Quantitative ubiquitylome analysis reveals specificity of RNF111/Arkadia E3 ubiquitin ligase for its degradative substrates SKI and SKIL/SnoN in TGF-β signaling pathway

Victor Laigle, Florent Dingli, Sadek Amhaz, Tiphaine Perron, Mouna Chouchène, Sabrina Colasse, Isabelle Petit, Patrick Poullet, Damarys Loew, Céline Prunier, et al.

Journal Pre-proof

Quantitative ubiquitylome analysis reveals specificity of RNF111/Arkadia E3 ubiquitin ligase for its degradative substrates SKI and SKIL/SnoN in TGF-β signaling pathway

Victor Laigle, Florent Dingli, Sadek Amhaz, Tiphaine Perron, Mouna Chouchène, Sabrina Colasse, Isabelle Petit, Patrick Poullet, Damarys Loew, Céline Prunier, Laurence Levy

PII: S1535-9476(21)00145-6

DOI: https://doi.org/10.1016/j.mcpro.2021.100173

Reference: MCPRO 100173

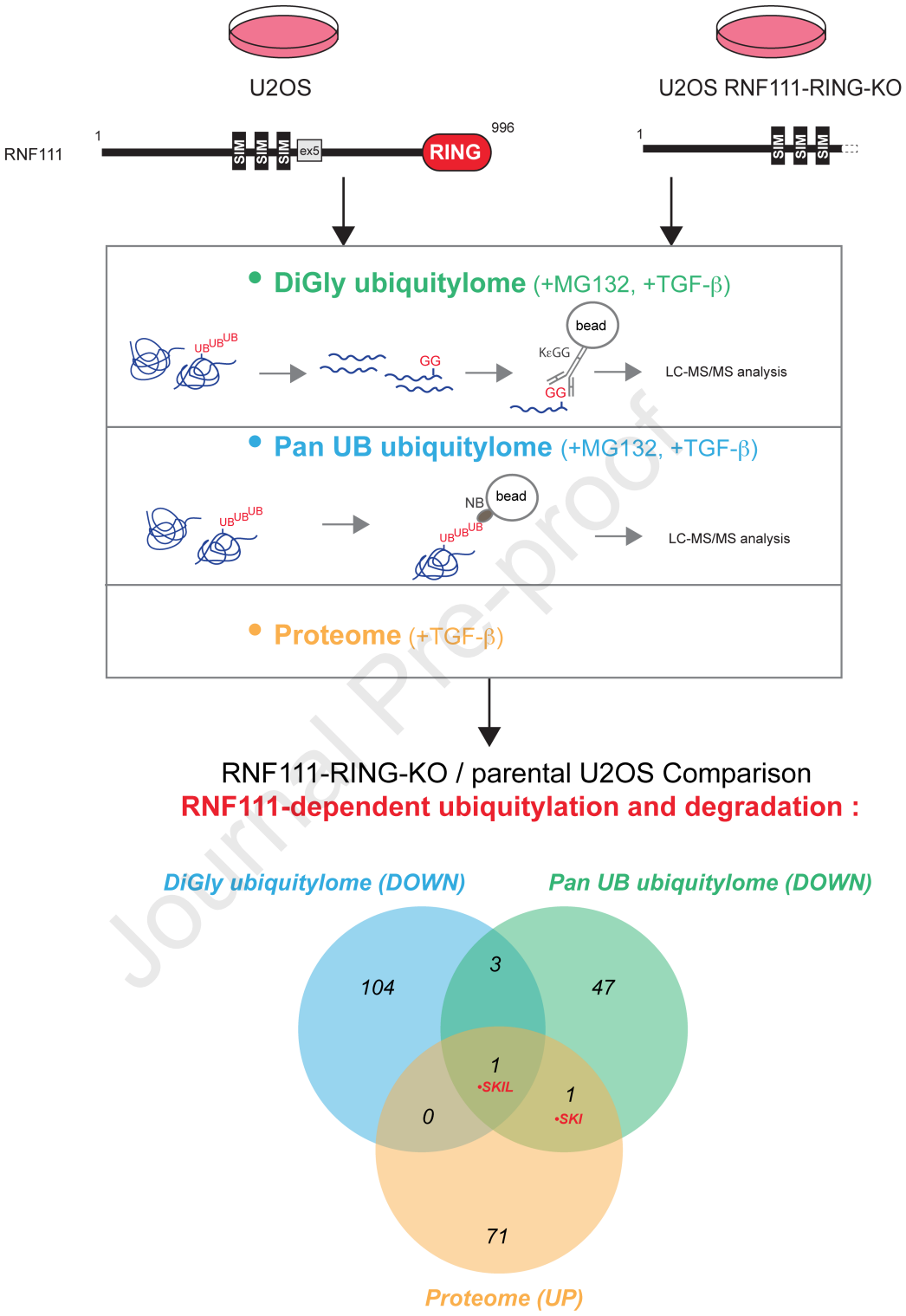
To appear in: *Molecular & Cellular Proteomics*

Received Date: 14 June 2021

Revised Date: 6 October 2021

Accepted Date: 1 November 2021

Please cite this article as: Laigle V, Dingli F, Amhaz S, Perron T, Chouchène M, Colasse S, Petit I, Poullet P, Loew D, Prunier C, Levy L, Quantitative ubiquitylome analysis reveals specificity of RNF111/ Arkadia E3 ubiquitin ligase for its degradative substrates SKI and SKIL/SnoN in TGF-β signaling pathway, *Molecular & Cellular Proteomics* (2021), doi: https://doi.org/10.1016/j.mcpro.2021.100173.



grahical abstract

TITLE: Quantitative ubiquitylome analysis reveals specificity of RNF111/Arkadia E3 ubiquitin ligase for its degradative substrates SKI and SKIL/SnoN in TGF-β signaling pathway

RUNNING TITLE : RNF111 substrates identified by integrative proteomics

AUTHORS : Victor Laigle [1], Florent Dingli [1], Sadek Amhaz [2], Tiphaine Perron [2], Mouna Chouchène [2], Sabrina Colasse [2], Isabelle Petit [2], Patrick Poullet [3], Damarys Loew [1], Céline Prunier [2], Laurence Levy [2]*

[1]Institut Curie, PSL Research University, Laboratoire de Spectrométrie de Masse Protéomique, 75005 Paris, France.

[2] Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, CRSA, F-75012 Paris, France.

[3] Institut Curie, PSL Research University, INSERM U900 France,

*Corresponding author: laurence.levy@inserm.fr

ABBREVIATIONS : TGF-β (Transforming Growth Factor-β); USP (ubiquitin-proteasome system); diGly (diGlycine); UB (Ubiquitin)

ABSTRACT

RNF111/Arkadia is an E3 ubiquitin ligase that activates the TGF-β pathway by degrading transcriptional repressors SKIL/SnoN and SKI, and truncations of the RING C-terminal domain of RNF111 that abolish its E3 function and subsequently TGF-β signaling are observed in some cancers. In the present study, we sought to perform a comprehensive analysis of RNF111 endogenous substrates upon TGF-β signaling activation using an integrative proteomic approach. In that aim we carried out label free quantitative proteomics after enrichment of ubiquitylated proteins (ubiquitylome) in parental U2OS cell line compared to U2OS CRISPR engineered clones expressing a truncated form of RNF111 devoid of its C-terminal RING domain. We compared two methods of enrichment for ubiquitylated proteins prior to proteomics analysis by mass spectrometry, the diGly remnant peptide immunoprecipitation with a K-ε-GG antibody (diGly) and a novel approach using protein immunoprecipitation with a ubiquitin pan nanobody (pan UB) that recognizes all ubiquitin chains and monoubiquitylation. While we detected SKIL ubiquitylation among 108 potential RNF111 substrates with the diGly method, we found that the pan UB method also constitutes a powerful approach since it enabled detection of 52 potential RNF111 substrates including SKI, SKIL and RNF111. Integrative comparison of the RNF111-dependent proteome and ubiquitylomes enabled identification of SKI and SKIL as the only targets ubiquitylated and degraded by RNF111 E3 ligase function in presence of TGF-β. Our results indicate that lysine 343 localized in the SAND domain of SKIL constitutes a target for RNF111 ubiquitylation and demonstrate that RNF111 E3 ubiquitin ligase function specifically targets SKI and SKIL ubiquitylation and degradation upon TGF-β pathway activation.

INTRODUCTION

The ubiquitin-proteasome system (UPS) plays an important role in the regulation of many cellular signaling pathways by controlling protein stability. The UPS involves ubiquitylation of proteins by E3 ubiquitin ligases that allow covalent attachment of the ubiquitin protein to a lysine residue on specific substrates, in cooperation with an E1 activating enzyme and an E2 conjugating enzyme. Polyubiquitylation can occur by polymerization of the ubiquitin molecules via one of its 7 lysine residues (K6, K11, K27, K29, K33, K48, K63) or its N-terminal methionine (M1), which generates as many different polyubiquitin linkages (1). This ubiquitylation code leads to distinct biochemical outcomes for the substrate, and it is admitted that only K48 polyubiquitylation, and to a lesser extent K11 polyubiquitylation, drive substrates towards degradation by the proteasome.

TGF-β pathway plays an important role in embryonic development and in tumor progression by inducing a large panel of target genes involved in cell cycle arrest, epithelial-mesenchymal transition and cell migration mainly through activation of the SMAD2/3-SMAD4 transcription complex. The TGF-β signaling pathway is highly regulated by various E3 ubiquitin ligases such as SMURF1/2, TRIM33, WWP1, and RNF111 (also named Arkadia) (2, 3). We and others have found that RNF111 harbors a C-terminal RING domain required for its E3 ubiquitin ligase function that activates SMAD-dependent transcription in response to TGF-β by inducing degradation of SKI and SKIL (also named SnoN) transcriptional repressors (4–6). Whereas the RNF111-dependent SKI and SKIL degradation induced by TGF-β is clearly established, the mechanism for this inducible degradation is still puzzling, in particular the ubiquitylation events that underlie this degradation. RNF111 has also been reported to regulate the stability of SMAD7, an inhibitor of TGF-β signaling that acts at the TGF-β receptor level (7). Mutations that disrupt the C-terminal RING domain of RNF111 can

