A robust method for identification of plant SUMOylation substrates in a library-based reconstitution system

Dear Editor,

SUMOylation, which transfers a small ubiquitin-like protein to the lysine residues of target proteins, is an important type of post-translational modification in eukaryotic cells. This modification plays a critical role in plant development and stress responses (Miura and Hasegawa, 2010); thus, the efficient identification of SUMOylation substrates in plant cells is a fundamental issue in the field. In Arabidopsis, hundreds of SUMOylated proteins have been identified by a mass spectrometry (MS) approach in which modified targets in transgenic plants harboring an H89R variant of SUMO1 are enriched via affinity chromatography and a four-residue footprint is left on the substrates after trypsin digestion for MS analysis (Miller et al., 2010). Given that antibodies generated against SUMO molecules are not very specific, it is impractical to catch endogenous SUMO moieties via antibody-based one-step purification. Thus, it is difficult to obtain SUMOylation substrates via existing approaches in plant species that cannot be efficiently transformed. Given that SUMOylation is highly reversible (Yates et al., 2016), another problem is that the modification may be removed quickly during sample preparation. In addition, based on the efficiency and sensitivity of affinity enrichment and MS, substrates with low protein levels may be missed in the analyses.

To overcome these technological difficulties, we aimed to establish a robust method for identifying SUMOylation substrates in plant species based on the use of a plant cDNA library for substrate screening in a reconstituted reaction system in bacteria (Figure 1A; the detailed procedure is included in the supplemental methods). SUMOylation has been shown to play critical roles in maize, including in development and stress responses (Augustine et al., 2016; Chen et al., 2018), but the molecular functions of SUMOylation remain to be investigated in this important cereal crop (Rosa and Abreu, 2019). Therefore, we used maize as a model system to set up a method to identify SUMOylation substrates and test the efficiency, specificity, and reliability of this approach.

SUMOylation is successively catalyzed via an E1 activating enzyme dimer, an E2 conjugating enzyme, and possibly an E3 ligase in plant cells (Miura and Hasegawa, 2010). However, in the reconstitution system in bacteria, co-expression of E1 and E2 is sufficient for SUMO conjugation to most substrates (Okada et al., 2009), providing a simple system for in vitro SUMOylation detection. In a previous study, multiple SUMO-related enzymes were characterized in maize (Augustine et al., 2016). Thus, we selected functional maize enzymes to establish a reconstituted SUMOylation system in Escherichia coli; these included ZmSUMO1aGG (the mature form of SUMO1a in maize; a His$_6$–FLAG tag was fused to its N terminus) and the SUMO E1 dimer (ZmSAE1/ZmSAE2a), with or without SUMO E2 (ZmSCE1b) (Figure 1A). AtADA2b, a well-characterized Arabidopsis substrate (Elrouby and Coupland, 2010), was used to verify SUMO conjugating activity in the reconstitution system (Supplemental Figure 1A). Then, a maize cDNA library fused with a Myc tag was constructed in the pCDF-Duet-1 vector for protein expression in bacteria. Given that heat stress induces SUMOylation in eukaryotic cells (Miller et al., 2013), the cDNA library was generated using an RNA mixture from different tissues and heat-treated seedlings of maize (Supplemental Figure 1B; Supplemental Table 1).

The cDNA library was then transformed into bacterial competent cells expressing SUMO1aGG and E1, with or without E2. The individual colonies were mixed for cell culture and inducible protein expression. Many protein bands with different molecular weights were detected in both bacterial strains in an anti-Myc immunoblot, but the anti-FLAG signals were much higher in the sample with E2 than in the sample without E2 (Figure 1B), suggesting that this system can potentially be used for identification of SUMOylation substrates. Therefore, the library transformation and protein expression were scaled up for tandem affinity purification. Proteins in the extract were mixed with nickel resin recognizing the His$_6$ tag on SUMO1aGG, and the enriched proteins were then eluted with imidazole and loaded on anti-Myc agarose to catch the Myc tag on the substrates from the library. Finally, the precipitated proteins were denatured with urea and reloaded onto the nickel resin for further reduction of contamination (Figure 1C). After this three-step purification, the enriched proteins were analyzed using MS to identify SUMOylation substrates.

