

Mechanisms and cellular responses to the loss of the intron-binding complex

Dissertation

zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Francesca Conte

geb. am 25.12.1993 in Mailand, Italien

Mainz, April 2024

Dekan:

1. Berichterstatterin:

2. Berichterstatter:

Tag der mündlichen Prüfung:

I. Preface

II. Summary

The intracellular environment is populated by hundreds of protein complexes, which are responsible for carrying out all cellular processes. Functionality of protein complexes is achieved through the tight, yet dynamic, regulation of each step of protein complex biogenesis, from production of stoichiometric amounts of subunits to the degradation of faulty assembly intermediates and unassembled subunits through the activity of protein quality control pathways.

In this thesis, we focused on the intron-binding complex, a known building block in spliceosome biogenesis. Besides its described role in splicing, the intron-binding complex is poorly characterized and the mechanisms behind its biogenesis and regulation are currently unknown. We observed that protein levels of intron-binding complex subunits display co-expression patterns, and that loss of one subunit is sufficient to trigger the collateral loss of the others. Using a combination of proteomic approaches, we investigated the protein quality control mechanisms responsible for the targeted degradation of unassembled intron-binding complex subunits, and identified E3 ubiquitin ligases UBR5 and TRIP12 as potential regulators.

In addition to regulating splicing as part of the intron-binding complex (IBC), individual intron-binding complex subunits are involved in a variety of biological processes. Informed by the knowledge that the levels of the subunits are interconnected as the result of the quality control mechanisms targeting the intron-binding complex, we set out to compare the cellular responses to the loss of three intron-binding complex subunits. By investigating similarities in the global proteomic alterations upon depletion of each subunit, we were able to define a core cellular response to the loss of the intron-binding complex. Features of this core response are irreparable DNA damage, p53 activation and cell cycle arrest. By complementing these phenotypes with specific signatures in our proteomic datasets, we concluded that loss of the intron-binding complex leads to cellular senescence.

To gain insight into the senescence phenotype, we looked into the mechanisms responsible for the observed genomic instability. Investigation of the interplay between double-strand break formation and nascent transcription revealed that loss of intron-binding complex subunit AQR causes transcription-associated genomic instability. In addition, general loss of intron-binding complex subunits causes increased levels of three-stranded nucleic acid structures known as R-loops: as formation of R-loops is a direct function of transcription, and their accumulation is a known source of genomic instability, we propose that R-loop-dependent DNA damage is another feature of the response to intron-binding complex loss.

Ultimately, this work underlies the importance of taking into account aspects of protein complex regulation when interpreting phenotypes arising from the loss of individual subunits.

III. Zusammenfassung

Die intrazelluläre Umgebung wird von Hunderten von Proteinkomplexen bevölkert, die für die Durchführung jeglicher zellulären Prozesse verantwortlich sind. Um die korrekte Funktionalität eines Komplexes zu gewährleisten, bedarf es einer strikten aber dynamischen Regulation der Proteinkomplexbiogenese. Diese Regulation umfasst unter anderem die Produktion stöchiometrischer Mengen von Untereinheiten und die Beseitigung fehlerhafter Zwischenprodukte oder überschüssiger Untereinheiten durch Protein-Qualitätskontrollmechanismen. In dieser Arbeit konzentrierten wir uns auf den intron-binding complex (IBC), einen bekannten Baustein der Spleißosomen-Biogenese. Abgesehen von seiner beschriebenen Rolle beim Spleißen ist der intron-binding complex nur unzureichend charakterisiert und die Mechanismen hinter seiner Biogenese und Regulierung sind derzeit unbekannt. Wir konnten beobachten, dass die Proteinkonzentrationen der Untereinheiten des intron-binding complex Koexpressionsmuster aufweisen und dass der Verlust einer Untereinheit ausreicht, um den kollateralen Verlust der anderen auszulösen. Mit einer Kombination aus proteomischen Ansätzen untersuchten wir die Mechanismen der Protein-Qualitätskontrolle, die für den gezielten Abbau überschüssiger Untereinheiten des intron-binding complex verantwortlich sind und identifizierten die E3-Ubiquitin-Ligasen UBR5 und TRIP12 als potenzielle Regulatoren.

Die einzelnen Untereinheiten des intron-binding complex haben auch abseits des Spleißens wichtige Funktionen für eine Vielzahl biologischer Prozesse. Ausgehend von dem Wissen, dass Qualitätskontrollmechanismen stochiometrische Verhältnisse der Untereinheiten gewährleisten, haben wir uns vorgenommen, die zellulären Reaktionen auf den Verlust drei verschiedener IBC-Untereinheiten zu vergleichen. Durch die Untersuchung der proteomischen Veränderungen nach Verlust der einzelnen Untereinheiten, waren wir in der Lage, eine allgemeine zelluläre Reaktion auf den Verlust des intron-binding complex zu definieren. Die zentralen Merkmale dieser Reaktion sind irreparable DNA-Schäden, p53-Aktivierung und Zellzyklus-Arrest. Indem wir diese Phänotypen mit spezifischen Signaturen aus unseren proteomischen Datensätzen ergänzten, kamen wir zu dem Schluss, dass der Verlust des intron-binding complex zu zellulärer Seneszenz führt.

Um einen Einblick in den Seneszenz-Phänotyp zu erhalten, untersuchten wir die Mechanismen, die für die beobachtete genomische Instabilität verantwortlich sind. Die Untersuchung des Zusammenhangs zwischen der Bildung von Doppelstrangbrüchen und aktiver Transkription ergab, dass der Verlust der IBC-Untereinheit AQR eine mit der Transkription verbundene genomische Instabilität verursacht. Des Weiteren konnten wir auf Verlust von allen drei untersuchten IBC-Untereinheiten hin eine Anhäufung von dreisträngigen Nukleinsäurestrukturen, sogenannten R-loops, beobachten. Da die Bildung von R-Loops eine direkte Funktion der Transkription ist und ihre Akkumulation eine bekannte Quelle genomischer Instabilität darstellt, nehmen wir an, dass die beobachtete genomische Instabilität auf durch R-loops verursachte Doppelstrangbrüche zurückzuführen ist.

Letztlich unterstreicht diese Arbeit, wie wichtig es ist, bei der Interpretation von Phänotypen, die sich aus dem Verlust einzelner Untereinheiten ergeben, dessen Funktion und Regulierung als Teil eines Komplexes zu berücksichtigen.

Table of Contents

Preface	III
Summary	V
Zusammenfassung	VI
Table of contents	IX
1. Introduction	13
1.1. Regulation of protein complexes	13
1.1.1. Assembly – a multi-step strategy to protein complex building	13
1.1.2. Stoichiometry – regulating the levels of protein complex components	15
1.1.3. “Orphan” subunits – removal of unassembled subunits by the ubiquitin-proteasome system	17
1.2. The intron-binding complex (IBC)	20
1.2.1. Aquarius intron-binding spliceosomal factor (AQR)	21
1.2.2. XPA binding protein A (XAB2)	23
1.2.3. ISY1 splicing factor homolog (ISY1)	24
1.2.4. Peptidylprolyl isomerase E (PPIE)	24
1.2.5. Zinc finger protein 830 (ZNF30)	25
1.3. Aim of the study	25
2. Results	27
2.1. Regulation of the intron-binding complex	27
2.1.1. Steady-state levels of IBC subunits follow co-expression patterns	27
2.1.2. Collateral loss of IBC subunits is observed upon IBC subunit depletion	29
2.1.3. The collateral loss of IBC subunits is resistant to inhibition of protein degradation pathways	30
2.2. Proteomic strategies for the identification of intron-binding complex regulators	32
2.2.1. Generation of a system for combined interactomics and proximity proteomics of endogenous AQR	32
2.2.2. Endogenous proximal proteome identifies known and novel potential interactors of AQR	34
2.2.3. E3 ubiquitin ligases are direct interactors of AQR	36
2.2.4. Integration of AQR interactome and proximal proteome nominates potential IBC regulators	38
2.2.5. TRIP12 interacts with IBC subunits	40
2.2.6. UBR5 knock-out opposes loss of IBC subunits	41
2.3. Cellular responses to the loss of intron-binding complex subunits	43

2.3.1. Proteomic characterization of IBC subunits-depleted cells	43
2.3.2. Integration of proteome and transcriptome of AQR-depleted cells	46
2.3.3. Nascent RNA sequencing sheds light on AQR-depleted cells transcriptional dynamics	48
2.3.4. Loss of IBC subunits alters cell cycle progression	50
2.3.5. Loss of IBC subunits causes p53 activation and senescence	52
2.3.6. Loss of IBC subunits affects factors involved in ribosome biogenesis	54
2.3.7. Loss of IBC subunits causes irreparable genomic instability	56
2.3.8. IBC subunits contribute to R-loops homeostasis	59
3. Discussion	63
3.1. Regulation of the stability of intron-binding complex subunits	63
3.1.1. Expression and stability of IBC subunits	63
3.1.2. The global interaction network of AQR	65
3.1.3. Looking for IBC regulators	66
3.1.4. The role of TRIP12 and UBR5 in the regulation of IBC subunits	67
3.2. Cellular responses to the loss of intron-binding complex subunits	69
3.2.1. “Omics” approaches to characterize the response to loss of IBC subunits	69
3.2.2. DNA damage, p53 activation and cell cycle arrest: the core response to the loss of IBC subunits	70
3.2.3. Senescence is the ultimate consequence of the of loss of IBC subunits	71
3.2.4. Interplay of transcription and DNA damage in AQR-depleted cells	72
3.2.5. The IBC in R-loop regulation	74
3.3. Concluding remarks	75
4. Material and Methods	77
4.1. List of all consumables, software and machines	77
4.2. Cell culture	84
4.2.1. Cell cultivation and passaging	84
4.2.2. siRNA transfection and treatments	84
4.3. Cell-based methods	85
4.3.1. Generation of whole cell lysates	85
4.3.2. SDS-PAGE and Western Blotting	85
4.3.3. Cell lysis for diGly-remnant enrichment experiments	85
4.3.4. Proximity-dependent biotinylation for mass spectrometry (PDB-MS)	85
4.3.5. Affinity purification for mass spectrometry (AP-MS)	86
4.3.6. Immunofluorescence and high-throughput confocal microscopy	86
4.4. Mass spectrometry-based proteomics	87

4.4.1. In-gel digestion	87
4.4.2. In-solution digestion	87
4.4.3. DiGly-remnant enrichment	87
4.4.4. Micro-tip based strong cation exchange chromatography (Micro-SCX)	88
4.4.5. Desalting and concentration of peptides	88
4.4.6. MS analysis	88
4.4.7. MS peptide quantification	88
4.4.8. Data processing and visualization	89
4.5. Genomics methods	89
4.5.1. mRNA-sequencing	89
4.5.2. TTChem-sequencing	90
4.5.3. Data analysis for mRNA-sequencing and TTChem-sequencing	90
4.5.4. sBLISS and data analysis	90
5. Abbreviations	93
6. Appendix	101
7. References	105
8. Acknowledgments	131
9. Curriculum vitae	133

1. Introduction

1.1 Regulation of protein complexes

The majority of proteins in the cell carry out their function as part of multiprotein complexes. In its easiest iteration, a protein complex can be homomeric, or formed by multiple copies of the same subunit. Around 45% of eukaryotic proteins can form homomeric complexes, indicating a strong propensity for proteins to self-assemble [1]. On the other hand, stable protein-protein interactions between polypeptide chains of different origin occurs in heteromeric complexes. The number of known existing protein complexes has steadily increased throughout the years, thanks both to intensive experimental efforts and to the systematization of prior and novel knowledge in open-access databases, such as CORUM 3.0, BioPlex, and the human reference interactome, HuRI [2], [3], [4], [5], [6], [7]. The general consensus from these studies is that, in eukaryotes, most protein complexes tend to be heteromeric.

The existence of heteromeric complexes poses some organizational challenges to the cell. For a complex to be functional, it requires timely assembly (and, later, disassembly). The identity and proportion of every subunit has to be controlled, in order to achieve correct stoichiometry. In addition, unassembled subunits need to be rapidly removed, to avoid deleterious consequences deriving from wrong interactions, misfolding or stoichiometric imbalances.

1.1.1 Assembly - a multi-step strategy to protein complex building

If single subunits of a protein complex had to find and interact with each other in the crowded intracellular environment, the chances of a protein complex to assemble in a biologically relevant timeframe would be extremely low. Indeed, numerous studies have shown that the assembly of both homomeric and heteromeric complexes occurs in an ordered and often evolutionary conserved fashion, which leads to the formation of different, energetically favorable assembly intermediates [1], [8], [9], [10]. The same logic applies to the disassembly of protein complexes. Multistep assembly was described for complexes ranging from homomers to large molecular machines, such as the spliceosome, the nuclear pore complex, and the 26S proteasome [11], [12], [13], [14]. Overall, the presence of parallel assembly pathways allows for efficient formation of multisubunit complexes.

The path to complex assembly begins at the ribosome, where nascent polypeptides are synthesized. Stability of the nascent polypeptide chain is dependent on the acquisition of a three-dimensional native conformation, which typically reduces the propensity of the protein to misfold or aggregate. Studies dating as far back as the 1960s have described protein folding as a co-translational process, which, depending on the physico-chemical properties of the amino acids forming the polypeptide chain, can start already in the ribosome tunnel [15], [16], [17]. Co-translational folding can be facilitated by ribosome-associated chaperone systems, such as the nascent polypeptide-associated complex (NAC), the ribosome-associated complex (RAC), prefoldin (PFD) and the TCP1 ring complex (TRiC) [18], [19], [20]. The concerted action of different families of chaperones prevents the formation of non-specific interactions during translation and maturation.

As mentioned before, assembly of multisubunit complexes requires a high level of coordination. To

an extent, this is achieved through the coupling of assembly and translation, in a process known as co-translational assembly [21], [22], [23]. In co-translational assembly, interactions between subunits of a protein complex can occur between nascent polypeptide chains emerging from the same mRNA (*in-cis*) or from different mRNAs (*in-trans*) (Figure 1A) [24]. *In-cis* co-translational assembly is typical of homomeric complexes, such as transcription factors working as homodimers [25]: one example is tumor suppressor p53, which first dimerizes co-translationally and subsequently assembles post-translationally in a homotetramer [26]. Heteromeric complexes can instead rely on *in-trans* co-translational assembly; in this case, the most common scenario involves a fully translated subunit interacting with the nascent polypeptide of a different subunit. Recent studies revealed an even more complex picture: for example, it was shown that multiple mature subunits are able to engage with the same emerging polypeptide at the same time [27]. Physical proximity of newly translated and nascent subunits from the same complex can be facilitated by compartmentalization of their respective mRNAs prior to translation, observed for example in the case mRNAs encoding for glycolytic enzymes, subunits of ion channels or proteasomal subunits [28], [29], [30]. In addition, other features of the encoding mRNA, such as length, codon usage and secondary structure, generally affects translation and, as a result, the rate of co-translational assembly of protein complexes [21]. Overall, *in-trans* mechanisms were shown to regulate the assembly of a number of multiprotein complexes, ranging from nuclear transcription complexes TFIID, SAGA and TREX-2, to the nuclear pore [31], [32], [33], [34], [35].

