether, we have analyzed 31 independent purifications from worms carrying the transgene for the tagged SUMO protein and 6 independent purifications from wild-type N2 strain (3 from non-treated and 3 from heat shock treated worms). Altogether, we have identified 874 proteins modified by SUMO in *C. elegans* in normal growth conditions and upon stress (Table S2). Unfortunately, variability of biological replicates prevented us from quantifying the differences between stress conditions in a statistically significant manner (Figs S4 and S5). Therefore, we analyzed only qualitative changes of presence/absence of an identified protein in a particular condition (Table S3). For a protein to be included in the difference list, it had to be identified in at least 2 out of 3 replicates in one condition and in none of the compared condition. Relatively low number of differentially
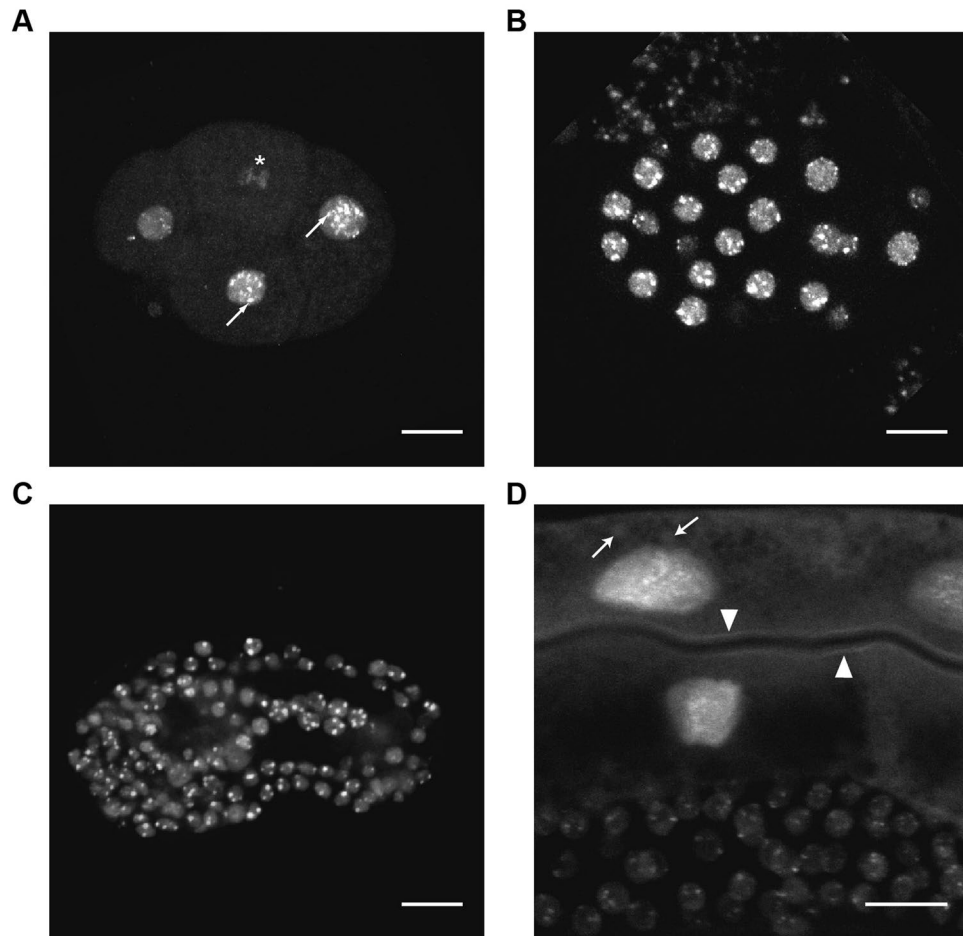
Figure 2. SUMO expression is enriched in the nucleus. (**A–D**) Stacked confocal images of subcellular localization of HIS-tagged GFP::*smo-1* in RU86. (**A**) 6-cell stage embryo. Arrow, nuclear speckles. Star, mitotic chromosomes (**B**) Embryo at gastrulation stage. (**C**) Comma stage embryo. (**D**) Zoom in into adult worm showing expression in the intestine and (upper part of image) and germline (lower part of image). Arrow, vesicles. Arrowhead, apical membrane. Scale bar 10 μm.

identified proteins suggests that targets of SUMOylation in normal and stress conditions are similar but the level of modification in stress conditions is highly increased. Furthermore, proteins found to be modified in normal conditions or shortly after stress (30 min) are often not modified 120 min after stress.

To validate our mass spectrometry data we used several antibodies to detect specific SUMO conjugated proteins. We confirmed SUMOylation of actin, alpha tubulin, mannosidase II, GRP94, catalase and cytochrome c1 which have been previously identified as SUMO targets in human cells (Figs 4 and S6).

Our study showed only a 25% (63 common, 810 specific this work, 185 specific for previously published targets[26]) overlap of SUMO conjugated proteins with the only present work on SUMO targets in *C. elegans*[26]. Thus, we conclude that the number of identified SUMO target proteins in worms is not saturated yet. Applying different purification strategies can yield in jet more comprehensive list of SUMO modified proteins in *C. elegans*.

Functional analysis of SUMOylated proteins in *C. elegans*. In our study we identified 874 proteins modified by SUMO in *C. elegans*. Gene ontology analysis showed that we identified proteins involved in early development and reproduction (Fig. S7C). The regulation of proteins during *C. elegans* development is in agreement with our data on the ubiquitous expression of SUMO in worms, especially high in developing embryo. Further, we identified nuclear proteins involved in genome stability, cell cycle progression, chromatin maintenance and modification, RNA splicing and ribosome biogenesis (Fig. 5A). Albeit, most of SUMOylation in the RU86 strain took place in the nucleus, the vast majority of the identified proteins were non-nuclear. Surprisingly, to a large extend these proteins are mitochondrial or extracellular (Figs 5B and S7A). None of the mitochondrial proteins identified was mitochondrially encoded though, suggesting involvement of SUMO in the biogenesis of mitochondrial proteins prior their translocation into the organelle.