In the MS data, 212 maize proteins were uniquely identified in the sample with E2 (referred to as +E2 candidates), 72 maize proteins were identified in both samples, with and without E2 (referred to as ±E2 candidates), and only 24 maize proteins were specifically identified in the control sample without E2 (referred to as –E2 candidates) (Figure 1D; Supplemental Table 2), supporting the potentially high specificity of this method. A Gene Ontology analysis indicated that the +E2 candidates are involved in multiple biological pathways, especially in metabolism and stress responses (Figure 1E; Supplemental Table 3). Given that SUMO is rapidly induced by heat and hydrogen peroxide (Miller et al., 2013), the enrichment of
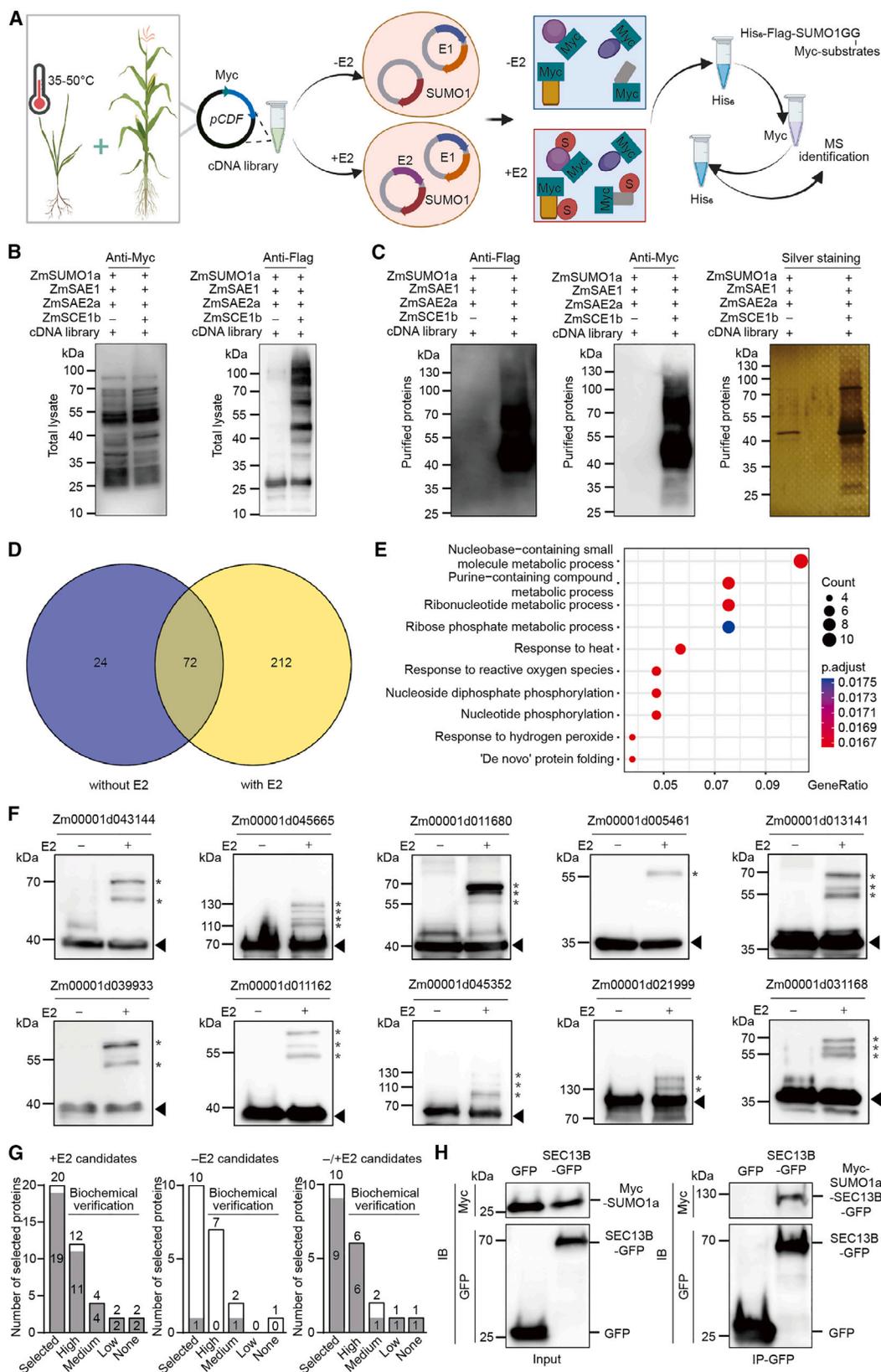
Figure 1. A robust method for identification of plant SUMOylation substrates in a library-based reconstitution system.
(A) Schematic design of the method. A maize cDNA library fused with a Myc tag was constructed in a *pCDF-Duet-1* vector using an RNA mixture from different tissues and heat-treated seedlings of maize. The cDNA library was transformed into bacterial competent cells expressing SUMO1aGG

proteins associated with heat and reactive oxygen response in the Gene Ontology analysis is consistent with the functions of this modification in plant cells.

The identified +E2 candidates were further analyzed with GPS-SUMO software (Zhao et al., 2014) to predict SUMOylation sites. As a result, 93.4% of these candidate proteins were predicted to be SUMOylation substrates (65.6% with high scores, 20.8% with medium scores, and 7.1% with low scores) (Supplemental Figure 1C; Supplemental Table 4). Surprisingly, 91.6% of the –E2 candidates (Supplemental Figure 2A; Supplemental Table 5) and 88.9% of the ±E2 candidates (Supplemental Figure 2B; Supplemental Table 6) were also predicted as potential SUMOylation substrates. Therefore, the identified proteins were verified in a biochemical assay to test the specificity of this approach.