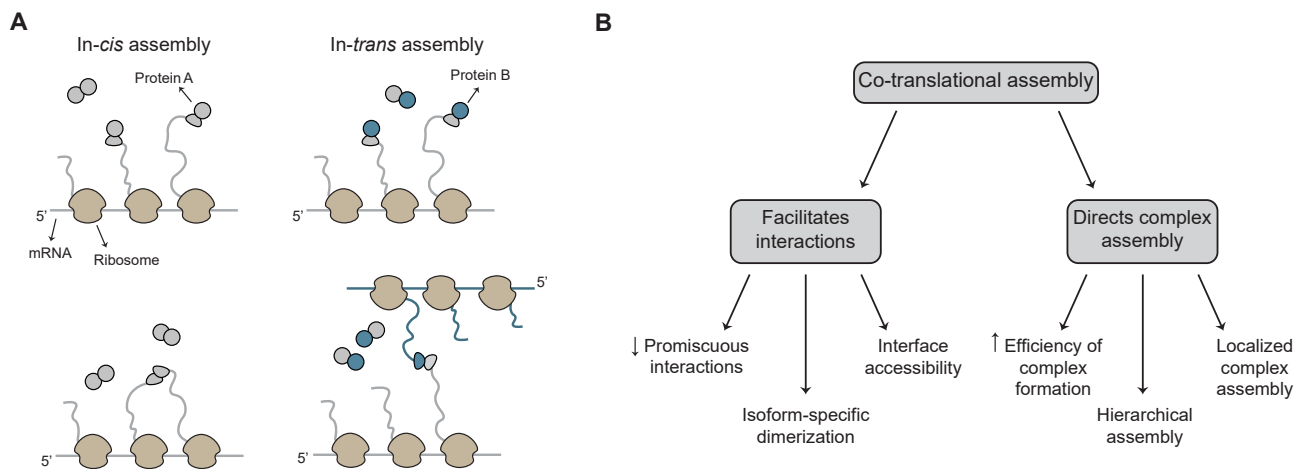


Figure 1. Modes and benefits of co-translational assembly

A) Representation of *in-cis* and *in-trans* mechanisms of co-translational assembly. *In-cis* assembly involves homomers, whilst *in-trans* assembly can occur between homomers or heteromers (adapted from Natan et al., 2017). B) Schematic of benefits of co-translational assembly (adapted from Khan & Fox, 2023).

The advantages of co-translational assembly are manifold (Figure 1B). First, engaging a nascent polypeptide in interactions with the correct binding partner reduces the likelihood of non-specific interactions, thus increasing the probability of producing a functional complex. Interestingly, N-terminal interfaces of protein complex subunits show a significant tendency to be larger than C-terminal interfaces and often harbor protein-protein interaction domains, which could favor the immediate binding of the nascent protein to the mature subunit [32], [36]. In addition to avoiding unwanted interactions,

immediate assembly can stabilize nascent proteins; in this scenario, the mature subunit would play a chaperon-like role, thus protecting nascent chains from premature misfolding and aggregation [32], [33], [37]. From the structural point of view, the unfolded state of a nascent polypeptide (or partially folded, if we consider the possibility of co-translational folding) can promote interactions between surfaces that would become sterically inaccessible upon folding, as seen in the case of ezrin and ezrin binding protein 50 (EBP50) [38].

As previously mentioned, complex assembly usually occurs through the formation of intermediate sub-complexes. In line with this, co-translational mechanisms can contribute to this multi-step process by coordinating the formation of such sub-complexes, which later interact post-translationally to form the mature complex. Among post-translational mechanisms of protein complex assembly, it is important to mention the role of chaperones. In addition to promoting folding, chaperones can protect subunits from degradation, often binding client proteins until the interaction is displaced by the correct interacting subunit. Chaperone-associated assembly is particularly important for bigger assemblies, such as the proteasome [39]. Therefore, the combination of co-translational and post-translational interactions orchestrates the assembly step, allowing the production of functional protein complexes.

1.1.2 Stoichiometry - regulating the levels of protein complex components

Assembly and function of heteromeric complexes relies on the presence of correct amounts of the individual monomers, which ideally are proportional to their stoichiometric coefficients. In bacteria, subunits of the same complex are often encoded by genes in the same operon, allowing for coordinated expression [40]. As operons are absent in eukaryotes, stoichiometric abundance of protein complex subunits is achieved through regulatory mechanisms acting at the level of gene expression, mRNA stability, protein synthesis and protein degradation (Figure 2A).

The general consensus is that, for subunits of protein complexes, co-expression is mainly achieved through regulation of post-transcriptional steps, rather than transcription itself. However, it is worth mentioning that genes encoding for subunits of stable protein complexes are found closer to each other on the genome than expected by chance, which could be an indication of transcriptional co-regulation [41]. After transcription, determinants of mRNA stability and localization could play a role in coordinating final expression levels. For example, the 5' end of mRNAs encoding for ribosomal proteins contain terminal oligo-pyrimidine tracts (5'TOP), which allow their translation to be repressed and de-repressed in concert [42]. Importantly, regulation at the mRNA level is tightly connected with regulation of the following step of translation.

Systematic studies employing ribosome profiling to investigate rates of protein synthesis in bacteria and yeast revealed that the synthesis of most protein complex subunits is proportional to their stoichiometry [43], [44]. Similar observations have been made in mammalian cells for complexes such as the ribosome and the proteasome [45], [46]. Importantly, it is likely that proportional synthesis applies mainly to obligate subunits, which are only stable when part of a complex. On the other hand, accessory and transiently associating subunits, or proteins with moonlighting functions that can be found in multiple complexes, are not necessarily expected to be stoichiometrically abundant. The limiting factor in the rate of protein synthesis is mRNA abundance, which is in turn affected by gene copy-number. However, high mRNA abundance does not always equate to an increase in translation or

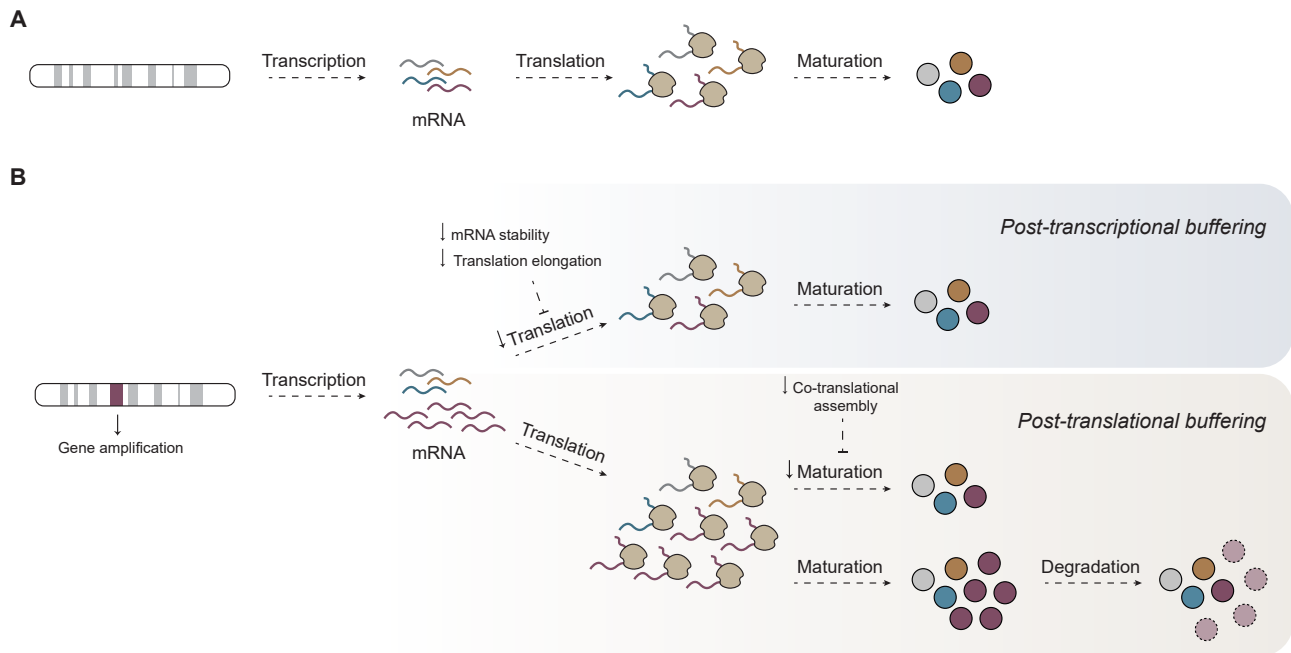


Figure 2. Regulation of protein complex assembly

Schematics displaying the multi-level organization behind the production of functional protein complexes. A) Coherent expression of subunits of the same complex is achieved through the regulation of transcription, translation and protein maturation. B) If this system is disrupted, for example by gene amplification events that cause overproduction of mRNA encoding for one subunit, different buffering mechanisms can intervene: post-transcriptional buffering inhibits the translation step, through increased mRNA decay or reduced elongation rate. On the other hand, post-translational buffering affects the stability of the maturing proteins, which are targeted to degradation as a result of reduced co- or post-translational assembly.

to elevated protein levels. For most genes, protein levels only modestly correlate with mRNA levels, and this is all the more evident in the case of protein complex subunits [47].

The maintenance of coherent protein expression levels between subunits of the same complex is observed even in extreme scenarios, such as the gain or loss of chromosomes in aneuploid cells: this highlights that a fraction of the proteome is very robust against dramatic alterations in gene dosage and mRNA abundance [44], [48], [49], [50], [51], [52], [53], [54], [55]. Such robustness is achieved through post-transcriptional and post-translational buffering mechanisms (Figure 2B). Post-transcriptional buffering acts at the level of translation. Interestingly, these mechanisms were found to act by modulating elongation rates, rather than constituting a negative feedback loop on translation initiation [44]. Codon usage, tRNA abundance and RNA modifications are some of the factors involved in this regulation [56][57]. Considering that protein complexes can assemble co-translationally, it is safe to speculate that post-transcriptional buffering could affect the levels of a subunit by interfering with its integration in a complex.

Despite some known examples of post-transcriptional buffering, evidence suggests that the bulk of the buffering occurs post-translationally through protein degradation pathways [57], [58]. For the majority of proteins, degradation at the steady-state is an exponential decay process [59]. However, it was observed that a fraction of newly synthesized proteins undergoes rapid degradation by the proteasome

immediately after synthesis [60]. Indeed, studies have shown that the degradation rate for a subset of proteins follows a non-exponential kinetic: in this case, proteins are either immediately degraded after synthesis, or are stabilized (and are later exponentially degraded) [61]. Non-exponentially degraded (NED) proteins included components of heteromeric protein complexes such as the proteasome or the ribosome; additionally, many NED proteins, such as ribosomal proteins, were found to be produced in super-stoichiometric quantities. This is in line with previous reports, and again substantiates the notion that complex formation stabilizes its subunits [62]. The mechanism by which a protein complex is stable, while its unassembled subunits are destabilized, is also known as cooperative stability. Regulation by cooperative stability has some clear benefits, such as allowing for the maintenance of high complex/monomer ratio, which is advantageous in the context of complex stoichiometry regulation [63]. Notably, the success of such a regulatory strategy relies on the existence of mechanisms that can sense and timely remove unassembled subunits and misassembled complexes.

1.1.3 “Orphan” subunits - removal of unassembled subunits by the ubiquitin-proteasome system

As previously mentioned, selective removal of unassembled subunits, also known as “orphan” subunits, is a key aspect of protein complex regulation. The bulk of this removal is carried out by the ubiquitin-proteasome system (UPS). Substrate degradation by the UPS is mediated by the concerted action of different enzymes, which attach and remove ubiquitin molecules from target proteins (Figure 3A). In the first step of this cascade, ubiquitin is activated through hydrolysis of adenosine triphosphate (ATP) by an E1 ubiquitin-activating enzyme; this is followed by the covalent attachment of the activated ubiquitin monomer to a catalytically active cysteine residue of an E2 ubiquitin conjugating enzyme. In the last step, the ubiquitin-bound E2 enzyme interacts with an E3 ubiquitin ligase, and ubiquitin is transferred to the acceptor lysine residue of a substrate protein with formation of an iso-peptide bond between its carboxyl

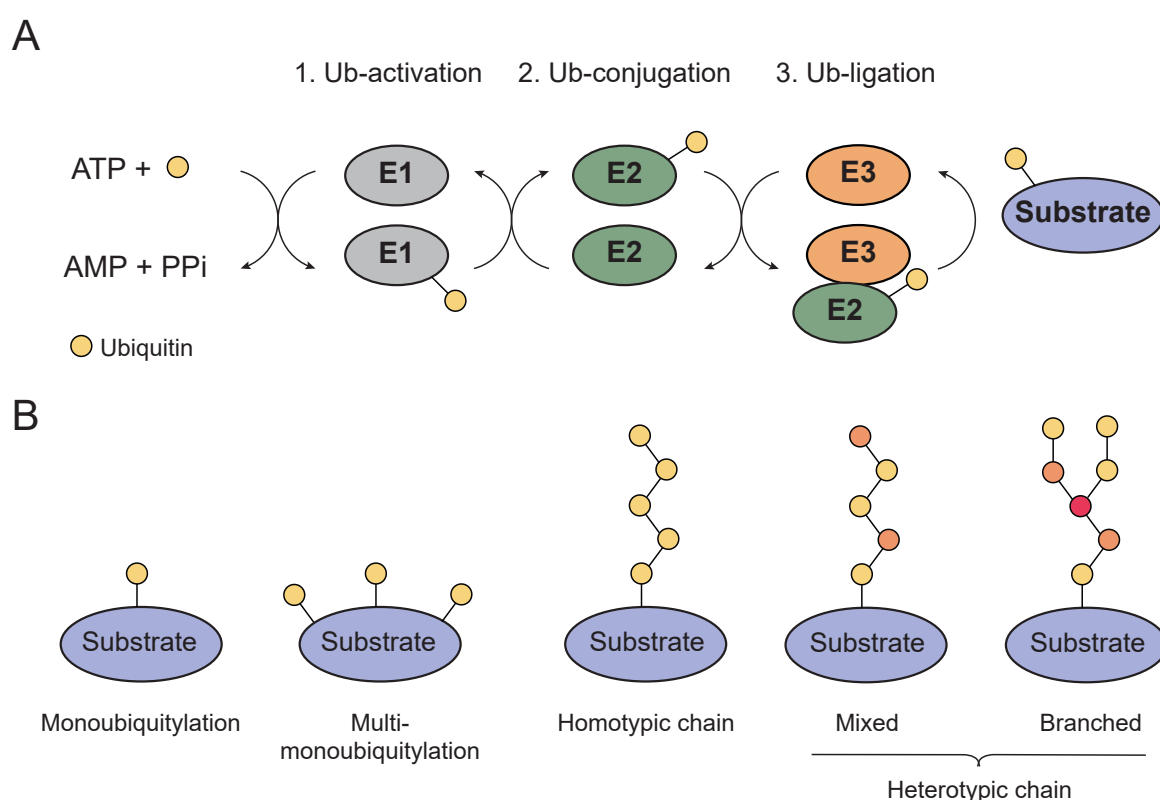


Figure 3. The ubiquitin cascade

A) Schematic representation of the ubiquitin cascade. Ubiquitin is first activated by a ubiquitin-activating enzyme (E1) through ATP hydrolysis, then covalently conjugated onto ubiquitin-conjugating enzyme (E2), and lastly attached to the substrate with the help of an E3 ligase. The removal of ubiquitin is mediated through deubiquitination enzymes (DUBs; not shown). B) Schematic representation of different ubiquitin modifications. Substrates can be modified with a single ubiquitin molecule on one or more acceptor sites, referred to as monoubiquitylation or multi-monoubiquitylation. Alternatively, ubiquitin monomers can be joined to each other to form different types of chains. In homotypic chains the same acceptor site of ubiquitin is used throughout the chain (e.g., K48-linked chains). Heterotypic chains contain multiple types of linkages (in yellow and orange) and are classified as either mixed or branched. In mixed chains, each ubiquitin monomer is modified once; on the other hand, branched chains contain at least one ubiquitin monomer (in red) linked to multiple ubiquitin molecules (adapted from French et al., 2021)

terminus and the epsilon-amino group (ϵ -NH₂) of lysine or, in case of linear ubiquitylation, a peptide bond with the C-terminal methionine [64]. E3 ubiquitin ligases can be divided into three families, the RING (really interesting new gene) ligases, the HECT (homologous to the E6AP carboxyl terminus) and the RBR (RING-between-RING) ligases [65].