Cytosolic proteins undergoing SUMOylation included proteasomal and ribosomal proteins; also proteins involved in metabolism, signaling, cell morphology and motility (cytoskeleton, microtubules, intermediate filaments and proteins involved in connecting cytoskeleton to the plasma membrane). Several membrane transport
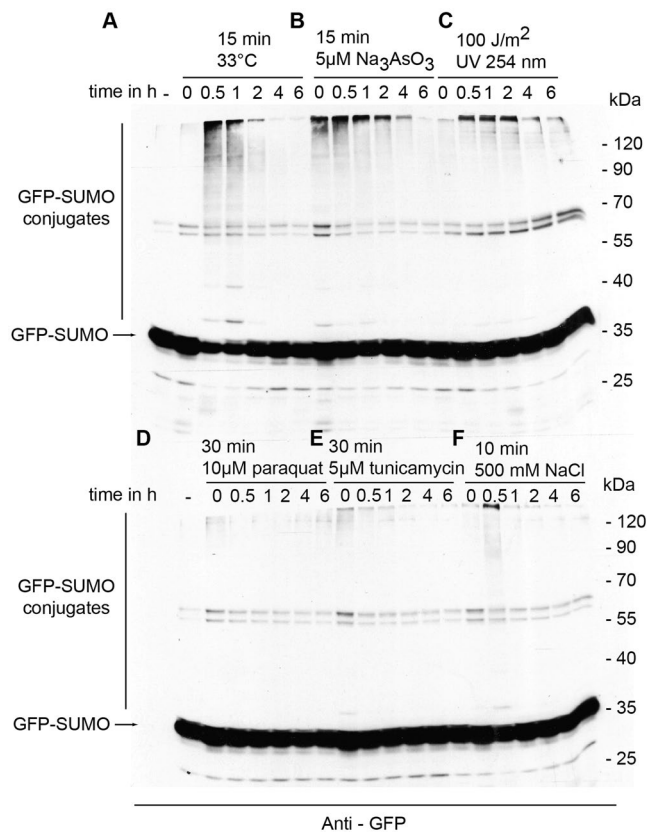
Figure 3. SUMO conjugation to target proteins response to stress conditions. (**A–F**) Mixed stage worms of RU86 strain were exposed to stress conditions: (**A**) heat shock, (**B**) Na$_3$AsO3, (**C**) ultraviolet light, (**D**) paraquat, (**E**) tunicamycin and (**F**) 0.5 M NaCl. Worms were harvested at indicated time points after exposure to the stress. Control, non-treated worms indicated by -.Total protein lysate of worms was separated by SDS-page and immunodecorated with anti-GFP antibody.
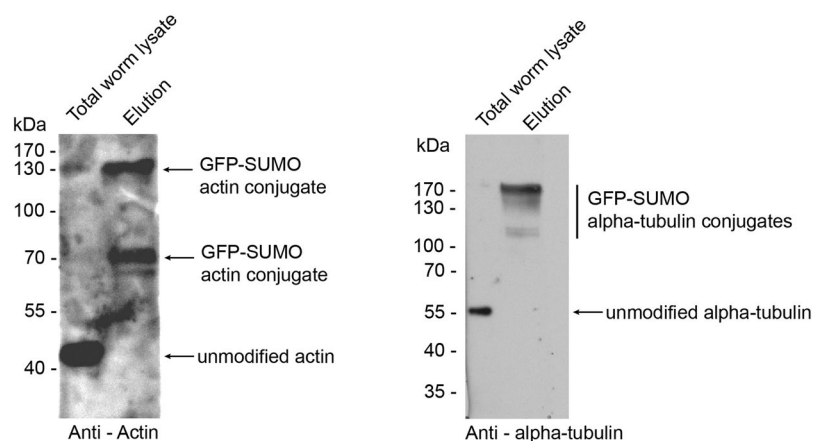

Figure 4. Validation of actin and alpha-tubulin as SUMO targets. Total worm lysate and elution of the purification of GFP-tagged SUMO was analysed by western blot and decorated with specific antibodies.

and vesicular transport proteins were also found to be modified by SUMO. Moreover, proteins involved in stress response, ROS homeostasis and proteostasis were regulated by SUMOylation (Figs 5A and S7).

In previous studies, proteins found to be modified by SUMO were highly interconnected[30]. We tested the interconnectedness of the identified SUMO targets in our study using the STRING database[35]. Indeed, the identified *C. elegans* proteins modified by SUMO are highly interconnected (Fig. 6A). The most prominent clusters included the proteasome, ribosome biogenesis and redox regulation (Fig. 6B). Such robust modification of proteasomal proteins was previously unidentified. Close inspection of proteasomal proteins modified by SUMO
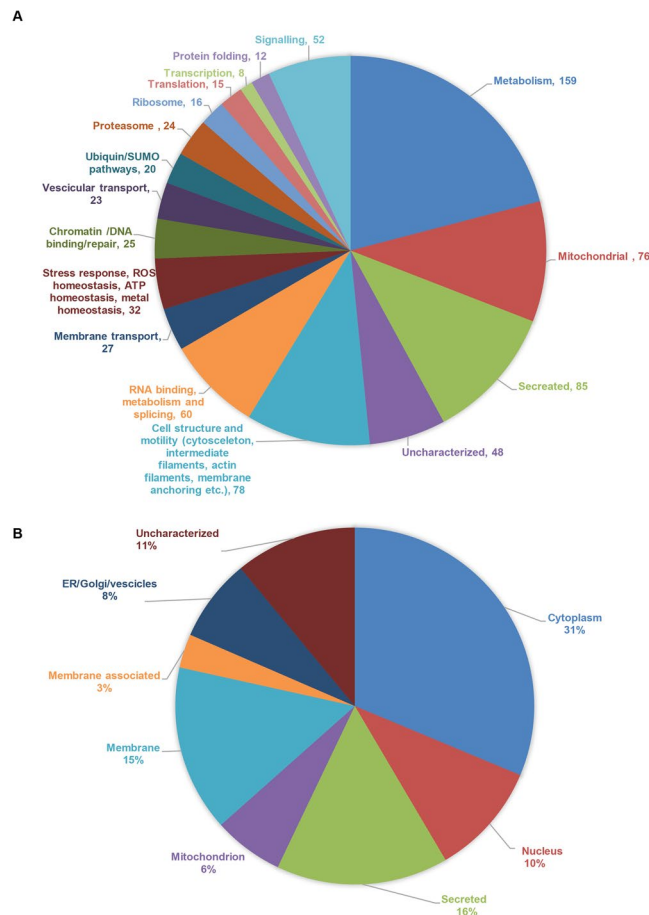
Figure 5. SUMO modifies targets involved in a broad range of cellular processes in the nucleus, cytosol and at membranes in *C. elegans*. (**A**) Grouping of identified SUMO targets in broad functional classes. (**B**) Grouping of identified SUMO targets based on cellular localization.

showed that most of proteins forming the core subunit as well as proteins from the base and the lid were SUMO conjugated.