To verify the substrates identified from MS, 20 proteins were selected from 212 +E2 candidates, 10 proteins were selected from 24 –E2 candidates, and 10 proteins were selected from 72 ±E2 candidates for SUMOylation detection in the bacterial reconstitution system. Selection was based on the percentages of proteins with different scores (high, medium, low, and none) in the GPS-SUMO prediction. The selected genes were cloned into *pCDF-Duet-1* with a Myc tag (the primer information is included in Supplemental Table 7) and detected in the reconstitution system (+E2 candidates are shown in Figure 1F and Supplemental Figure 1D; –E2 candidates are shown in Supplemental Figure 2C; ±E2 candidates are shown in Supplemental Figure 2D). As a result, 95% (19/20) of the selected +E2 candidates were modified by SUMO1aGG in the presence of E2 (the left graph in Figure 1G), whereas only 10% (1/10) of the selected –E2 candidates were SUMOylated in the

biochemical verification (the middle graph in Figure 1G), supporting the high specificity of our method. Interestingly, in the selected ±E2 candidates, 90% (9/10) of proteins were SUMOylated in the reconstituted system (the right graph in Figure 1G). Because some substrates non-covalently interact with SUMO for subsequent covalent conjugation, the ±E2 candidates may result from this type of substrate. Therefore, most proteins in both +E2 and ±E2 candidates were SUMOylation substrates, supporting the high reliability of our approach. From among these verified substrates, a candidate protein, transport protein SEC13 Homolog B (Zm00001d005461), was also fused with GFP and expressed in maize protoplasts. The immunoprecipitation result confirmed that SEC13 Homolog B–GFP was modified by SUMO1a in maize cells (Figure 1H). Thus, verification of SUMOylation on the candidates in bacteria and maize cells provided evidence that the current method is efficient, specific, and reliable for the identification of SUMOylation targets.

In summary, we have established a robust screening method to identify plant SUMOylation substrates based on a cDNA library in a bacterial reconstitution system. Our identification of novel maize SUMOylation targets will provide an important resource for further functional studies. Given that generation of transgenic plants and specific antibodies are not required, this approach can be extended to other plant species. Massive amounts of plant material must be used to identify endogenous SUMOylation substrates via affinity-based MS (Miller et al., 2010), but our method is based on a cDNA library and may be combined with single-cell and spatial transcriptomics to identify SUMOylation targets in a specific tissue and even in a single cell. Furthermore, because SUMO proteases are not present in *Escherichia coli*, the SUMO conjugates will be more stable in the current system for

(His$_6$–FLAG tag) and the E1 complex (ZmSAE1/ZmSAE2a), with or without E2 (ZmSCE1b). The SUMOylated proteins were enriched via a three-step (His$_6$–Myc–His$_6$) purification and subjected to MS analysis.

(B) SUMOylation of substrates expressed from the cDNA library in bacterial cells. The cDNA library was transformed into competent cells with or without E2 for SUMOylation detection. Representative immunoblots with anti-Myc (for substrates) and anti-FLAG (for SUMO1a) signals are shown.

(C) Purification of SUMOylated proteins for MS analysis. In the scale-up experiment, proteins were purified by a three-step (His$_6$–Myc–His$_6$) affinity approach. Before being subjected to MS identification, the enriched proteins were detected in immunoblots with anti-FLAG (left graph) or anti-Myc (middle graph) antibodies and in a silver staining gel (right graph).

(D) Venn diagram comparing the numbers of identified proteins. The numbers of proteins identified in the purification from bacterial cells without (in blue) or with E2 (in yellow) are shown. 212 proteins were uniquely identified in the sample with E2 (+E2 candidates), 24 proteins were specifically identified in the sample without E2 (–E2 candidates), and 72 proteins were identified in both samples with and without E2 (±E2 candidates). The data are from three biologically independent experiments (only the proteins identified in all three replicates are shown in the sample with E2; once a protein exists in one replicate of the sample without E2, it is shown in the diagram without E2).

(E) Biological process Gene Ontology analysis of the +E2 candidates.

(F) Verification of SUMOylation of selected +E2 candidates in the bacterial reconstitution system. Twenty proteins were selected from the +E2 candidates. SUMOylation of 20 candidates fused with a Myc tag was detected in the immunoblots using an anti-Myc antibody. Ten of them are shown in **(F)**, and the other 10 proteins are shown in Supplemental Figure 1D. The unmodified proteins are indicated by black triangles, and the SUMOylated forms are indicated by asterisks.

