Ubiquitinated Proteome Analysis by Mass Spectrometry

Effect of R505C mutation in FBXW7 E3 ligase on protein ubiquitination

Author: Nadzeya Staliarova (7838799)

MSc programme 'Drug Innovation', Faculty of science

Supervisor: Dr. Wei Wu

Reviewer: Dr. Marc Baggelaar

Utrecht University

Abstract

Ubiquitination is a reversible post-translational modification (PTM) that governs many critical processes in eukaryotes. In turn, defects in the ubiquitin system can lead to various diseases. Targeted strategies to correct aberrations in the ubiquitin pathway could provide new therapeutic opportunities. The specificity of ubiquitination is determined by the substrate preference of E3 ligases. Here we investigated R505C mutation in a substrate receptor subunit FBXW7 of E3 ligase in colon organoids. We identified ubiquitin sites and FBXW7 substrates using mass spectrometry (MS) with ubiquitin enrichment by a commercial antibody recognizing di-Gly remnant left after trypsin digestion. A number of experimentally-identified FBXW7 targets were cross-validated by an established degron prediction method. By compiling evidence from the transcriptome, proteome, phosphoproteome, ubiquitinome and cell-surface staining, we demonstrated that impaired degradation via FBXW7 is the mechanism behind elevated EGFR level, MAPK signalling, and EGF-independent proliferative survival. Collectively, these illustrate the utility of our ubi-site identification workflow and further implicate FBXW7 R505C mutation in highly-proliferative cancers.

Table of Contents

List of abbreviations

AGC	Automated Gain Control
DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
DUB	Deubiquitinating Enzyme
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
FBXW7	F-Box and WD Repeat Domain Containing 7
FC	Fold Change
FDR	False Discovery Rate
GO	Gene Ontology
HCD	Higher energy Collisional Dissociation
IAP	Immuno-Affinity Purification
IP	Immunoprecipitation
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Liquid Chromatography
LFQ	Label-Free Quantification
MAPK	Mitogen-Activated Protein Kinase
MS	Mass Spectrometry
PTM	Post-Translational Modification

1. Introduction

Ubiquitination is a dynamic post-translational modification (PTM) of proteins with ubiquitin, a small 76 amino acid globular protein with a mass of 8.5-kDa[1]. Ubiquitin was named after the word ubiquitous since it is found across eukaryotes and plays a role in almost all processes in the cell, like cell cycle progression, cell signalling cascades, differentiation, and growth[2]. As such, defects in the ubiquitin system can lead to dysregulation of many cellular processes and the development of pathologies and diseases, such as cancer, metabolic disorders and neurogenerative diseases[3]. Therefore, understanding the mechanism of ubiquitination, sites of ubiquitin modification, and the consequence of dysregulated ubiquitination has the potential to general novel therapeutic strategies and help us to treat such diseases.

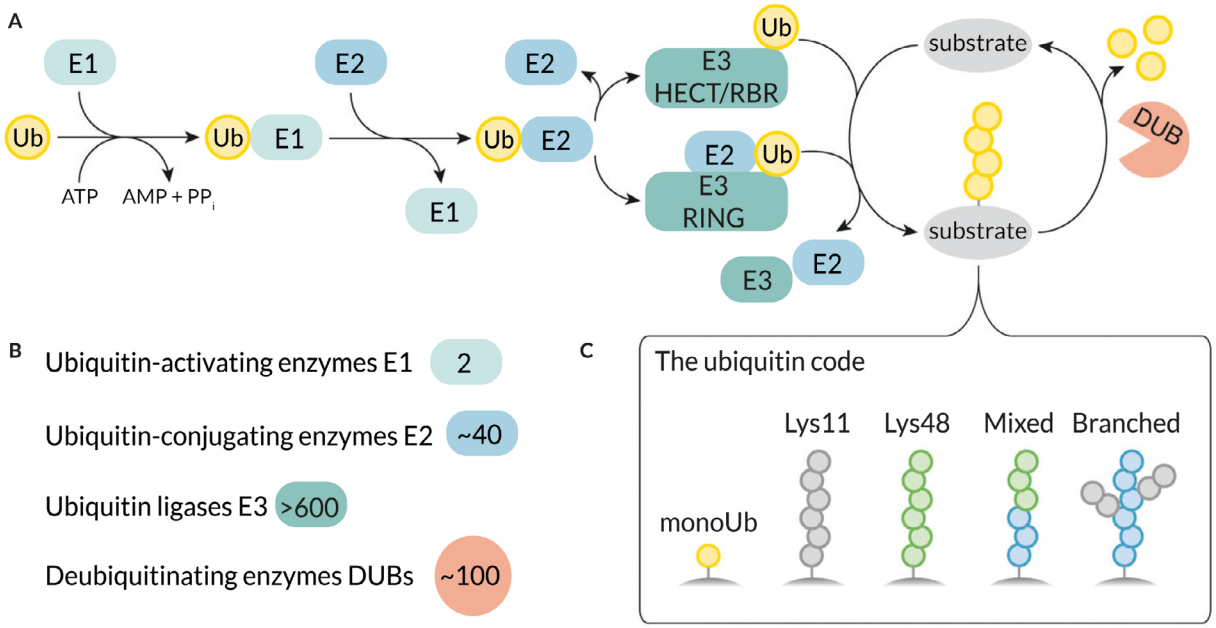

Fig. 1 | Ubiquitin system. **A** Ubiquitination requires a cascade of three consecutive adenosine 5'-triphosphate–dependent enzymatic reactions. The ubiquitin-activating enzyme E1 attaches ubiquitin (Ub), activating it. The activated ubiquitin is then transferred to the ubiquitin-conjugating enzyme E2. Finally, the ubiquitin ligase E3 facilitates the ubiquitin's attachment to a substrate through an isopeptide bond. Ubiquitination is a reversible process. Ubiquitin is removed from protein by deubiquitinating enzymes DUBs - ubiquitin isopeptidases, and can be used for the next ubiquitination cycle. **B** In the human genome, two ubiquitin-activating enzymes E1, about 40 ubiquitin-conjugating enzymes E2, more than 600 ligases E3 and about 100 deubiquitinases are encoded. **C** Ubiquitin binds to proteins either as a monomer or as a polyubiquitin chain, formed through any of seven internal lysine residues or N-terminal methionine. This collectively referred to as 'ubiquitin code'. Some of the possible ubiquitin modifications are illustrated. Modified from Fig. 1 in Ref.[3]

INTRODUCTION

Ubiquitination requires a cascade of three consecutive adenosine 5'-triphosphate–dependent enzymatic reactions. First, the ubiquitin binds to the ubiquitin-activating enzyme E1, and then ubiquitin is transferred from E1 to the ubiquitin-conjugating enzyme E2. Finally, the ubiquitin ligase E3 facilitates the ubiquitin's attachment to the ε-amino group of a lysine residue of a substrate through an isopeptide bond. Ubiquitination is a reversible process where deubiquitinating enzymes (DUBs) have the opposite function of E3 ligase, namely removing the attached ubiquitin from the substrate[2] (Fig. 1A).

Ubiquitin can be attached as polyubiquitin chains of various structures in diverse arrangements by attaching subsequent ubiquitin molecules to one of seven lysine residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63)[3,4] or the N-terminal methionine of the previous ubiquitin[3]. Different polyubiquitin chains can then send distinct signals[5], and the variety of all possible ubiquitin combinations is called a "ubiquitin code"[2,6,7] (Fig. 1C). The protein lifespan is determined by proteasomal degradation, which is usually encoded by polyubiquitin chains of at least four ubiquitin molecules linked through isopeptide bond at Lys48 or Lys11[5,7–9]. However, monoubiquitination as a degradation signal is also known for small protein substrates[10].

In the human genome, there are only two ubiquitin-activating enzymes E1, about 40 ubiquitin-conjugating enzymes E2, more than 600 ligases E3 and about 100 deubiquitinases (DUBs)[2,9] (Fig. 1B). Such a large amount of E3 ligases provides specificity for ubiquitination. Furthermore, to form a complex with E3 ligase substrates themselves undergo PTM changes in the binding motif, for example, degron phosphorylation by kinases[11]. All this ensures precise regulation of ubiquitination.

E3 ligases are divided into two main classes: HECT (homologous to E6-AP C terminus) and RING (really interesting new gene)[12]. RING E3 ligase superfamily comprises about 95% of the known E3 ligases and is represented by various multisubunit ligase complexes, one of which is CRL1 or SCF from cullin-RING ligase (CRL) family[13]. As illustrated in Fig. 2A Cullin-RING E3 ligase complex (CRL1 or SCF) consist of four components: RING domain-containing E3 ligase (RBX1), Cullin domain-containing scaffold protein (Cullin1), substrate adaptor subunit (SKP1) and substrate receptor subunit (F-box protein)[13]. Cullin1 serves as a scaffold to which RBX1 and SKP1 are attached. RBX1 further engage E2 with attached ubiquitin, while SKP1 engage F-box protein. Substrate receptor subunit F-box protein recognises the substrate by its distinct degron motif. 69 specific F-box protein subunits are known in the human genome[12]. Binding occurs between the substrate's degron and WD40 domain of the F-box protein subunit, allowing ubiquitin transfer from the ubiquitin-conjugating enzyme (E2) to the substrate[12]. As a result, a ubiquitin chain is formed on the substrate, signalling its proteasomal degradation.