In summary, we identified the largest number of SUMOylated proteins in *C. elegans* up to now including previously unnoticed classes of potential SUMO targets such as proteasomal proteins as well as mitochondrial and secreted ones.

Conservation of SUMO target proteins. To address the question of conservation of SUMO target proteins and relevancy of research on SUMO in *C. elegans* we analyzed how evolutionary conserved are the SUMO targets in the worm in comparison to the overall protein conservation between species. We compared the list of SUMO modified proteins in the worm to the yeast[36], *Drosophila*[37], *Xenopus*[38], *Arabidopsis*[39], human[33] and mouse[40] proteomes using a strict definition of an ortholog requiring BLASTP e-value below 10^{-10} along at least 80% length of the proteins (Table 1). Applying these strict conditions, approximately 65% of SUMO targets in *C. elegans* have homologues in human proteome. This represents high enrichment compared to approximately 8% of homology of entire proteomes using the same conditions of comparison.

Since *C. elegans* carry only one SUMO gene and vertebrates 4, we asked if worm SUMO targets are more conserved with human SUMO1 or SUMO2/3 targets. We found that in humans, homologues of worm SUMO targets are modified both by SUMO1 and SUMO2/3 with no statistical significant bias.

SUMO target predication by extrapolation from other species. Our and others'[30] analyses showed remarkable evolutionary conservation of SUMO targets throughout species. We assumed that if SUMO targets are so evolutionary conserved and the processes they are involved are also conserved, regulation of this proteins by SUMOylation might also be conserved. Thus, we attempted to predict occurrence of SUMO modification based on the experimental knowledge about the modification in any of the previously analyzed species. We applied a strict definition of an ortholog requiring BLASTP e-value below 10^{-10} along at least 80% length of the proteins. This extrapolation predicts the number of SUMOylated proteins in *C. elegans* to be approximately 4735, in humans 11115, in mouse 11777, in *A. thaliana* 5663, in *X. laevis* 6579, in *D. melanogaster* 2051 and in *S. cerevisiae* 1436 (Table S4). Our prediction suggests that at least 15–20% of the eukaryotic proteome can be modified by SUMO.
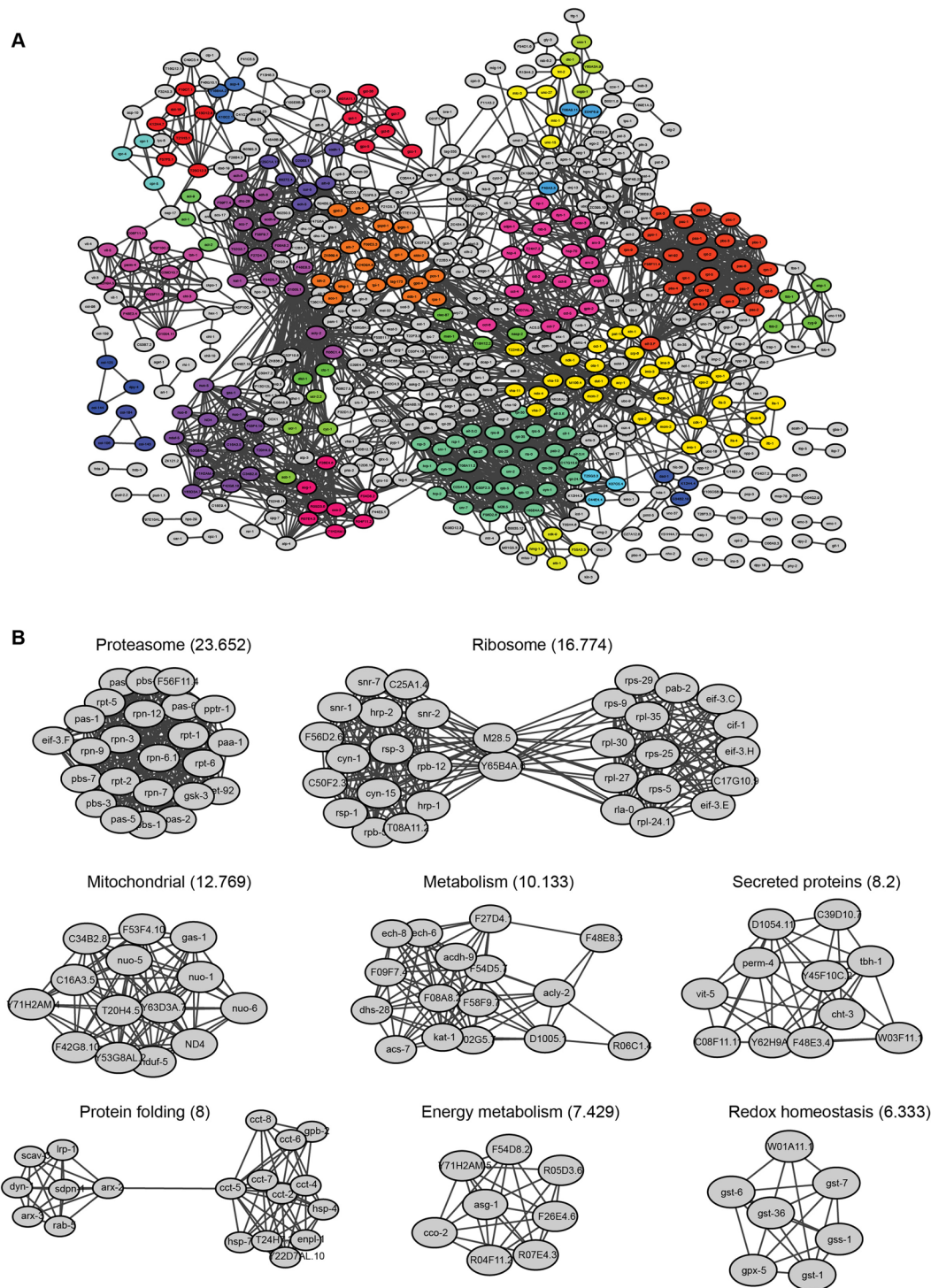
Figure 6. SUMO modifies highly interconnected networks of proteins in *C. elegans*. (**A**) Entire network of identified proteins analyzed by STRING database and visualized by Cytoscape. Clusters were labeled with different colors (**B**). Individual, most prominent clusters are shown. MCODE score is in parentheses.