Depending on the type of E3 ubiquitin ligase involved, ubiquitin is first transferred to the E3 enzyme and then to the substrate; this is the case for HECT and the RBR ligases. As ubiquitin monomers can themselves be modified with ubiquitin, this allows the formation of polyubiquitin chains (Figure 3B) [66]. More specifically, ubiquitin contains seven lysine residues that can accept other ubiquitin molecules: K6, K11, K27, K29, K33, K48 and K63. The trimming of ubiquitin chains is mediated by deubiquitinating enzymes, or DUBs; this action modulates the stability of the substrate, as well as replenishing the pool of free ubiquitin molecules. The most extensively studied function of ubiquitin chains is to target substrates to degradation by the 26S proteasome; however, not all ubiquitin chains act as signals for proteolysis. A key factor in determining the fate of the substrate is linkage, that is which lysine residue on the ubiquitin monomer accepts the new ubiquitin molecule [67]. Ubiquitin chains can consist of one or multiple linkage types (homotypic and heterotypic chains); in addition, depending on whether ubiquitin is modified on one or multiple lysine residues, a chain can be mixed or branched. Usually, roughly half of the ubiquitin chains in unchallenged cells contain K48 linkages, alone or in combination with other and less abundant linkages [68]. K48-linked ubiquitin chains, as well as K11/K48, K29/K48 and K48/K63 heterotypic branched chains, are responsible for targeting substrates for degradation [68], [69], [70], [71]. The human genome encodes for more than 600 E3 ubiquitin ligases [72]; for many of these enzymes, substrate specificity can be determined by the presence of stretches of aminoacids in the sequence of the target protein, which therefore act as degradation signals. These sequences, known as degrons, are typically found at the N-terminus or C-terminus of a protein: for this reasons, proteolytic pathways involving the recognition of degron sequences are known as N-degron and C-degron pathways [73]. In a broader interpretation, any sequence motif that can destabilize a protein could be considered a degron, independently of its localization.

The process that takes care of degrading unassembled subunits is termed orphan quality control. Depending on the protein in question, degradation can occur through different mechanisms [74], [75], [76]. For some proteins, misfolding is the driving cause of degradation. As the interaction with partner subunits is often required for proper folding, absence of this interaction can lead to aberrant conformation and exposure of hydrophobic portions of the polypeptidic chain, which act as degradation signals [77]. A similar strategy is observed in the case of mislocalized proteins: for example, the hydrophobic transmembrane domains of membrane proteins are recognized in the cytosol by a surveillance complex

containing, among other proteins, chaperone BAG6, and targeted to degradation [78], [79]. These similarities reflect a more general overlap between quality control pathways for misfolded orphan proteins and mislocalized proteins.

Notably, failure in targeting a protein to the correct cellular compartment can be a source of orphan subunits in itself. Newly synthesized ribosomal subunits, for example, are trafficked from the cytosol to the nucleolus, where ribosomes assemble. As the binding to ribosomal RNA is a main stabilizing factor, the subunits that do not reach the nucleolus are extremely unstable, and rapidly degraded by E3 ubiquitin ligase Tom1/HUWE1 and/or E2/E3 ubiquitin ligase UBE2O [62], [80], [81], [82]. Interestingly, it was found that HUWE1 binds orphan ribosomal subunits at sites that would normally mediate interaction with ribosomal RNA: this provides a mechanism through which the E3 ubiquitin ligase would target for degradation only unassembled subunits, while sparing the ones that are part of the ribosome.

Fully folded, unassembled subunits can be destabilized and targeted to degradation through the exposure of degron signals that would otherwise be buried inside the complex (Figure 4A and B). A good example is the removal of N-terminally acetylated orphan subunits. Around 80% of eukaryotic proteins have their N-terminus co-translationally modified with an acetyl group [83], [84]; among other functions, N-terminal acetylation constitutes a degradation signal, which is recognized by a branch of the N-degron pathway known as the Ac/N-degron pathway [85]. As interactions in a protein complex often form between the acetylated N-termini of its subunits, the assembly itself leads to shielding the degron sequences, thus protecting the subunits from degradation [86]. However, this is not the case for unassembled subunits, which are therefore rapidly degraded. As a result, this type of regulation allows monitoring of the assembly state of a complex, as well as maintaining its stoichiometry.

In summary, the production and maintenance of a functional protein complex is a challenging task, requiring the concerted action of regulatory mechanisms and protein quality control pathways. The importance of this multi-level organization becomes all the more clear when some of these mechanisms go awry. Reduced translation fidelity and a decline in proteostasis are typically observed in aged organisms, leading to pervasive protein aggregation and often profound proteomic alterations [87], [88], [89], [90]. As a result, the homeostasis of protein complexes is also altered: this can cause changes in the composition and overall abundance of specific complexes, such as the proteasome or the nuclear pore complex, as well as the unchallenged production of superstoichiometric quantities of subunits [87], [91], [92], [93]. The general inability of the system to cope with overabundance and misfolding of subunits can in turn fuel the formation of toxic protein aggregates, which is a feature of numerous neurodegenerative and age-associated diseases [88], [94]. Regulation (and misregulation) of protein complexes homeostasis also has important implications in the context of cancer. Ryan and colleagues showed that numerous subunits of protein complexes are hemizygously deleted in breast cancer [95]; interestingly, proteins belonging to the same complexes as the mutated subunits also displayed reduced levels, although not accompanied by alterations in copy-number or mRNA level. This collateral loss phenomenon is the consequence of the activity of post-transcriptional or post-translational buffering mechanisms, which respond to the absence of one subunit by bringing down the overall levels of the complex. Overall, these findings underlie the importance of taking into account the behavior of protein complexes when interpreting phenotypes arising from the loss of individual subunits, in both physiological and pathological contexts.

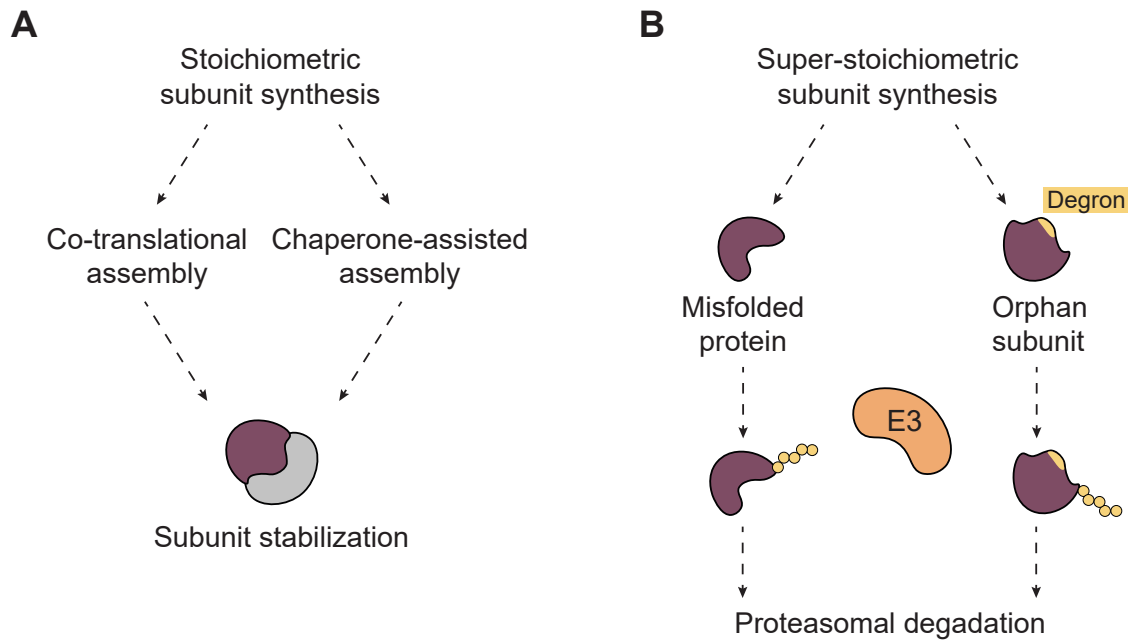


Figure 4. Elimination of unassembled protein complex subunits

Schematics showing the fate of assembled versus unassembled subunits of protein complexes. A) Stoichiometrically produced subunits engage in a complex by co-translational assembly or chaperone-assisted assembly, and are stabilized as a result. B) Excess subunits can undergo misfolding or expose degron sequences: in both cases, E3 ubiquitin ligases polyubiquitinate the unassembled subunits and target them to proteasomal degradation.

1.2 The intron-binding complex (IBC)

The intron-binding complex, IBC (for simplicity, this acronym will be used throughout this thesis), is a protein complex with a theoretical mass of 384 kDa composed by one single copy of each of five subunits: AQR (Aquarius, also known as intron-binding spliceosomal factor 160), XAB2 (XPA binding protein 2, also known as hSyf1), ISY1 (Isy1 splicing factor homolog, or hIsy1), PPIE (Peptidylprolyl isomerase E, also known as CypE or Cyp33) and ZNF830 (Zinc finger protein 830, also known as CCDC16). The complex was characterized in 2015, first by co-fractionation experiments of proteins co-immunoprecipitating with endogenous AQR, and secondly by purification of the full complex [96]. However, the idea of stable interactions being established between these proteins predates this study. Studies had already proven the existence of a multimeric complex, named XAB2 complex, involving the five subunits and the protein PRP19, a spliceosome-associated protein [97]. PRP19 itself and its associated complex, the PRP19-associated complex in humans or NineTeen Complex (NTC) in yeast, have been thoroughly investigated in relation to their fundamental role in spliceosome activation [98], [99]. IBC subunits Syf1 and Isy1 have been described as subunits of *S. cerevisiae* NTC complex [100], [101]; in addition, all IBC subunits associate with the human PRP19 complex, with the exception of ZNF830. The formal characterization of the IBC in 2015 came with important findings that shed light on the nature of the complex. A key finding is that AQR is efficiently integrated in the spliceosome only as part of the IBC: this suggests that the IBC formation is an event preceding the integration of the

single subunits in the spliceosome, making it a pre-formed spliceosomal building block. The molecular mechanisms behind IBC formation and the following integration in the pre-catalytic B spliceosome (B^{act}) are currently not known. Once delivered to the spliceosome, AQR plays a fundamental role in the remodeling events necessary to the catalytic activation of the spliceosome (or B^{act} to B^* transition). Notably, integration in the spliceosome of an IBC carrying an ATPase-dead AQR mutant (K829A) caused stalling of spliceosomes after B^{act} formation; on the other hand, integration of an helicase-dead AQR mutant (Y1196A) did not cause a similar phenotype, suggesting that only the ATPase activity is required for the catalytic activation (although it could be necessary in other steps of the splicing cycle) [96], [102]. Taken together, these results highlight a crucial role for AQR in splicing in metazoans, and more broadly for correct IBC assembly and timely integration in the spliceosome.

1.2.1 Aquarius intron-binding spliceosomal factor (AQR)

AQR (also known as intron-binding protein 160, IBP160) is a 160 kDa protein, the largest of the IBC subunits. It is conserved from yeast to human, with the exception of *S. cerevisiae* where no homolog was found. The crystal structure of AQR was resolved in 2015 and provided great insight into the protein's domain architecture and associated functions (Figure 5) [96]. The C-terminus of the protein harbours two RecA-like domains, named RecA1 and RecA2, responsible for the helicase activity of AQR; based on sequence motifs in these domains, AQR was classified as a member of the SF1 superfamily, and more specifically of the Upf1-like family [103]. This sets AQR apart from all other spliceosomal helicases, which all belong to the SF2 family. In vitro assays showed that these domains confer 3'-to-5' RNA-unwinding activity, which is dependent on ATP hydrolysis. The "SF1 core" extends outside the RecA-like domains to include more structural components, the "stalk" and a β -barrel domain, which mediate interactions with IBC subunits. The N-terminal part of the protein is occupied by a domain which is found uniquely in AQR among all helicases, the ARM domain (aminoacids 1-416). The key feature of this domain is a structural motif called Armadillo (ARM) repeat; these are highly conserved sequences of 50 aminoacids, phylogenetically related to HEAT repeats and known to mediate protein-protein interactions [104]. The prototype ARM repeat protein is the Armadillo protein of *D. melanogaster*, and its human homolog β -catenin. ARM repeats are usually present in multiple tandem copies: from the structural point of view, this allows neighboring ARM repeats, each one consisting of three α -helices, to organize into a single domain, which adopts a superhelical or solenoid-like conformation [105]. AQR contains only five ARM repeats, plus additional α -helices which cannot be formally classified as repeats.

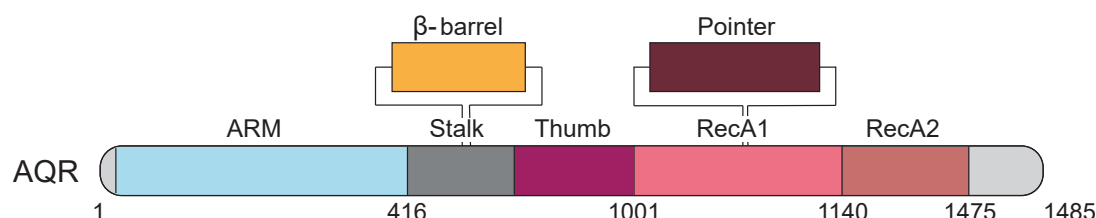


Figure 5. Schematic representation of human AQR

Schematics displaying AQR domain organization (N-terminus to C-terminus). Numbers indicate amino acid positions. The β -barrel domain extends from aminoacid 493 to 661, the stalk encompasses aminoacids 416–492 and 660–710, the thumb domain extends from aminoacid 681 to 804, the pointer domain extends from aminoacid 881 to 1034 (adapted from De et al., 2015).