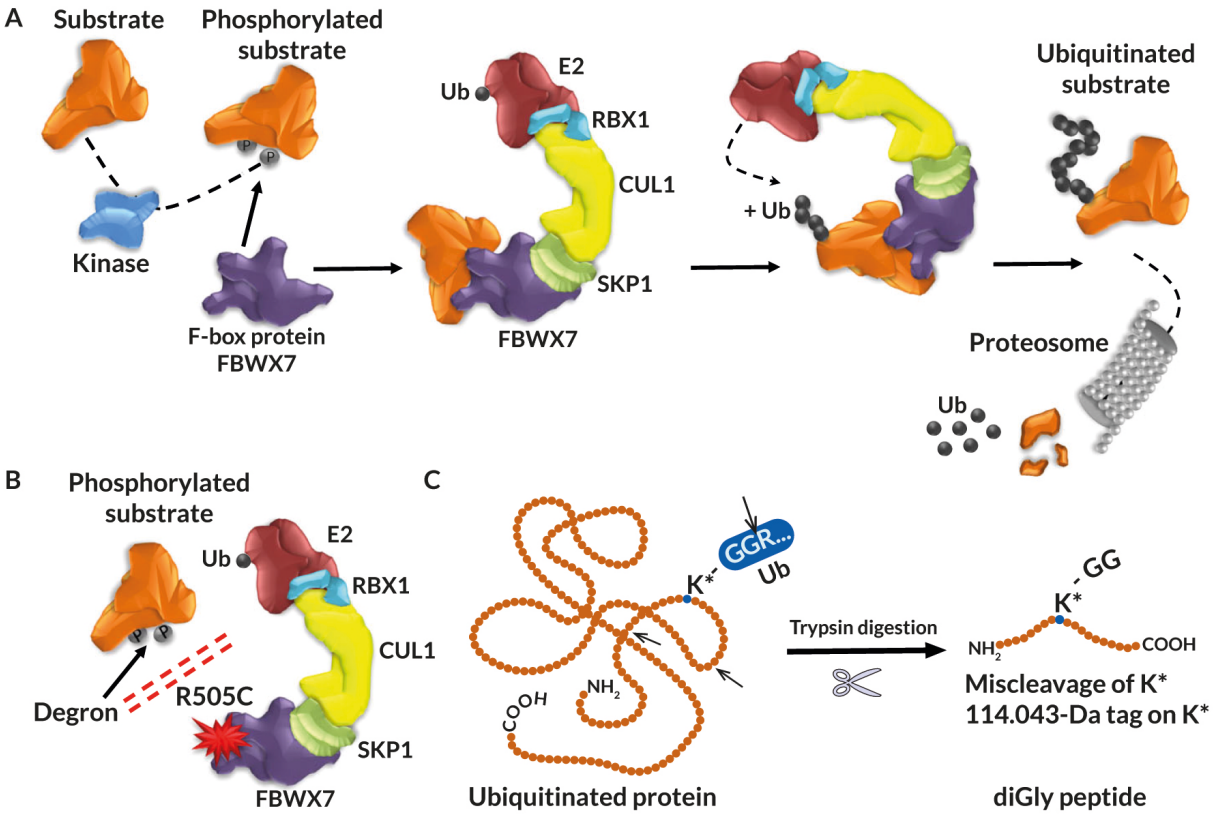

Fig. 2 | **A** Cullin-RING E3 ligase (CRL1 or SCF) complex consist of four components: RING domain-containing E3 ligase (RBX1), Cullin domain-containing scaffold protein (CUL1), substrate adaptor subunit (SKP1) and substrate receptor subunit (F-box and WD repeat domain containing 7 (FBXW7)[13]. Substrate receptor subunit FBXW7 recognises the substrate by its specific degron motif, phosphorylated by a kinase (phosphodegron). Binding occurs between the substrate's degron and WD40 domain of the FBXW7 subunit, allowing ubiquitin (Ub) transfer from the ubiquitin-conjugating enzyme (E2) to the substrate (in orange). As a result, a ubiquitin chain (in black) is formed on the substrate, signalling its proteasomal degradation. **B** R505C mutation in the WD40 binding domain of FBXW7 hampers binding of the substrate to the ubiquitin ligase complex. As a result, ubiquitin is not transferred to the substrate, and the substrate does not degrade. Modified from Fig. 2 in Ref. [14] **C** After trypsin digestion of ubiquitinated protein, diGly stab remains at the site of ubiquitin attachment to a lysine residue of the modified peptide. On this site, lysine is not cleaved by trypsin. This diGly remnant (black arrow) adds a 114.043-Da mass shift. Based on ions after MS/MS fragmentation, the localisation of the diGly remnant is determined. The black arrows indicate the C-terminal to arginine and lysine residues where trypsin trims protein.

Ge *et al.* analysed genes involved in the ubiquitin pathway across 9,125 patients from 33 cancer types from The Cancer Genome Atlas (TCGA). R505C mutation in the binding domain of E3 substrate receptor subunit F-box and WD repeat domain containing 7 (FBXW7) (Fig. 2B) was enriched among many cancers (in particular endometrial and colorectal) and in both the hotspot (recurrently mutated DNA positions in cancer[15]) and loss-of-function criteria[1].

INTRODUCTION

Although a large number of ubiquitination sites have been identified over the years, as demonstrated by 110,124 sites reported in the PhosphositePlus knowledgebase[16] (as of December 2021), it should be noted that this vast number remains a compendium from many different experiments. At the same time, little is known about the function and consequences of these modifications and the E3 ligase-substrate complexes formed.

To advance the goal of relating ubiquitination to phenotype, we focused on FBXW7 E3 ligase targets. In this project, we simplified the variety of ubiquitin configurations by performing tryptic digestion. This would cleave all forms and arrangements of ubiquitin after two glycine residues that remain on the lysine (Fig. 2C). With a specific antibody targeting this di-glycine stub, we could experimentally retrieve all peptides altered by the digly-modification. By comparing the peptide retrieved in wild type and FBXW7 mutant material, bona fide FBXW7 substrates can be identified. The final goal of this project is to understand how faulty ubiquitination due to FBXW7 mutation at R505C can lead to aggressive tumour phenotypes.

2. Results

2.1 Robust immunoenrichment of ubiquitinated peptides.

To assess the extent of ubiquitination in the cellular proteome, we used an antibody-based method to identify intracellular ubiquitinated peptides (Fig. 3). This approach relies on antibody recognition of a di-glycine stub remaining after trypsin digestion of ubiquitin chains (Fig. 2C). Enriched di-glycine peptides have the addition of 114.043-Da to their mass[5,17], resulting in the distinct mass shift at the MS2 spectrum (Fig. 4A, B). This antibody-enrichment step is necessary as ubiquitin modifications are generally low abundant[2,17,18], and found on diverse substrates with low sequence consensus. We performed the enrichment on the peptide level, since approaches for the enrichment of endogenously ubiquitylated proteins were reported as less efficient than anti-diGly antibody strategy[19].

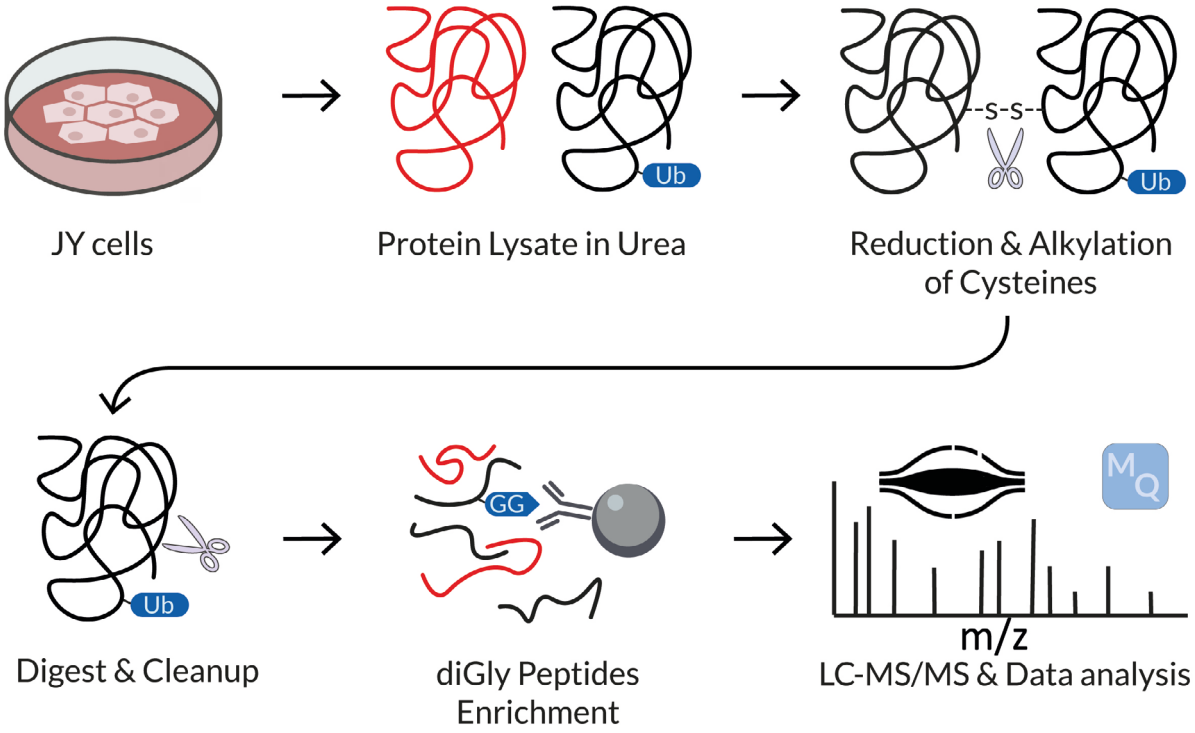

Fig. 3 | Schematic workflow to identify ubiquitinated proteins. JY cells were lysed. Protein disulfide bridges were reduced and alkylated, followed by digest. K-ε-GG stab was left from ubiquitin chains after trypsin digestion. Ubiquitinated peptides were enriched by immunoprecipitation with K-ε-GG antibody, which recognises K-ε-GG remnant. Pulled down peptides were analysed by LC-MS/MS. Protein identification and quantification was performed using MaxQuant software[20].