occur in cancer (8), and we have shown that the NCI-H460 lung cancer cell line exhibits a S432* nonsense mutation that leads to the expression of a truncated form of RNF111 lacking its C-terminal RING domain. Such truncation abolishes SKI and SKIL degradation and subsequent SMAD-dependent transcription in response to TGF-β in this cancer cell line (9). While RNF111 has mainly been involved in the activation of TGF-β signaling, its E3 ubiquitin ligase function is not restricted to this pathway. RNF111 also contains 3 SUMO interacting motifs in its N-terminal region which confer to RNF111 a SUMO-Targeted Ubiquitin Ligase (STUBL) function involved in PML degradation in response to arsenic treatment (10) and in XPC ubiquitylation during DNA damage repair induced by UV (11). It has also been proposed that RNF111 is involved in Histone H4 neddylation during DNA damage repair induced by ionizing radiation (12), and in endocytosis by targeting the micro2 subunit of Clathrin adaptor 2 (AP2) complex (13). Hence, RNF111, like most E3 ligases, might target different substrates involved in different biological processes. However all these substrates were characterized by protein interaction approaches, which are not the most relevant considering that E3 ubiquitin ligases interaction with their substrates tend to be labile and could lead to substrate degradation. Moreover, in most studies, ubiquitylation was detected by overexpression of RNF111 and ubiquitin, which could lead to forced ubiquitylation. To prevent such biases, in the present study, we have sought to use an endogenous approach to comprehensively identify the substrates of RNF111. Since it represents a small proportion of a protein pool in the cell, the ubiquitylated proteins can be challenging to detect at the endogenous level. However, in the past years, different methods of enrichment for ubiquitylated proteins have been developed that allow profiling of ubiquitylated proteins by mass spectrometry (ubiquitylome) (14–16). The breakthrough came with the use of K-ε-GG antibody that immmunoprecipitates the di-glycin (diGly)

remnant peptides obtained after trypsic digestion of the ubiquitin linked to its targeted lysine on a substrate (17, 18). Yet this method requires a large amount of starting material since it only focuses on specific peptides at ubiquitylation sites and not the whole proteins. Thus, development of alternative approaches is still needed to increase sensitivity and to simplify ubiquitylome analysis.

In order to profile substrates of RNF111, we have generated U2OS osteosarcoma CRISPR modified cell lines that express a truncated form of RNF111 devoid of the C-terminal RING domain (RNF111-RING-KO) that mimics the truncation observed in the cancer cell line NCI-H460. To detect degradative substrates of RNF111, we have performed label free quantitative proteomics to compare the proteome of the RNF111-RING-KO clones to the parental U2OS upon TGF-β induction, which enabled detection of SKI and SKIL among 73 candidates. Further analysis of selected candidates indicates that regulation of their protein level by RNF111 occurs at the transcriptional level. To identify more precisely ubiquitylated substrates of RNF111 we performed label free quantitative comparison of the ubiquitylome of RNF111-RING-KO clones to the parental U2OS upon TGF-β induction by two means. We used diGly enrichment with the K-ε-GG antibody and developed a method using a pan UB nanobody that strongly interacts with monoubiquitinated and all linkage polyubiquitinated proteins. With the diGly antibody approach, among the 3641 proteins corresponding to the ubiquitylation sites quantified, we identified 108 proteins that are potentially ubiquitylated by RNF111, including SKIL on lysine 343; while the pan UB nanobody approach enabled the detection of 54 potential substrates including SKI, SKIL and RNF111 among the 8547 proteins quantified, demonstrating that this new method is very robust for substrate identification of E3 ubiquitin ligases. Moreover comparison of the two ubiquitylomes leads to detection of SKIL as the only validated common RNF111 substrate, and integrative

comparison of the ubiquitylomes and proteome identified SKI and SKIL as the only proteins both ubiquitylated and degraded by RNF111 upon TGF-β pathway activation, among the 7746 proteins quantified in the proteome analysis. Altogether, our findings indicate a strong specificity of RNF111 E3 ubiquitin ligase function for degradative ubiquitylation of SKI and SKIL in response to TGF-β.

EXPERIMENTAL PROCEDURES

Cell lines and plasmids

U2OS human osteosarcoma, HEK-293 human embryonic kidney, and NCI-H460 human non-small cell lung carcinoma cell lines were cultured in DMEM (U2OS, HEK-293) or RPMI (NCI-H460) medium containing 10% fetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin at 37°C in 5% CO_2. The pMLM3636 expression vector for S.pyogenes Cas9 sgRNA was a gift from JK Joung laboratory (Addgene plasmid # 43860). $Cas9^{D10A}$ expression vector was a gift from G Church (addgene plasmid #41816) (19), the $CAGA_{12}$-Luc plasmid was described in (20) and the Renilla expression vector pRL-TK vector is from Promega. The pCMV10-3xFlag-RNF111-WT (Flag-RNF111-WT) expression vector was generated by PCR subcloning of human RNF111 cDNA (corresponding to isoform 3) from PcDNA4/TO-SFS-RNF111 (11) in PCMV10-3xFlag (Invitrogen). The pCMV-3xHA-SKIL-WT (HA-SKIL-WT) expression vector was generated by PCR subcloning of human SKIL cDNA form PCMV5B-HA-SnoN (21) in PCMV-3xHA. PCMV10-3xFlag-RNF111-C933A (Flag-RNF111-C933A) and pCMV-3xHA-SKIL-342/43-KR (HA-SKIL-342/43-KR) mutants were generated by site-directed mutagenesis respectively on pCMV10-3xFlag-RNF111-WT and pCMV-3xHA-SKIL-WT by using the QuickChange Lightning kit (Agilent).

CRISPR cell lines

SgRNA-rev CACTGTGGAAGGTTGGCTAC and SgRNA-fw CTTACAAGCAATAGTACCAC targeting exon 5 of the human RNF111 gene were designed using the CRISPOR software (crispor.org) (22) in order to perform double-nicking. Double-stranded oligonucleotides with overhangs were cloned into BsmBI digested MLM3636 vector and 0.5 million of U2OS cells were co-transfected with 2µg of $Cas9^{D10A}$ expression vector and 2µg of each SgRNA-rev and SgRNA-fw MLM3636 expressing vector, using Amaxa Nucleofector V kit (Lonza), program X-001. Single cells were individually seeded in 96 well plates and clones were amplified and assessed by western-blot for full length RNF111 depletion. RNF111 Exon5 targeted region was PCR amplified from clone#1 and #2 genomic DNA with primers CATCTACCTCTGAGCAGGCC and TCATGCTTTTGGTGTCAGCC and PCR products were subcloned into pCR2.1 vector using TOPO-TA Cloning kit (Invitrogen). For each CRISPR clones, a total of 10 cloned PCR products were sequenced in order to determine the genomic modification on the different alleles.

Immunoprecipitation and Western blot

Whole-cell extracts were prepared from 6-well plates and treated or not for 1h with TGF-β (2ng/ml) before lysis with RIPA buffer (50 mM Tris [pH 8], 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate) supplemented with EDTA-free protease Inhibitor (Roche), 50µM NaF and 50µM β-glycerophospate. Cleared lysates were quantified by BCA protein assay (Pierce) and 30µg of proteins were analyzed by western blotting using standard procedures.

For pan UB nanobody immunoprecipitation on endogenous proteins, cells were grown to 90% confluence in a 150 mm plates for each condition and treated with 10µM MG132 for 4h followed by 1h TGF-β (2 ng/ml) treatment before lysis in RIPA buffer. Lysates were sonicated (10 seconds ON, 10 seconds OFF, four times) and cleared lysates were quantified

using BCA quantification. 6 mg of proteins were immunoprecipitated with 50µl of Ubiquitin Pan Selector beads slurry (Nanotag Biotechnologies, #N2510) on a rotator for 1h at 4°C. After 3 washes with RIPA buffer, proteins were eluted in laemmli buffer before subsequent analysis by western blotting using standard procedures.

For pan UB nanobody immunoprecipitation on transfected cells, cells grown in 6-well plates were transfected with 2µg of the appropriate plasmids using X-tremGENE HP (Roche) and were treated 24h hour later with 10µM MG132 for 4h followed by 1h with 2 ng/ml TGF-β (U2OS) or 20 ng/ml Activin A (HEK-293) before lysis in RIPA buffer. Lysates were immunoprecipitated with 20µl of Nanotag Ubiquitin Pan Selector resin for 1h and washed 3 times with RIPA buffer before analysis by western blotting using standard procedures.

The following antibodies were used for Western blotting: anti-Flag-HRP (Sigma), anti-hemagglutinin (anti-HA-HRP; Roche), anti-RNF111 (M05, Abnova), anti-SKI (G8, Santa Cruz), anti-SKIL (19218-1-AP, Proteintech), anti-SMAD2/3 (BD), anti-UB (P4D1, Santa Cruz), anti-KYNU (E5, Santa Cruz), anti-GDF15 (G5, Santa Cruz), anti-FABP3 (Proteintech).

Luciferase assay

For luciferase assay, cells grown in 24-well plates were cotransfected with 0.3 µg $CAGA_{12}$-Luc and 0.2 µg pRL-TK (Promega). In the -TGF-β condition, 10 µM of TGF-β inhibitor (SB-431542, Torcis) was added at the time of transfection to inhibit autocrine TGF-β signaling in NCI-H460 cells. 24 h post-transfection, TGF-β (2ng/ml) was added for 8 h before lysis in passive lysis buffer (Promega) and successive measurements of Luciferase and Renilla activity with the dual-luciferase reporter assay system (Promega) were performed. Luciferase activities were normalized to Renilla activities in triplicate experiments.

Quantitative RT-PCR (qRT-PCR)

Total RNA was extracted with Trizol (Invitrogen) according to standard procedure from cells grown at 90% confluence in a 10 mm dish. cDNA were synthetized from 1.5 µg of RNA using the iScript cDNA synthesis kit (Bio-Rad). QPCR was performed in triplicate using the 2XSYBR Green qPCR master mix (Biotools) according to the manufacturer's protocol in a Light Cycler 96 (Roche). Expression of each gene was calculated by the $2^{-\Delta\Delta Ct}$ methods using GAPDH as a control. All data represent mean +/- standard deviation for at least 3 independent experiments. The following primers were used: GAPDH-F TGCACCACCAACTGCTTAGC, GAPDH-R GGCATGGACTGTGGTCATGAG, GDF15-F ACTCACGCCAGAAGTGCGG, GDF15-R AGATTCTGCCAGCAGTTGGTC, FABP3-F CTTCCCCCTACCCTCAGGTG, FABP3-R CAGTGTCACAATGGACTTGACC, SKIL-F CAGCCTGATGCTCCGTGTAT, SKIL-R TGATGGTGCATCTGTCTTGGA, KYNU-F TTGCGGCTGAACTCAAATGC, KYNU-R GCTTCCCCACTTCATGACCA.