(G) A summary of the prediction and verification results of identified SUMOylation substrates. The proteins from +E2 candidates, –E2 candidates, and ±E2 candidates were analyzed with GPS-SUMO software. The numbers of proteins with high, medium, and low scores, as well as the number of proteins without predicted sites (none), are shown in the right graphs of Supplemental Figures 1C, 2A, and 2B. On the basis of these percentages, proteins (20 +E2 candidates, 10 –E2 candidates, and 10 ±E2 candidates) from the groups with different predicted scores were selected for verification. The summarized result indicated 95% (19/20) of selected proteins from the +E2 candidates were confirmed as SUMOylation substrates (11/12 in the high group, 4/4 in the medium group, 2/2 in the low group, and 2/2 in the none group); 10% (1/10) of selected proteins from the –E2 candidates were confirmed as SUMOylation substrates (0/7 in the high group, 1/2 in the medium group, 0/0 in the low group, and 0/1 in the none group); and 90% (9/10) of selected proteins from the ±E2 candidates were confirmed as SUMOylation substrates (6/6 in the high group, 1/2 in the medium group, 1/1 in the low group, and 1/1 in the none group).

(H) Verification of the SUMOylation of the candidate SEC13 homolog B (SEC13B; Zm00001d005461) in maize cells. GFP-fused SEC13B or free GFP was co-expressed with Myc–SUMO1aGG in maize protoplasts, and the cells were collected for immunoprecipitation using an anti-GFP antibody; the input and precipitated samples were analyzed in immunoblots with anti-GFP or anti-Myc antibodies.

subsequent purification. Compared with the efficient method based on transgenic plants in Arabidopsis (Rytz et al., 2018), cDNA library size and colony number for protein expression may be further scaled up in our approach to catch more SUMOylation substrates. Because our assay is performed in bacterial cells, the candidates need to be verified in plant cells before further functional characterization. Given that SUMOylation is conserved in eukaryotic cells, our method may also be used for the identification of SUMOylation targets in other species including humans.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

AUTHOR CONTRIBUTIONS

C.Y. and J.L. designed the project and supervised the research; R.L. and W. Li conducted experiments; Z.X., W. Liu, Q.Z., W. Lin, and J.J. provided technological support; R.L., W. Li, C.Y., and J.L. analyzed the data and wrote the manuscript.

Ruiqiang Lai[1,2,3], Wenliang Li[1,3], Zewei Xu[1],
Wen Liu[1], Qianrui Zeng[1], Wenxiong Lin[1],
Jieming Jiang[1], Jianbin Lai[1,*] and
Chengwei Yang[1,*]

[1]Guangdong Provincial Key Laboratory of Biotechnology for Plant Development, School of Life Science, South China Normal University, Guangzhou 510631, China
[2]Guangzhou Key Laboratory for Research and Development of Crop Germplasm Resources, College of Agriculture & Biology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[3]These authors contributed equally to this article.
*Correspondence: Jianbin Lai (20141062@m.scnu.edu.cn), Chengwei Yang (yangchw@scnu.edu.cn)
https://doi.org/10.1016/j.xplc.2023.100573

REFERENCES

Augustine, R.C., York, S.L., Rytz, T.C., and Vierstra, R.D. (2016). Defining the SUMO system in maize: SUMOylation is up-regulated during endosperm development and rapidly induced by stress. Plant Physiol. **171**:2191–2210. https://doi.org/10.1104/pp.16.00353.

Chen, J., Kalinowska, K., Müller, B., Mergner, J., Deutzmann, R., Schwechheimer, C., Hammes, U.Z., and Dresselhaus, T. (2018). DiSUMO-LIKE interacts with RNA-Binding proteins and affects cell-cycle progression during maize embryogenesis. Curr. Biol. **28**:1548–1560.e5. https://doi.org/10.1016/j.cub.2018.03.066.

Elrouby, N., and Coupland, G. (2010). Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify Arabidopsis proteins implicated in diverse biological processes. Proc. Natl. Acad. Sci. USA **107**:17415–17420. https://doi.org/10.1073/pnas.1005452107.

Miller, M.J., Barrett-Wilt, G.A., Hua, Z., and Vierstra, R.D. (2010). Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in Arabidopsis. Proc. Natl. Acad. Sci. USA **107**:16512–16517. https://doi.org/10.1073/pnas.1004181107.

Miller, M.J., Scalf, M., Rytz, T.C., Hubler, S.L., Smith, L.M., and Vierstra, R.D. (2013). 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.

Miura, K., and Hasegawa, P.M. (2010). Sumoylation and other ubiquitin-like post-translational modifications in plants. Trends Cell Biol. **20**:223–232. https://doi.org/10.1016/j.tcb.2010.01.007.

Okada, S., Nagabuchi, M., Takamura, Y., Nakagawa, T., Shinmyozu, K., Nakayama, J.i., and Tanaka, K. (2009). Reconstitution of Arabidopsis thaliana SUMO pathways in E. coli: functional evaluation of SUMO machinery proteins and mapping of SUMOylation sites by mass spectrometry. Plant Cell Physiol. **50**:1049–1061. https://doi.org/10.1093/pcp/pcp056.