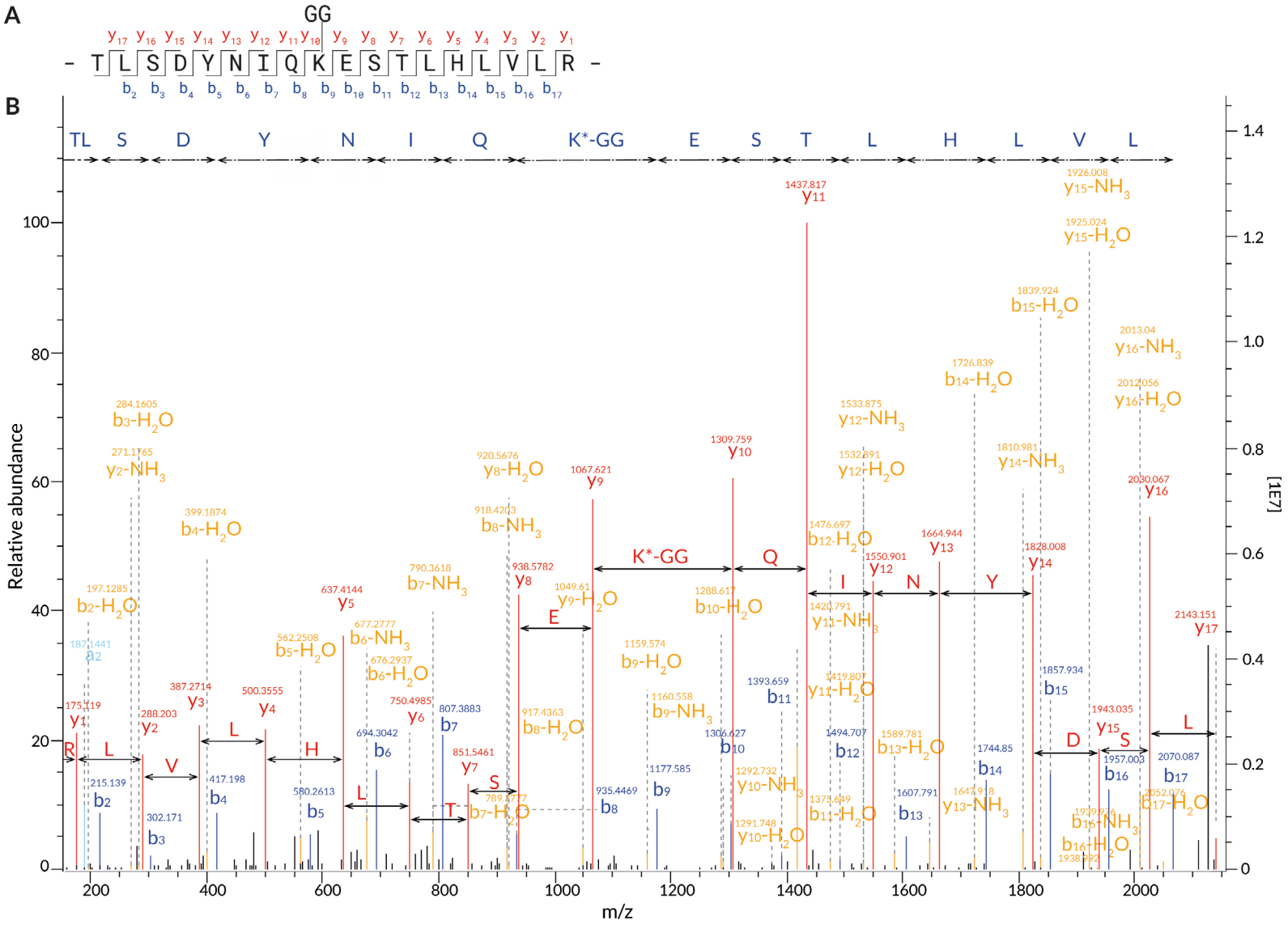

Fig. 4 | **A** Ubiquitinated peptide sequence with b and y ions. Amino acids are shown in one-letter code. **B** A sample annotated diGly peptide MS/MS spectrum obtained in this study. It illustrates ubiquitin with Lys-63 site identified in WT organoid line.

Considering the low stoichiometry of ubiquitinated peptides, we tested a commercial enrichment kit[21] on a JY cell lysate of 10 mg input. From 10 mg of trypsin-digested lysate, we identified a total of 6,713 peptides, of which 1,035 were modified by ubiquitin (Fig. 5A). These ubiquitinated peptides originated from 766 distinct protein groups (Fig. 5B), averaging out to be about 1.35 ubiquitinated peptides per protein. Although a considerable number of ubiquitinated peptides and ubiquitination sites were detected, only 15.4% of the peptides analysed by LC-MS/MS carried ubiquitin modification. This was not yet ideal since the mass spectrometer primarily measured non-ubiquitinated peptides during the 3h analysis time.

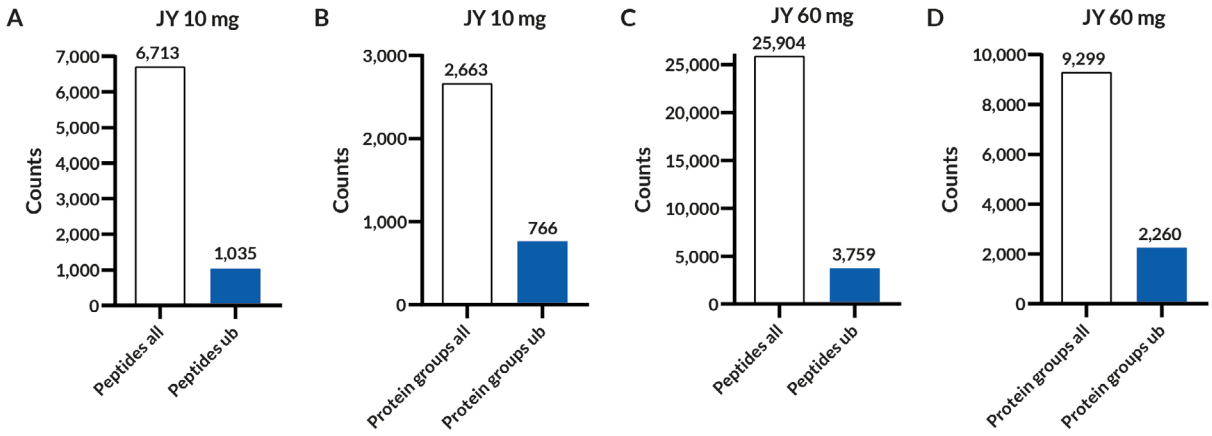

Fig. 5 | **A** Number of all and ubiquitinated peptides identified in 10 mg of total protein in JY cell line. **B** Number of all and ubiquitinated unique protein groups identified in 10 mg of total protein in JY cell line. **C** Number of all and ubiquitinated peptides identified in 60 mg of total protein in JY cell line. **D** Number of all and ubiquitinated unique protein groups identified in 60 mg of total protein in JY cell line.

To improve the specificity of ubiquitin-pulldown, we considered the factors that could cause low enrichment specificity. Since ubiquitination is a rare post-translational modification, we reasoned that there might be only a relatively limited amount of ubiquitinated peptides, such that it was not sufficient to saturate the antibody-coated beads fully. On the other hand, increasing the input material might provide more ubiquitinated peptides for resin binding. This may out-compete unmodified peptides with lower affinity, thereby increasing the proportion of ubiquitinated peptides in our preparation.

To test this hypothesis, we repeated the ubiquitin-pulldown with a larger input load of 60 mg digested peptides. We identified a total of 25,904 peptides, of which 3,759 were modified by ubiquitin (Fig. 5C). These ubiquitinated peptides originated from 2,260 distinct protein groups (Fig. 5D), averaging about 1.66 ubiquitinated peptides per protein.

Although the number of ubiquitinated peptides increased with the larger input load, the proportion of ubiquitinated peptides remained almost unchanged (14.5%). Thus, we reasoned saturation was not the cause of low specificity of antibody retrieval of ubiquitinated peptides. It is possible that other parameters such as duration of pulldown and wash procedures may further improve specificity, but due to time constraints and limitation in cell lysate material, we could not optimize this further, and instead chose to apply the same worklow to our FBXW7 R505C biological system.

2.2 FBXW7 R505C mutation severely impairs intracellular ubiquitination

To evaluate changes in the ubiquitinated proteome elicited by FBXW7 R505C, we applied the workflow optimized above (Fig. 6A) to three colon organoid lines (WT parental line, single mutant CRISPR-edited FBXW7 harbouring mutation R505C, and double mutant FBXW7 R505C/TP53KO). In the WT line, we identified 7,222 diGly peptides quantified in at least one technical replicate and 5,484 diGly peptides quantified in at least two out of three technical replicates in 2,028 unique ubiquitinated proteins (Fig. 6B). It is evident that the numbers of identified ubiquitination dropped in both mutant organoid lines, in line with the function of FBXW7 in protein ubiquitination, as FBXW7 is F-box substrate receptor subunit of E3 ligase complex, and R505C mutation is located in its WD40 domain at the substrate-binding interface, disrupting the interaction with a substrate. Therefore, the effect of FBXW7 R505C mutation on ubiquitination was dramatic. In addition, the phenotype (loss of ubiquitination) is fully recapitulated in a second double-mutant organoid line generated separately, further confirming that the biological phenotype of hampered protein ubiquitination is reproducible.

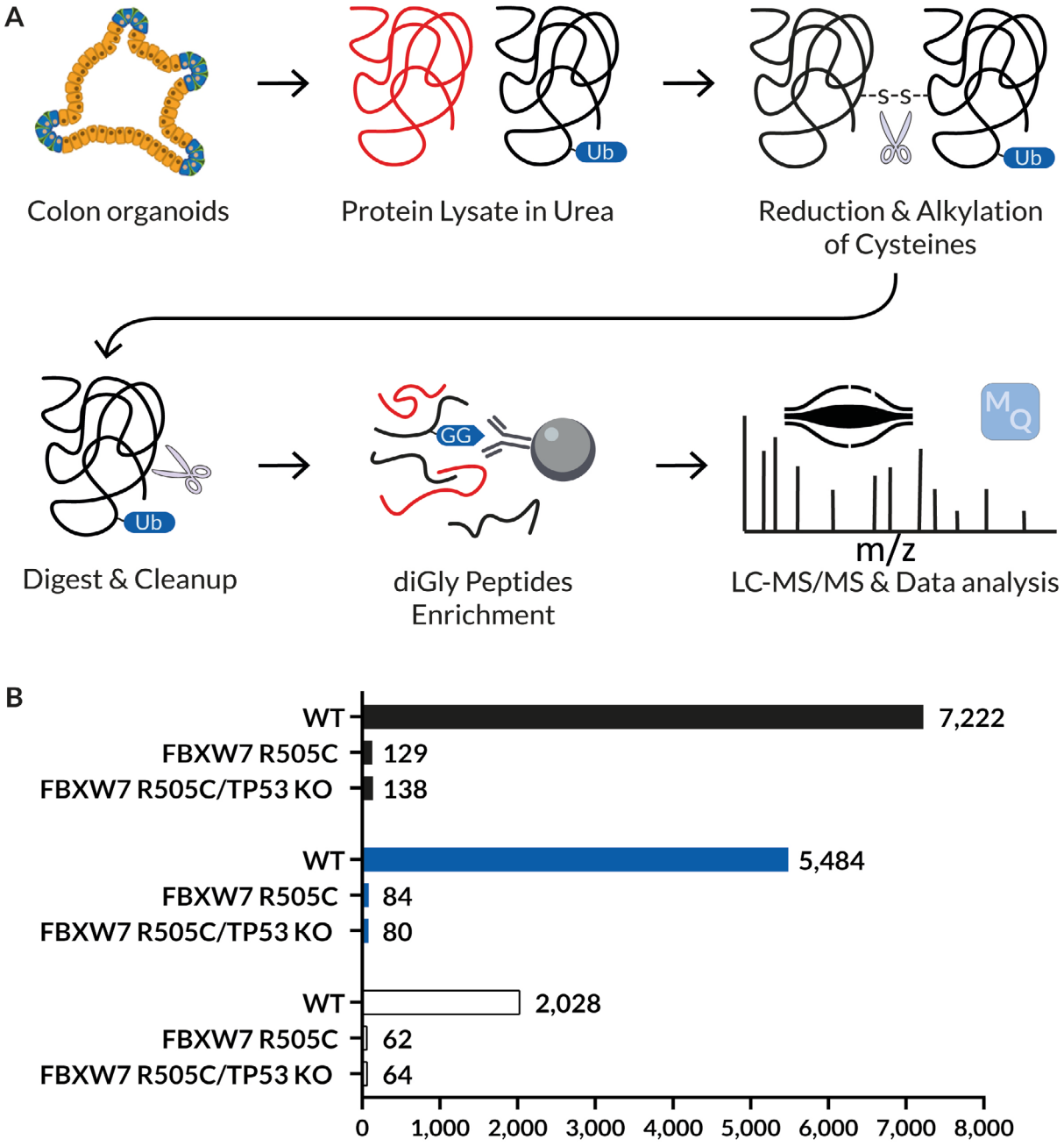

Fig. 6 | **A** Schematic workflow to identify ubiquitinated proteins. Colon organoids were lysed. Protein disulfide bridges were reduced and alkylated, followed by digest. K-ε-GG stab was left from ubiquitin chains after trypsin digestion. Ubiquitinated peptides were enriched by immunoprecipitation with K-ε-GG antibody, which recognises K-ε-GG remnant. Pulled down peptides were analysed by LC-MS/MS. Protein identification and quantification were performed using MaxQuant software[20]. **B** Bar plot of the total number of diGly-modified peptides quantified in at least 1 out of 3 replicates (black), diGly-modified peptides quantified in at least 2 out of 3 replicates (blue) and diGly-modified proteins quantified in at least 2 out of 3 replicates (white) in 25 mg of total protein in each organoid line. We did not use a stringent cut of 3 out of 3 replicates because of possible gaps in identification using data-dependent acquisition (DDA) method due to variation in top N precursor selection by MS instrument.