We tested our prediction by analyzing the human proteins that we have predicted to be SUMOylated but were not identified in the published experimental identifications. We have analyzed if these proteins carry any of the SUMO consensuses described by Hendriks *et al.*[30]. 88% of proteins predicted carried the consensus site compared to 94% of the experimentally identified ones. In contrast, in the entire human proteome, 65% of proteins carry the consensus site and 60.5% in the part of the proteome not shown or predicted to be SUMOylated (Table 2). The p-value of this enrichment is below 2.2×10^{-16}.

Species	Proteome	Homologues in *C. elegans* proteome	Homologues in *C. elegans* SUMO targets	Enrichment p-value
C. elegans	26723			
H. sapiens	72874	6010	570	$>2.2 \times 10^{-16}$
M. musculus	50311	5977	580	$>2.2 \times 10^{-16}$
X. laevis	16378	4939	503	$>2.2 \times 10^{-16}$
D. melanogaster	21977	5066	529	$>2.2 \times 10^{-16}$
A. thaliana	31392	2993	393	$>2.2 \times 10^{-16}$
S. cerevisiae	6838	1886	275	$>2.2 \times 10^{-16}$

Table 1. Evolutionary conservation of *C. elegans* SUMO targets with other species. *C. elegans* proteome and SUMO targets from *C. elegans* were compared to whole proteomes and to orthologues of *C. elegans* proteins in other species. Only strict homologues with BLASTP e-value smaller than 1×10^{-10} and over 80% of residues in both sequences included in the BLASTP alignment, were included in the comparison.

	Sequences	[IV]KE	[VF]K[QTEP][ED][PKE]	KE[ED]5	[ED]K	All motifs	Containing SUMOylation motif
Identified SUMO targets	4146	1343	293	41	3862	3905	94.19%
Predicted SUMO targets	7161	1348	234	24	6219	6298	87.95%
Proteome	70956	8184	1300	148	45462	46286	65.23%
Proteome minus (identified + predicted)	59649	5493	773	83	35381	36083	60.49%

Table 2. Testing of SUMOylation prediction by analysis of SUMO motifs in predicted human proteins. Percentage of human proteins predicted in this study but not previously experimentally identified, carrying experimentally defined SUMOylation consensus; compared to the presence of consensus in identified proteins, the entire human proteome and the proteome lacking identified and predicted SUMO targets. The p-value of the enrichment is below 2.2×10^{-10}.

SUMO target predication web server. To facilitate access to predictions of SUMO modifications based on evolutionary conservation of known targets, we have established a publicly available web server (sumobase. mslab-ibb.pl). Sequence alignments of known and predicted SUMOylated proteins is shown is Fig. S8B. The database contains known and predicted SUMO targets in *C. elegans*, *H. sapiens*, *M. musculus*, *X. laevis S. cerevisiae*, *D. melanogaster* and *A. thaliana*. Screenshots of the web page are presented in Fig. S8.

Discussion

Every dynamic protein modification depends on the developmental stage, cell type and cell cycle stage. Analysis on the level of the entire organism enables a search throughout the entire spectrum of targets and modifications. *Caenorhabditis elegans* is an excellent model to decipher the universe of the proteins and processes regulated by SUMO. *C. elegans* enables identification of targets in all tissues and developmental stages. Recent high throughput proteomics identifications from human cells (for review see Hendriks and Vertegaal[33]) could not decipher the SUMO proteome in embryonic development and many tissues, for example neurons, germline, and liver or kidney cells.

We have analyzed SUMO modifications in *C. elegans* in all developmental stages and in various stress conditions: heat shock, DNA damage, ER stress, osmotic stress and complex inhibition of many enzymes and pathways by arsenite. Other studies have shown that SUMOylation is increased in ischemia[17,41] and drought in plants[42]. All these stressors elicit different defense responses; nevertheless all result in increased levels of SUMOylation suggesting that SUMO is a key stress response protein on which many stress resistance mechanisms converge.

We have identified 874 proteins modified by SUMO in *C. elegans*, more than 3 times more than previously reported[26]. We compared the list of proteins identified in this study with the previously published data in *C. elegans*[26]. Surprisingly, there is little overlap between the set of proteins identified by Kaminsky *et al*. This might be due to a completely different purification strategy. We found that two step purification in native conditions, using IMAC followed by antibody purification resulted in much false positive identification of proteins that were purified from wild-type N2 worms. This is, probably due to high number of poly-His and carbohydrate binding proteins in worms as well as co-purification of SUMO interacting proteins and proteins interacting with SUMO targets. In contrast, our strategy of lysing worms in boiling 1% SDS should exclude all SUMO interacting proteins as well as proteins interacting with SUMO modified proteins.

We identified several known targets of SUMOylation for example proteins involved in cell cycle regulation, transcription, translation and cytoskeleton confirming the conservation of these processes being regulated by SUMO. We have identified several new targets of SUMO modification related to cell homeostasis and metabolism. Metabolic processes have been shown to be regulated by SUMOylation of transcription factors governing expression of metabolic enzymes[43]. In previous proteomics studies, proteins involved in metabolism and homeostasis were also identified but at much smaller proportion than in our study. Our analysis suggests also that other cellular processes are regulated by SUMOylation. Beside the ribosome, the other macromolecular complexes heavily modified by SUMO were the proteasome and proteins involved in maintenance of cell shape and motility.