Ubiquitylome and proteome sample preparation

For proteome analysis, cells were grown to 90% confluence in a 6 well plates and treated with 2 ng/ml TGF-β for 1h before lysis in freshly prepared urea buffer (8M urea, 200mM ammonium bicarbonate, EDTA-free protease Inhibitor). After sonication, lysates were quantified by BCA and 300µg of proteins were reduced with 5 mM dithiothreitol (DTT) for 1h at 37 °C and alkylated with 10 mM iodoacetamide (IAA) for 30 min at room temperature in the dark. Samples were then diluted in 200mM ammonium bicarbonate (ABC) to reach a final concentration of 1 M urea and digested overnight at 37 °C with Trypsin (Worthington #LS003750) at a ratio of 1/50. 150 µg of each sample were separated with the High pH Reversed-Phase peptide fractionation kit (Pierce #84868). Peptides were eluted successively into six fractions using elution buffers containing the following percentages of acetonitrile: 10, 12.5, 15, 17.5, 20 and 50%. Eluted peptides were vacuum concentrated to

dryness and resuspended in 20 µl of 0.1% formic acid (FA) / 3% acetonitrile (CH3CN) (vol/vol) prior to LC-MS/MS analysis.

For diGly ubiquitylome analysis, samples were prepared according to the protocol described in (18). Briefly, for each condition, cells were grown to 90% confluence in 8 x 150 mm plates and treated with 10µM MG132 for 4h followed by 1h TGF-β (2 ng/ml) treatment before lysis in freshly prepared urea buffer (8M urea, 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 2 µg/ml aprotinin, 10 µg/ml leupeptin, 50 µM PR-619, 1 mM chloroacetamide, 1 mM PMSF). Cleared lysates were quantified by BCA and 10mg of proteins were reduced and alkylated by adding successively 5mM DTT and 10mM IAA respectively for 1h and 30 min at room temperature. Protein samples were diluted 4 times with 50mM Tris-HCL to obtain a concentration at 2 M urea. Trypsin (Worthington # LS003750) digestion was then performed overnight at 37°C at a ratio of 1/50. After centrifugation at 3000g for 5 min, supernatant containing the digested peptides were desalted on a 500-mg tC18 SepPak cartridge and eluted peptides were dried by vacuum centrifugation. Peptides were resuspended in 1.4 ml IAP buffer (50 mM MOPS (pH 7.2), 10 mM sodium phosphate $Na2HPO_4$, 50 mM NaCl), cleared at 20000g for 5 min, and incubated on a rotator for 2h at 4°C with 50µl of PTMScan® Ubiquitin Remnant Motif (K-ε-GG) Antibody Beads slurry (Cell Signaling #5562) equilibrated in IAP buffer. After 2 washes in IAP buffer followed by 3 washes with milliQ water, two successive elutions of the K-ε-GG peptides with 55 µl of TFA 0.15% for 10min were combined and desalted using C18 StageTips. The final Peptide eluates were dried and resuspended in 8µl 0.1% FA/3% (vol/vol) CH3CN prior to LC-MS/MS analysis.

For pan UB nanobody Ubiquitylome, immunoprecipitation on endogenous proteins was performed as described in the immunoprecipitaiton section, except that the 3 washes with

RIPA buffer were followed by 2 washes with washing buffer (150 mM NaCL, 50 mM TRIS pH 7.5) and proteins were subsequently eluted twice with 150µl of freshly prepared solution of 1.4% triethylamide (TEA) for 5 min at room temperature under agitation. The 300µl TEA eluates were neutralized with 100µl of Tris 1M pH 7.5 and dried by vacuum centrifugation. Proteins were reduced and alkylated by adding successively 5mM DTT and 10mM IAA as previously described. Samples were then diluted in 400µl of 25mM ABC, digested at 37°C for 2h with 0.4µg of trypsin/LysC (#V5073 Promega) before overnight digestion by adding 1µg of trypsin/LysC. Sample were then loaded onto homemade SepPak C18 Tips packed by stacking one AttractSPE® disk (#SPE-Disks-Bio-C18-100.47.20, Affinisep) and 2mg beads (SepPak C18 #186004521, Cartridge Waters) into a 200 µL micropipette tip for desalting. Peptides were eluted using 40/60 MeCN/H2O in 0.1% FA and vacuum concentrated to dryness. Sample were resuspended in 10µl of TFA 0.3% prior to LC-MS/MS analysis

Ubiquitylome and proteome analysis by LC-MS/MS

Peptides for proteome analyses were separated by reversed phase liquid chromatography (LC) on an RSLCnano system (Ultimate 3000, Thermo Scientific) coupled online to an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Peptides were trapped on a C18 column (75 µm inner diameter × 2 cm; nanoViper Acclaim PepMap^{TM} 100, Thermo Scientific) with buffer A (2/98 MeCN/H2O (vol/vol) in 0.1% FA) at a flow rate of 4.0 µL/min over 4 min. Separation was performed on a 50 cm x 75 µm C18 column (nanoViper Acclaim PepMap^{TM} RSLC, 2 µm, 100Å, Thermo Scientific) regulated to a temperature of 55°C with a linear gradient of 5 to 25 % buffer B (100% MeCN, 0.1% FA) at a flow rate of 300 nL/min over 100 min. Peptides were ionized by a nanospray ionization (NSI) ion source at 2.2 kV. Full-scan MS in the Orbitrap was set at a scan range of 400-1500 with a resolution at 120,000 (at 200 m/z) and ions from each full scan were fragmented in higher-energy

collisional dissociation mode (HCD) and analyzed in the linear ion trap in rapid mode. The fragmentation was set in top speed mode in data-dependent analysis (DDA). We selected ions with charge state from 2+ to 6+ for screening. Normalized collision energy (NCE) was set to 30, automatic gain control (AGC) target to 20,000 ions with a dynamic exclusion of 30s.

For diGly ubiquitylome analyses, LC was performed as previously with an RSLCnano system (same trap column, column and buffers), coupled online to a Q Exactive HF-X mass spectrometer (Thermo Scientific). Peptides were trapped onto the C18 column with buffer A at a flow rate of 2.5 µL/min over 4 min. Separation was performed at a temperature of 50°C with a linear gradient of 2 to 30% buffer B at a flow rate of 300 nL/min over 91 min. Peptides were ionized by a NSI ion source (voltage was 2.2 kV). MS full scans were performed in the ultrahigh-field Orbitrap mass analyzer in ranges m/z 375–1500 with a resolution of 120 000 (at 200 m/z) and detected in the Orbitrap analyzer after accumulation of ion at 3E6 target value with a maximum injection time (IT) of 50ms. For every full scan, the top 20 most intense ions were isolated (isolation width of 1.6 m/z) and fragmented (NCE of 27) by HCD with an IT of 60ms, AGC target set to 1E5, and 15 000 resolution. Charge state from <2+ and >6+ were excluded, and dynamic exclusion was set to 40s.

For pan UB ubiquitylome analyses, LC was performed as previously with an RSLCnano system coupled online to an Orbitrap Exploris 480 mass spectrometer (Thermo Scientific). Peptides were trapped on a C18 column with buffer A at a flow rate of 3.0 µL/min over 4 min. Separation was performed at a temperature of 40°C with a linear gradient of 3% to 32% buffer B at a flow rate of 300 nL/min over 211 min. MS full scans were performed in the ultrahigh-field Orbitrap mass analyzer in ranges m/z 375–1500 with a resolution of 120 000 at m/z 200, AGC target value set at 300 % and with a maximum IT of 25ms. The top

30 most intense ions were isolated (isolation width of 1.6 m/z) and fragmented with a NCE set at 30%, a resolution of 15 000 and AGC target value set to 100%. We selected ions with charge state from 2+ to 6+ for screening and dynamic exclusion of 40s.

Mass Spectrometry Data Analysis

For identification, the raw MS files were searched against the Homo sapiens UniProt database (UP000005640, downloaded 11/2017 with 20239 entries for the diGly ubiquitylome, 01/2018 with 20231 entries for the proteome, and 12/2019 with 20364 entries for the pan UB ubiquitylome), combined with common contaminants (245 sequences, downloaded from http://www.coxdocs.org/doku.php?id=maxquant:start_downloads.htm the 27/07/2016) for the diGly ubiquitylome analyses. The proteome and pan UB ubiquitylome samples being sufficiently complex, the non-human contaminants were negligible and were not added to the search. The search was conducted using SEQUEST-HT through Proteome Discoverer (version 2.1 for the diGly ubiquitylome, 2.2 for the proteome and 2.4 for the pan UB ubiquitylome) after the Spectrum Selector node with default settings. Enzyme specificity was set to trypsin (full) and a maximum of two miscleavage sites were allowed for the proteome and pan UB ubiquitylome and three for the diGly ubiquitylome. Oxidized methionine, Carbamidomethyl cysteines, N-terminal acetylation, were set as variable modifications and GlyGly on lysine (+114.0429) was added for the ubiquitylomes analyses. Methionine-loss and Methionine-loss + N-terminal acetylation were also added to the pan UB variable modifications. For all analyses, the maximum allowed mass deviation was set to 10 ppm for monoisotopic precursor ions. For fragment ions, it was set respectively to 0.6 Da and 0.02 Da for the proteome and the ubiquitylomes. For proteome analyses the Top N peaks filter of Proteome Discoverer was set to the 6 most intense peaks every 100 Da. FDR

calculation used Percolator (23) and was set to the conventional threshold of 1% at the peptide level for the whole study. The resulting files were further processed using myProMS v3.9 (24) (code available at **https://github.com/bioinfo-pf-curie/myproms**). The label free quantification was performed using peptide Extracted Ion Chromatograms (XICs) computed with MassChroQ (25), version 2.2.21. For proteome and ubiquitylome quantifications, XICs from all proteotypic peptides shared between compared conditions (TopN matching) were used and missed cleavages were allowed. All peptides data, including quantification, are available from ProteomeXchange with identifier PXD025890 and the proteins and ubiquitination sites data (identification and quantification) are available in Tables S1 to S4.

Experimental Design and Statistical Rationale

Proteome analyses were performed on 3 biological replicates for the parental U2OS cells, treated with TGF-β and used as controls (n=3), and 3 for each of RNF111-RING-KO clones #1 and #2 (n=6 biological replicates in total for RNF111-RING-KO), to accommodate biological variability. Each replicate was divided into 6 fractions for MS analysis, and XICs were summed across fractions for each peptide before statistical analysis. DiGly ubiquitylome analyses were performed on the parental U2OS cells, and on the RNF111-RING-KO clone #2, with 3 biological replicates each. Each sample was analysed twice (2 technological replicates) and the obtained XICs were merged (averaged when measured twice) to improve the number of ubiquitylation sites identified and quantified. For the pan UB ubiquitylome analysis, 4 biological replicates for parental U2OS cell line (n=4) were compared to 4 biological replicates for each of RNF111-RING-KO clones #1 and #2 (n=8 for the RNF111-RING-KO condition).