2.2.1 FBXW7 mutant shows dysregulated signalling mediated through impaired protein degradation

Since precursor peptide selection in the mass spectrometer is data-dependent and can be sporadic even when measuring the same samples thrice. It is therefore challenging to ascertain beyond doubt if a peptide is truly absent. Nonetheless, the likelihood of a peptide actually present but is repeatedly missed during precursor selection in all three injection replicates is much lower. For this reason, we applied stringent filtering to identify which proteins were likely to be no longer ubiquitinated in FBXW7 R505C organoids. Out of 2,028 protein detected with ubiquitin modification in WT organoids, 1,644 proteins were quantified in all three replicates in WT and none of the replicates in both mutants: single mutant FBXW7 R505C and double mutant FBXW7 R505C/TP53 KO (Supplementary Table 1). We expect hampered ubiquitination to be a phenotype that is also conserved in FBXW7 R505C/TP53 KO double mutant, since the ability to ubiquitinate proteins depends on FBXW7 and not directly on TP53. Indeed the drastic loss of protein ubiquitination was also recapitulated strongly in the double-mutant, giving further confidence that these proteins are certainly ubiquitination targets of FBXW7. Based on this observation, it is likely that the 1,644 proteins uniquely ubiquitinated in WT organoids are bona fide targets of FBXW7 ubiquitination.

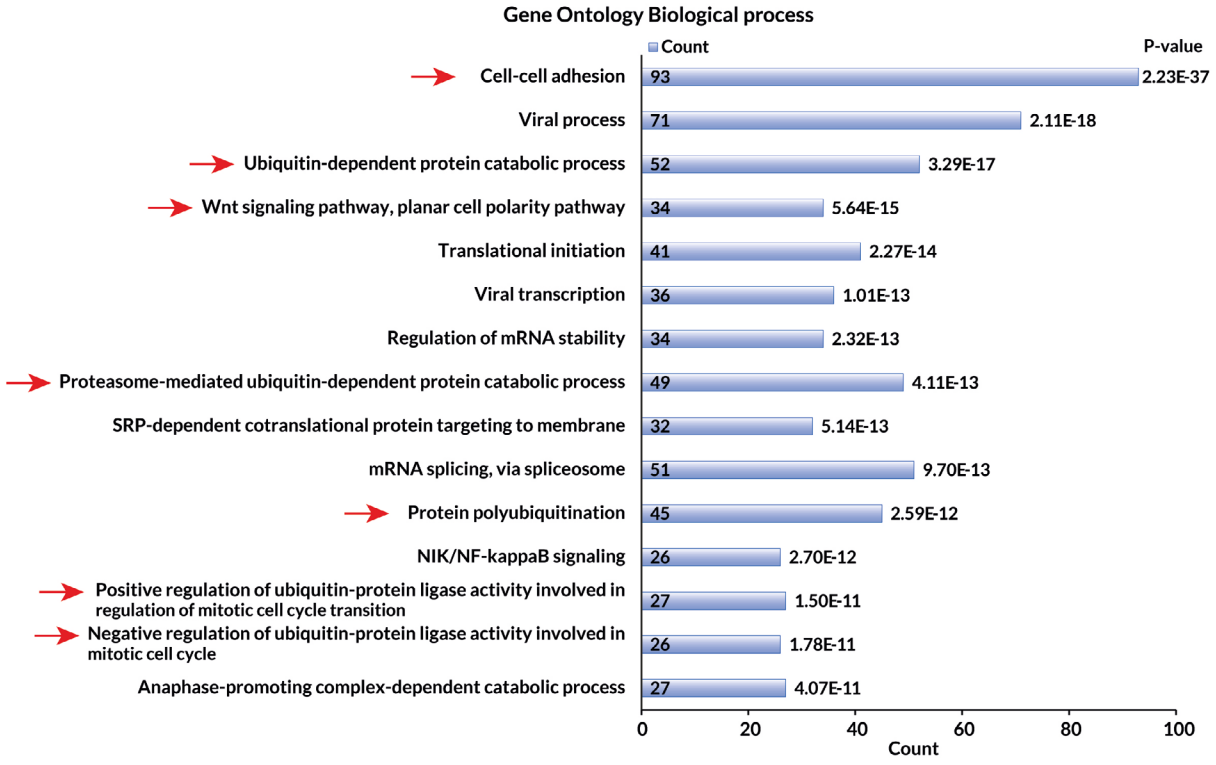

Fig. 7 | Gene ontology biological process resulting from DAVID[22,23] analysis of 1,644 proteins ubiquitinated in WT and not ubiquitinated in both FBXW7 R505C single mutant and FBXW7 R505C /TP53 KO double mutant organoid lines. 15 the most significantly ($p \leq 4.07 \times 10^{-11}$) enriched biological processes are given. *Count – gene counts belonging to annotation term.

RESULTS

To dissect biological processes significantly affected by FBXW7 R505C mutation, these 1,644 proteins (Supplementary Table 1) were categorized by Gene Ontology to obtain a hint of cellular processes that may be changed as a result of failure to ubiquitinate (Fig. 7). 92 significantly enriched biological processes (p<0.05 and FDR<0.05) were identified (Supplementary Table 2). GO analysis revealed that proteins with ubiquitination loss in mutants participate in cell-cell adhesion, Wnt signalling pathway, planar cell polarity pathway. As anticipated, many ubiquitin-involved processes, such as ubiquitin-dependent protein catabolic process, protein polyubiquitination, regulation of ubiquitin-protein ligase activity(Fig. 7), were also extensively altered. This suggests possible auto-regulation of the ubiquitination pathway by controlling the steady-state provision of enzymes and factors needed for ubiquitination. We reasoned that inability to ubiquitinate induces cells to preserve components of the ubiquitination machinery longer, and hence loss in degradation. This could likely expand and potentiate the impact of FBXW7 R505C mutation on proteostasis within the cell.

Visualising proteins distinctly ubiquitinated in WT compared to both mutants on Kyoto Encyclopedia of Genes and Genomes (KEGG)[24–26] pathway maps, it became clear that terms 'Ubiquitin mediated proteolysis' (hsa04120), 'Protein processing in endoplasmic reticulum' (hsa04141) and 'Cell cycle' (hsa04110) were impacted by the loss of degradation (Fig. 8). Overall, 24 significantly enriched pathway terms (p<0.05 and FDR<0.05) were identified (Supplementary Table 3). These results demonstrate that FBXW7 R505C mutation account for the retention of proteins involved in every step of ubiquitin mediated proteolysis (Fig. 8A). Moreover, FBXW7 R505C mutation contributes to the accumulation of many proteins of 'MAPK signaling pathway' (has04010), especially the receptor, from which the entire pathway is subsequently activated (Fig. 8B).

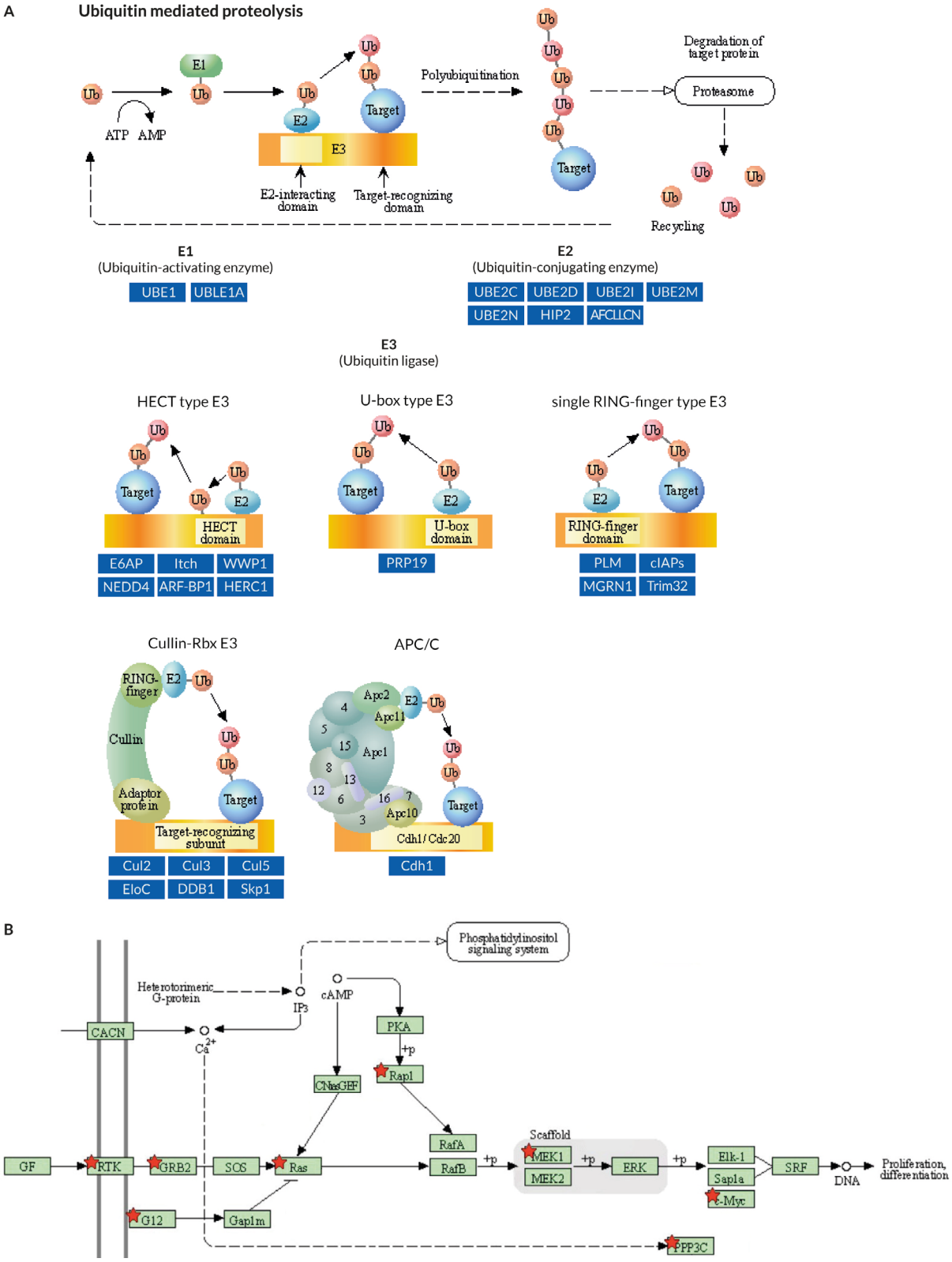

Fig. 8 | Mapping of proteins to the KEGG pathways. **A** 'Ubiquitin mediated proteolysis' (hsa04120) where blue rectangles contain proteins uniquely ubiquitinated in WT compared to both mutants: FBXW7 R505C single mutant and FBXW7 R505C /TP53 KO double mutant. **B** A portion of 'MAPK signaling pathway' (has04010) where proteins uniquely ubiquitinated in WT compared to both mutants: FBXW7 R505C single mutant and FBXW7 R505C /TP53 KO double mutant, flagged with red stars.