In this study, we have identified several mitochondrial proteins as well as extracellular proteins. It is surprising because SUMOylation is a strictly nuclear and cytosolic process. However, recently it was shown that several inner mitochondrial proteins undergo ubiquitin mediated proteasomal degradation[44]. Furthermore, proteins can be exported from the mitochondrion for proteasomal degradation[45]. SUMO tagging is also known to indirectly target protein for proteasomal degradation[46]. Moreover, it was shown that accumulation of newly synthetized, misfolded proteins triggers SUMO conjugation response[47]. Thus, we propose that misfolded mitochondrial and secreted proteins are first SUMOylated, subsequently polyubiquitinated and degraded by the proteasome.

C. elegans SUMO does not contain SUMOylation consensus site. Yeast SUMO (Smt3) and human SUMO2/3 can form poly SUMO chains. Sequence analysis suggests that *C. elegans* SUMO is not more similar to SMT3 or to human SUMO2/3 than to human SUMO1 (Fig. S9). However, our functional analysis of identified SUMO targets, i.e. in response to stress suggests that worm SUMO can be functional related to SUMO2/3. The presence of high molecular weight SUMO-conjugates on the immunoblot might suggest that SUMO protein forms chains in *C. elegans* and thus, can be a signal for polyubiquitination and subsequent proteasomal degradation. *C. elegans* Y47G6A.31 protein show weak homology to *S. pombe* Rfp1 and human RNF4 proteins responsible for SUMO chain directed polyubiquitiantion. Another argument for *C. elegans* SUMO being a functional homologue of SUMO2/3 is that SUMO2/3, in contrast to SUMO1, in normal conditions is present as a pool of free protein and is being attached to target proteins upon stress[48] what is also true in case of the worm SUMO protein. Thus, it is reasonable to assume that the only *C. elegans* SUMO performs all functions reserved for both SUMO1 and SUMO2/3 in vertebrates.

SUMOylation is a very dynamic process depending on developmental state of an organism as well as internal and external state of the cells. Precise mass spectrometry identification of SUMO modified proteins in vertebrates would require isolating and identifying SUMO conjugates from all cells and tissues in all developmental stages of a mouse, for example. Knowledge of close to complete list of SUMO targets is essential if SUMO pathway is to become a target of therapeutic interventions in humans. Thus, several groups attempted to predict potential SUMOylation based on the presence of SUMO modification consensus sites and other sequence features[49,50]. These attempts are usually hampered by the fact that 65% of proteins in the human proteome carry the SUMOylation consensus. In recent years almost 4000 human proteins have been experimentally identified to be SUMOylated. A major limitation of these identifications is that SUMO conjugates are isolated from few cell lines. Thus, our approach to identify SUMO modified proteins at the level of entire higher eukaryotic organism, at all stages of development; combined with prediction based on evolutionary conservation of targets is a useful step towards deciphering the entire SUMOylonome.

We predict that at least 15–20% of the eukaryotic proteome can be SUMOylated and suggest that SUMO functions in 3 main areas: regulation of activity of individual proteins, biogenesis of macromolecular complexes and SUMO directed proteasomal degradation.

Methods

C. elegans and bacterial strains.
Worms were maintained and handled in standard conditions as described by Brenner[51]. Worms were cultured at 20 °C unless stated otherwise on plates seeded with *E. coli* HB101 bacteria. Strains used in this study: N2 (wild type) and VC186[*smo-1*(ok359)/szT1 [*lon-2*(e678)] I; +/szT1 X] (provided by the *Caenorhabditis* Genetics Center), RU86[*smo-1*(ok359); Is(p*smo-1::smo-1gfp*::3′UTR*smo-1*].

Constructs and transformation.
The 8-HisGFP tagged SUMO was constructed as follows: The SUMO coding sequence and 3′UTR was amplified from the worm genomic DNA with the primers 5′ ACTCCCGGCTAGCACGATGGCCGATATGC and 3′ GGACGGAGAAGGCCTTCGAATCTCGTGTC and cloned into NheI and StuI sites of the pPD117.01 vector. The 8-His tag was introduced into pPD117.01 vector (Fire kit) by PCR together with the SUMO promoter. The SUMO promoter (1 kb) was amplified from the genomic DNA with 5′-ATTTTATCACGGGCATGCTGGCCTTCCTC, 3′-CATGCTACCACACC ATCACCATCACCATGCCGATGATGCAGCTCAAATTC primers and cloned into SphI and KpnI siteS of the pPD117.01 vector (Fire). The 3′ primer introduced 8His tag at the N-terminus of the GFP.

VC186 SUMO knock-out strain was transformed by biolistic transformation as previously descried[52]. Integrated transformers were selected by repeated sorting of GFP positive animals with the COPAS worm sorter (Union Biometrica) for 40 generations. Integrated strain was back-crossed 10 times with wild type N2 worms.

Sumo conjugate purification.
SUMO conjugates were purified under SDS denaturing condition as described[34,53]. Briefly, worms were grinded in liquid nitrogen and resuspended in boiling SDS lysis buffer (25 mM NaPO4 pH 7.5, 300 mM NaCl, 1%SDS) and sonicated at 20% for 2 min with Branson sonifier (Branson). Next, the lysate was spun down for 30′ at 20000 g at 4 °C. Supernatant was transferred to new tubes and cooled on ice for at least 30′ to precipitate the SDS. The lysate was cleared of precipitated SDS by centrifugation at 10000 g for 30′ at 0 °C. The cleared lysate was filtered through 0.22 μ filter and applied to IMAC column (GE Healthcare) equilibrated with the buffer (25 mM NaPO4 pH 7.5, 300 mM NaCl, 20 mM imidazole. 0.1% sarcosyl). The column was washed with at least 20 volumes of the wash buffer (25 mM NaPO4 pH 7.5, 300 mM NaCl, 50 mM imidazole. 0.1% sarcosyl). Proteins were eluted with 2 column volumes of the elution buffer (25 mM NaPO4 pH 7.5, 300 mM NaCl, 300 mM imidazole. 0.1% sarcosyl).