For each experiment, statistical analysis was then performed inside myProMS v3.9 (24), after checking for normal distribution. Identified contaminants were excluded from the analysis at this point for the diGly ubiquitylome samples. Median and scale normalization was applied on the total signal to correct the XICs for each biological replicate. Outlier peptides were removed, within each condition and for each protein, with the Tukey's fences method. To estimate the significance of the change in protein abundance, a linear model (adjusted on peptides and biological replicates) was used and p-values were adjusted with the Benjamini–Hochberg FDR procedure. In addition, proteins/sites specific to a single condition were also included in the analysis. Candidates for RNF111 ubiquitylation substrates or sites were retained under the criteria of a minimum 2-fold change between RNF111-RING-KO and the parental U2OS cells (increase in proteome, decrease in diGly and pan UB ubiquitylomes) which must be statistically significant as reported by an adjusted p-value under 0.05 and the candidates must have at least one corresponding peptide identified in three replicates of any of the two conditions.

RESULTS

CRISPR engineered U2OS RNF111-RING-KO cell lines are not responsive to TGF-β

In order to identify endogenous substrate of RNF111 E3 ubiquitin ligase function, we generated CRISPR engineered cell lines devoid of RNF111 RING domain by mimicking the stop mutation S432* observed on exon 5 of the RNF111 gene in the NCI-H460 carcinoma cell line. Considering the role of RNF111 in both TGF-β signaling and DNA repair, we set our study in the U2OS osteosarcoma cell line that exhibits an intact functional response to both TGF-β and DNA damage. We used the CRISPR/Cas9^{D10A} double nicking system that reduces

off-target cleavage by 50 to 1500-fold in cell lines (26) to generate highly specific double strand break with paired single guide RNA (sgRNA-rev and sgRNA-fw) located on either side of the targeted region in exon 5 of RNF111 in U2OS cell line **(Fig. 1A)**. CRISPR single clones were selected by western blotting with an antibody that recognizes the N-terminal region (amino acids 1-108) of RNF111 (Abnova M05). Two independent U2OS clones that express truncated forms of RNF111 with size range equivalent to the one observed in NCI-H460 cells, and no full length RNF111, were selected (referred as RNF111-RING-KO clones #1 and #2) **(Fig. 1B)**. Further sequencing of the genomic region of exon 5, indicates that RNF111-RING-KO clone #1 carries 2 alleles with a frameshift mutation at position 434 followed by 15 extra amino acids before stop mutation, and 1 allele with a frameshift mutation at position 437 followed by 46 extra amino acids; while RNF111-RING-KO clone #2 carries the same frameshift mutation at position 426 followed by 3 extra amino acids on all alleles **(Fig. S1)**.

As observed in NCI-H460 cell line (9), RNF111-RING-KO clones #1 and #2 undergo an intact SMAD2 phosphorylation upon TGF-β signaling after 1h treatment with TGF-β, but the degradation of the SMAD transcriptional repressors SKI and SKIL is abolished in these two clones as compared to the parental U2OS cell line **(Fig. 1B)**. As a consequence, the SMAD-dependent transcription in response to TGF-β assessed by luciferase assay with the $CAGA_{12}$-LUC reporter is completely abolished in the two RNF111-RING-KO clones, as observed in the NCI-H460 cell line **(Fig.1C)**. Altogether, our results confirm that the 2 independent CRISPR engineered U2OS RNF111-RING-KO clones #1 and #2 have lost the E3 ubiquitin ligase function responsible for SKI and SKIL degradation upon TGF-β and are subsequently not responsive to TGF-β.

Identification of the RNF111-dependent proteome

To identify proteins that are degraded by RNF111 E3 ligase function, we performed label free quantitative proteomics on U2OS parental and RNF111-RING-KO clones #1 and #2 **(Fig. 2A)**. We compared 3 biological replicates for each RNF111-RING-KO clones #1 and #2 (RNF111-RING-KO clones, n=6) to 3 biological replicates for U2OS parental cells (U2OS, n=3) to quantify RNF111-RING-KO / U2OS protein ratios. In order to detect RNF111-induced protein degradation that depends on active TGF-β pathway, as for SKI and SKIL degradation, cells were treated with TGF-β for 1h before lysis. This timing of TGF-β induction allows Phospho-SMAD2 accumulation and SKI and SKIL degradation but is too early to enable detection of TGF-β target genes induction at the protein level. Significant increase of protein quantity in RNF111-RING-KO clones compared to parental U2OS cells (RNF111-RING-KO/U2OS fold increase ≥2; p-value ≤0.05) was detected for 73 proteins including SKI and SKIL among the 7746 proteins quantified **(Fig. 2A, 2B, Table S1)**. We further validated by western-blot that Kynureninase (KYNU), GDF15 and FABP3, some of the strongest candidates with commercially available antibodies, are indeed increased in both RNF111-RING-KO Clones #1 and #2 as compared to parental U2OS cells, but unlike SKI and SKIL, we found that this increase is independent of TGF-β **(Fig. 2C)**. Quantitative PCR analysis of these candidates indicate that their RNA levels are also increased in RNF111-RING-KO Clones #1 and #2 compared to parental U2OS cells independently of TGF-β, whereas, as expected, SKIL RNA level is not. These results suggest that RNF111 downregulates KYNU, GDF15 and FABP3 directly or indirectly at the transcriptional level, rather than affecting their protein stability as for SKI and SKIL **(Fig. 2D).**

Identification of RNF111-dependent substrates by label free quantitative comparison of diGly Ubiquitylome

Since the RNF111-dependent proteome is crowed with proteins that are modified at the transcriptional level, and considering that RNF111 might also perform non-degradative ubiquitination (11, 27), we went on to investigate directly global endogenous substrates ubiquitylated by the RING domain of RNF111. In that aim, we first performed diGly peptides profiling. Lysates were digested with trypsin and peptides carrying diGly modified lysine remnant of ubiquitylation were immunoprecipitated with the K-ε-GG antibody (cell signaling) **(Fig. 3A)**. Three biological replicate experiments were performed in order to compare global ubiquitylation in parental U2OS (n=3) and RNF111-RING-KO clone #2 (n=3). In order to identify RNF111 substrates that are dependent of an active TGF-β, we treated the cells with TGF-β for 1h prior to lysis, which corresponds to the peak of SKIL degradation before detection of TGF-β target gene induction at the protein level. To prevent the proteasomal degradation of ubiquitylated substrates, cells were treated with MG132 for 4h prior to TGF-β induction. We assumed that these experimental conditions enable detection of both TGF-β dependent and independent ubiquitylation events, as well as degradative and non-degradative ubiquitylation.

Label free quantification of ubiquitylation sites by mass spectrometry enabled the identification of 160 sites within 108 proteins, among 12675 sites quantified within 3641 proteins, that display a significant decrease in RNF111-RING-KO clone #2 compared to U2OS parental cells (RNF111-RING-KO/U2OS fold decrease ≥2; p-value ≤0.05; **Fig. 3A and 3B; Table S2 and S3**).

Among the 160 sites of ubiquitylation that significantly decrease in RNF111-RING-KO clone #2, we identified lysine 343 (K343) as an RNF111-dependent ubiquitylation target for SKIL **(Fig. 3B)**. Interestingly, this lysine is localized in the SAND domain of SKIL (aa 258-353) that interacts with SMAD4 (28) and RNF111 (5) **(Fig. 3C)**. Of note, one other ubiquitylation site

was detected for SKIL on lysine 432 that do not display any significant changes between RNF111-RING-KO clone #2 and U2OS cells **(Table S3).** In order to determine if lysine 343 of SKIL is the lysine responsible for the degradation of SKIL induced by RNF111 in response to TGF-β signaling, we have mutated lysine 343 on a HA-SKIL expression vector. Lysine 343 is adjacent to lysine 342, it is therefore likely that lysine 342 could be ubiquitylated by RNF111 in absence of lysine 343. To overcome this possibility, we performed lysine to arginine mutation of the 2 lysines 342 and 343 to generate a HA-SKIL-342/43-KR mutant. We have compared the ability of Flag-RNF111-WT or its inactive RING mutant Flag-RNF111-C933A to ubiquitylate HA-SKIL-WT or the HA-SKIL-342/43-KR mutant in presence of TGF-β/Activin signal in HEK-293 cells. To immunoprecipitate ubiquitylated proteins we used the newly developed ubiquitin pan nanobody (pan UB) (Nanotag biotechnologies) that exhibits strong affinity for polyubiquitylated and monoubiquitylated proteins. As expected, because RNF111 auto-ubiquitylates, Flag-RNF111-WT was immunoprecipitated, but not its catalytic inactive mutant C933A **(Fig.3D)**. Moreover, RNF111-WT, but not the C933A mutant, induces ubiquitylation of HA-SKIL-WT. However we could not detect any decrease in SKIL ubiquitylation when lysines 342/43 were mutated, indicating that RNF111 might also ubiquitylate other lysines on SKIL that were not detected in the diGly ubiquitylome.

Identification of RNF111-dependent substrates by label free quantitative comparison of pan UB nanobody ubiquitylome.

Since the diGly approach led to the identification of numerous putative targets, we decided to refine our analysis by employing the pan UB nanobody to analyze the RNF111-dependent ubiquitylome. We first assessed whether immunoprecipitation with the pan UB nanobody enables detection of endogenous SKIL ubiquitylation. Although TGF-β-induced SKIL degradation has been well documented (4–6, 29), no study has provided evidence of

449 an increased ubiquitylation of SKIL upon TGF-β signaling, presumably due to the difficulty to
450 detect endogenous ubiquitylation and because overexpression experiments temper this
451 inducible effect. We found that immunoprecipitation with the pan UB nanobody enables
452 detection of increased endogenous SKIL ubiquitylation in the parental U2OS cells after 1h
453 TGF-β treatment in presence of proteasome inhibitor MG132, but not in the two RNF111-
454 RING-KO clones despite the presence of equivalent amount of SKIL protein in the input of
455 each condition **(Fig. 4A)**. This result clearly demonstrates that TGF-β induces SKIL
456 ubiquitylation and that RNF111 E3 ubiquitin ligase activity is required in this process.
457 Moreover, ubiquitylated RNF111 was also immunoprecipitated in the parental cell lines but
458 not in the two RNF111-RING-KO clones that lack RNF111 auto-ubiquitylation ability. These
459 results highlight that the pan UB nanobody constitutes a very efficient tool for endogenous
460 purification of ubiquitylated proteins and we then decided to carry out label free
461 proteomics on pan UB immunoprecipitated parental U2OS, RNF111-RING-KO clones #1 and
462 #2 lysates. The easy workflow and the low amount of starting material of such an approach
463 compared to the diGly approach, allowed us to increase the number of biological replicates.
464 We then compared 4 biological replicates of each RNF111-RING-KO clones #1 and #2
465 (RNF111-RING-KO clones n=8) to 4 biological replicates of parental cells U2OS (U2OS n=4)
466 in the same conditions as the diGly experiments (4h MG132, 1h TGF-β). Among the 8547
467 quantified proteins, differential analysis enabled the detection of 52 proteins that display a
468 statistically significant decrease of ubiquitylation in RNF111-RING-KO clones compared to
469 parental U2OS cells (RNF111-RING-KO/U2OS fold decrease ≥2; p-value ≤0.05; **Fig. 4B and**
470 **4C, Table S4)**. Among these 52 putative RNF111 substrates, we identified SKI, SKIL and
471 RNF111, which confirms that the pan UB approach is very efficient for ubiquitylome

analysis. Curiously, we also identified SMAD4 among the 52 hits but were not able to further validate that RNF111 indeed increases SMAD4 ubiquitylation (data not shown).