2.2.2 Cross-validation by degron prediction

Using our sensitive approach to map ubiquitination sites, we found 1,644 proteins where ubiquitination is FBXW7-dependent and abolished in the FBXW7 R505C mutants. To assess the confidence in these identifications, we cross-validated our target list against degron prediction that had been created for FBXW7 previously[11]. This prediction method was designed by a random forest classifier trained on known degrons and was needed because experimental detection of E3-substrate complexes has been challenging and problematic[11,13,27]. We used this publicly available approach, in consultation with Francisco Martínez-Jiménez, the first author leading the publication[11], to check for FBXW7 internal degrons.

Search across 1,644 proteins (Supplementary Table 1) estimated 125 proteins (a total of 413 protein isoforms) match any of the two known FBXW7 degron motifs (Supplementary Table 4). Furthermore, 79 of these 125 proteins match the motif and a set of 11 favourable biochemical features that resemble those of annotated degron instances. Finally, 16 of these proteins are known to be involved in cancer (known cancer drivers). It is surprising to note that amongst the 1,644 putative FBXW7 ubiquitination targets we identified in this study, less than 10% can be verified by this prediction algorithm. This may be rationalized by the fact that this prediction algorithm can only handle linear phospho-degrons (of continuous primary degron sequence), whereas degron recognition in cells works based on a three-dimensional groove. In this respect, the novel FBXW7 putative targets identified in this study can be used to refine the prediction algorithm further.

2.2.3 Epidermal Growth Factor Receptor (EGFR) is a FBXW7 substrate

An observable phenotype of FBXW7 R505C colon organoids is mitogen independent growth. It has been shown that the R505C mutation eliminates the requirement for EGF to proliferate (personal communication via Matteo Boretto, Hubrecht Institute). To link the FBXW7 R505C degradation phenotype to aggressive growth, we focused on the degradation of EGFR in our model system. EGFR is a receptor tyrosine kinase and known oncogene[13,28]. It is a transmembrane protein that responds to various mitogen growth factors, particularly EGF[13,29], making it easy for experimental phenotyping, and may be directly responsible for the aggressive growth phenotype of FBXW7 R505C organoids.

Using our established workflow, we observed that EGFR ubiquitination was hampered when FBXW7 was mutated (Fig. 9A). This observation implies that the FBXW7 R505C mutant organoids may have a higher EGFR protein level since it is less degraded. Indeed, from Fig. 9B, it can be seen that EGFR abundances in proteome increased >2-fold change (FC) in FBXW7 R505C mutant compared to WT, whereas ubiquitinated EGFR peptides were not detected in FBXW7 R505C mutant (Fig. 9A). EGFR function is determined by receptor abundance, but also by phosphorylation, since intracellular cross-phosphorylation ultimately governs activity of the EGFR dimer. Using phosphoproteomics (assisted by Ziliang Ma), we found that the increase in total EGFR phosphorylation was four-fold in the FBXW7 R505C mutant compared to WT (Fig. 9C), and 3.9-fold at Thr693, a key activating site (Fig. 9D). This increase in phosphorylation remains upon normalization to a two-fold stabilisation of the total EGFR protein level. Collectively, this indicates that EGFR signaling is likely increased in FBXW7 R505C organoids, and that the threshold requirement for EGF ligand stimulation may be lower or even not necessary. This is coherent with the EGF-independent growth of FBXW7 R505C organoids over months in culture, whereas WT organoids cannot survive beyond a week without EGF supplementation in culture.

In addition, via degron prediction, we confirmed the presence of the FBXW7 degron motif in EGFR, surrounding the phosphorylation site Thr693 (Supplementary Table 4). The degron prediction was performed according to established methods[11]. Next to EGFR, 124 additional FBXW7 substrates identified by ubiquitin pulldown were also independently predicted to be under the ubiquitin modification of FBXW7.

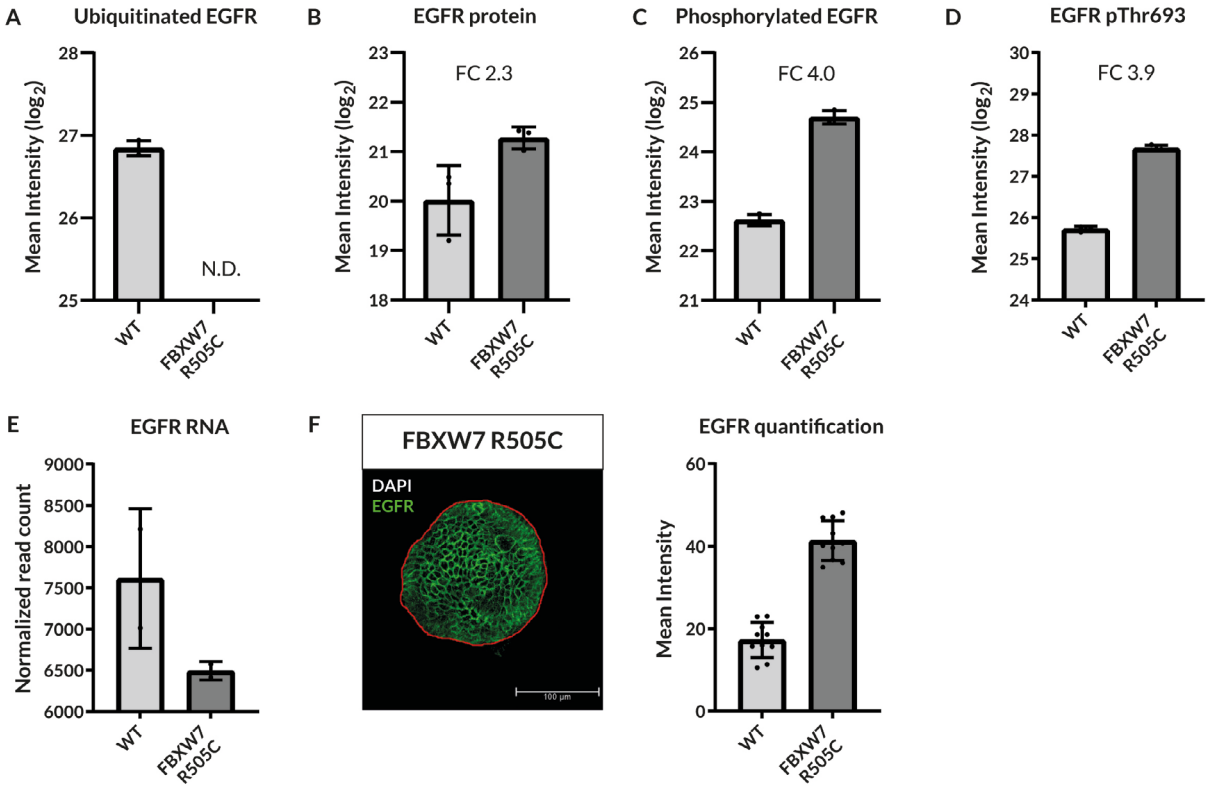

Fig. 9 | **A** Intensity of EGFR ubiquitinated peptides in WT, and FBXW7 R505C mutated organoid lines. *N.D. - not detected. **B** Intensity of EGFR protein in WT and FBXW7 R505C mutated organoid lines. **C** Intensity of EGFR phosphorylated peptides in WT, and FBXW7 R505C mutated organoid lines. **D** Intensity of EGFR phosphorylated peptide with Thr693 phosphorylation site in WT, and FBXW7 R505C mutated organoid lines. **E** Normalised read count in WT and FBXW7 R505C mutated organoid lines. **F** Immunofluorescence of EGFR and DAPI staining in WT and FBXW7 R505C organoids. Green, EGFR; white, DAPI. The white bar is the scale bar, 100 µm. N=10 organoids imaged per clone. Immunofluorescence signal intensity of EGFR in WT and FBXW7 R505C mutated organoid lines.

The remaining question is whether the EGFR transcript level is already elevated in FBXW7 R505C cells, or if EGFR protein accumulation is primarily driven by the lack of FBXW7-dependent ubiquitination and degradation. To elucidate this, we compiled multiple levels of evidence, from EGFR RNA level to EGFR protein, ubiquitinated peptide, phosphopeptide to cell surface staining of EGFR protein on organoids, for combined analysis. As shown in Fig. 9E, the EGFR RNA level was actually reduced in FBXW7 R505C organoids, making it unlikely that elevated EGFR (Fig. 9B) can result from transcriptional regulation. On the other hand, all the readouts on EGFR protein levels were increased (Fig. 9B, F), and ubiquitination (Fig. 9A) was not observed. Collectively, these more confidently establish EGF-independent growth in FBXW7 R505C organoids, as a phenotype resulting from loss of EGFR degradation.

Predicted degron at EGFR spans 691 to 697 amino acids, including the phosphorylation site T693 with several flanking lysines (708, 713, 714, 716)[11]. Phosphorylation adds an additional level of regulation to the existing primary degron motif sequence and con-

trols target recognition by E3 ubiquitin ligase. As illustrated by Fig. 9D, the abundance of threonine 693 phosphorylation site in EGFR increased 3.9 fold in FBXW7 R505R mutant compared to WT, suggesting that EGFR is an FBXW7 substrate that possesses the necessary degron. The ubiquitination sites, phospho-peptide and predicted degron surrounding T693 are annotated in Fig. 10.

Fig. 10 | EGFR (P00533) sequence. Predicted degron motif coloured in red. Peptides identified during ubiquitin pulldown in WT organoid line coloured in blue. Identified modifications annotated below corresponding amino acid: diGly – ubiquitination site (in blue), P - phosphorylation site (in red). Phosphorylated peptide marked with a yellow line on top. Amino acids are shown by single-letter code. No diGly sites were identified in 1-659 amino acids residues.