Introduction

This domain mediates interactions both within AQR and with other binding partners, including IBC subunit ISY1 and spliceosomal core protein SF3B1.

Interestingly, studies showed that AQR binds to introns –33 to –40 nucleotides upstream of the intron branch site and in a sequence-independent manner [106]. The determinants targeting AQR to this position are not known, but it is likely that ARM domain-mediated interactions play a role. Finally, as previously mentioned, lack of the ARM domain completely inhibits IBC formation in vitro, further strengthening AQR role as a scaffolding protein.

The involvement of AQR in mRNA processing is not limited to splicing. Reports showed that AQR acts as a bridge between the spliceosome and the exon-junction complex (EJC), a multiprotein complex controlling the fate of spliced mRNAs through its regulation of mRNA export and localization, nonsense-mediated mRNA decay (NMD) and translation [107], [108]. Strikingly, AQR loss in HeLa cells was sufficient to impair EJC assembly on ligated exons, causing a downstream NMD arrest [109]. On a similar note, loss of AQR or of EJC factors were both sufficient to cause the leakage of unspliced transcripts from nucleus to cytoplasm in *C. elegans*, suggesting AQR is part of a surveillance mechanism orchestrating the export of properly spliced RNAs [110]. In addition to this global role as an EJC-associated factor, AQR is required for the maturation and assembly of specific type of small nucleolar ribonucleoproteins (snoRNP), namely box C/D snoRNPs [106]. As in the case of all RNPs, snoRNPs consist of an RNA component, a small nucleolar RNA (snoRNA), and associated proteins; as snoRNA genes for the most part lack independent promoters (at least in mammalian cells), they are encoded within the introns of host genes and are with them co-transcribed [111], [112]. AQR was found to assist with the folding of box C/D snoRNAs, an event required for the association with core snoRNP proteins. Once in complex with their protein partners, snoRNAs guide post-transcriptional modifications of precursor rRNAs, such as 2'-O-methylation in the case of boxC/D snoRNPs, which ultimately affect ribosome biogenesis and translation [113], [114]. Interestingly, a recent study nominated AQR as a factor involved in the synthesis of the 40s ribosomal subunit, suggesting a role in the early steps of ribosome biogenesis [115].

Last but not least, in recent years AQR emerged as a potential player in the regulation of R-loops, three-stranded nucleic acid structures that form co-transcriptionally [116], [117]. AQR depletion was seen to cause a spontaneous genomic instability in mammalian cells which could be partially rescued by overexpression of ribonuclease H1 (RNaseH1), a nuclease able to cleave the RNA:DNA moiety of R-loops [118]. This first observation was followed by the finding that AQR depletion leads to increased R-loop levels, which are cleared by factors of the transcription-coupled nucleotide excision repair (TC-NER) pathway causing R-loop-dependent genomic instability [119]. A second study observed a similar phenotype and speculated a role for AQR in mediating homology recombination at R-loop-dependent double-strand breaks (DSBs), albeit not deciphering the molecular mechanisms behind these observations [120]. AQR was also found to be proximal to R-loops in R-loop-proximal proteome previously published by our lab [121], again strengthening its implication in the homeostasis of these structures. This said, the answer to how AQR regulates R-loop levels remains elusive. The most obvious way AQR could oppose R-loops accumulation is through its helicase activity. In vitro assays showed that AQR is indeed able to unwind double-stranded RNA [96], but nor its binding or unwinding activity towards RNA:DNA hybrids or R-loop substrates has been tested. Notably, members of the same SF1 helicase family UPF1 and Senataxin (SETX) can unwind R-loops and RNA–DNA hybrids with a 5' to

3' polarity [122], [123], [124]. It is therefore reasonable to speculate that AQR could have a similar activity. However, it is equally possible that AQR impact on R-loop levels is not dependent on its ability to unwind the structures directly, but is rather the consequence of the misregulation of the pathways AQR is involved in. Indeed, AQR was the RNA binding protein (RBP) whose depletion caused the highest number of alternative splicing events (13,000 identified events) among 356 human RBPs analyzed in a recent study by Van Nostrand and colleagues [125], which might cause substantial alterations in transcription and R-loop formation.

1.2.2 XPA binding protein 2 (XAB2)

XAB2 is a nuclear protein of roughly 100 kDa conserved between higher and lower eukaryotes. The structure of XAB2 encompasses 14 HAT repeats (Half A TPR, or tetratricopeptide), degenerate repeats of 34 aminoacids known to mediate protein-protein interactions and often found in scaffolding proteins (Figure 6) [126], [127].

As hinted by its name, human XAB2 was initially characterized as an interactor of XPA, a factor central to both transcription-coupled and global genome nucleotide excision repair (TC-NER and GG-NER respectively). Interaction with TCR-specific factors CSA and CSB and with RNA polymerase II further implicated XAB2 in TC-NER, although later reports showed that these interactions are most likely transient and context-specific, and that a stable complex between these proteins is not formed [97],



Figure 6. Schematic representation of human XAB2

Schematics displaying XAB2 domain organization (N-terminus to C-terminus). Numbers indicate amino acid positions. HAT repeats are localized as follows: HAT 1 aa 15-47, HAT 2 aa 48-80, HAT 3 aa 90-122, HAT 4 aa 124-158, HAT 5 aa 160-192, HAT 6 aa 198-230, HAT 7 aa 235-268, HAT 8 aa 270-305, HAT 9 aa 369-407, HAT 10 aa 498-530, HAT 11 aa 532-566, HAT 12 aa 571-605, HAT 13 aa 643-677, HAT 14 aa 679-713.

[128]. XAB2 involvement in DNA repair pathways extends to the homologous recombination (HR) pathway, where XAB2 regulates the end resection step [129], [130]. XAB2 homologs in yeast, respectively Cwf3 and Syf1 in *S. pombe* and *S. cerevisiae*, are both spliceosome-associated protein and subunits of the NineTeen Complex [131], [132]. Human XAB2, however, was linked to pre-mRNA splicing only later, when studies showed the existence of a XAB2 complex involving PRP19, AQR, and other subunits of what later will be called intron-binding complex [133]. Additionally, proteomic characterization of purified human spliceosomes at different stages of the splicing cycle revealed XAB2 is present in the B [134], B^{act} [135], [136], [137], C [138] and C* complex [139]. As most splicing-associated proteins, XAB2 can regulate transcription; most notably, XAB2 loss was found to alter splicing of POLR2A pre-mRNA, leading to loss of the protein, global decrease in transcription and senescence in HeLa cells [140]. Last but not least, XAB2 recently emerged as factor favoring R-loop processing, in virtue of its ability to interact with TC-NER factors XPG and XPF and recruit them to RNA:DNA hybrids [141].

1.2.3 ISY1 splicing factor homolog (ISY1)

ISY1 is a small nuclear protein homolog of the non-essential Isy1p protein in yeast. With a molecular weight of 33kDa, it harbors no notable feature or annotated domain (Figure 8A). In *S. cerevisiae*, Isy1p was first identified as a component of the NineTeen Complex and studied for its role in controlling splicing fidelity [142], [143]. In more recent years, additional splicing-unrelated functions for ISY1 have emerged. ISY1 was found to assist CPSF3 (Cleavage and polyadenylation specificity factor subunit 3) in the maturation of a novel miRNA biogenesis intermediate named progenitor miRNA (pro-miRNA) in ESCs [144]. By mediating this post-transcriptional regulatory step for miRNA expression and the downstream changes in gene expression, ISY1 can program ESCs to a specific state of poised pluripotency, which is required for ESC transition from naïve to primed state [145]. These findings suggest an important role for ISY1 in cell differentiation, warranting further investigation.

A third role for ISY1 emerged in the field of DNA repair, as a regulator of the base excision repair pathway (BER). Here ISY1 was found to form a complex with the DNA-(apurinic or apyrimidinic site) endonuclease APEX1, promoting the binding to and 5'–3' endonucleolytic cleavage of abasic DNA lesions by APEX1, and therefore protecting cells from oxidative stress [146]. These observations open up opportunities to study the role of ISY1 in fields ranging from tumor biology, where high expression of APEX1 is a hallmark of neoplasia, to aging, which usually comes with a decrease in APEX1 tissue levels [147], [148].

1.2.4 Peptidylprolyl isomerase E (PPIE)

PPIE is a nuclear protein belonging to the family of cyclophilins, a group of ubiquitous proteins highly conserved throughout evolution and equipped with peptidyl-prolyl *cis-trans* isomerase (PPIase) activity. Briefly, peptidyl-prolyl isomerization is the process by which the *cis* peptide bond that can naturally occur between a proline residue and the preceding aminoacid (peptidyl-prolyl bonds) isomerizes in the lower energy *trans* state. Most peptide bonds exist in the *trans* state, which is the favored state for de novo folding and refolding of proteins; however, the spontaneous isomerization of peptidyl-prolyl bonds is a slow process, which can prove to be rate-limiting for protein folding [149]. By catalyzing proline isomerization, proteins with PPIase activity can therefore facilitate this process in vivo (Figure 7A) [149], [150], [151]. The structure of PPIE comprises a C-terminal Cyp domain, responsible for the PPIase activity, and an N-terminal RNA recognition motif (RRM) domain that mediates binding to AU-rich mRNAs (Figure 7B) [152], [153]. Interestingly, recent studies have demonstrated that the binding interface with RNA extends outside the RRM domain and into the Cyp domain, thus causing RNA binding to regulate the isomerase activity of the protein, as well as the interaction with other proteins [154]. Most studies on PPIE revolve around its physical interaction with the mixed lineage

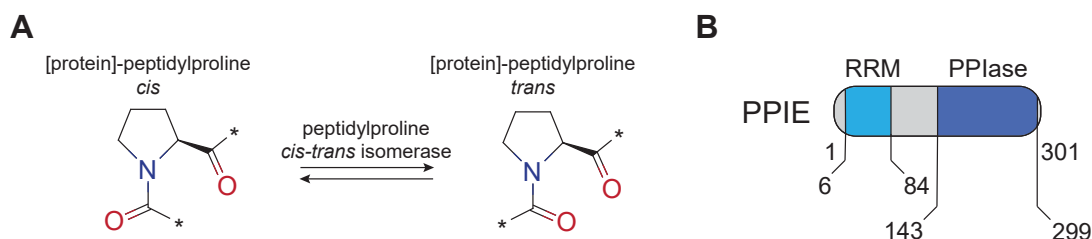


Figure 7. Schematic representation of human PPIE

A) Peptidyl-prolyl *cis-trans* isomerization reaction catalyzed by PPIE (adapted from RHEA:16237). B) Schematics displaying PPIE domain organization (N-terminus to C-terminus). Numbers indicate amino acid positions.

leukemia protein 1 (MLL1, also known as KMT2A), a histone methyltransferase and transcriptional regulator that is actively repressed by PPIE binding and isomerase activity, and whose misregulation drives leukemogenesis [155]. Despite the important efforts in dissecting the molecular details of PPIE-mediated MLL1 regulation, other possible roles for PPIE, for example in splicing and mRNA processing, remain understudied.

1.2.5 Zinc finger protein 830 (ZNF830)

ZNF830, also known as CCDC16 or OMCG1, is a nuclear 42 kDa protein harboring an N-terminal C2H2-type zinc finger domain and a structural coiled-coil domain at its C-terminus (Figure 8B). This protein is conserved from *H. sapiens* to *D. melanogaster*, but it has no known homolog in yeast, and was mostly characterized in mouse models, where it was shown to be essential for embryonic development [156]. In recent years, putative roles for ZNF830 have been suggested in several biological processes, ranging from cell cycle regulation as a mitosis regulator to maintenance of genomic stability as homology recombination-associated factor [157], [158], [159]. Similarly to other IBC subunits, ZNF830 was identified as a subunit of the human spliceosome; however, the mechanisms by which it participates in pre-mRNA splicing are currently unknown.

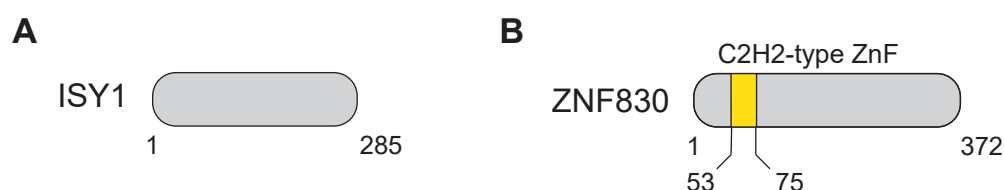


Figure 8. Schematic representation of human ISY1 and ZNF830

Schematics displaying ISY1 (A) and ZNF830 (B) domain organization (N-terminus to C-terminus). Numbers indicate amino acid positions.

1.3 Aim of this study

The intron-binding complex, IBC, is a pentameric complex and a building block in spliceosome biogenesis. As with any other protein complex, the assembly and function of the IBC requires the production of stoichiometric amounts of its subunits, as well as the rapid clearance of unassembled components. To this day, the quality control mechanisms responsible for this regulation are unknown. In this thesis, we set out to investigate how the IBC is regulated, both at the steady state and upon challenging the integrity of the complex by removing one of its subunits. By applying different mass spectrometry-based proteomics approaches, we plan to identify and characterize potential players in the quality control pathways for the IBC, with particular emphasis on the role of the ubiquitin-proteasome system.

In addition, we will look downstream the formation of the IBC and investigate how cells react to the loss of the complex. By using a combination of “omics” approaches, we will outline similarities and differences in the phenotypes arising from the depletion of three IBC subunits. This will allow us to define and functionally characterize the key features of the cellular response to the loss of the IBC.

2. Results

2.1 Regulation of the intron-binding complex

To gain insight into IBC regulation, we investigated whether IBC subunits are the target of post-translational mechanisms responsible for maintaining coherent levels of all components through elimination of excess subunits. We then used quantitative proteomics to unveil the interactome and proximal proteome of IBC subunit AQR and discover potential regulators of the complex. This allowed us to nominate E3 ubiquitin ligases UBR5 and TRIP12 as potential regulators of the stability of IBC subunits.

2.1.1 Steady-state levels of IBC subunits follow co-expression patterns

The IBC is a pentameric protein complex and a putative spliceosome building block [96], [102]. IBC formation is thought to occur before the integration in the spliceosome; in this scenario, it is likely that the IBC is the target of dedicated quality control mechanisms, regulating its assembly and ensuring its functionality before its association with the spliceosome. Such mechanisms would also ensure the production of stoichiometric amounts of IBC subunits, as well the rapid removal of excess subunits.