Rosa, M.T., and Abreu, I.A. (2019). Exploring the regulatory levels of SUMOylation to increase crop productivity. Curr. Opin. Plant Biol. **49**:43–51. https://doi.org/10.1016/j.pbi.2019.04.009.

Rytz, T.C., Miller, M.J., McLoughlin, F., Augustine, R.C., Marshall, R.S., Juan, Y.T., Charng, Y.Y., Scalf, M., Smith, L.M., and Vierstra, R.D. (2018). SUMOylome profiling reveals a diverse array of nuclear targets modified by the SUMO ligase SIZ1 during heat stress. Plant Cell **30**:1077–1099. https://doi.org/10.1105/tpc.17.00993.

Yates, G., Srivastava, A.K., and Sadanandom, A. (2016). SUMO proteases: uncovering the roles of deSUMOylation in plants. J. Exp. Bot. **67**:2541–2548. https://doi.org/10.1093/jxb/erw092.

Zhao, Q., Xie, Y., Zheng, Y., Jiang, S., Liu, W., Mu, W., Liu, Z., Zhao, Y., Xue, Y., and Ren, J. (2014). GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res. **42**:W325–W330. https://doi.org/10.1093/nar/gku383.

Supplemental information

A robust method for identification of plant SUMOylation substrates in a library-based reconstitution system

Ruiqiang Lai, Wenliang Li, Zewei Xu, Wen Liu, Qianrui Zeng, Wenxiong Lin, Jieming Jiang, Jianbin Lai, and Chengwei Yang

A robust method for identification of plant SUMOylation substrates in a library-based reconstitution system

Ruiqiang Lai,[1,2,3] Wenliang Li,[1,3] Zewei Xu,[1] Wen Liu,[1] Qianrui Zeng,[1] Wenxiong Lin,[1] Jieming Jiang,[1] Jianbin Lai,[1,4] Chengwei Yang[1,4]

[1]Guangdong Provincial Key Laboratory of Biotechnology for Plant Development, School of Life Science, South China Normal University, Guangzhou 510631, China

[2]Guangzhou Key Laboratory for Research and Development of Crop Germplasm Resources, College of Agriculture & Biology, Zhongkai University of Agriculture and Engineering, Guangzhou, 510225, China

[3]These authors contributed equally to this work.

[4]Corresponding authors: yangchw@scnu.edu.cn (C.Y.); 20141062@m.scnu.edu.cn (J.L.).

SUPPLEMENTARY METHODS AND FIGURES

SUPPLEMENTARY METHODS

Construction of the maize cDNA library

The seedlings of *Zea mays* cultivar (B73) were grown under a cycle with 16 h of light at 28°C and 8 h of dark at 25°C. The seedlings under a variety of heat stress treatments and different tissues from mature leaves, stems, roots, stamens, unfertilized ears of corn, and corn bracts from adult maize plants were collected. Detailed information on the material treatment is included in Table S1. The mixed samples were sent to Shanghai Biogene Biotech Company for RNA preparation and cDNA library construction. The mRNA was purified by Oligotex mRNA Kit (QIAGEN) and reverse transcription products were inserted into the second cloning site of the *pCDFDuet-1* vector with a 5×Myc tag (at N terminals) to generate a cDNA library for expression of maize proteins.

Reconstitution of a maize SUMOylation system in bacteria cells

The CDS of the E1 dimer genes (*ZmSAE1* and *ZmSAE2a*) were inserted into two multiple cloning sites of *pACYCDuet-1* and the fragment containing the expression cassettes of these two genes were amplified by PCR and cloned into *pET28a* to generate a *pET28a-ZmSAE1-ZmSAE2a* expression construct. The CDS of *ZmSUMO1aGG* (encoding the mature form of ZmSUMO1a) and E2 (*ZmSCE1b*) were cloned into *pACYCDuet-1* to obtain a *pACYC-His$_6$-FLAG-ZmSUMO1aGG-ZmSCE1b* (SUMO1a+E2) construct. The construct *pACYC-His$_6$-FLAG-ZmSUMO1aGG* (SUMO1a) was used as a control. The construct pairs were transformed into *E. coli* BL21(DE3) to generate the competent cells expressing the E1 dimer and SUMO1aGG (-E2) or the competent cells expressing the E1 dimer and SUMO1aGG with E2 (+E2). The High-efficiency Competent Cell Preps Kits (Sangon Biotech, Shanghai) were used for preparation. SUMO conjugation activity was verified by a well-characterized substrate AtADA2b.