Furthermore, these trends are in line with the previously described KEGG analysis showed the MAPK pathway is highly affected by FBXW7 R505C mutation, since many proteins involved were not degraded in FBXW7-mutated organoids, which enhances MAPK signalling (Fig. 8B). Collectively, these may further support the aggressive growth phenotype of FBXW7 R505C organoids.

2.2.4 Ubiquitination inversely correlates with protein abundance

Failure to degrade a protein may lead to intracellular accumulation. By this rationale, actual degradation substrates of FBXW7 could accumulate to higher protein levels in FBXW7 R505C. By overlapping experimentally-detected FBXW7 substrates (peptide pulldown, MS) and proteins with predicted FBXW7 degrons, we retained 125 proteins to further correlate the steady-state proteome abundance with the FBXW7 defect. As shown in Fig. 11A, a total of 12 proteins were identified as putative FBXW7 substrates that accumulate in organoids by at least 50% in R505C mutants. Amongst these, GPA33[30,31], VPS13C[32,33], TGOLN2[34] and IMMT[35,36] genes are cancer-related, and shown to accumulate to higher levels in both mutants, providing additional biological reliability.

A cell-surface glycoprotein A33, encoded by GPA33 gene, has shown to be present in >95% of primary and metastatic tumours in colorectal cancer[31]. A therapeutic strategy targeting GPA33 cell-surface expression has been in phase I Clinical trial (NCT02248805)[37]. An anti-glycoprotein A33 (gpA33)/anti-CD3 bispecific humanized monoclonal antibody MGD007 has been developed to recruit CD3-expressing T cells to target gpA33-expressing tumour cells[30]. Antitumor activity is provided through cancer cell lysis by cytotoxic T-lymphocytes. If FBXW7 is mutated in a tumour, it may then be possible to treat such tumours with anti-GPA33 antibodies. Another protein accumulated in FBXW7 mutants is inner membrane mitochondrial protein encoded by IMMT gene. IMMT has shown to be highly expressed in lung adenocarcinomas and correlated with a more advanced stage, larger tumor size and a poorer prognosis[35]. IMMT as mitochondrial marker could also be applied to predict poor clinical outcome in gastric cancer patients[36].

Conversely, the abundance of proteins not regulated by FBXW7-dependent degradation should not change significantly in the FBXW7 mutant models. The ubiquitination of 23 proteins could still be detected in FBXW7 mutants as well in WT. The steady-state abundance of these proteins remains largely unchanged as shown in Fig. 11B, C.

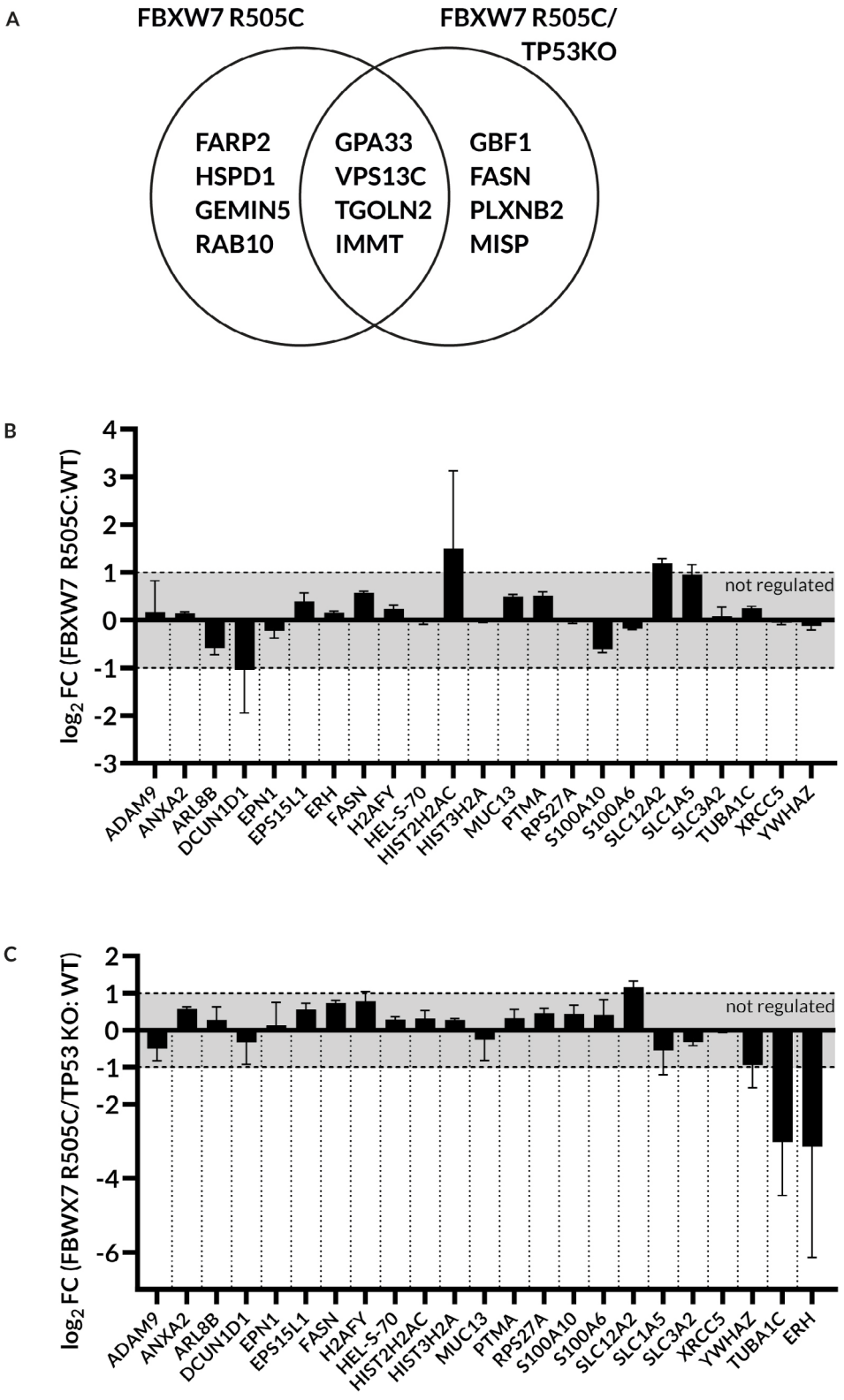

Fig. 11 | **A** Venn diagram of proteins significantly (q-value<0.05) upregulated +1.5 FC in FBXW7-mutated organoids compared to WT organoids. Proteins listed fit the following filtering criteria: not ubiquitinated in both FBXW7-mutated organoids (single and double mutant in every replicate) while ubiquitinated in WT organoids (every replicate); have predicted FBXW7 degron motif. **B, C** Proteome level fold change analyses of proteins quantified across all three organoid lines in ubiquitinomics (at least 2 out of 3 replicates). Fold changes are statistically significant by paired t-test (q-value<0.05).

3. Discussion and future outlook

The ubiquitin machinery can determine a protein's lifespan inside the cell; the ability to ubiquitinate proteins can regulate many critical pathways inside a cell, like cell cycle progression, signalling, differentiation, growth and antigen presentation[5,38]. As such, dysregulation of ubiquitin pathways can contribute to the development of many diseases, including cancer[1,3,9,39,40].

Despite the importance of the ubiquitin system, it remains to be completely decoded[4,40]. The ubiquitin pathway consists of a cascade of three consecutive enzymatic reactions[2,3,38]. We decided to focus on a more detailed study of the last step - recognition of the substrate by the E3 ligase and the attachment of the ubiquitin molecule to substrate, for several reasons. The E3 family is the least studied and most extensive (around 600 E3 ubiquitin ligases)[13,41,42], and each E3 ligase is more target-specific[13] than the ubiquitin-activating enzyme E1 or the ubiquitin-conjugating enzyme E2, with only 2 and 30-50 genes in humans respectively[2,42,43]. This E3 specificity can be used in the future to find potential biomarkers or develop targeted drugs with fewer side effects and understand the resistance mechanism of already approved treatments[2,5,9,44]. In addition, in theory, defects at an earlier stage, such as ubiquitin itself or a ubiquitin-activating enzyme E1, should lead to lethal consequences since these would affect not only the entire ubiquitin cascade, but also ubiquitin signaling. As expected, E1 and E2 mutations in the human genome are not tolerated[13].

In this work, we focused on the FBXW7 E3 ligase since it is known to be a frequently mutated tumour suppressor[14]. In the literature, several oncoproteins are already known FBXW7 substrates, for instance, JUN, cyclin E, c-Myc, Mcl-1, mTOR, Notch and AURKA[1,13,14]. To focus only on the enzymatic activity of FBXW7, we introduced an R505C point mutation instead of a full FBXW7 knockout. Amongst the known mutations in FBXW7, R505C was found in many cancers and in both hotspot and loss of function criteria[1]. Using a ubiquitinated peptide pulldown strategy that was optimized in this project, we detected a drastic loss of protein ubiquitination in FBXW7 R505C organoids, as well as FBXW7 R505C/TP53 KO organoids. Collectively these establish FBXW7 as an important and dominant E3 ligase in the colorectal system.

Although E3 enzymes are generally considered major determinants of substrate specificity, we hypothesised that in our case, FBXW7 was a highly dominant E3 enzyme in the colon organoids, which could explain the striking loss of most ubiquitinations. This is in general agreement with previous studies where the number of proteins that inter-

act with a particular type of E3 ligase differs depending on the tissue type[13]. Amongst 1,644 experimentally detected FBXW7 substrates, 125 were cross-verified by degron motif prediction using established methods. Although important substrates like EGFR emerged from this double data-filter, only 8% of FBXW7 experimental substrates were predicted with FBXW7 degron sequence. This may be due to the limitation of prediction in a linear sequence, i.e. 3D docking between the substrate and E3 subunit is not taken into account[13]. Moreover, the random forest classifier on which degron prediction is based, was trained only on ~200[11] the previously reported degron instances. Collectively, these limit the utility of degron prediction as the only basis for FBXW7 substrate identification. On the contrary, the extensive experimental FBXW7 substrate dataset obtained in this project may be extremely useful to refine degron prediction further.

Ubiquitin modifications depend on phosphorylation in degron motifs[13]. Ubiquitination also decreases the protein abundance at the proteome level through substrate degradation in the proteasome[12]. Therefore, joint and combined analyses of these intertwined protein states makes it possible to reveal intricate patterns and interactions within the cell. As we demonstrate here with the EGFR example, it would be beneficial to acquire and analyse all these datasets together from the same sample to get a complete picture.