To gain insight into how the intron complex is regulated, we first investigated whether steady-state protein levels of IBC subunits display co-expression patterns, which is commonly observed between subunits of the same complex [95]. To this aim, we looked at the steady-state levels of IBC subunits across human cancer cell lines in the Cancer Cell Line Encyclopedia (CCLE). The unperturbed proteome of 375 cancer cell lines in the CCLE was profiled by Nusinow and colleagues [160]. From this dataset, we first extracted protein levels of AQR, XAB2, ISY1, PPIE and ZNF830 and tested for data normality and homoscedasticity. The levels of all subunits across all cancer cell lines were normally distributed, with the exception of PPIE (Shapiro-Wilk test, $p\text{-value} \geq 0.05$), and no significant difference was observed between the mean expression of the five subunits (Kruskal-Wallis test) (Figure 9A). Visually, expression levels of ISY1, PPIE and ZNF830 appeared to have a larger spread than the ones of AQR and XAB2; indeed, testing the different populations for equality of variances with Fligner-Killeen's test confirmed that expression levels of ISY1, PPIE and ZNF830 showed a higher variability. On the other hand, AQR and XAB2 expression levels had a more similar distribution, which could indicate a closer relationship or a similar regulation (Figure 9A). We went on to look more closely at the relationship between the single subunits. Pairwise Pearson correlation coefficients (PCC) calculated for every possible pair of IBC subunits across the whole dataset of 375 cancer cell lines revealed some clear trends (Figure 9B): the highest positive correlation was observed between subunits AQR and XAB2 ($r = 0.94$), followed by protein pairs ISY1-ZNF830 ($r = 0.783$), ISY1-PPIE ($r = 0.702$) and PPIE-ZNF830 ($r = 0.697$). On the other hand, we observed little to no correlation between the remaining protein pairs (AQR/XAB2-ISY1/PPIE/ZNF830). To check how the correlation between IBC subunits levels differed in a cancer cell line-specific way, we grouped the 375 cancer cell lines according to their lineage. To ensure PCC computation, only lineages with at least 3 cell lines were considered for this analysis. As expected the above-mentioned protein pairs (AQR-XAB2, ISY1-PPIE, ISY1-ZNF830 and PPIE-ZNF830) maintained their high positive correlation throughout all the 18 examined lineages (Figure 9C). Consistent with this observation,

Results

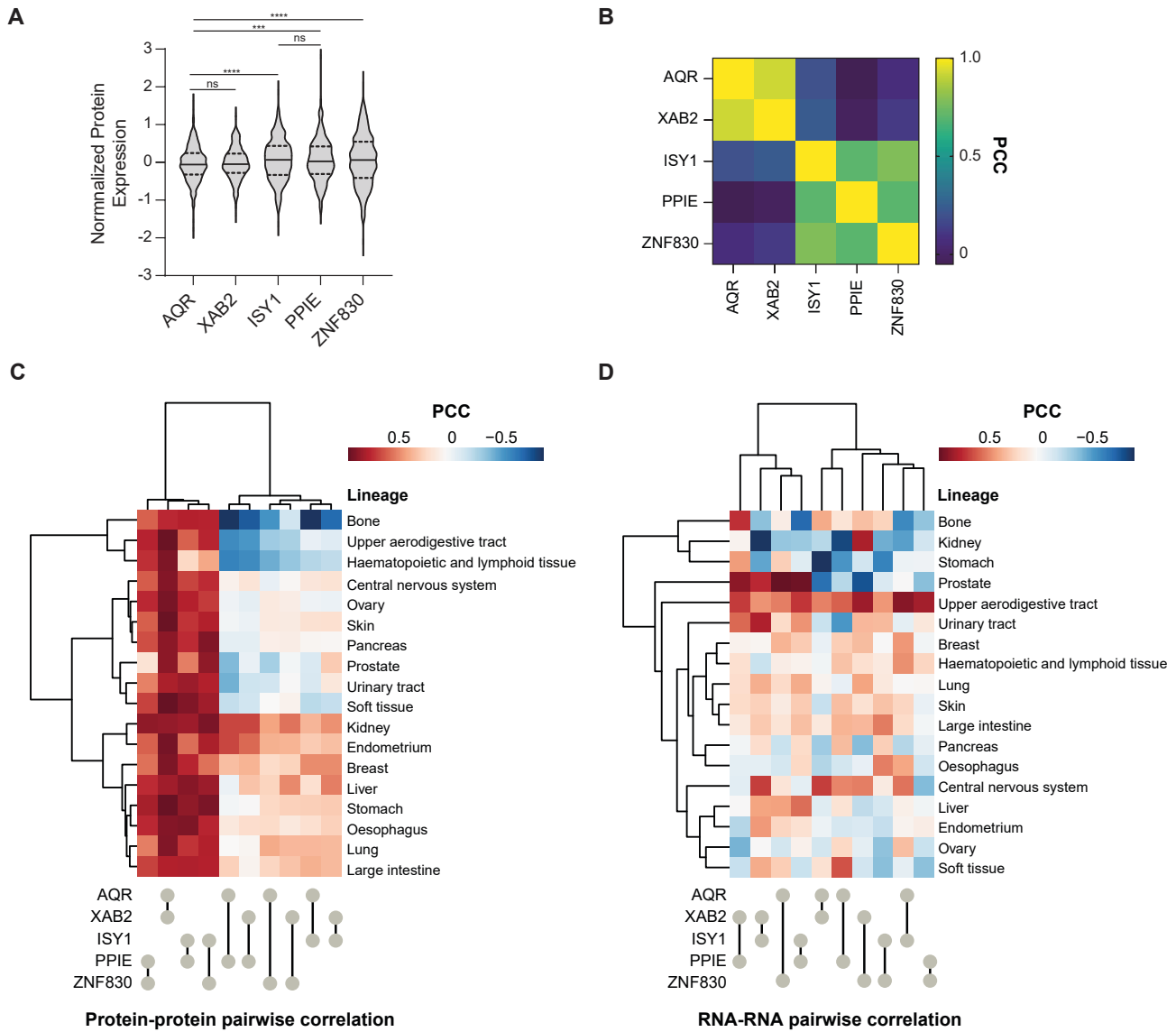


Figure 9. Steady-state protein levels of IBC subunits follow co-expression patterns

A) Violin plot of distribution of IBC subunits protein levels across cell lines in CCLE. Asterisks show the statistical significance of the Fligner-Killeen's test for equal variances (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). B) Heatmap of pairwise Pearson correlation coefficients calculated for protein levels of each pair of IBC subunits across all cell lines in CCLE. C and D) Cell lines in CCLE were grouped by lineage and pairwise Pearson correlation coefficients for protein or RNA levels of each pair of IBC subunits were calculated and visualized in a heatmap. Hierarchical clustering using Ward's method and Euclidean distance was performed on both columns and rows.

hierarchical clustering showed that these protein pairs cluster together and away from the remaining pairs. For the remaining protein pairs the PCC varied in a lineage-dependent way, most likely reflecting differences in the biology of the individual cancer lineages.

To assess if this correlation was present also at the mRNA level, we extracted publicly available mRNA expression levels of the five proteins from the same set of cell lines and, using the approach described above, computed the pairwise PCC between each possible pair of subunits [161]. Unlike what we observed at the protein level, the correlation coefficients for mRNA were variable across all lineages and subunits pairs, not mirroring what we observed at the protein level (Figure 9D). This is in line with the

general observation that levels of protein complex subunits cannot be predicted from mRNA expression or gene copy-number, as the final abundance of each component is mostly determined through co- and post-translational mechanisms [47], [95]. Taken together, this data shows that IBC subunits show specific co-expression patterns suggesting the existence of two subgroups of subunits, one including AQR and XAB2, and a second one composed of ISY1, PPIE and ZNF830. These two subgroups could reflect different steps of complex biogenesis, and be target of separate quality control mechanisms.

2.1.2 Collateral loss of IBC subunits is observed upon IBC subunit depletion

Studies have shown that the loss of a subunit of a protein complex can trigger a reduction in the levels of other members of the same complex [95], [162]. Considering the high correlation observed between levels of IBC subunits, we went on to check if this collateral loss mechanism affects IBC subunits. We used a siRNA-mediated approach to deplete three IBC subunits AQR, XAB2 and ISY in the osteosarcoma cell line U2OS and monitored protein levels after 48h (Figure 10A, upper left). Two single siRNA sequences for each target protein were tested and showed similar effects in terms of depletion of their respective target; all the experiments shown here were carried out with one siRNA sequence (siAQR#1, siXAB2#2 and siISY1#4) unless stated otherwise. In addition to the depletion of each target protein, western blot

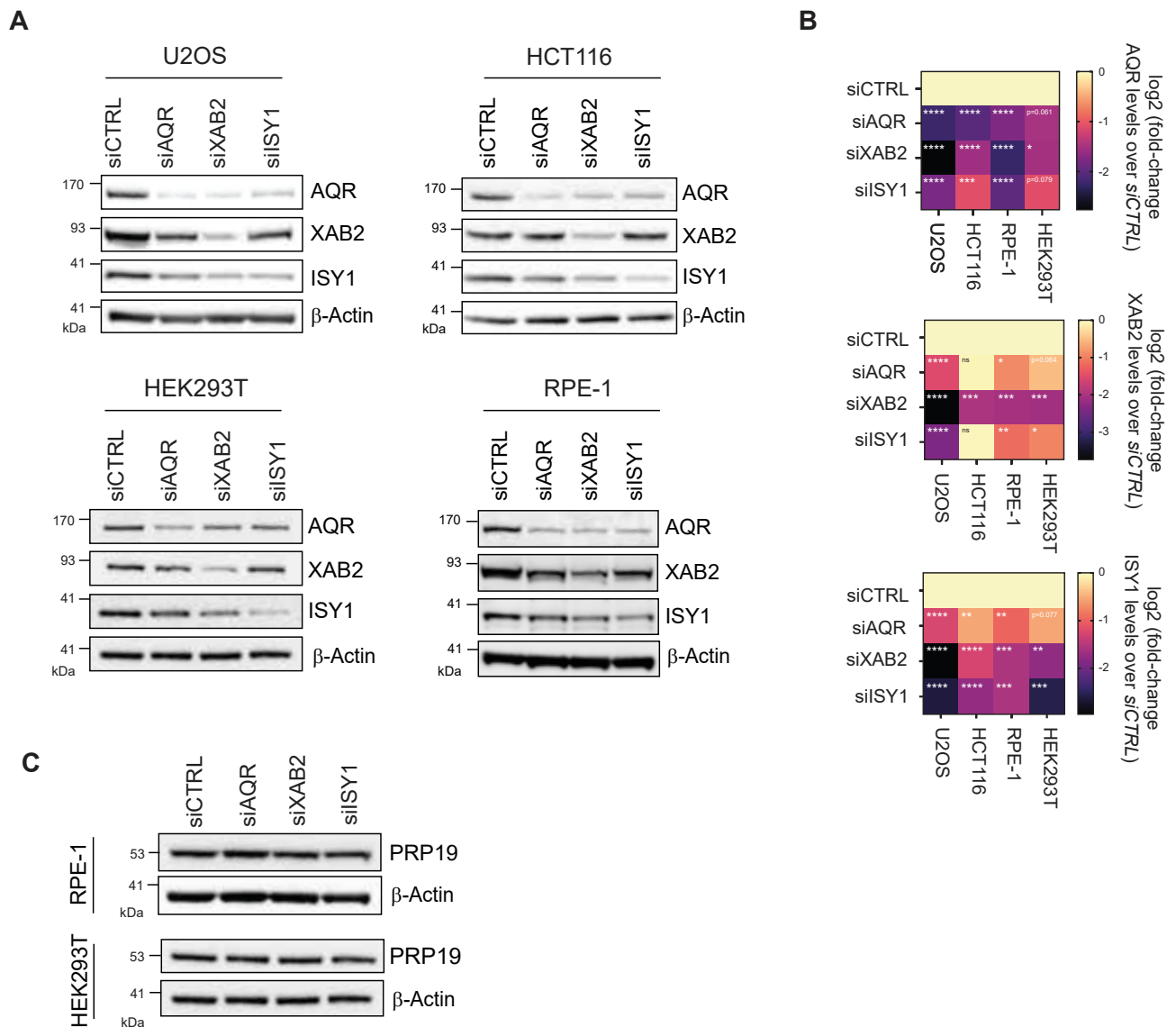


Figure 10. Collateral loss of IBC subunits is observed upon IBC subunit depletion

A) Western blot analysis of IBC subunits levels in AQR, XAB2 or ISY1-depleted U2OS, HCT116, RPE-1 and HEK293T cells. Levels were monitored after 48h of knock-down. B) Quantification of AQR, XAB2 or ISY1 levels upon depletion of IBC subunits in U2OS, HCT116, RPE-1 and HEK293T cell lines (n=3 or 4). Protein levels were normalized on endogenous β -Actin levels and represented as log2 fold-change over siCTRL. Bar plots show mean and SD. Statistical significance was calculated with one-way ANOVA with Tukey's post hoc test for multiple comparisons (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). C) Western blot analysis of PRP19 levels in AQR, XAB2 or ISY1-depleted RPE-1 and HEK293T.

analysis showed that each knock-down caused a significant reduction in protein levels of the other IBC subunits (Figure 10A, upper left). We checked if this phenotype was conserved across different cell lines by depleting AQR, XAB2 or ISY1 in another cancer cell line, HCT116, and in two non-transformed cell lines, RPE-1 and HEK293T (Figure 10A). We observed similar effects across all tested cell lines, with some variability. Specifically, we observed XAB2 levels to be more resistant to the fluctuations in levels of the other subunits. Conversely, AQR levels were more sensitive to the knock-down of other subunits across all tested cells lines (Figure 10B). In the spliceosome, IBC subunits establish dynamic interactions with a large number of proteins, including the E3 ubiquitin ligase PRP19 and subunits of the homonymous PRP19 complex. Interestingly, we found that PRP19 levels were not affected by any of the knock-downs (Figure 10 C), suggesting that the collateral loss is restricted to IBC subunits.