Protein expression and purification

3 μL (550 ng.μL^{-1}) of generated *pCDFDuet-1* cDNA library was transformed into 300 μL of two types of competent cells (-E2 and +E2), respectively. The colonies were selected on the LB medium with three kinds of antibiotics (Kanamycin, Chloramphenicol, and Streptomycin). The colonies from each type of cells were respectively combined and incubated in 500 mL LB medium at 37°C. 2.5 mL of 24 mg.mL^{-1} IPTG was added when OD$_{600}$ reached 0.6-0.8, and the protein expression was induced at 25°C for 18 h. The bacteria cells were collected by centrifugation and resuspended in 45 mL of lysis buffer (1×PBS pH7.4, 10 mM imidazole, 100 μM PMSF), subsequently broken with high pressure (1000 Mpa) using JG-1A High-pressure Cells Press. After centrifugation at 4°C for 30 min, the supernatant was mixed with 1 mL of Ni-NTA Agarose (QIAGEN, 30210) and incubated in a rotator at 4°C for 4 h. Then the beads were spun down and rinsed 3 times using washing buffer (1×PBS pH7.4, 20 mM imidazole). 1 mL of elution buffer (1×PBS pH7.4, 250 mM imidazole) was used to release the purified proteins from Ni-NTA Agarose. The

elution was further incubated with 100 µL of anti-MYC nanobody magarose beads (Alpa-life, KTSM1306) at 4°C for 2 h. After spun down, the anti-Myc agarose was rinsed 3 times and the enriched proteins were eluted with 8 M urea and incubated again with 100 µL of Ni-NTA Agarose (QIAGEN, 30210) at room temperature for 1 h. The beads were rinsed 3 times and elution buffer (1×PBS pH=7.4, 250 mM imidazole) was used for the final release of SUMOylated proteins. Silver staining was performed using Protein Stains K (Sangon Biotech). Three biologically independent experiments of protein expression and purification were performed and the samples were sent to the Wininnovate Bio Company (Shenzhen, China) for MS identification.

Protein digestion with in-solution enzymes

Aliquots of proteins were mixed with 200 µL of 8 M urea in Nanosep Centrifugal Devices (PALL). The filter tubes were spun at 12,000 g at 20°C for 20 min. Then, 200 µL of 8 M urea solution with 10 mM DTT was added and maintained at 37°C for 2 h. After solution removal by centrifugation, 200 µL of 8 M urea with 50 mM iodoacetamide was added. The sample was incubated in the dark for 15 min at room temperature. The filter tube was washed with 200 µL of 8 M urea 3 times and 200 µL of 25 mM ammonium bicarbonate 3 times by centrifugation. Then, 100 µL of 25 mM ammonium bicarbonate containing 0.01 µg/µL trypsin was added and the tubes were incubated at 37°C for 12 h. The filter tubes were washed twice with 100 µL of 25 mM ammonium bicarbonate by centrifugation and the flow-through fractions were collected and lyophilized.

LC-MS/MS analysis

The lyophilized peptide fractions were resuspended in ddH$_2$O containing 0.1% formic acid, and 2 µL aliquots were loaded into a nanoViper C18 (Acclaim PepMap 100, 75µm×2 cm) trap column. Online Chromatography separation was performed on the Easy nLC 1200 system (ThermoFisher). The trapping and desalting procedures were carried out with a volume of 20 µL 100% solvent A (0.1% formic acid). Then, an elution gradient of 5-38% solvent B (80% acetonitrile, 0.1% formic acid) in 60 min

was used on an analytical column (Acclaim PepMap RSLC, 75μm×25 cm C18-2 μm 100 Å). DDA (data-dependent acquisition) MS techniques were used to acquire tandem MS data on a ThermoFisher Q Exactive mass spectrometer (ThermoFisher, USA) fitted with a Nano Flex ion source. Data were acquired using an ion spray voltage of 1.9 kV, and an interface heater temperature of 275°C. For a full MS survey scan, the target value was 3×106 and the scan ranged from 350 to 2,000 m/z at a resolution of 70,000 and a maximum injection time of 100 ms. For the MS2 scan, only spectra with a charge state of 2-5 were selected for fragmentation by higher-energy collision dissociation with a normalized collision energy of 28. The MS2 spectra were acquired in the ion trap in rapid mode with an AGC target of 8,000 and a maximum injection time of 50 ms. Dynamic exclusion was set for 25 s.

Database retrieval and analysis

The MS/MS data were analyzed for protein identification and quantification using PEAKS Studio 8.5. The local false discovery rate at PSM was 1.0% after searching against the Zea mays database with a maximum of two missed cleavages. The following settings were selected: Oxidation (M), Acetylation (Protein N-term), Deamidation (NQ), Pyro-glu from E, Pyro-glu from Q for variable modifications as well as fixed Carbamidomethylation of cysteine. Precursor and fragment mass tolerance were set to 10 ppm and 0.05 Da, respectively.

Bioinformatics analyses

For GO analysis, the Ensembl ID was converted to ENTREZID using the R software AnnotationHub package. Gene enrichment analysis was performed using the R software clusterProfiler package. The significant enrichment in the GO results was analyzed (p < 0.05). For the prediction of SUMOylation sites, the GPS-SUMO software (http://sumosp.biocuckoo.org/online.php) was used.