A distinctive feature of cancer is rapid and uncontrolled cell growth, and dysregulation of homeostasis in proteins that govern proliferation is a clear link to the clinical phenotype of FBXW7 loss of function. The MAPK pathway is one of the main mechanisms regulating cell proliferation[45], that is also often disrupted in many types of cancer[45,46]. We demonstrated the value of our multi-omics method using EGFR. Hampered ability to ubiquitinate EGFR in the FBXW7 R505C mutant leads to an increase in EGFR receptors on the membrane surface. This, in turn, creates conditions for the activation of the MAPK pathway even when a small amount of EGF ligand is present, and in turn, proliferative signals are relayed for uncontrolled cell growth. In this regard, EGFR inhibitors[29] may help to functionally counteract the accumulation of cell-surface receptors resulting from hindered degradation, and aid in normalizing the growth signals downstream of lower EGFR activation threshold.

Despite the successful implementation of this ubiquitinated peptide pulldown, and the identification of essential mechanisms that support FBXW7 R505C driven cancer, the strategy described here cannot yet distinguish the myriad of ubiquitin modifications. Enzymatic digestion with trypsin could almost entirely trim away the ubiquitin tree

except for a di-glycine stub on the modified lysine. In practice, this significantly reduces the sample complexity to map only ubiquitination sites, but inevitably collapses biological variation in ubiquitin branching. In addition, ubiquitin-like modifiers ISG15 and NEDD8 have the same diGly tag at the junction lysine on the substrate[4,19,47], this makes the strategy here also unsuitable to distinguish sites of ubiquitin modifications from sites of NEDD8 or ISG15. Nonetheless, the percentage of such modifications is documented to be much lower (<6%) compared to ubiquitin[18,19,47,48].

Since ubiquitinated peptides are retrieved by antibody pulldown in our strategy, the specificity depends heavily on antibody quality and characteristics. The attempt to optimize ubiquitinated peptide retrieval was made, but specificity remained low (at ~15%) presumably due to antibody characteristics, which could not be overcome by experimental handling. It might, however, be possible to raise an antibody in-house with better specificity. Although slight peptide sequence biases have been previously described in commercially available di-Gly antibodies[47], a motif logo of antibody PTMScan® showed no distinct selectivity for specific amino acids in the proximity of lysine residues with attached ubiquitin[21], which makes identification of all existent ubiquitination sites feasible.

Although the dataset presented here is already extensive compared to most existing high throughput studies of protein ubiquitination, it remains possible to expand this further with prefractionation and other forms of MS acquisition. For instance, Data Independent Acquisition (DIA) methods have been applied successfully to phosphoproteomics[49] and lately to ubiquitinomics[48]. This could ameliorate the problem of sporadic precursor selection in Data Dependent Acquisition (DDA)[2].

In conclusion, understanding the causes and consequences of defects that can arise in the ubiquitin system could lead to new therapeutic opportunities[1,9,13]. Although the small surface area between E3 ubiquitin ligase and degron binding domain of substrate is considered a difficult target for drugs due to low-affinity interaction, progress is still being made[13]. Medicines that modulate or re-target E3 ubiquitin ligase activity, such as Lenalidomide (Revlimid®), Thalidomide (Thalomid®), Pomalidomide (Pomalyst®) are already on the market[5].

With a complete understanding of the interaction mechanism of the E3 ubiquitin enzymes and their targets, it might be feasible to create molecules that could connect these two interacting agents, acting as an additional linker[13]. Ultimately the attachment of the ubiquitin molecule to the substrate will be restored. In addition, defining the

ubiquitination substrates impacted by E3 ubiquitin ligase mutations could also aid in understanding how diseases if impaired ubiquitination can arise, as we demonstrate with EGFR in this work.

4. Materials and methods

4.1. Cell culture and organoid lines

The B-lymphoblastoid cell line JY and three human colon organoid lines were used in this study. Normal (WT) colon organoids were derived and maintained as described previously[50]. From the WT organoid line, FBXW7 R505C mutation was introduced by the CRISPR-Cas9 system in collaboration with Matteo Boretto (Hubrecht Institute, Utrecht). Another double mutant organoid harbouring both FBXW7 R505C mutation and TP53 knockout was also included to evaluate if the absence of P53 has an influence over protein degradation phenotype.

4.2. Sample preparation

Organoid lysis and digestion. Organoid material was lysed in 8 M urea, 50 mM ammonium bicarbonate, supplemented with EDTA-free protease inhibitor (cOmplete Mini, Roche Diagnostics) and phosphatase inhibitor (PhosSTOP, Roche Diagnostics), then was vortexed several times and ultra-sonicated for 5 min with 50% duty of 5 cycles. Lysates were cleared by centrifugation at 20,000 g for 1 hour at 15 °C. Protein content of the supernatant was determined using the Bradford protein assay (Bio-Rad, USA). Proteins of each sample were reduced in 10 mM dithiothreitol (DTT) at 20 °C for 1 hour and then alkylated in 20 mM iodoacetamide (IAA) at 20 °C for 30 min in the dark. Samples were diluted with 50 mM ammonium bicarbonate to a final concentration of urea <2 M. Lys-C (Wako, Japan) was added for digestion at an enzyme/protein ratio of 1/75 and incubated for 4 hours at 37 °C, followed by trypsin (Sigma) at an enzyme/protein ratio of 1/75 and incubated overnight at 37 °C.

Digested peptides were purified using reverse-phase Sep-Pak 3cc (200mg) tC18 cartridges (Waters, Ireland). Columns were activated with 100% acetonitrile (ACN), and equilibrated with 0.1% formic acid in water (loading buffer). Samples were diluted to pH<3 with loading buffer and applied to the column with one additional reload of the column flowthrough. tC18 cartridges were washed once with loading buffer, and peptides were eluted twice with 500 µl 60% ACN in 0.1% formic acid. The eluates were subsequently combined and dried by vacuum centrifugation using refrigerated CentriVap Concentrator (Labconco, USA).

Immuno-affinity purification (IAP) of ubiquitinated peptides. To isolate ubiquitinated peptides, 25 mg of digested peptides were dissolved in 10 ml of phosphate-buffered saline (PBS) for immunoprecipitation (IP) with ubiquitin branch motif (K-ε-GG) antibody bead conjugate (Cell Signaling, kit #5562). After 7 hours of IP at 4 °C with end-to-end

rotation, agarose beads were separated from the supernatant by gentle centrifugation at 50 g for 1 min. The resin was washed twice with 12 ml cold PBS, and immunoprecipitated peptides were eluted for 2 min using 500 µl of 0.15% trifluoroacetic acid (TFA) in water. The beads were then equilibrated with PBS once, and re-incubated with the IAP flowthrough, for two additional rounds of ubiquitinated peptide pulldown. The three elutions were pooled and desalted using Sep-Pak Vac 1cc (50mg) tC18 cartridges (Waters) as described before. All samples were dried by vacuum centrifugation and stored at -80 °C for mass spectrometry analyses.

Western Blotting for p53 protein validation. 10 µg of protein from each sample lysate were denatured by boiling in XT Sample buffer (Bio-Rad) with 100 mM DTT. Proteins were resolved on 12% Bis-Tris Criterion XT precast SDS-PAGE gel (Bio-Rad) at a constant voltage of 80 V for 20 min, followed by 150 V for 80 min. Proteins were transferred to a PVDF membrane (Bio-Rad) by wet-tank transfer at 100 V for 1 hour at 4 °C in Towbin buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol (pH 8.3), Bio-Rad). The membrane was then blocked for 1 hour at room temperature with 5% milk in tris-buffered saline containing 0.1% Tween 20 (TBS-T), rinsed in TBS-T and then incubated overnight at 4 °C with an anti-p53 antibody (SC Biotech, DO-1) diluted 1:1,000 with 1% milk in TBS-T. After overnight incubation, the membrane was washed 3x 5 minutes in TBS-T on the Orbital VWR digital shaker and then incubated 1 hour at room temperature with HRP-conjugated anti-mouse secondary antibody (Agilent Technologies, Denmark) diluted 1:2,500 1% milk in TBS-T. SuperSignal West Dura (Thermo Scientific, USA) chemiluminescent substrate was added to membrane and signal was detected using an Amersham Imager 600 (GE healthcare, USA). To assess loading, the blot was further incubated in α-Actin antibody (Sigma, A2066; 1:1,000 dilution) for 2 hours at room temperature. The secondary antibody was goat anti-rabbit (Dako, Denmark; 1:2,500 dilution) with 1 hour incubation at room temperature.

4.3. LC-MS/MS analyses

Peptides were reconstituted in 2% formic acid and triplicate injections were analysed on an Orbitrap Exploris 480 mass spectrometer (Thermo Scientific), coupled to an UltiMate 3000 UHPLC system (Thermo Scientific). Solvent A was 0.1% formic acid in water, Solvent B was 0.1% formic acid in 80% acetonitrile, 20% water. Peptides were first trapped on an µ-precolumn (C18 PepMap100, 5 µm, 100 Å, 5 mm × 300 µm; Thermo Scientific) in 9% Solvent B, and then separated on an analytical column (120 EC-C18, 2.7 µm, 50 cm × 75 µm; Agilent Poroshell). The flow rate was kept at 300 nl/min. Peptides from K-ε-GG pulldown were separated by 155 min gradient from 10 to 38% solvent B, followed by washout with a 3-min increase to 99%, kept at 99% for 4

min and finally equilibration back to 9% solvent B for 1 min, remaining the same for the last 10 minutes. For the proteome analysis, the gradient was as follows: 9-13% in 1 min, 13-44% in 155 min, 44-99% in 3 min, 99% for 4 min, and finally 99-9% in 1 min.

Eluting peptides were online-injected into the mass spectrometer for data-dependent acquisition. The temperature of the ion transfer tube was set to 275 °C and a RF lens voltage of 40%. MS scans were acquired at a resolution of 60,000 within the m/z range of 375-1600, accumulating to 'Standard' pre-set automated gain control (AGC) target. Multiply charged precursor ions starting from m/z 120 were selected for further fragmentation. Higher energy collisional dissociation (HCD) was performed with 28% normalised collision energy (NCE), at a resolution of 30,000, with 1.4 m/z isolation window and dynamic exclusion of 24 s.

4.4. Raw data processing

MS data were acquired with Thermo Scientific Xcalibur (version 4.4.16.14), and raw files were processed using MaxQuant software[20] (version 1.6.17.0) with the integrated Andromeda search engine[51]. Data were searched against the human UniProt database including common contaminants (downloaded in April 2021, containing 196,111 entries - for PTM analysis; in August 2018, containing 173,324 entries - for the proteome analysis). For all files, standard parameter settings were used with enabled the label-free quantification (LFQ) algorithm. Trypsin/P was set as the digestion enzyme (cleaves after lysine and arginine also if a proline follows), and up to two missed cleavages were tolerated. The match-between-run feature was enabled for identification, with a match time window of 0.7 min and an alignment time window of 20 min. A false discovery rate (FDR)[52,53] of 1% was used for peptide and protein identification. Cysteine carbamidomethylation was included as a fixed modification. Protein N-terminal acetylation and methionine oxidation were allowed as variable modifications for proteome searches, with added GlyGly (K) for the ubiquitinated proteome analysis.