2.1.3 The collateral loss of IBC subunits is resistant to inhibition of protein degradation pathways

Quality control of protein complexes heavily relies on post-translational regulation to remove unassembled subunits or faulty assembly intermediates. In general, post-translational regulation of protein levels involves the activation of the two main protein degradation pathways, the ubiquitin-proteasome system (UPS) and autophagy. To assess the role of these pathways in regulating the stability of IBC subunits, we monitored protein levels in CTRL and AQR-depleted U2OS cells upon inhibition of the proteasome or of autophagy by western blot analysis (Figure 11A and B). Inhibition of the chymotrypsin-like activity of the 26S proteasome by treatment with MG132 in a time-course from 0h to 6h caused the expected accumulation of ubiquitinated proteins, as seen by blotting for total ubiquitin [163]. Furthermore, we observed increased level of transcription factor c-MYC, which is tightly regulated by the ubiquitin-proteasome system and was used as positive control [164]. To inhibit autophagy, cells were treated with Bafilomycin A1 (BafA) for 4h. This chemical targets the vacuolar-type H^+ -ATPase (V-ATPase) and inhibits the acidification of lysosomes, which is needed for activation of degradative enzymes such as cathepsins [165]. Additionally, BafA hinders the maturation of autophagosomes to autolysosomes [166]. To assess the efficacy of BafA treatment we monitored the levels of the lipidated form of MAP1LC3 (also known as LC3B-II), whose accumulation is marker for autophagy inhibition [167], [168], [169]. As expected, we observed a significant BafA-induced increase in LC3B-II levels, consistent with reduced turnover of the protein. As an additional control, we probed for the autophagy receptor SQSTM1 (also known as p62), whose levels are expected to increase as a result of inhibition of the autophagic flux [170]. In our case, the increase of p62 levels was mild and not statistically significant in both control cells and AQR-depleted cells. Interestingly, levels of p62 were significantly reduced in absence of AQR. Overall, we concluded that the both treatments worked in the expected way.

Consistent with our previous results, siRNA-mediated depletion of AQR caused the collateral loss of XAB2 and ISY1. We observed a similar trend for the levels of another subunit, PP1E, but changes were

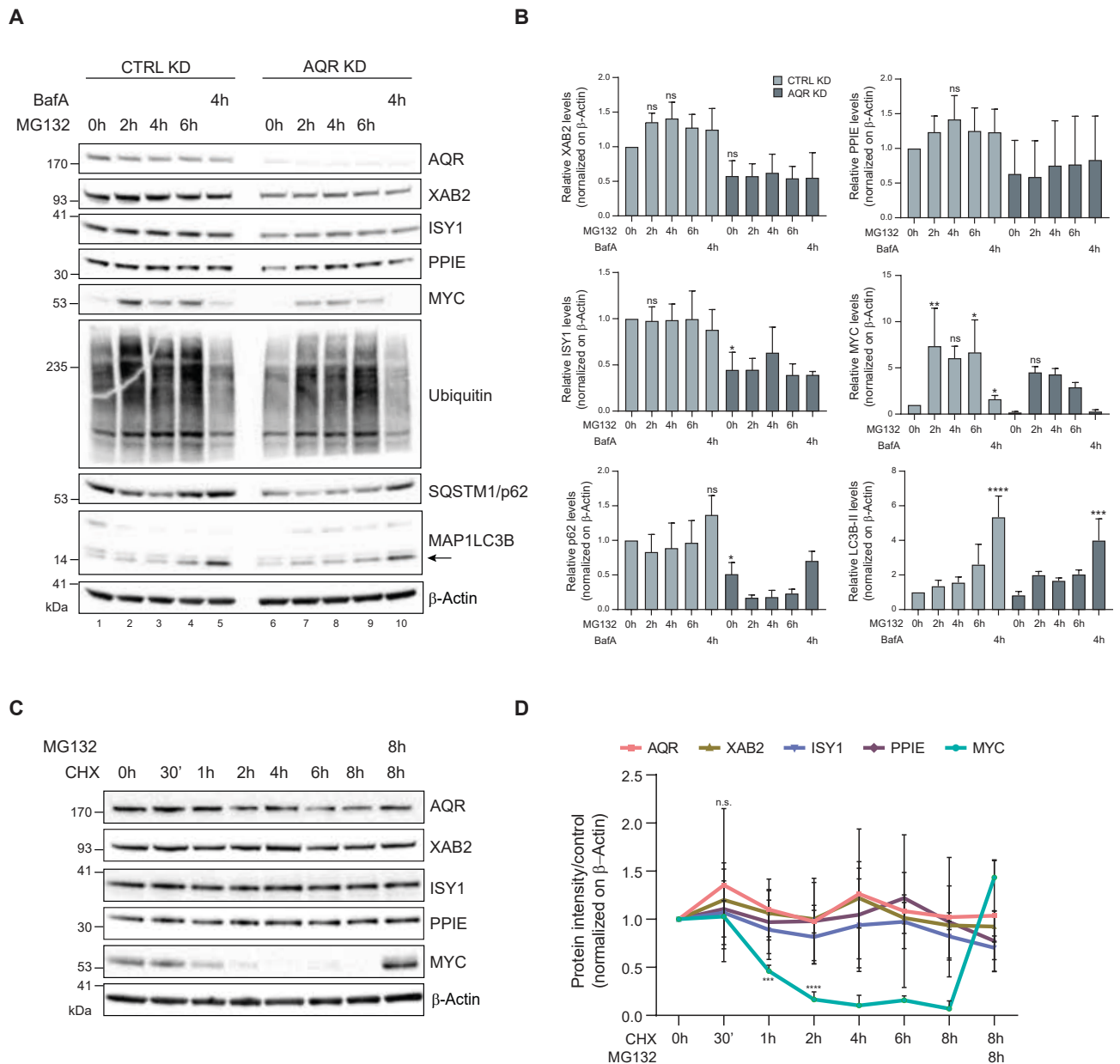


Figure 11. Collateral loss of IBC subunits is resistant to inhibition of protein degradation pathway

A) Western blot analysis of levels of indicated proteins in CTRL or AQR-depleted U2OS cells. Cells were treated with 10 μ M MG132 or 200nM Bafilomycin A for the indicated time-points. Levels were monitored after 48h of knock-down. Arrow indicates the correct band for MAP1LC3B. B) Quantification of levels of indicated proteins in CTRL or AQR-depleted U2OS cells (n=3). Protein levels were normalized on endogenous β -Actin levels and represented as fold-change over siCTRL. Bar plots show mean and SD. Statistical significance was calculated with one-way ANOVA with Tukey's post hoc test for multiple comparisons (* P <0.05, ** P <0.01, *** P <0.001, **** P <0.0001). C) Western blot analysis of levels of indicated proteins in U2OS cells treated with 100 μ g/mL CHX or 10 μ M MG132 for the indicated time-points. D) Quantification of levels of indicated proteins normalized on endogenous β -Actin levels and represented as fold-change over the 0h time-point. Statistical significance was calculated with one-way ANOVA with Tukey's post hoc test for multiple comparisons (* P <0.05, ** P <0.01, *** P <0.001, **** P <0.0001). Error bars represent the standard deviation.

not statistically significant. When comparing treated and untreated cells, we observed that the levels of IBC subunits in CTRL KD cells were not affected by MG132 or BafA treatment (Figure 11A, lanes 1-5). Surprisingly, neither proteasome nor autophagy inhibition could rescue the levels of IBC subunits (Figure 11A, lanes 6-10). We observed a similar outcome upon proteasomal inhibition with the drug Bortezomib, as well as upon depletion of a different IBC subunit (Appendix, Figure 30A and B). Lastly, a side-by-side comparison of our lysis protocol with high salt RIPA buffer with a denaturing lysis protocol using 8M urea gave comparable results, indicating that IBC subunits are not present in the insoluble fraction (Appendix, Figure 30C). The inability of our treatments to influence steady-state levels or rescue the collateral loss phenotype could have both biological and technical reasons.

For example, it is possible that IBC subunits are long-lived proteins, and therefore the length of the MG132 and BafA treatments was too short to observe visible changes in their steady-state levels. To test this hypothesis, we assessed the stability of IBC subunits by inhibiting the elongation step of translation in a cyclohexamide (CHX) chase assay (Figure 11C and D) [171]. Wild-type U2OS cells were treated with CHX at time-points from 0h to 8h, and levels of IBC subunits were monitored by western blot. Interestingly, levels of AQR, XAB2, ISY1 and PPIE remained mostly stable throughout the length of the time-course. Conversely, levels of a protein with short half-life such as c-MYC dropped drastically after 2h inhibition of protein synthesis. In line with previous observations, co-treatment with MG132 rescued c-MYC levels but did not affect the ones of IBC subunits. A study by McShane and colleagues investigating the kinetics of protein degradation categorized AQR, XAB2 and PPIE as exponentially degraded proteins, with their abundance slowly decreasing over the course of an 8h pulse-chase experiment [61]. Given the similarity of our results, we concluded that IBC subunits are stable proteins that undergo exponential degradation. As mentioned before, this could partially explain the unchanged protein levels upon inhibition of protein degradation pathways. However, at the moment we cannot explain the inability in rescuing the collateral loss of IBC subunits.

To sum up, we have shown that IBC dynamics display features consistent with the regulation of a multiprotein complex, such as co-expression patterns and collateral loss phenotype. Some outstanding questions remain, for example which post-translational mechanisms are responsible for these phenomena. Additional work and different strategies are needed to answer these questions: for example, identifying the interactors of the IBC could give us a clue as to how the complex is regulated.

2.2 Proteomic strategies for the identification of intron-binding complex regulators

2.2.1 Generation of a system for combined interactomics and proximity proteomics of endogenous AQR

In order to gain further insight into the mechanisms behind IBC dynamics, we focused on the identification of potential regulators of the complex by mapping the AQR interactome and proximal proteome in a combined approach. For this we employed a genetically modified cell line, previously generated in our lab using a tailored CRISPR/Cas9-based knock-in workflow to C-terminally tag the endogenous *AQR* locus in U2OS cells with an APEX2-3×FLAG cassette. The 3×FLAG

peptide was used to probe the direct interactors of AQR by affinity purification and MS (AP-MS); on the other hand, the engineered ascorbate peroxidase APEX2 allowed for identification of AQR proximal proteins in vivo by proximity-dependent biotinylation coupled with mass spectrometry (PDB-MS) (Figure 12). C-terminal tagging was favored over N-terminal based on the previously described role of the N-terminal ARM domain in mediating protein-protein interactions [96], as well as on the lower amount of post-translational modifications annotated at the C-terminus (as seen in PhosphoSite Plus [172]). After transfection with the Cas9-RNP complex and antibiotic selection, single cell clones were screened based on FLAG expression, biotinylation activity of APEX2 and subcellular localization by high-content microscopy.

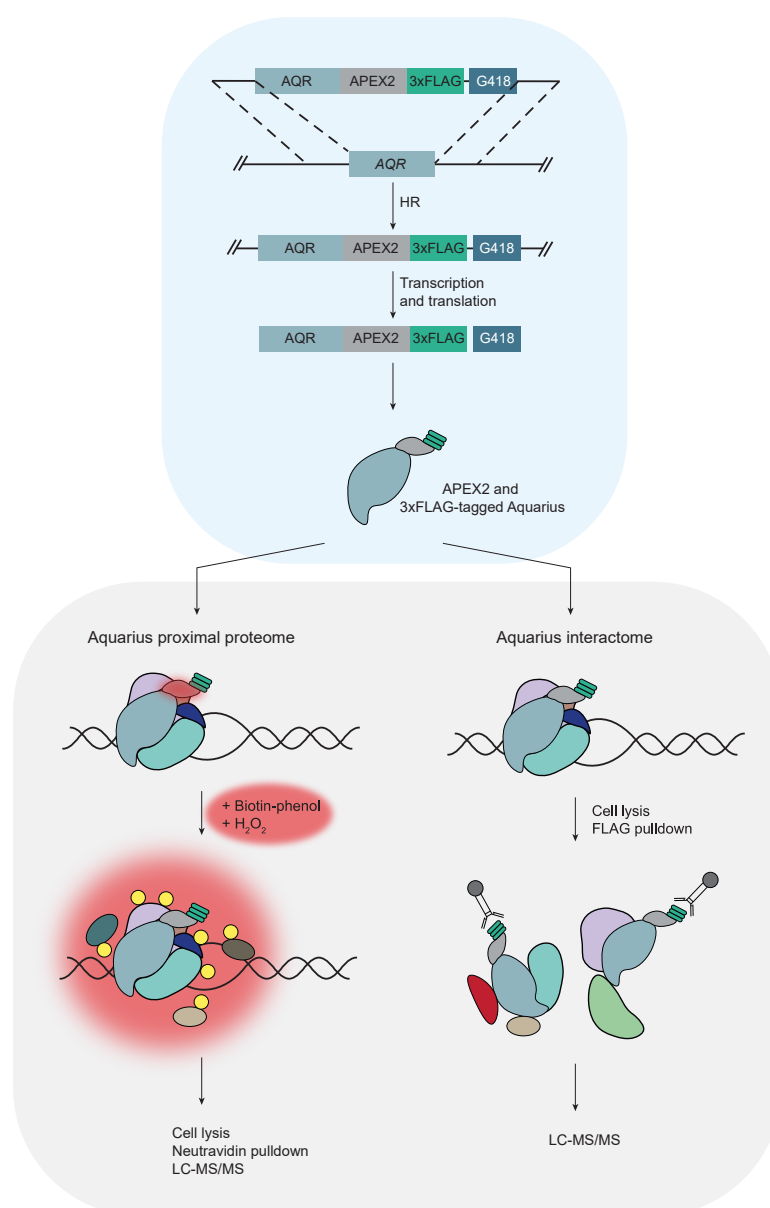


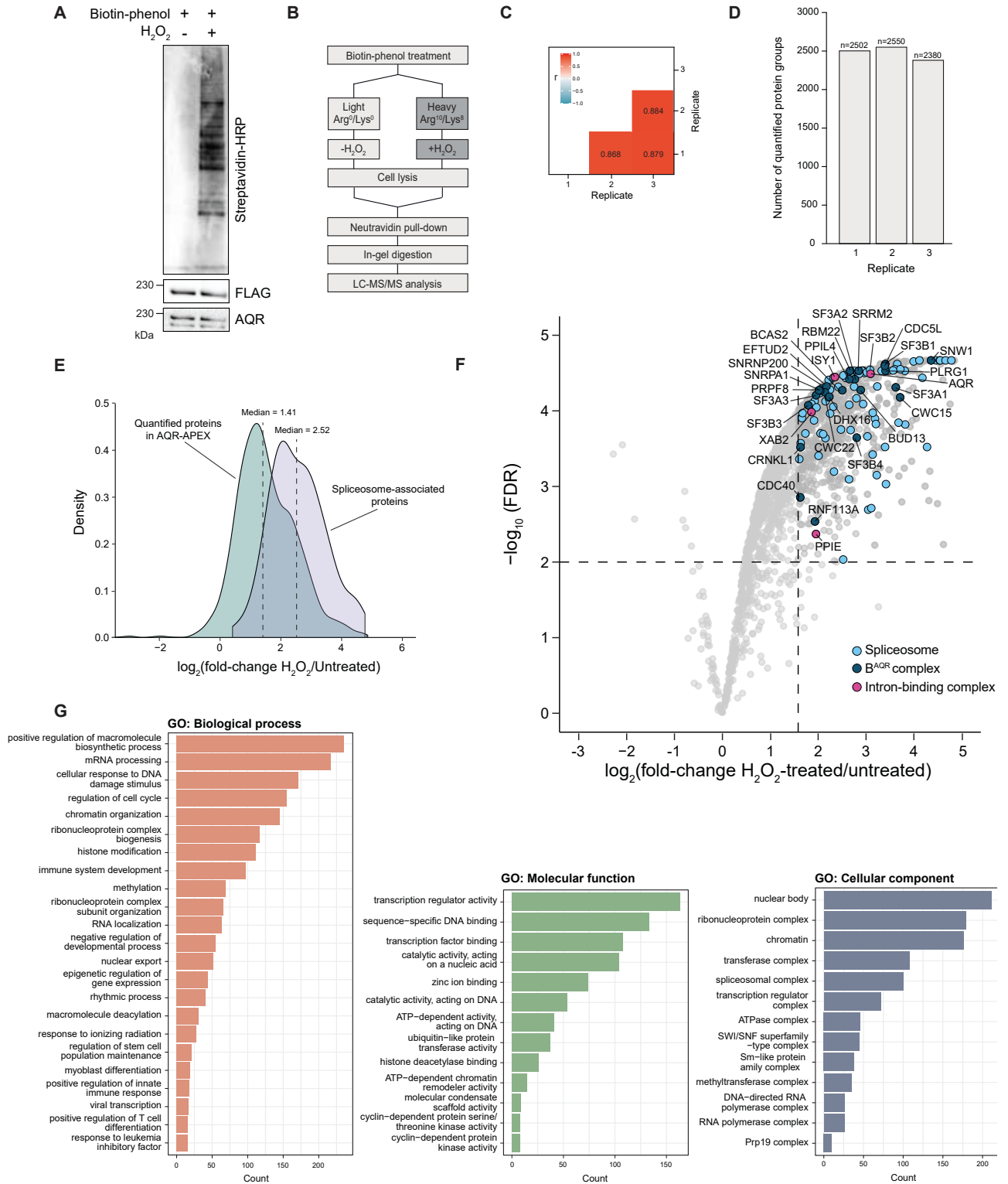
Figure 12. Workflow of generation of the AQR-APEX2-3xFLAG cell line and applications in proteomics