Integrative comparison of RNF111-dependent ubiquitylomes and proteome

To pinpoint the most robust RNF111 putative substrates, we next compared the 108 and 52 candidates obtained respectively with the diGly and pan UB ubiquitylome analyses **(Fig. 5A)**. Strikingly, we identified only 4 common hits including SKIL and 3 other proteins PDK4, MED10 and ZFAND2A, **(Fig 5B)**. Since we were not able to validate these 3 candidates by western-blot after pan UB immunoprecipitation neither at the endogenous level, nor after co-expression of an HA-tagged cDNA expression vector with Flag-RNF111 as for SKIL (data not shown), we concluded that SKIL constitutes the only common validated candidate for the diGly and pan UB ubiquitylome.

Finally, in order to identify substrates for RNF111 ubiquitin ligase function that are ubiquitylated and degraded by RNF111 in response to TGF-β, we performed an integrative comparison of the proteome with the two RNF111-dependent ubiquitylomes **(Fig. 6A, Table S5).** We identified SKIL as the only candidate with both a decreased ubiquitylation and increased protein level in absence of RNF111 ubiquitin ligase function when we compared the proteome to the diGly ubiquitylome, and SKI and SKIL as the only candidates when we compared the proteome to the pan UB ubiquitylome **(Fig 6B).** Although it cannot be ruled out that other proteins than SKI and SKIL that were not detected can be ubiquitylated and degraded by RNF111, these results strongly argue that RNF111 specifically targets its substrates SKI and SKIL for degradative ubiquitylation in response to TGF-β.

DISCUSSION

Lack of RNF111 E3 ubiquitin ligase activity affects protein expression at the transcriptional level

In the present work, we performed a comprehensive analysis of RNF111 substrates by using different quantitative proteomics approaches to compare parental U2OS cells to RNF111-RING-KO clones devoid of RNF111 E3 ubiquitin ligase activity. A proteome analysis enabled identification of 73 proteins, including SKI and SKIL, with an increased level in absence of RNF111 RING domain, suggesting that these hits could constitute new substrates of RNF111. However, further western-blot and Q-PCR analysis of selected candidates indicates that RNF111 regulates these proteins at the transcriptional level independently of TGF-β, rather than affecting their protein stability. Although this could be an indirect effect, this finding raises the possibility that RNF111 could act as a transcriptional repressor. This would be in agreement with the study of Sun et al, which performed transcriptomic analysis on MEF RNF111 -/- and identified a panel of genes repressed by RNF111, independently of TGF-β, that are also regulated by the Polycomb complex (31). Our study further indicates that RNF111 could regulate transcription in a RING dependent manner, suggesting that this effect might depend on a ubiquitylation event. However, we did not identify any components of the Polycomb complex in the ubiquitylome or proteome that would be a target for RNF111 ubiquitylation and could explain this effect. Future investigations in this direction will be required to decipher the ability of RNF111 to directly repress transcription and to understand the relevance of its E3 ubiquitin ligase activity in this effect.

Advantages and limitations of the diGly and pan UB ubiquitylome analyses

Development of efficient approaches to enrich ubiquitylated proteins is a prerequisite to enable identification of endogenous substrates of E3 ubiquitin ligases. Here, we have

compared two methods of enrichment by performing immunoprecipitation with the diGly antibody and pan UB nanobody. DiGly remnant peptides enrichment followed by mass spectrometry analysis constitutes the most powerful method used to profile endogenous ubiquitylation and gives the major advantage to profile ubiquitylated lysines. However, because it involves detection of a single peptide for each ubiquitylation site, this method displays a low sensitivity that requires considerable amount of starting material and can be quite challenging to set up. In our study we used a recently developed pan UB nanobody that empowers enrichment of ubiquitylated proteins at the endogenous level to perform ubiquitylome analysis. We found that this method is highly sensitive and much easier to implement that the diGly approach, and can therefore constitute an easy workflow alternative to investigate E3 ubiquitin ligases substrates profiling. Interestingly, the current development of nanobodies targeting the different polyubiquitin chain linkages will provide the opportunity in the future to investigate identification of substrates for an E3 ubiquitin ligase according to the different ubiquitin linkages.

Comparison of the candidates obtained from our diGly and Pan UB ubiquitylome analyses indicates that they poorly overlap, with only 4 common hits including SKIL, that, apart from SKIL, could not be validated. This suggests that ubiquitylome analyses might be significantly poised with false positive candidates due to biological noise. However, the discrepancy between the two ubiquitylomes approaches could also be explained by the difference in the two antibodies used for ubiquitin enrichment. Indeed, while the pan UB nanobody is specific of monoubiquitin and all polyubiquitin linkage, the K-ε-GG antibody used in the diGly approach is unable to distinguish between attachment of ubiquitin and two other ubiquitin-like proteins NEDD8 and ISG15 that also leave a diGly remnant after trypsic digestion. Therefore, it is possible that some of the RNF111-dependent diGly peptides

identified could correspond to a neddylation or ISGylation event. Indeed, it has been proposed that RNF111 could act as a NEDD8 E3 ligase (12, 32) and it will therefore be interesting in a future work to determine if RNF111 could neddylate some of the different candidates identified in the diGly ubiquitylome.

SKI and SKIL are the only identified degradative substrates for RNF111 ubiquitin ligase function in TGF-beta signaling

In the present work, we showed that comparison of the RNF111-dependent proteome and ubiquitylomes leads to the identification of SKI and SKIL as the only substrates ubiquitylated and degraded by RNF111 upon TGF-β signaling. Notably we provide clear evidence for the first time that SKIL ubiquitylation is increased by TGF-β stimulation and that RNF111 is absolutely required for this inducible effect. We are aware that our finding is not a proof that SKI and SKIL constitute the only endogenous substrates of RNF111, since we cannot rule out that some other substrates have not been detected in our screen. However, the important representativeness of our study with more than 7700 proteins identified in the proteome and the pan UB ubiquitylome and 12 000 ubiquitylation sites identified in the diGLy ubiquitylome, indicates that despite the paradigm of the ability for an E3 ubiquitin ligase to target many different substrates, RNF111 ubiquitin ligase function seems to display a very stringent specificity for its degradative substrates SKI and SKIL upon TGF-β pathway activation. On the other hand, the comparison of the RNF111-dependent ubiquitylomes to the proteome indicates that most of the putative ubiquitylated substrates of RNF111 are not associated with a significant increase in protein level, which could suggest that such proteins are non-degradative ubiquitylation substrates for RNF111. The fact that RNF111 has been reported to trigger K27 and K63 non-degradative ubiquitylation corroborates this possibility (11, 27). Further investigation will be required for the

validation and significance of these potential non-degradative ubiquitylation events. However, our comparison of the pan UB and diGly ubiquitylomes which identifies SKIL as the only common validated substrate, further supports the idea that SKIL constitutes the major ubiquitylated substrate for RNF111. Importantly, our study has been performed in the presence of an active TGF-β signaling pathway and it is likely that RNF111 affects ubiquitylation only in a stimuli dependent manner. In order to depict the ubiquitylation network of RNF111, it would be interesting to employ the same approach to identify targets of RNF111 in response to other stimuli where RNF111 has also been involved, such as response to arsenite treatment (33), UV (11) or IR irradiation (12).

Lysine K343 of SKIL is a ubiquitylation target for RNF111

The advantage of the diGly approach is that it enables detection of the lysines that are ubiquitylated on the substrate. We found that ubiquitylation of lysine 343 of SKIL is dependent of RNF111. However, mutation of lysine 343, together with the adjacent lysine 342 did not result in attenuated SKIL ubiquitylation. A previous study on overexpressed SKIL lysine mutants has shown that SKIL degradation in response to TGF-β depends on lysine 440, 446 and 449, but no ubiquitylation experiments have confirmed this observation (29). It is, however, possible that RNF111 ubiquitylates multiple lysines on SKIL including lysine 343, and other lysines such as lysine 440, 446 and 449 that were not detected in the ubiquitylome. A more targeted approach by SKI and SKIL immunoprecipitation followed by diGly enrichment would enable to map precisely all the lysines ubiquitylated by RNF111 on SKI and SKIL. Alternatively, it is possible that lysine 343 is indeed the only endogenous ubiquitylation target for RNF111 and that overexpression of SKIL and RNF111 triggers ubiquitylation towards other lysines. Intriguingly, the lysine 343 is located in the SAND domain of SKIL known to interact with SMAD4 (28) and RNF111 (5). It has been shown

recently that besides its interaction with phospho-SMAD2/3 (4), interaction of SKIL with SMAD4 is also absolutely required for SKIL degradation (34). Altogether, these findings point out that RNF111 ubiquitylation of SKIL might occur at the interface of a SKIL/SMAD4/phospho-SMAD2/3 heteromeric complex, and future investigation in this direction would be critical to unravel the molecular mechanism of RNF111-dependent SKIL degradation in response to TGF-β.

In conclusion, by showing that SKI and SKIL constitute the main degradative substrates for RNF111 E3 ubiquitin ligase function in TGF-β signaling, our integrative proteomics analysis indicates that RNF111 displays a strong specificity toward its substrates. This finding further suggests that drugs targeting RNF111 E3 ubiquitin ligase function would specifically enable inactivation of TGF-β signaling by preventing SKI and SKIL degradation.

DATA AVAILABILITY

The mass spectrometry proteomics data, including peptide quantification, have been deposited to the ProteomeXchange Consortium via the PRIDE (35) partner repository identified with the dataset identifier PXD025890.

ACKNOWLEDGEMENTS

We thank J.K. Joung lab for pMLM3636 plasmid, G.Church Lab for hCas9_D10A plasmid and N. Mailand for the PcDNA4/TO-SFS-RNF111 plasmid. We thank Jean-Paul Concordet for help with the CRISPR, and Raphaël Margueron for help with the diGly ubiquitylome. We thank rotation students Alexandre Gournay and Sarah Hjjaji for participation to the experiments and all members of the lab for useful discussions.

FUNDING

This work was funded by a grant from Canceropole R18056DD.

CONFLICT OF INTEREST

Authors declare no competing interests.