Detection of SUMOylation in maize protoplasts

ZmSUMO1aGG was cloned into a *35S:5×Myc pBluescript*-based vector for

expressing Myc-SUMO1aGG; the CDS of *SEC13B* (Zm00001d005461) was cloned into the vector （*pCambia1300221-UBQ: GFP* ） for expressing SEC13B-GFP. The indicated plasmid pairs were transformed into maize protoplasts by a PEG-mediated method. 14 h after transformation, the cells were collected for centrifugation at 60 g at 4°C for 10 min, protein extraction in immunoprecipitation (IP) buffer (50 mM Tris–HCl pH7.4, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 10% (v/v) glycerol, and 0.5 % (v/v) Nonidet P-40, 1×protease inhibitor (PI) cocktails (P9599; Sigma-Aldrich), 25 μM MG132 (HY-13259; MedChemExpress, Monmouth Junction, NJ, USA), and 20 mM N-ethylmaleimide (NEM, E3876; Sigma-Aldrich). After spun down at 18,000 g for 20 min, the supernatant was incubated with anti-GFP nanobody magarose beads (Alpa-life, KTSM1334) for 2 h at 4°C. The beads were rinsed 3 times with washing buffer (50 mM Tris–HCl pH7.4, 150 mM NaCl, 1 mM $MgCl_2$, 10% (v/v) glycerol, and 0.4% (v/v) Nonidet P-40), and 20 mM NEM) and then mixed with protein sample buffer and boiled for elution. The released proteins were subjected to SDS-PAGE and the immunoblots were detected using anti-GFP (HT801-01; TransGen Biotech) and anti-Myc (HT101-01; TransGen Biotech) antibodies.
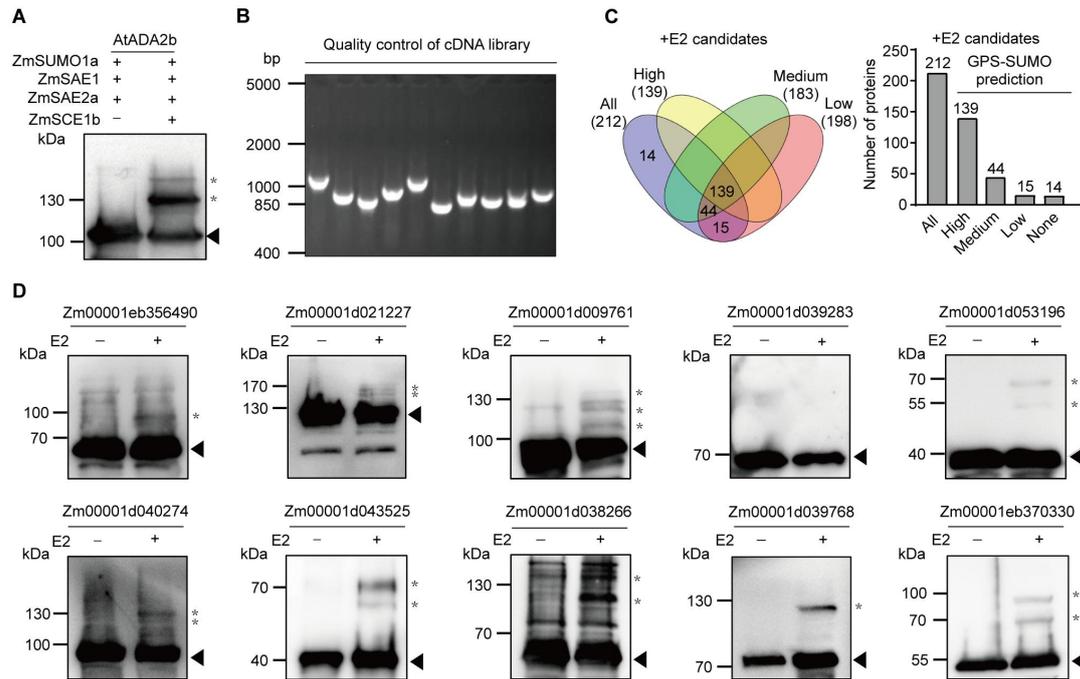
SUPPLEMENTARY FIGURE