2.2.2 Endogenous proximal proteome identifies known and novel potential interactors of AQR

To identify the AQR proximal proteome, we first treated AQR-APEX2-3×FLAG cells with biotin-phenol and induced biotinylation by treating with H₂O₂ for 2 minutes. Western blot analysis confirmed efficient biotinylation, which was specifically induced by the H₂O₂ pulse with minor background in untreated cells (Figure 13A). Due to the covalent bond between biotin and acceptor proteins, the protocol allows for harsh lysis conditions to extract nuclear and chromatin-bound proteins. Biotinylated proteins were enriched via streptavidin pulldown and analyzed by liquid chromatography (LC)-tandem mass spectrometry (MS/MS). Stable isotope labelling with amino acids in cell culture (SILAC) was used to differentiate between proteins that were specifically biotinylated upon H₂O₂ treatment from endogenously biotinylated proteins (Figure 13B). We performed the experiment in three replicates which showed good reproducibility ($r > 0.85$ across all replicates), and quantified 2502, 2550 and 2380 proteins respectively (Figures 13C and D). After applying cut-offs for statistical significance (fold-change ≥ 3 ; FDR ≤ 0.01 ; proteins quantified in at least $n=2$ replicates) we obtained a total of 921 proteins enriched in AQR proximity with high confidence (760 proteins were quantified across all three replicates). As AQR is a component of the spliceosome, we reasoned that other spliceosome-associated proteins should be preferentially enriched in AQR proximity. Accordingly, we found a four-fold enrichment of spliceosomal proteins compared to all other non-spliceosomal proteins in our dataset (Figure 13E). A recent study defined the existence and composition of the BAQR complex, a transitional state of the spliceosome occurring during the step of catalytic activation [102]. Accordingly, we could identify 17 out of the 20 subunits listed as key components of the BAQR complex in our dataset, including all the IBC subunits with the exception of ZNF830, which was not measured in this experiment (Figure 13F). Functional enrichment analysis showed that AQR-proximal proteins are primarily nuclear and chromatin-bound proteins, often belonging to protein complexes involved in co-transcriptional processes and chromatin remodeling (Figure 13G). In addition to the expected transcription and splicing-related GO terms, we found an enrichment of proteins associated to the GO term “Ubiquitin-like protein transferase activity”. In our dataset, we identified 37 proteins associated to this term, encompassing one E2 conjugating enzyme, 25 E3 ubiquitin ligases, 5 E3 SUMO ligases, and the NEDD8-activating enzyme E1 catalytic subunit UBA3. Further evaluation of our dataset revealed additional members of the ubiquitin system as potential AQR-proximal proteins. Among these, we identified 8 deubiquitinating enzymes and the E1 ubiquitin activating enzyme UBA1 and UBA6 (Figure 14A). Overall, this approach allowed the identification of potential novel AQR interactors, some of which might play a role in IBC regulation.

Figure 13. Endogenous proximal proteome identifies known and novel potential interactors of AQR

A) Western blot analysis of H₂O₂-induced biotinylation in the AQR-APEX2-3XFLAG cell line. Cells were treated with biotin-phenol for 2h and with H₂O₂ for 2 minutes. Biotinylated proteins were detected with Streptavidin-HRP. B) Workflow of the SILAC-based endogenous proximal proteome of AQR. C) Pearson correlation coefficients across three replicates of the AQR proximal proteome. D) Quantified proteins in three replicates of the AQR proximal proteome. E) Density plot of fold-change distribution of all proteins quantified in the AQR proximal proteome or of spliceosome-associated proteins only. Dashed lines indicate the median of the distributions. F) Volcano plot of the AQR proximal proteome. Dashed lines indicate cut-offs for statistical significance (fold-change ≥ 3 ; FDR ≤ 0.01). G) Functional enrichment analysis performed with ClusterProfiler showing the enriched GO terms (FDR ≤ 0.05) for AQR-proximal proteins.



A

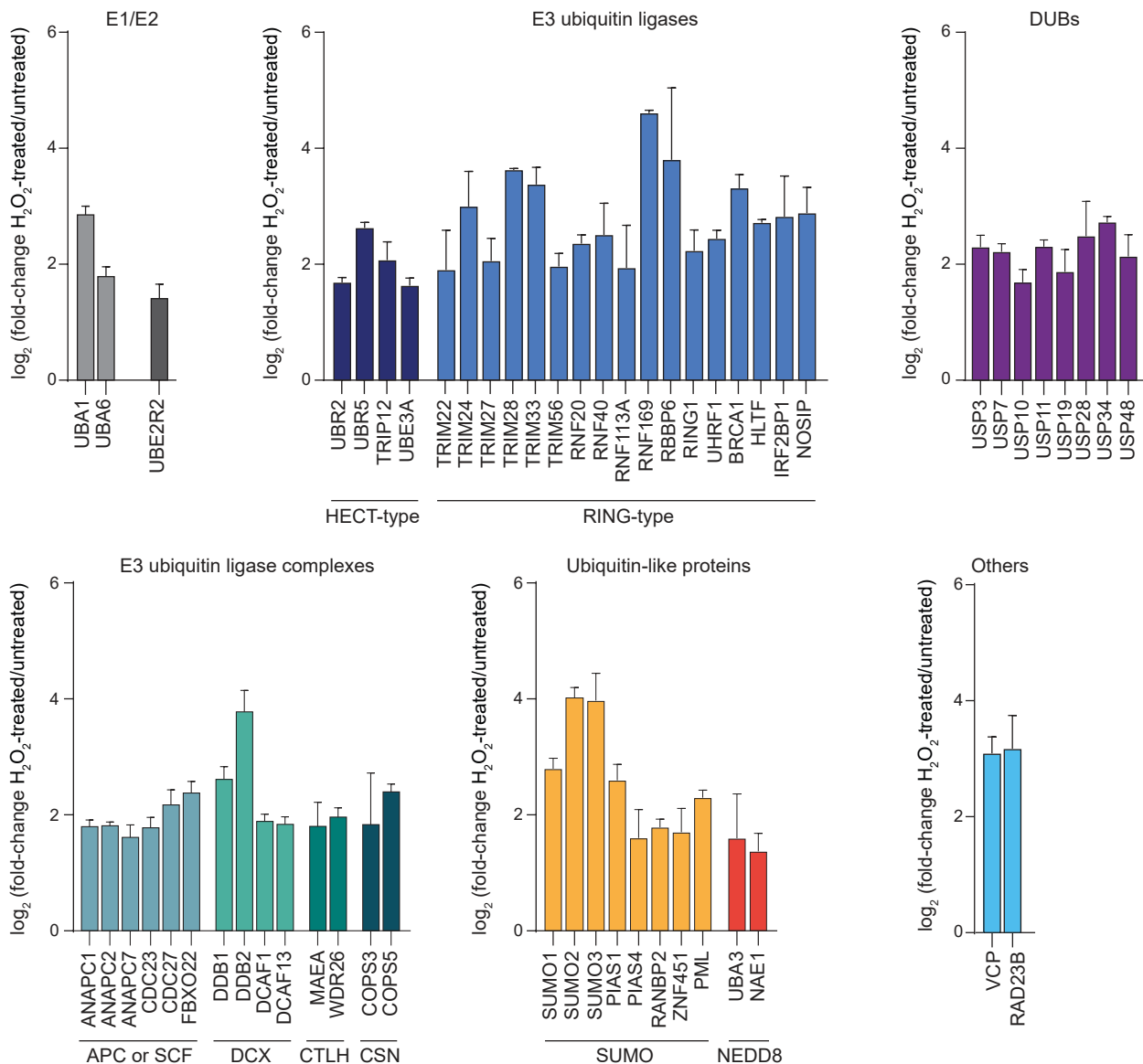


Figure 14. Ubiquitin-proteasome system-associated hits in AQR endogenous proximal proteome

A) Bar plots showing the mean and SD of log₂ (non-MQ normalized) SILAC ratios for H₂O₂-treated versus untreated cells for different groups of enzymes associated to the ubiquitin-proteasome system. All proteins pass an FDR cutoff of 0.01.

2.2.3 E3 ubiquitin ligases are direct interactors of AQR

As previously described, the AQR-APEX2-3×FLAG cell line allowed us not only to enrich for proteins proximal to AQR, but also to probe for AQR direct interactors. To this aim, we performed a FLAG pull-down and quantified co-immunoprecipitating proteins by LC-MS/MS. SILAC was employed to obtain quantitative information on enriched proteins (Figure 15A). We performed two separate experiments with four replicates each, including two SILAC label switches to account for label-specific effects. After application of quality filters, we quantified more than 1500 proteins across all replicates (Figure 15B). As expected, IBC subunits consistently co-immunoprecipitated with AQR (Figure 15C). After applying cut-offs (fold-change ≥ 1.5; p-value ≤ 0.05; proteins quantified in at least n=4 replicates) we obtained a total of 128 AQR interactors (Figure 15D). To investigate the biological processes associated with our

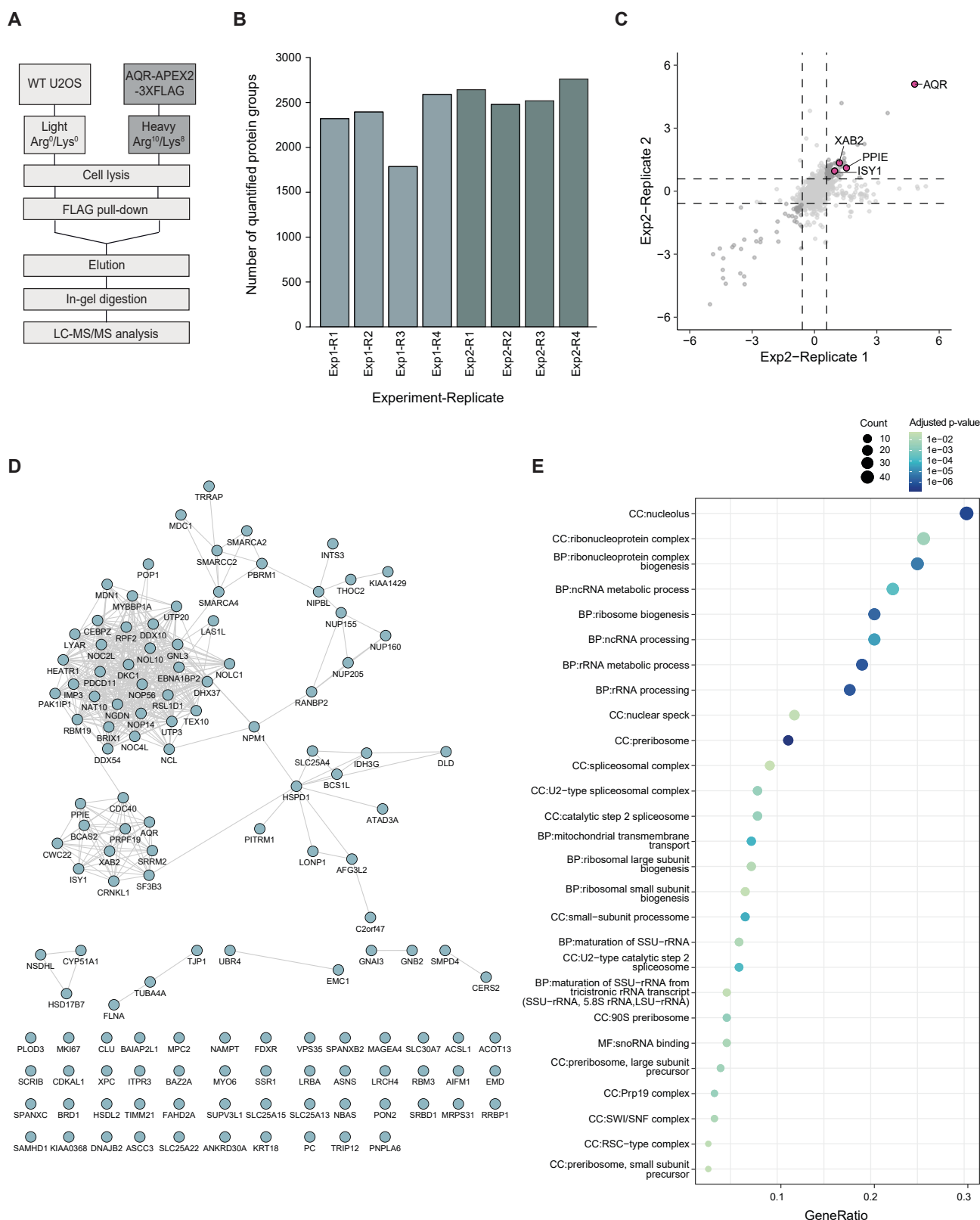


Figure 15. AQR-3xFLAG AP-MS successfully enriches endogenous AQR interactors

A) Workflow of the SILAC-based AQR-3xFLAG AP-MS. B) Quantified proteins in two experiments (four replicates each experiment) of the AQR-3xFLAG AP-MS. C) Scatterplot of two replicates of the AQR-3xFLAG AP-MS. Axis show the log₂ fold-change of SILAC ratios (AQR-3xFLAG cells over parental WT cells). Dashed lines indicate fold-change cut-offs (fold-change ≥ 1.5). D) STRING analysis and Cytoscape visualization of significant hits in the AQR-3xFLAG AP-MS (fold-change ≥ 1.5 ; p-value ≤ 0.05). E) Functional enrichment analysis performed with ClusterProfiler showing the enriched GO terms (FDR ≤ 0.05) for AQR-3xFLAG interactors.

hits, we performed functional enrichment analysis for the 128 significantly enriched proteins against all the proteins quantified in our dataset. As expected, GO term analysis showed enrichment for splicing-associated terms such as “U2-type catalytic step 2 spliceosome” and “Prp19 complex”, matching the array of spliceosomal proteins with which AQR is known to interact with (Figure 15E). In line with the proximal proteome, we found enrichment of chromatin remodeling complexes in our dataset: specifically, we identified 5 members of the SWI/SNF complex, namely SMARCA2, SMARCA4, ARID2, PBRM1 and SMARCC2. To identify potential upstream regulators of the complex stability, we inspected our dataset for proteins known to function in the ubiquitin-proteasome system. Interestingly, E3 ligases TRIP12 and UBR4, and the E3 SUMO-protein ligase RANBP2 passed the significance cut-offs for high-confidence AQR interactors. Together with our previous results, this could hint at a role for these proteins in regulating the stability of IBC subunits.