REFERENCES

1. Yau R, Rape M. 2016. The increasing complexity of the ubiquitin code. 6. Nature Cell Biology 18:579–586.
2. Imamura T, Oshima Y, Hikita A. 2013. Regulation of TGF-β family signalling by ubiquitination and deubiquitination. J Biochem 154:481–489.
3. Sinha A, Iyengar PV, ten Dijke P. 2021. E3 Ubiquitin Ligases: Key Regulators of TGFβ Signaling in Cancer Progression. 2. International Journal of Molecular Sciences 22:476.
4. Levy L, Howell M, Das D, Harkin S, Episkopou V, Hill CS. 2007. Arkadia activates Smad3/Smad4-dependent transcription by triggering signal-induced SnoN degradation. Mol Cell Biol 27:6068–6083.
5. Nagano Y, Mavrakis KJ, Lee KL, Fujii T, Koinuma D, Sase H, Yuki K, Isogaya K, Saitoh M, Imamura T, Episkopou V, Miyazono K, Miyazawa K. 2007. Arkadia induces degradation of SnoN and c-Ski to enhance transforming growth factor-beta signaling. J Biol Chem 282:20492–20501.
6. Le Scolan E, Zhu Q, Wang L, Bandyopadhyay A, Javelaud D, Mauviel A, Sun L, Luo K. 2008. Transforming growth factor-beta suppresses the ability of Ski to inhibit tumor metastasis by inducing its degradation. Cancer Res 68:3277–3285.
7. Koinuma D, Shinozaki M, Komuro A, Goto K, Saitoh M, Hanyu A, Ebina M, Nukiwa T, Miyazawa K, Imamura T, Miyazono K. 2003. Arkadia amplifies TGF-beta superfamily signalling through degradation of Smad7. EMBO J 22:6458–6470.
8. Sharma V, Antonacopoulou AG, Tanaka S, Panoutsopoulos AA, Bravou V, Kalofonos HP, Episkopou V. 2011. Enhancement of TGF-β signaling responses by the E3 ubiquitin ligase Arkadia provides tumor suppression in colorectal cancer. Cancer Res 71:6438–6449.
9. Briones-Orta MA, Levy L, Madsen CD, Das D, Erker Y, Sahai E, Hill CS. 2013. Arkadia regulates tumor metastasis by modulation of the TGF-β pathway. Cancer Res 73:1800–1810.
10. Erker Y, Neyret-Kahn H, Seeler JS, Dejean A, Atfi A, Levy L. 2013. Arkadia, a novel SUMO-targeted ubiquitin ligase involved in PML degradation. Mol Cell Biol 33:2163–2177.
11. Poulsen SL, Hansen RK, Wagner SA, van Cuijk L, van Belle GJ, Streicher W, Wikström M, Choudhary C, Houtsmuller AB, Marteijn JA, Bekker-Jensen S, Mailand N. 2013. RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. J Cell Biol 201:797–807.
12. Ma T, Chen Y, Zhang F, Yang C-Y, Wang S, Yu X. 2013. RNF111-dependent neddylation activates DNA damage-induced ubiquitination. Mol Cell 49:897–907.

13. Mizutani A, Saitoh M, Imamura T, Miyazawa K, Miyazono K. 2010. Arkadia complexes with clathrin adaptor AP2 and regulates EGF signalling. J Biochem 148:733–741.
14. Akimov V, Barrio-Hernandez I, Hansen SVF, Hallenborg P, Pedersen A-K, Bekker-Jensen DB, Puglia M, Christensen SDK, Vanselow JT, Nielsen MM, Kratchmarova I, Kelstrup CD, Olsen JV, Blagoev B. 2018. UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. Nat Struct Mol Biol 25:631–640.
15. Vere G, Kealy R, Kessler BM, Pinto-Fernandez A. 2020. Ubiquitomics: An Overview and Future. Biomolecules 10.
16. Hu Z, Li H, Wang X, Ullah K, Xu G. 2021. Proteomic approaches for the profiling of ubiquitylation events and their applications in drug discovery. Journal of Proteomics 231:103996.
17. Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, Gygi SP. 2011. Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. Molecular Cell 44:325–340.
18. Udeshi ND, Mertins P, Svinkina T, Carr SA. 2013. Large-scale identification of ubiquitination sites by mass spectrometry. Nat Protoc 8:1950–1960.
19. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. Science 339:823–826.
20. Dennler S, Itoh S, Vivien D, ten Dijke P, Huet S, Gauthier JM. 1998. Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. EMBO J 17:3091–3100.
21. He J, Tegen SB, Krawitz AR, Martin GS, Luo K. 2003. The transforming activity of Ski and SnoN is dependent on their ability to repress the activity of Smad proteins. J Biol Chem 278:30540–30547.
22. Concordet J-P, Haeussler M. 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res 46:W242–W245.
23. The M, MacCoss MJ, Noble WS, Käll L. 2016. Fast and Accurate Protein False Discovery Rates on Large-Scale Proteomics Data Sets with Percolator 3.0. J Am Soc Mass Spectrom 27:1719–1727.
24. Poullet P, Carpentier S, Barillot E. 2007. myProMS, a web server for management and validation of mass spectrometry-based proteomic data. PROTEOMICS 7:2553–2556.
25. Valot B, Langella O, Nano E, Zivy M. 2011. MassChroQ: a versatile tool for mass spectrometry quantification. Proteomics 11:3572–3577.
26. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F. 2013. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell 154:1380–1389.
27. Xia T, Lévy L, Levillayer F, Jia B, Li G, Neuveut C, Buendia M-A, Lan K, Wei Y. 2013. The four and a half LIM-only protein 2 (FHL2) activates transforming growth factor β (TGF-β) signaling by regulating ubiquitination of the E3 ligase Arkadia. J Biol Chem 288:1785–1794.
28. Wu JW, Krawitz AR, Chai J, Li W, Zhang F, Luo K, Shi Y. 2002. Structural mechanism of Smad4 recognition by the nuclear oncoprotein Ski: insights on Ski-mediated repression of TGF-beta signaling. Cell 111:357–367.
29. Stroschein SL, Bonni S, Wrana JL, Luo K. 2001. Smad3 recruits the anaphase-promoting complex for ubiquitination and degradation of SnoN. Genes Dev 15:2822–2836.
30. Dupont S, Mamidi A, Cordenonsi M, Montagner M, Zacchigna L, Adorno M, Martello G, Stinchfield MJ, Soligo S, Morsut L, Inui M, Moro S, Modena N, Argenton F, Newfeld SJ,

Piccolo S. 2009. FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. Cell 136:123–135.

31. Sun H, Liu Y, Hunter T. 2014. Multiple Arkadia/RNF111 structures coordinate its Polycomb body association and transcriptional control. Mol Cell Biol 34:2981–2995.

32. Li C, Zhang L, Qian D, Cheng M, Hu H, Hong Z, Cui Y, Yu H, Wang Q, Zhu J, Meng W, Xu J-F, Sun Y, Zhang P, Wang C. 2021. RNF111-facilitated neddylation potentiates cGAS-mediated antiviral innate immune response. PLoS Pathog 17:e1009401.

33. Erker Y, Neyret-Kahn H, Seeler JS, Dejean A, Atfi A, Levy L. 2013. Arkadia, a novel SUMO-targeted ubiquitin ligase involved in PML degradation. Mol Cell Biol 33:2163–2177.

34. Gori I, George R, Purkiss AG, Strohbuecker S, Randall RA, Ogrodowicz R, Carmignac V, Faivre L, Joshi D, Kjær S, Hill CS. 2021. Mutations in SKI in Shprintzen–Goldberg syndrome lead to attenuated TGF-β responses through SKI stabilization. eLife 10:e63545.

35. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaíno JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res 47:D442–D450.

721 **FIGURES LEGENDS**

722 **Figure 1. CRISPR engineered U2OS RNF111-RING-KO cell lines**

723 (A) Upper diagram: Design of the reverse and forward sgRNA (sgRNA-rev and sgRNA-fw, in
724 red) used to target exon 5 of RNF111 gene in U2OS cells. Red arrows indicate the breaking
725 sites. Amino acids of the corresponding codons are annotated below; nonsense mutation
726 on serine 432 observed in NCI-H460 cell line is indicated (S432*). Lower diagram: Schematic
727 representation of wild type RNF111 in parental U2OS cell line compared to RNF111
728 truncation in U2OS CRISPR engineered RNF111-RING-KO clones #1 and #2 and NCI-H460
729 cell line. (B) U2OS cells, NCI-H460 cells and RNF111-RING-KO clones #1 and #2 were treated
730 or not with TGF-β for 1h. Whole cell extracts were analyzed by Western blotting using
731 antibodies against RNF111, SMAD2/3, P-SMAD2, SKIL, SKI and GAPDH. (C) U2OS cells, NCI-
732 H460 cells and RNF111-RING-KO clones #1 and #2 were cotransfected with the $CAGA_{12}$-Luc
733 and pRL-TK reporters and were treated or not with TGF-β for 8h. Data represent means +/-
734 standard deviation of luciferase activities normalized to Renilla in triplicate experiments.

735 **Figure 2. RNF111-dependent Proteome**

736 (A) Schematic representation of the proteome experiment with a summary table of the
737 RNF111-RING-KO/U2OS protein ratio quantification results. (B) Volcano Plot representation
738 of the differential analysis with the $\log_2$ fold change RNF111-RING-KO/U2OS versus the
739 negative $\log_{10}$ p-value. Each dot represents a protein. Green and red lines represent the
740 cut-off applied (respectively fold increase or decrease ≥ 2 and p-value ≤ 0.05). Volcano Plot
741 is zoomed in the fold increase ≥ 2 and p-value ≤ 0.05 area. (C) Western blot analysis of the
742 proteome candidates. U2OS cells and RNF111-RING-KO clones #1 and #2 were treated or
743 not with TGF-β for 1h. Whole cell extracts were analyzed by Western blotting using
744 antibodies against RNF111, SKIL, SKI, GDF15, KYNU, FABP3 and GAPDH. (D) Quantitative

PCR analysis of the proteome candidates. U2OS cells and RNF111-RING-KO clones #1 and #2 were treated or not with TGF-β for 1h. Levels of mRNA for SKIL, GDF15, KYNU and FABP3 were analyzed by qPCR and normalized to GAPDH using the $2^{-\Delta\Delta Ct}$ methods. Data represent mean +/- standard deviation for at least 3 independent experiments.