In-solution digestion by FASP.
Isolated proteins were prepared for mass spectrometry analysis by FASP (Filter Assisted Sample Preparation)[54]. Briefly, 50 μg of each sample was reduced with 50 mM DTT at 55 °C for 30 min, diluted 1:8 in 8 M urea and applied to Vivacon 500 ultrafiltration unit MWCO 30000 (Sartorius). Samples were washed twice with 8 M urea, free cysteins were blocked with 50 μM iodoacetamide in 8 M urea and the filters

were washed twice with 8 M urea and twice with 50 mM NH_4HCO_3. Next, the samples were digested in a wet chamber, O/N with trypsin (Promega) in 50 mM NH_4HCO_3, protein: trypsin ratio 100:1.

Mass spectrometry. Peptide mixtures were applied to RP-18 precolumns (nanoACQUITY Symmetry® C18—Waters) using water containing 0.1% trifluoroacetic acid as mobile phase and then transferred to nano-HPLC RP-18 columns (nanoACQUITY BEH C18—Waters) using an acetonitrile gradient (5–35%) for 180 min in the presence of 0.05% formic acid with a flow rate of 250 nl min^{-1}. The column outlet was directly coupled to the ion source of Q Exactive™ Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Electron Corp). The mass spectrometer was operated in positive ion mode with a selected mass range of 300–2000 mass/charge (m/z).

Data processing analysis. Raw files were processed, including peak list generation, using the MaxQuant (v1.5.7.4) computational proteomics platform and default parameters were used. The fragmentation spectra were searched using Andromeda search engine integrated into the MaxQuant platform against a custom, non-redundant database of *C. elegans* proteome that included the sequence of the SUMO-GFP protein. The error ranges for the first and main searches were 20 ppm and 6 ppm, respectively, with 2 missed cleavages. Carbamidomethylation of cysteines was set as a fixed modification, and oxidation and protein N-terminal acetylation were selected as variable modifications for database searching. The minimum peptide length was set at 6 aa. Both peptide and protein identifications were filtered at a 1% false discovery rate. Enzyme specificity was set to trypsin.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[55] partner repository with the dataset identifier PXD006644.

Stress assays in *C. elegans*. Heat shock was induced by 15′ incubation in at 33 °C in a water bath. UV stress was induced by irradiation of 100 J/m² UV 254 nm in a Stratalinker (Stratagene). Arsenite treatment 30′ 5 µM sodium arsenite (Sigma) in M9. Osmotic stress was induced by incubation in 500 mM NaCl in M9 for 10′. ER stress was introduced by 30′ incubation in 5 µM tunicamycin in M9. Oxidative stress was introduced by 30 min incubation in 10 µM paraquat in M9.

Microscopy. Worms were anesthetized in 20 mM sodium azide in M9 and immobilized on 2% agarose pads. Fluorescent images were taken with Zeiss Observer D1 microscope 40 × 0.6 lens, excitation 480 nm LED light source, Confocal images were takes with Olympus FV1000 microscope (Olympus) 60 × 1.4 oil immersion lens. Excititation 440 nm. The images were processed with ImageJ v 1.51n.

Immunoblot analysis. Worms were collected at different time points in M9 buffer, spun down and equal volume 2x SDS lysis buffer (50 mM NaPO4 pH 7.5, 600 mM NaCl, 2%SDS) was added. Worms were frozen in liquid nitrogen, thawed and sonicated at 10% for 2 min with Branson sonifier (Branson). The lysates were spun down. The protein content was measured with BCA kit (Thermo Scientific). 20 µg of protein lysate was loaded on each lane and separated by SDS-PAGE followed by electroblotting on PVDF membrane. The membrane was blocked with 5% non-fat milk and incubated with anti-GFP antibodies (Roche) o/n followed by horse radish conjugated secondary antibody. Rabbit antibody against *C. elegans* SUMO was generated against the N-terminal peptide MADDAAQAGDNAEYIKIK (Eurogentec). Antibodies used to confirm the SUMO modification included: anti actin (MAB1501 from Chemicon), anti alpha tubulin clone DM1A (Sigma T9026), anti mannosidase II (AbDSerotec AHP674), anti cytochrome C1 (custom made), anti catalase (SantaCruz sc-365738) anti GRP94 (SantaCruz sc-11402). The secondary antibody was detected with ECL system.

Bioinformatics analysis. Network analysis was performed using STRING database[56], minimal interactions score was set to "high confidence". STRING generated networks were further analyzed by Cytoscape[57].

Proteins were considered as orthologues when the BLASTP E-value was smaller than 1×10^{-10} and over 80% of residues in both sequences included in the BLASTP alignment[58]. Enrichment analysis p-value was calculated using Pearson's Chi-squared test with Yates' continuity correction. The multiple sequence alignment was performed by MUSCLE software[59].

Evolutionary analyses were conducted in MEGA7[60]. The evolutionary history was inferred by using the Maximum Likelihood method based on the Whelan And Goldman model.

References

1. Johnson, E. S. Protein modification by SUMO. *Annu Rev Biochem* **73**, 355–382, https://doi.org/10.1146/annurev.biochem.73.011303.074118 (2004).
2. Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**, 159–180, https://doi.org/10.1146/annurev.cellbio.22.010605.093503 (2006).
3. Hay, R. T. Decoding the SUMO signal. *Biochem Soc Trans* **41**, 463–473, https://doi.org/10.1042/BST20130015 (2013).
4. Hickey, C. M., Wilson, N. R. & Hochstrasser, M. Function and regulation of SUMO proteases. *Nat Rev Mol Cell Biol* **13**, 755–766, https://doi.org/10.1038/nrm3478 (2012).
5. Hannoun, Z., Greenhough, S., Jaffray, E., Hay, R. T. & Hay, D. C. Post-translational modification by SUMO. *Toxicology* **278**, 288–293, https://doi.org/10.1016/j.tox.2010.07.013 (2010).
6. Psakhye, I. & Jentsch, S. Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell* **151**, 807–820, https://doi.org/10.1016/j.cell.2012.10.021 (2012).
7. Park-Sarge, O. K. & Sarge, K. D. Detection of sumoylated proteins. *Methods Mol Biol* **464**, 255–265, https://doi.org/10.1007/978-1-60327-461-6_14 (2009).
8. Zhu, S., Sachdeva, M., Wu, F., Lu, Z. & Mo, Y. Y. Ubc9 promotes breast cell invasion and metastasis in a sumoylation-independent manner. *Oncogene* **29**, 1763–1772, https://doi.org/10.1038/onc.2009.459 (2010).
9. Wang, L. & Banerjee, S. Differential PIAS3 expression in human malignancy. *Oncol Rep* **11**, 1319–1324 (2004).