4.5. Data analysis and availability

The provided data was analysed using Microsoft Excel and Perseus software[54,55] (version 1.6.15.0). LFQ intensities[56,57] of proteins were $\log_2$ transformed. Proteins quantified in two out of three replicates in one organoid line were retained for further analysis, after imputation based on the normal distribution. To filter for significant changes between experimental conditions two-sided unpaired Student's t test was performed. FDR-corrected *p*-values (*q*-values) were calculated from 250 randomisations and considered significant if they were 0.05 or less. Gene ontology analysis was performed with Database for Annotation, Visualization, and Integrated Discovery (version 6.8/

DAVID)[22,23]. Data visualisation was produced with GraphPad Prism software (version 9.2.0).

Data availability. The mass spectrometry proteomics data about FBXW7 have been deposited to the ProteomeXchange Consortium via the PRIDE[58] partner repository with the dataset identifier PXD028905.

5. Acknowledgements

First and foremost, I am incredibly grateful to my supervisor Wei Wu, for the excellent opportunity to work on the fascinating project with the top-of-the-line equipment, opening to me, a naive student from Belarus, the world of international science. Thank you for your tutelage, invaluable advice, continuous support, being an inspiring and demanding supervisor.

I would like to thank Matteo Boretto (Hubrecht Institute, Utrecht) for generating CRISPR-edited FBXW7-mutated organoid lines, immunofluorescence staining of organoids, Ziliang Ma (Utrecht University) for providing the phosphoproteomics dataset, and Francisco Martínez-Jiménez (UMC Utrecht) for degron prediction.

In addition to the people who directly worked on the project, I would like to thank all technicians without whom the work would not be so smooth. Special appreciation to Mirjam Damen, the first person to introduce me to the laboratory and teach techniques. Thank you for the excellently organised stock ordering and equipment setup.

My sincere thanks to Marc Baggelaar for allowing me to enrol on the Biomolecular Mass Spectrometry course. I learned so much from you during the course.

My gratitude extends to Monique Slijper, without whom I would not be where I am today. Thank you for inviting me to pursue a Master's in such a stimulating environment and for the guidance you have given me throughout my studies.

Finally, I would like to thank the entire Heck Lab.

6. References

1. Ge, Z. *et al.* Integrated Genomic Analysis of the Ubiquitin Pathway across Cancer Types Resource Integrated Genomic Analysis of the Ubiquitin Pathway across Cancer Types. 213–226 (2018).

2. Vere, G., Kealy, R., Kessler, B. M. & Pinto-Fernandez, A. Ubiquitomics: An overview and future. *Biomolecules* **10**, 1–22 (2020).

3. Damgaard, R. B. The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. *Cell Death Differ.* **28**, 423–426 (2021).

4. Akimov, V. *et al.* Ubisite approach for comprehensive mapping of lysine and n-terminal ubiquitination sites. *Nat. Struct. Mol. Biol.* **25**, 631–640 (2018).

5. Dunphy, K., Dowling, P., Bazou, D. & O'Gorman, P. Current methods of post-translational modification analysis and their applications in blood cancers. *Cancers (Basel).* **13**, (2021).

6. Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat. Cell Biol.* **18**, 579–586 (2016).

7. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).

8. Thrower, J. S., Hoffman, L., Rechsteiner, M. & Pickart, C. M. Recognition of the polyubiquitin proteolytic signal. **19**, 94–102 (2000).

9. Huang, X. & Dixit, V. M. Drugging the undruggables: Exploring the ubiquitin system for drug development. *Cell Res.* **26**, 484–498 (2016).

10. Braten, O. *et al.* Numerous proteins with unique characteristics are degraded by the 26S proteasome following monoubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E4639–E4647 (2016).

11. Martínez-Jiménez, F., Muiños, F., López-Arribillaga, E., Lopez-Bigas, N. & Gonzalez-Perez, A. Systematic analysis of alterations in the ubiquitin proteolysis system reveals its contribution to driver mutations in cancer. *Nat. Cancer* **1**, 122–135 (2020).

REFERENCES

12. Bassermann, F., Eichner, R. & Pagano, M. The ubiquitin proteasome system - Implications for cell cycle control and the targeted treatment of cancer. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 150–162 (2014).

13. Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B. & Dosztányi, Z. Degrons in cancer. *Sci. Signal.* **10**, (2017).

14. Yeh, C. H., Bellon, M. & Nicot, C. FBXW7: A critical tumor suppressor of human cancers. *Mol. Cancer* **17**, 1–19 (2018).

15. Trevino, V. HotSpotAnnotations - A database for hotspot mutations and annotations in cancer. *Database* **2020**, 1–8 (2020).

16. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015).

17. Hendriks, I. A., Akimov, V., Blagoev, B. & Nielsen, M. L. MaxQuant.Live Enables Enhanced Selectivity and Identification of Peptides Modified by Endogenous SUMO and Ubiquitin. *J. Proteome Res.* **20**, 2042–2055 (2021).

18. Bard, J. A. M. & Martin, A. *Ubiquitin Proteasome System Methods and Protocols. Methods in Molecular Biology* vol. 1844 (Humana Press, New York, NY, 2018).

19. Hu, Z., Li, H., Wang, X., Ullah, K. & Xu, G. Proteomic approaches for the profiling of ubiquitylation events and their applications in drug discovery. *J. Proteomics* **231**, 103996 (2021).

20. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

21. Cell Signaling Technology, I. PTMScan® Ubiquitin Remnant Motif (K-ε-GG) Kit #5562. https://media.cellsignal.com/pdf/5562.pdf

22. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

23. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

24. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2020).

25. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).

26. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551 (2021).

27. Chen, D. *et al.* A Multidimensional Characterization of E3 Ubiquitin Ligase and Substrate Interaction Network. *iScience* **16**, 177–191 (2019).

28. Du, Z. & Lovly, C. M. Mechanisms of receptor tyrosine kinase activation in cancer. *Mol. Cancer* **17**, 1–13 (2018).

29. Sabbah, D. A., Hajjo, R. & Sweidan, K. Review on Epidermal Growth Factor Receptor (EGFR) Structure, Signaling Pathways, Interactions, and Recent Updates of EGFR Inhibitors. *Curr. Top. Med. Chem.* **20**, 815–834 (2020).

30. Moore, P. A. *et al.* Development of MGD007, a gpA33 x CD3-bispecific DART protein for T-cell immunotherapy of metastatic colorectal cancer. *Mol. Cancer Ther.* **17**, 1761–1772 (2018).

31. Lopes, N. *et al.* A panel of intestinal differentiation markers (CDX2, GPA33, and LI-cadherin) identifies gastric cancer patients with favourable prognosis. *Gastric Cancer* **23**, 811–823 (2020).

32. Wu, X. & Zhao, J. Novel oxidative stress-related prognostic biomarkers for melanoma associated with tumor metastasis. *Medicine (Baltimore).* **100**, e24866 (2021).

33. Chang, Y. S., Huang, H. Da, Yeh, K. T. & Chang, J. G. Identification of novel mutations in endometrial cancer patients by whole-exome sequencing. *Int. J. Oncol.* **50**, 1778–1784 (2017).

34. Nohata, N. *et al.* Tumour suppressive microRNA-874 regulates novel cancer networks in maxillary sinus squamous cell carcinoma. *Br. J. Cancer* **105**, 833–841 (2011).

35. Hiyoshi, Y. *et al.* Prognostic significance of IMMT expression in surgically-resected lung adenocarcinoma. *Thorac. Cancer* **10**, 2142–2151 (2019).

36. Sotgia, F. & Lisanti, M. P. Mitochondrial biomarkers predict tumor progression and poor overall survival in gastric cancers: Companion diagnostics for personalized medicine. *Oncotarget* **8**, 67117–67128 (2017).

37. Azadi, A. *et al.* Recent Advances on Immune Targeted Therapy of Colorectal Cancer Using bi-Specific Antibodies and Therapeutic Vaccines. *Biol. Proced. Online* **23**, 1–13 (2021).

38. Cichanover, A. Intracellular protein degradation: From a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Cell Death Differ.* **12**, 1178–1190 (2005).

39. Lecker, S. H., Goldberg, A. L. & Mitch, W. E. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J. Am. Soc. Nephrol.* **17**, 1807–1819 (2006).

40. Vu, P. K. & Sakamoto, K. M. Ubiquitin-mediated proteolysis and human disease. *Mol. Genet. Metab.* **71**, 261–266 (2000).

41. Li, W. *et al.* Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 that Regulates the Organelle's Dynamics and file:///C:/Users/Nadezhda/Downloads/18213395.nbibSignaling. *PLoS One* **1**, 62–64 (2016).

42. George, A. J., Hoffiz, Y. C., Charles, A. J., Zhu, Y. & Mabb, A. M. A comprehensive atlas of E3 ubiquitin ligase mutations in neurological disorders. *Front. Genet.* **9**, 1–17 (2018).

43. Clague, M. J., Heride, C. & Urbe, S. The demographics of the ubiquitin system. **25**, (2015).

REFERENCES

44. Salami, J. & Crews, C. M. Waste disposal - An attractive strategy for cancer therapy. *Science (80-.).* **355**, 1163–1167 (2017).

45. Downward, J. Targeting RAS signalling pathways in cancer therapy. *Nat. Rev. Cancer* **3**, 11–22 (2003).

46. Sebolt-Leopold, J. S. Advances in the development of cancer therapeutics directed against the RAS-mitogen-activated protein kinase pathway. *Clin. Cancer Res.* **14**, 3651–3656 (2008).

47. Sylvestersen, K. B., Young, C. & Nielsen, M. L. Advances in characterizing ubiquitylation sites by mass spectrometry. *Curr. Opin. Chem. Biol.* **17**, 49–58 (2013).

48. Hansen, F. M. *et al.* Data-independent acquisition method for ubiquitinome analysis reveals regulation of circadian biology. *Nat. Commun.* **12**, (2021).

49. Bekker-Jensen, D. B. *et al.* Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nat. Commun.* **11**, 1–12 (2020).

50. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009).

51. Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V & Mann, M. Andromeda : A Peptide Search Engine Integrated into the MaxQuant Environment. 1794–1805 (2011).

52. Benjamini, Y. & Hochberg, Y. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* **25**, 60–83 (2000).

53. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 (1995), Publi. *J. R. Stat. Soc.* **57**, 289–300 (1995).

REFERENCES

54. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

55. Tyanova, S. & Cox, J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. in *Cancer Systems Biology: Methods and Protocols* (ed. von Stechow, L.) vol. 1711 133–148 (Springer New York, 2018).

56. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).

57. Sinitcyn, P., Rudolph, J. D. & Cox, J. Computational Methods for Understanding Mass Spectrometry–Based Shotgun Proteomics Data. *Annu. Rev. Biomed. Data Sci.* **1**, 207–234 (2018).

58. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).