Figure 3. Identification of RNF111-dependent substrates by diGly ubiquitylome quantitative comparison

(A) Schematic representation of the experimental design for RNF111-dependent ubiquitylome quantification with the diGly approach using immunoprecipitation of ubiquitylated trypsic remnant peptides with the K-ε–GG antibody. The table summarizes the results for the ubiquitylation sites quantification of the RNF111-RING-KO/U2OS ratio. (B) Volcano Plot representation of the differential analysis with the $\log_2$ fold change RNF111-RING-KO/U2OS versus the negative $\log_{10}$ p-value. Each dot represents a ubiquitylation site. Green and red lines represent the cut-off applied (respectively fold increase or decrease ≥ 2 and p-value ≤ 0.05). Volcano Plot is zoomed in the fold decrease ≥ 2 and p-value ≤ 0.05 area. (C) Schematic representation of the SKIL protein with its domains DHD, SAND and CC (coild-coild) and the localization of lysine 343. Lysines 342 and 343 that were subsequently mutated are indicated in red (D) HEK-293 cells were transfected with Flag-RNF111-WT or Flag-RNF111-C933A catalytic inactive mutant along with HA-SKIL-WT, HA-SKIL-342/43-KR mutant or empty vector. After 1h Activin/TGF-β treatment, whole cell extracts were immunoprecipitated with the pan UB nanobody and analyzed by western blotting (pan UB, right panel) along with the corresponding whole cell extract (input, left panel) using HA and Flag antibodies.

Figure 4. Identification of RNF111-dependent substrates by pan UB ubiquitylome quantitative comparison

(A) U2OS cells and RNF111-RING-KO clones #1 and #2 were treated with MG132 for 4h prior to induction or not with TGF-β for 1h. Whole cell lysates immunoprecipitated with the pan UB nanobody were analyzed by Western blotting (pan UB, right panel) along with whole cell lysates (Input, left panel) using antibodies against RNF111, SKIL, and UB. (B) Schematic representation of the experimental design for RNF111-dependent ubiquitylome quantification with the pan UB approach using immunoprecipitation of ubiquitylated proteins with the pan UB nanobody. The table summarizes the results of the differential quantification for the RNF111-RING-KO/U2OS ratio. (C) Volcano Plot representation of the differential analysis with the $\log_2$ fold change RNF111-RING-KO/U2OS versus the negative $\log_{10}$ p-value. Each dot represents a protein. Green and red lines represent the cut-offs applied (respectively fold increase or decrease ≥ 2 and p-value ≤ 0.05). Plot is zoomed in the fold change decrease ≥ 2 and p-value ≤ 0.05 area.

Figure 5. Comparison of diGly and pan UB ubiquitylome

(A) The heatmap represents the ratios for the significant RNF111 substrate candidates in the pan UB ubiquitylome, compared to the diGly ubiquitylome ratios. Ratios are annotated as fold increase or decrease, indicated by the arrows. Asterisks indicate their p-value: *=p≤0.05, **=p≤0.01, ***=p≤0.001. Note that we compared whole protein quantification in the pan UB case against ubiquitylation sites quantification for diGly (protein values may then be repeated when corresponding to multiple sites). (B) Venn diagram comparison shows 4 common hits including SKIL between the pan UB and diGly ubiquitylomes significant RNF111 substrate candidates.

Figure 6. Identification of RNF111 degradative substrates by comparison of RNF111-dependent proteome to diGly and pan UB ubiquitylomes

792 (A) The heatmap represents the ratios for the significant RNF111 substrate candidates in
793 the proteome, compared to pan UB and diGly ubiquitylomes ratios. Ratios are annotated as
794 fold increase or decrease, indicated by the arrows. Asterisks indicate their p-value:
795 *=p≤0.05, **=p≤0.01, ***=p≤0.001. Note that we compare whole protein quantification in
796 the proteome and pan UB cases against ubiquitylation sites quantification for diGly (protein
797 values may then be repeated when corresponding to multiple sites). (B) Venn diagram
798 shows SKI and SKIL as the only common hits between the proteome and pan UB or diGly
799 ubiquitylomes significant candidates for RNF111 substrates.
800

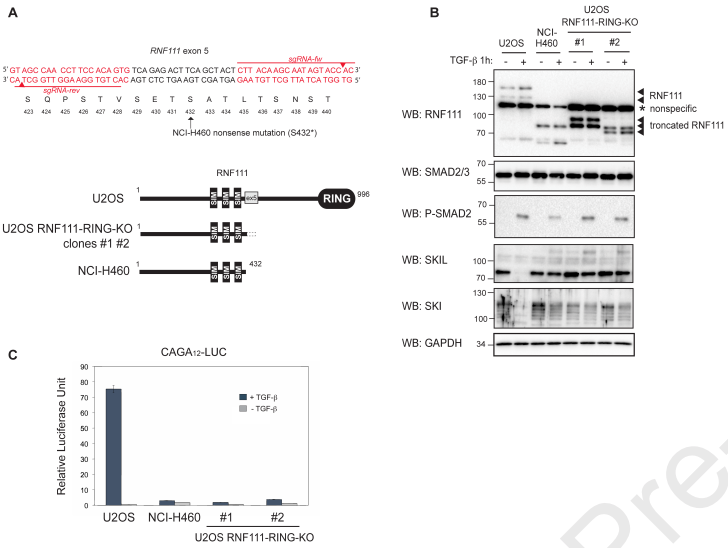


Figure 1

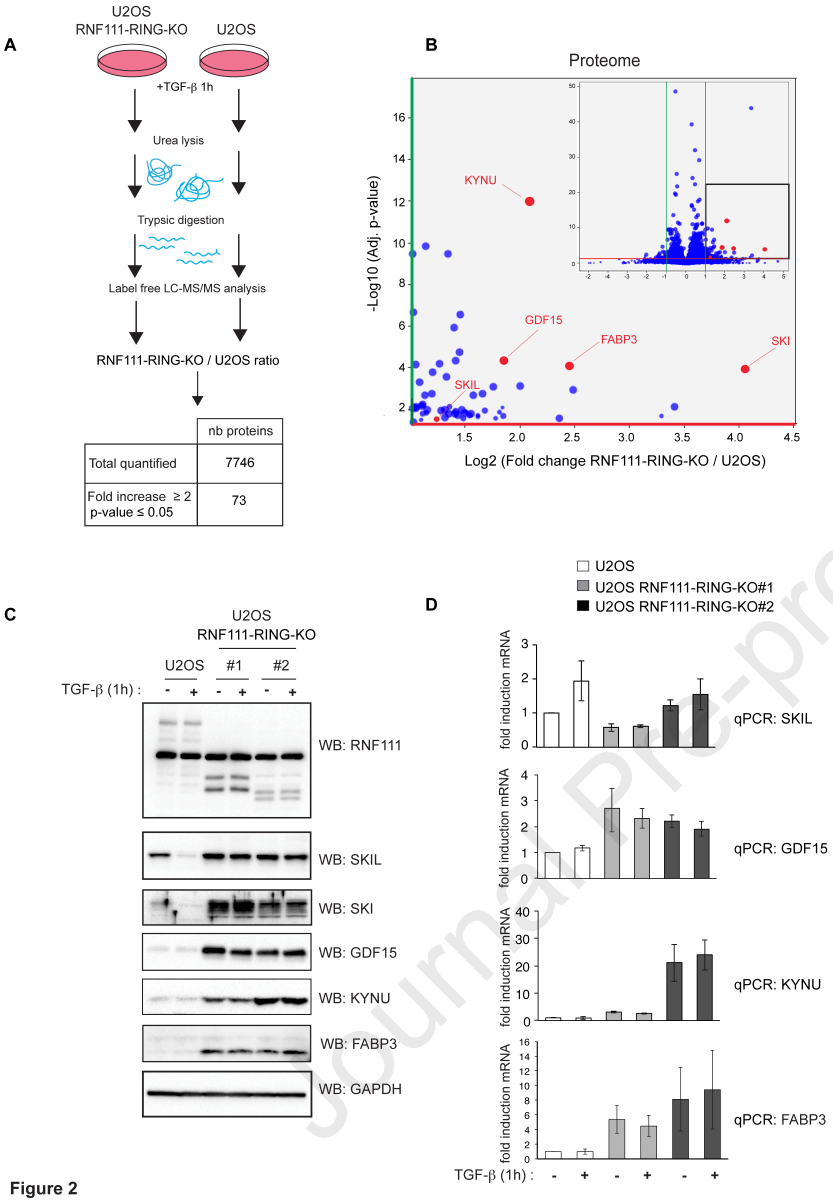
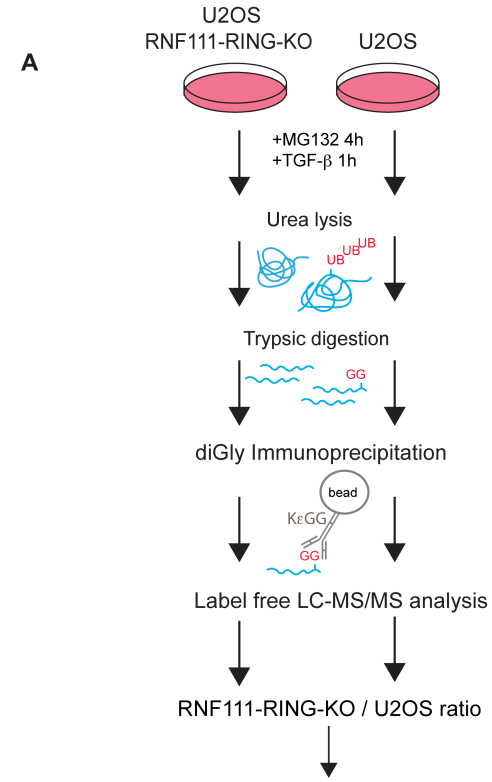
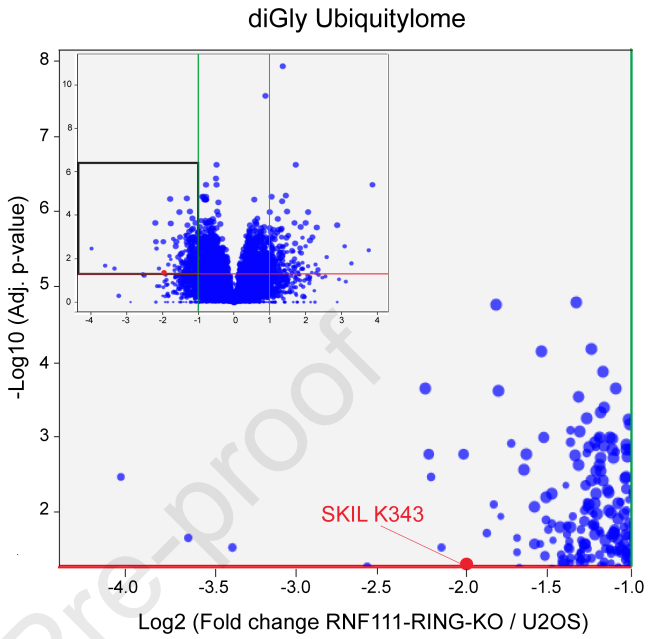
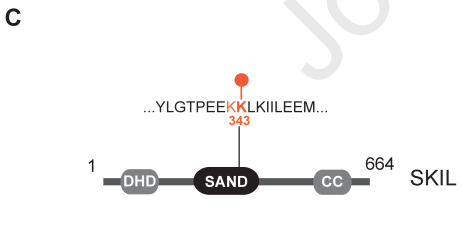