10. Stehmeier, P. & Muller, S. Regulation of p53 family members by the ubiquitin-like SUMO system. *DNA Repair (Amst)* **8**, 491–498, https://doi.org/10.1016/j.dnarep.2009.01.002 (2009).
11. Seeler, J. S., Bischof, O., Nacerddine, K. & Dejean, A. SUMO, the three Rs and cancer. *Curr Top Microbiol Immunol* **313**, 49–71 (2007).
12. Dorval, V. & Fraser, P. E. SUMO on the road to neurodegeneration. *Biochim Biophys Acta* **1773**, 694–706, https://doi.org/10.1016/j.bbamcr.2007.03.017 (2007).
13. Woo, C. H. & Abe, J. SUMO–a post-translational modification with therapeutic potential? *Curr Opin Pharmacol* **10**, 146–155, https://doi.org/10.1016/j.coph.2009.12.001 (2010).
14. Golebiowski, F. *et al.* System-wide changes to SUMO modifications in response to heat shock. *Sci Signal* **2**, ra24, https://doi.org/10.1126/scisignal.2000282 (2009).
15. Nie, M., Xie, Y., Loo, J. A. & Courey, A. J. Genetic and proteomic evidence for roles of Drosophila SUMO in cell cycle control, Ras signaling, and early pattern formation. *PLoS One* **4**, e5905, https://doi.org/10.1371/journal.pone.0005905 (2009).
16. Miller, M. J. & Vierstra, R. D. Mass spectrometric identification of SUMO substrates provides insights into heat stress-induced SUMOylation in plants. *Plant Signal Behav* **6**, 130–133 (2011).
17. Lee, Y. J., Mou, Y., Klimanis, D., Bernstock, J. D. & Hallenbeck, J. M. Global SUMOylation is a molecular mechanism underlying hypothermia-induced ischemic tolerance. *Front Cell Neurosci* **8**, 416, https://doi.org/10.3389/fncel.2014.00416 (2014).
18. Yang, W. *et al.* Analysis of oxygen/glucose-deprivation-induced changes in SUMO3 conjugation using SILAC-based quantitative proteomics. *J Proteome Res* **11**, 1108–1117, https://doi.org/10.1021/pr200834f (2012).
19. Bossis, G. & Melchior, F. Regulation of SUMOylation by reversible oxidation of SUMO conjugating enzymes. *Mol Cell* **21**, 349–357, https://doi.org/10.1016/j.molcel.2005.12.019 (2006).
20. de la Vega, L. *et al.* A redox-regulated SUMO/acetylation switch of HIPK2 controls the survival threshold to oxidative stress. *Mol Cell* **46**, 472–483, https://doi.org/10.1016/j.molcel.2012.03.003 (2012).
21. Shrivastava, V., Pekar, M., Grosser, E., Im, J. & Vigodner, M. SUMO proteins are involved in the stress response during spermatogenesis and are localized to DNA double-strand breaks in germ cells. *Reproduction* **139**, 999–1010, https://doi.org/10.1530/REP-09-0492 (2010).
22. Lim, Y. *et al.* Sumoylation regulates ER stress response by modulating calreticulin gene expression in XBP-1-dependent mode in Caenorhabditis elegans. *Int J Biochem Cell Biol* **53C**, 399–408, https://doi.org/10.1016/j.biocel.2014.06.005 (2014).
23. Pinto, M. P. *et al.* Heat shock induces a massive but differential inactivation of SUMO-specific proteases. *Biochim Biophys Acta* **1823**, 1958–1966, https://doi.org/10.1016/j.bbamcr.2012.07.010 (2012).
24. Enserink, J. M. Regulation of Cellular Processes by SUMO: Understudied Topics. *Adv Exp Med Biol* **963**, 89–97, https://doi.org/10.1007/978-3-319-50044-7_5 (2017).
25. Seifert, A., Schofield, P., Barton, G. J. & Hay, R. T. Proteotoxic stress reprograms the chromatin landscape of SUMO modification. *Sci Signal* **8**, rs7, https://doi.org/10.1126/scisignal.aaa2213 (2015).
26. Kaminsky, R. *et al.* SUMO regulates the assembly and function of a cytoplasmic intermediate filament protein in C. elegans. *Dev Cell* **17**, 724–735, https://doi.org/10.1016/j.devcel.2009.10.005 (2009).
27. Pelisch, F. *et al.* Dynamic SUMO modification regulates mitotic chromosome assembly and cell cycle progression in Caenorhabditis elegans. *Nat Commun* **5**, 5485, https://doi.org/10.1038/ncomms6485 (2014).
28. Tsur, A., Bening Abu-Shach, U. & Broday, L. ULP-2 SUMO Protease Regulates E-Cadherin Recruitment to Adherens Junctions. *Dev Cell* **35**, 63–77, https://doi.org/10.1016/j.devcel.2015.08.019 (2015).
29. Lamoliatte, F. *et al.* Large-scale analysis of lysine SUMOylation by SUMO remnant immunoaffinity profiling. *Nat Commun* **5**, 5409, https://doi.org/10.1038/ncomms6409 (2014).
30. Hendriks, I. A. *et al.* Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927–936, https://doi.org/10.1038/nsmb.2890 (2014).
31. Impens, F., Radoshevich, L., Cossart, P. & Ribet, D. Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli. *Proc Natl Acad Sci USA* **111**, 12432–12437, https://doi.org/10.1073/pnas.1413825111 (2014).
32. Becker, J. *et al.* Detecting endogenous SUMO targets in mammalian cells and tissues. *Nat Struct Mol Biol* **20**, 525–531, https://doi.org/10.1038/nsmb.2526 (2013).
33. Hendriks, I. A. & Vertegaal, A. C. A comprehensive compilation of SUMO proteomics. *Nat Rev Mol Cell Biol* **17**, 581–595, https://doi.org/10.1038/nrm.2016.81 (2016).
34. Schlager, B., Straessle, A. & Hafen, E. Use of anionic denaturing detergents to purify insoluble proteins after overexpression. *BMC Biotechnol* **12**, 95, https://doi.org/10.1186/1472-6750-12-95 (2012).
35. Szklarczyk, D. *et al.* STRINGv10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447–452, https://doi.org/10.1093/nar/gku1003 (2015).
36. Nie, M., Vashisht, A. A., Wohlschlegel, J. A. & Boddy, M. N. High Confidence Fission Yeast SUMO Conjugates Identified by Tandem Denaturing Affinity Purification. *Sci Rep* **5**, 14389, https://doi.org/10.1038/srep14389 (2015).
37. Handu, M. *et al.* SUMO-Enriched Proteome for Drosophila Innate Immune Response. *G3 (Bethesda)* **5**, 2137–2154, https://doi.org/10.1534/g3.115.020958 (2015).
38. Ma, L. *et al.* Identification of small ubiquitin-like modifier substrates with diverse functions using the Xenopus egg extract system. *Mol Cell Proteomics* **13**, 1659–1675, https://doi.org/10.1074/mcp.M113.035626 (2014).
39. Miller, M. J. *et al.* Quantitative proteomics reveals factors regulating RNA biology as dynamic targets of stress-induced SUMOylation in Arabidopsis. *Mol Cell Proteomics* **12**, 449–463, https://doi.org/10.1074/mcp.M112.025056 (2013).
40. Cai, L. *et al.* Proteome-wide Mapping of Endogenous SUMOylation Sites in Mouse Testis. *Mol Cell Proteomics* **16**, 717–727, https://doi.org/10.1074/mcp.M116.062125 (2017).
41. Silveirinha, V., Stephens, G. J. & Cimarosti, H. Molecular targets underlying SUMO-mediated neuroprotection in brain ischemia. *J Neurochem* **127**, 580–591, https://doi.org/10.1111/jnc.12347 (2013).
42. Catala, R. *et al.* The Arabidopsis E3 SUMO ligase SIZ1 regulates plant growth and drought responses. *Plant Cell* **19**, 2952–2966, https://doi.org/10.1105/tpc.106.049981 (2007).
43. Kamynina, E. & Stover, P. J. The Roles of SUMO in Metabolic Regulation. *Adv Exp Med Biol* **963**, 143–168, https://doi.org/10.1007/978-3-319-50044-7_9 (2017).
44. Wrobel, L. *et al.* Mistargeted mitochondrial proteins activate a proteostatic response in the cytosol. *Nature* **524**, 485–488, https://doi.org/10.1038/nature14951 (2015).
45. Bragoszewski, P. *et al.* Retro-translocation of mitochondrial intermembrane space proteins. *Proc Natl Acad Sci USA* **112**, 7713–7718, https://doi.org/10.1073/pnas.1504615112 (2015).
46. Schimmel, J. *et al.* The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Mol Cell Proteomics* **7**, 2107–2122, https://doi.org/10.1074/mcp.M800025-MCP200 (2008).
47. Tatham, M. H., Matic, I., Mann, M. & Hay, R. T. Comparative proteomic analysis identifies a role for SUMO in protein quality control. *Sci Signal* **4**, rs4, https://doi.org/10.1126/scisignal.2001484 (2011).
48. Bursomanno, S. *et al.* Proteome-wide analysis of SUMO2 targets in response to pathological DNA replication stress in human cells. *DNA Repair (Amst)* **25**, 84–96, https://doi.org/10.1016/j.dnarep.2014.10.011 (2015).
49. Beauclair, G., Bridier-Nahmias, A., Zagury, J. F., Saïb, A. & Zamborlini, A. JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics* **31**, 3483–3491, https://doi.org/10.1093/bioinformatics/btv403 (2015).