Figure S1. The verification of supporting experiment data for the established method and SUMOylation detection of 10 other selected +E2 candidates. (A) Verification of the SUMO conjugation activity of the reconstituted system using a well-known Arabidopsis substrate AtADA2b. (B) Quality measurement of the generated cDNA library. The inserted fragments amplified by PCR were detected in agarose electrophoresis. Different lanes mean cDNA inserts from individual colonies. (C) The summary of SUMOylation site prediction via GPS-SUMO in the proteins uniquely identified in the sample with E2 (+E2 candidates). The numbers of candidates with high, medium and low scores are shown in the left graph. The non-overlapping numbers are shown in the right graph. None: proteins without predicted sites. (D) Confirmation of the selected +E2 candidates in the reconstituted system. SUMOylation of 20 candidates fused with a Myc tag was detected in the immunoblots using an anti-Myc antibody. Ten of them are shown and the other 10 proteins are shown in Figure 1F. The unmodified proteins are indicated by black triangles, the SUMOylated forms are indicated by asterisks.
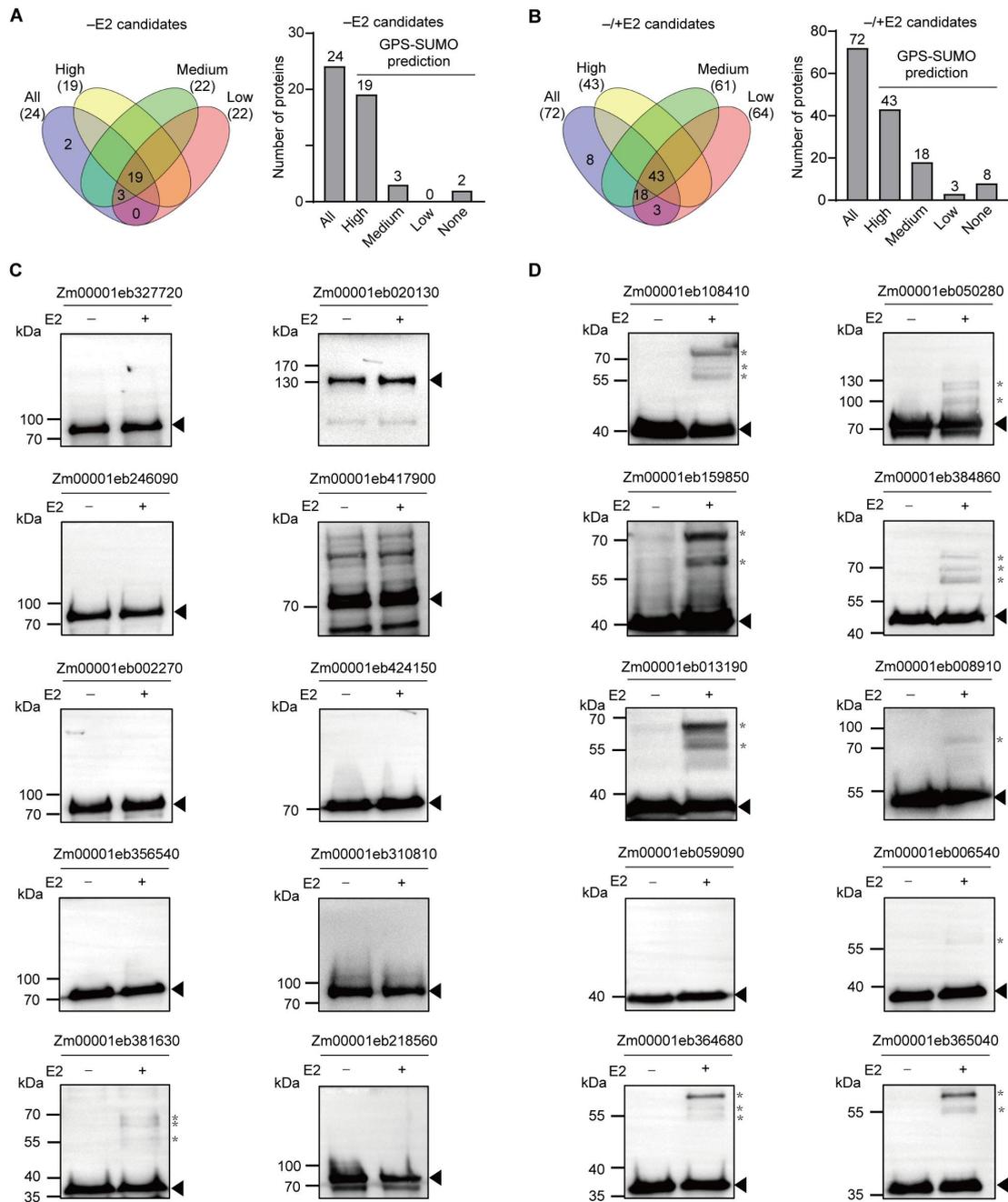
Figure S2. Prediction and biochemical verification of –E2 and –/+E2 candidates.
(A, B) The summary of SUMOylation site prediction via GPS-SUMO from the proteins specifically identified in the sample without E2 (–E2 candidates) in (A) and from the proteins identified in both samples with and without E2 (–/+E2 candidates) in (B). The numbers of candidates with high, medium and low scores are shown in the left graph. The non-overlapping numbers are shown in the right graph. None: proteins

without predicted sites. (C, D) Verification of the identified substrates from the –E2 and –/+E2 candidates in the reconstituted system. SUMOylation of 10 –E2 candidates (C) and 10 –/+E2 candidates (D) fused with a Myc tag was detected in the immunoblots using an anti-Myc antibody. The unmodified proteins are indicated by black triangles, the SUMOylated forms are indicated by asterisks.