Figure 2



	nb Ubi sites	nb proteins
Total quantified	12675	3641
Fold decrease ≥ 2 p-value ≤ 0.05	160	108





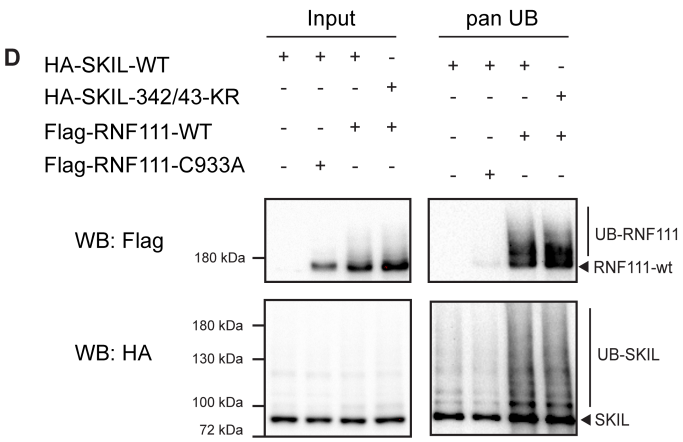
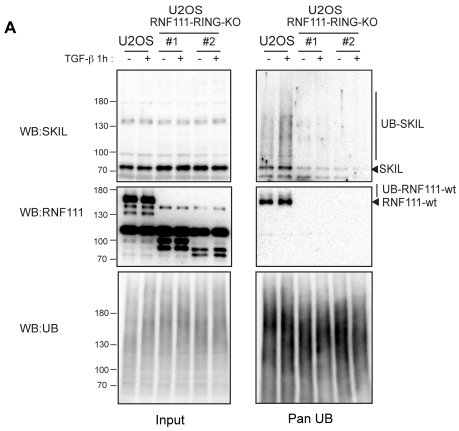
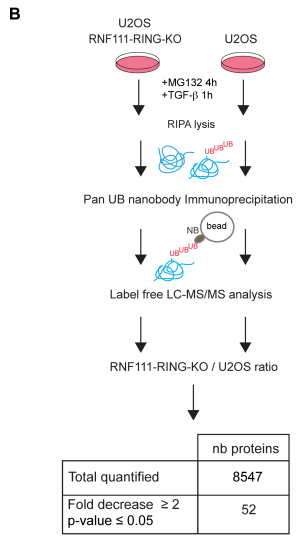


Figure 3





	nb proteins
Total quantified	8547
Fold decrease ≥ 2 p-value ≤ 0.05	52

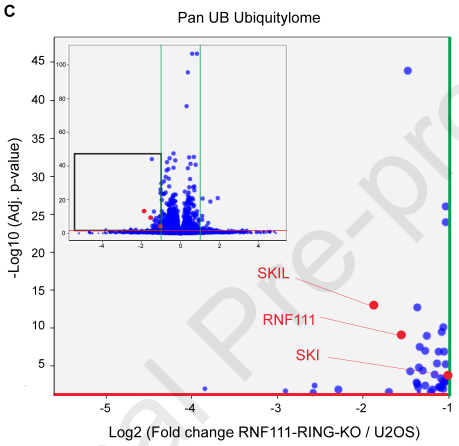


Figure 4

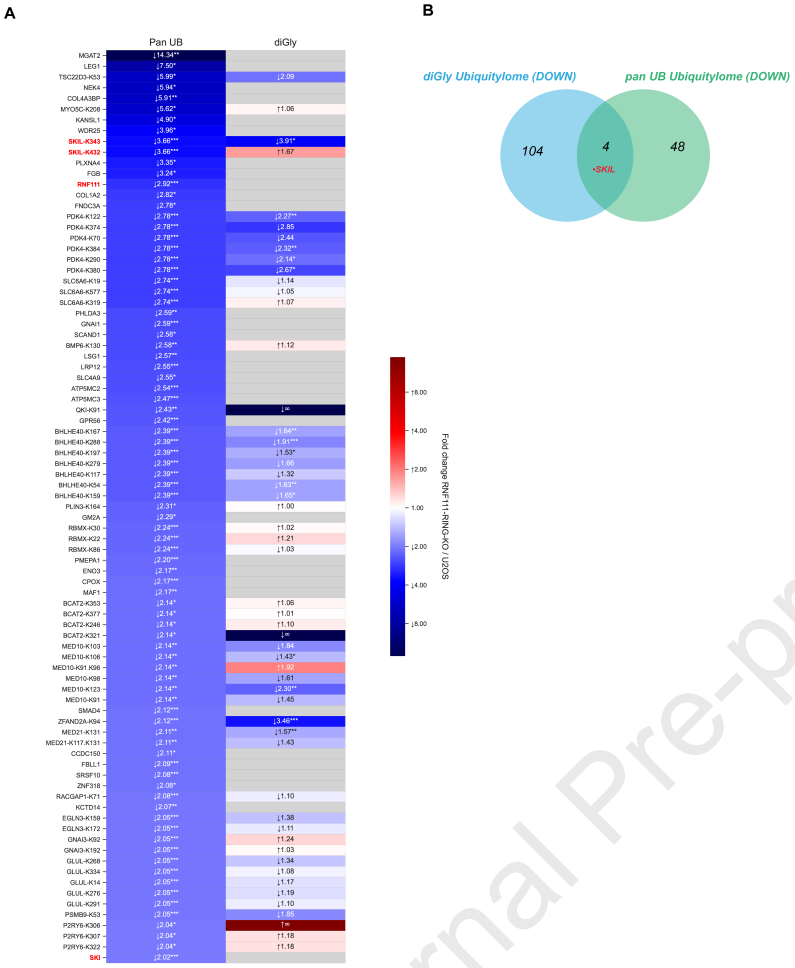


Figure 5

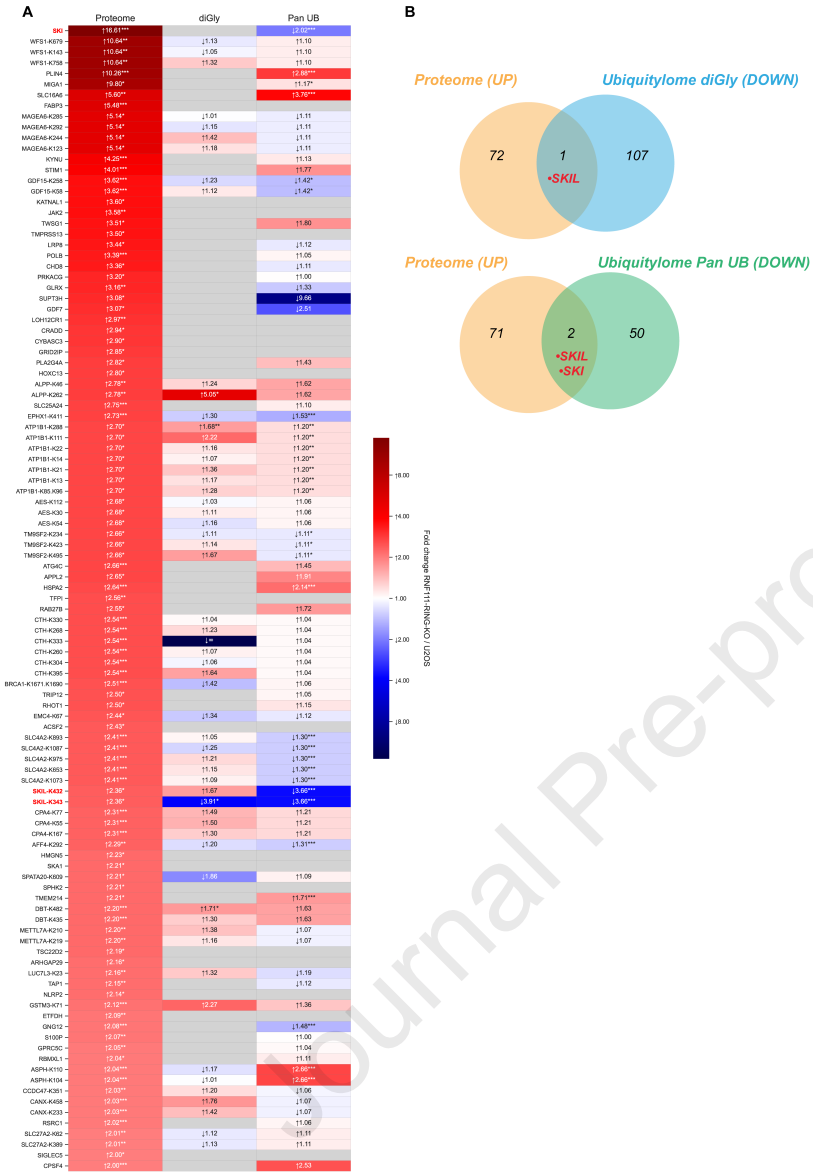


Figure 6

- ***Identification of endogenous RNF111 ubiquitylated substrates***
- ***A new powerful method to profile protein ubiquitylation using Pan Ub nanobody***
- ***Lysine K343 as a target for RNF111 ubiquitylation***
- ***SKI and SKIL are the only identified RNF111 degradative targets in TGF-β signaling***

Victor Laigle : Formal analysis, Methodology, Software, Conceptualization, Writing - Original Draft. **Florent Dingli :** Investigation, Methodology, Conceptualization, Writing - Original Draft. **Sadek Amhaz** : Investigation. **Tiphaine Perron :** Investigation. **Mouna Chouchène:** Investigation. **Sabrina Colasse :** Investigation. **Isabelle Petit :** Investigation. **Patrick Poullet :** Software. **Damarys Loew :** Methodology, Conceptualization. **Céline Prunier :** Conceptualization, Funding acquisition. **Laurence Levy :** Formal analysis, Investigation, Methodology, Conceptualization, Writing - Original Draft, Funding acquisition, Supervision.

In this study we aimed to identify exhaustively the substrates of the E3 ubiquitin ligase RNF111 that activates TGF-β signaling. We performed quantitative ubiquitylome comparison of parental U2OS cells to CRISPR modified clones that express a truncated RNF111 devoid of RING domain using two approaches based on enrichment of ubiquitylated proteins. Integrative proteomics comparison of ubiquitylome and proteome identifies SKI and SKIL as the only targets ubiquitylated and degraded by RNF111 upon TGF-β stimulation.