50. Zhao, Q. *et al*. GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Res* **42**, W325–330, https://doi.org/10.1093/nar/gku383 (2014).
51. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94 (1974).
52. Praitis, V., Casey, E., Collar, D. & Austin, J. Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. *Genetics* **157**, 1217–1226 (2001).
53. Drabikowski, K. & Dadlez, M. Dissecting SUMO Dynamics by Mass Spectrometry. *Methods Mol Biol* **1449**, 291–298, https://doi.org/10.1007/978-1-4939-3756-1_18 (2016).
54. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 359–362, https://doi.org/10.1038/nmeth.1322 (2009).
55. Vizcaíno, J. A. *et al*. update of the PRIDE database and its related tools. *Nucleic Acids Res* **44**(11033), 2016, https://doi.org/10.1093/nar/gkw880 (2016).
56. Franceschini, A. *et al*. STRINGv9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**, D808–815, https://doi.org/10.1093/nar/gks1094 (2013).
57. Su, G., Morris, J. H., Demchak, B. & Bader, G. D. Biological network exploration with cytoscape 3. *Curr Protoc Bioinformatics* **47**(8), 13.11–18.13.24, https://doi.org/10.1002/0471250953.bi0813s47 (2014).
58. Yu, H. *et al*. Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res* **14**, 1107–1118, https://doi.org/10.1101/gr.1774904 (2004).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797, https://doi.org/10.1093/nar/gkh340 (2004).
60. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870–1874, https://doi.org/10.1093/molbev/msw054 (2016).

Acknowledgements

Author Contributions

K.D. conceived the project, performed experiments, analyzed the data and wrote the manuscript; J.F. performed experiments; M.K. wrote scripts and performed bioinformatics analysis; J.O. performed mass spectrometry, M.D. supervised the project; R.C.E. conceived and supervised the project.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-19424-9.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.