2.2.4 Integration of AQR interactome and proximal proteome nominates potential IBC regulators

The overlap of our interactome and proximal proteome datasets showed that 45% of interactors were also identified in the AQR proximal proteome (Figure 16A). To evaluate the similarity of the biological processes in which AQR interactors and proximal proteins might be involved in, we compared the results of the functional enrichment analysis of both datasets. Of note, functional enrichment analysis of the AQR proximal proteome yielded a significantly higher number of enriched terms than the interactome dataset. This can be explained by the fact that the number of quantified proteins in the AQR proximal proteome is 7.2 times higher than the number of *bona fide* AQR interactors. We found enrichment of 27 and 391 GO terms across all ontologies for AQR AP-MS and PDB-MS datasets respectively. Similarly, AQR interactors and proximal proteins associated to 5 and 22 KEGG pathways respectively. Despite the difference in the total number of GO terms, the shared terms between the two dataset indicated a similar signature: in addition to the expected splicing-associated terms, such as “U2-type spliceosomal complex” and “Prp19 complex”, we found an enrichment of proteins belonging to “nucleolus” and “SWI/SNF complex”. In line with this, the top three enriched KEGG pathways of AQR interactome, “Spliceosome”, “ATP-dependent chromatin remodeling” and “Ribosome biogenesis in eukaryotes”, were shared with the AQR proximal proteome (Figure 16B). In agreement with the functional enrichment analysis, the overlap between the two datasets encompassed IBC subunits XAB2, ISY1 and PPIE, as well as components of the B^{act} and B^{AQR} complex, SF3B3, CWC22, BCAS2, SRRM2, CDC40 and CRNKL1 (Figure 16C). Consistent with the potential role of AQR as a factor involved in the early steps of ribosomal biogenesis, a number of shared interactors mapped to the “nucleolus” ontology [115]. Lastly, subunits of the SWI/SNF complex SMARCA2, SMARCA4, SMARCC2 and PBRM1 appeared in both datasets. When looking at potential regulators of AQR stability, we found that the E3 ubiquitin ligase TRIP12 and the E3 SUMO-protein ligase RANBP2 were present in both datasets. Previous studies have shown a role for TRIP12 in targeting for degradation unbound subunits of protein complexes: examples include BAF complex subunit BAF57, neddylation E1 enzyme subunit APP-BP1, and transcription factor ARF when not in complex with nucleophosmin and TRADD [173], [174], [175], [176], [177]. This knowledge prompted us to look into TRIP12 as a possible candidate for IBC regulation.

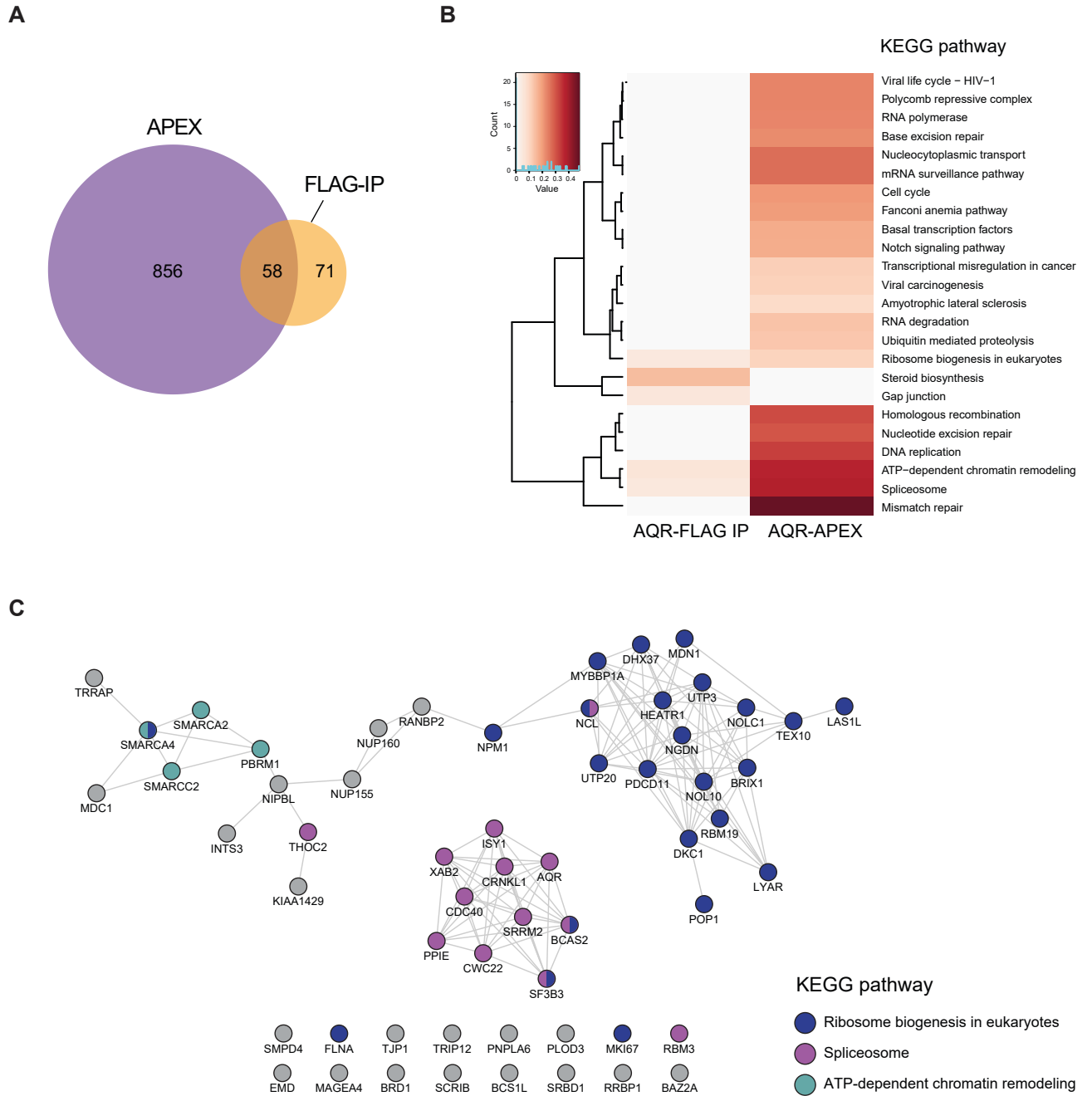


Figure 16. Integration of AQR interactome and proximal proteome nominates potential IBC regulators

A) Venn Diagram showing the overlap between significant hits in the AQR PDB-MS (fold-change ≥ 3 ; FDR ≤ 0.01) and the AQR-3 $\times$ FLAG AP-MS (fold-change ≥ 1.5 ; p-value ≤ 0.05) datasets. B) Heatmap of enriched KEGG pathways (FDR ≤ 0.05) from functional enrichment analysis of the AQR proximal proteome or the AQR interactome with Clusterprofiler. C) STRING analysis and Cytoscape visualization of the overlap between the AQR proximal proteome and the AQR interactome.

2.2.5 TRIP12 interacts with IBC subunits

To investigate the possible role of TRIP12 in regulating the stability of IBC subunits, we first set out to confirm the interaction between AQR and TRIP12. To this aim, we immunoprecipitated endogenous TRIP12 in U2OS cells and identified co-immunoprecipitated proteins by SDS-PAGE and western blot analysis (IP-WB). To achieve a more efficient extraction of chromatin-bound proteins, the NaCl concentration in the lysis buffer was increased up to 450mM, and then brought back down to the original 150mM for the immunoprecipitation and washes steps. The IP-WB approach confirmed the interaction between endogenous TRIP12 and the IBC subunits AQR and XAB2; on the other hand, we did not detect co-immunoprecipitation of ISY1 and PPIE with TRIP12 (Figure 17A). We independently confirmed the interaction between TRIP12 and XAB2 by performing an endogenous XAB2 IP-WB in U2OS cells (Figure 17B). Taken together, these results indicate that TRIP12 might play a role in IBC regulation by acting directly on either AQR, XAB2, or both.

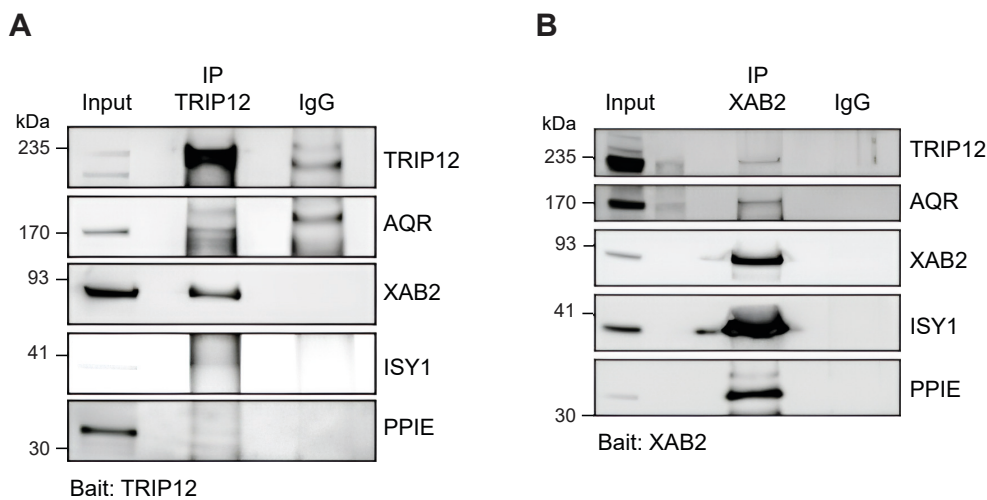


Figure 17. TRIP12 interacts with IBC subunits

A) IP-Western blot of endogenous TRIP12. B) IP-Western blot of endogenous XAB2.

In some instances, TRIP12 requires the cooperation with other E3 ubiquitin ligases to efficiently target substrates to degradation. For example, TRIP12 homolog in *S. cerevisiae*, Ufd4, interacts and cooperates with the RING-type Ubr1 E3 ubiquitin ligase in the degradation of Arg/N-end rule substrates [85]. In humans, four E3 ubiquitin ligases, also known as UBR box-containing E3 ligases, participate in the N-end rule pathway, namely UBR1, UBR2, UBR4 and UBR5 [178]. Of these proteins, UBR5 and UBR2 appeared enriched in proximity of AQR; additionally, UBR4 was significantly enriched in the AQR interactome and was identified in the AQR proximal proteome, showing a strong trend toward enrichment albeit not meeting our criteria for statistical significance. Notably, functional interaction between TRIP12 and UBR5 has been previously observed, for example in the context of the regulation of RNF168 levels at double-strand breaks [179]. Taken together, these findings prompted us to test whether TRIP12 and UBR box-containing E3 ligases cooperate in controlling the levels and stability of IBC subunits.

2.2.6 UBR5 knock-out opposes loss of IBC subunits

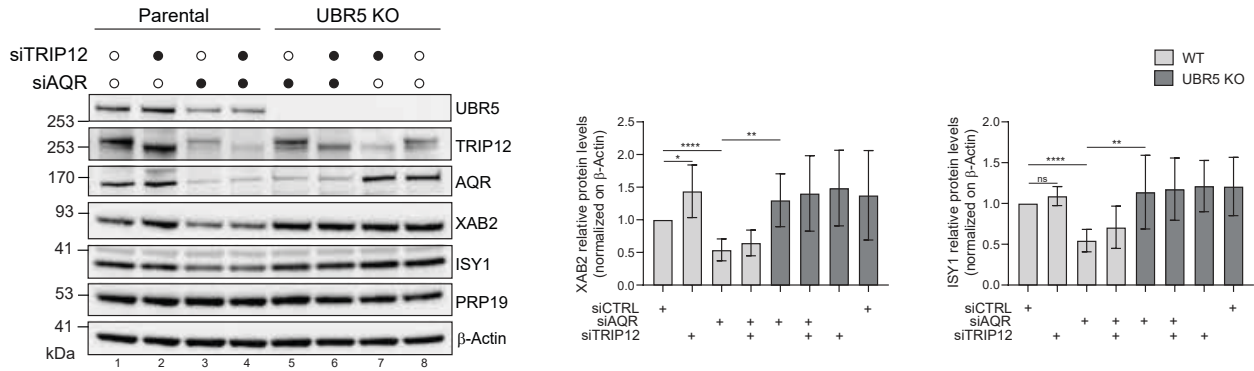
To assess the role of UBR box-containing E3 ligases in regulating levels of IBC subunits, we decided to first focus on UBR5, and took advantage of an RPE-1 UBR5 knock-out cell line previously generated in our lab. Thorough characterization of two independent clones from this cell line at the genomic, transcriptomic and proteomic levels was carried out previously and will not be presented here (unpublished data). Consistent with our previous results, siRNA-mediated depletion of AQR in parental wild-type RPE-1 cells led to loss of XAB2 and ISY1 (Figure 18A, lane 3). Strikingly, levels of XAB2 and ISY1 remained unaffected by AQR depletion in two independent UBR5 knock-out clones, suggesting that absence of UBR5 abolishes the collateral loss of IBC subunits (Figure 18A, lane 5). We repeated the same experiment upon depletion of either XAB2 or ISY1 (Figure 18B and C): interestingly, we observed different outcomes according to the subunit we monitored. We found that absence of UBR5 opposed the loss of AQR in both XAB2 and ISY1-depleted cells. However, this did not apply to XAB2 or ISY1, which underwent a reduction in protein levels similar to the one observed in wild-type cells (Figure 18B, central and lower panels, right). In other words, absence of UBR5 was not always sufficient to block the collateral loss of IBC subunits. This suggests a subunit-dependent regulation, in which UBR5 is possibly only one of multiple players.

Considering the potential interplay between UBR5 and TRIP12, we monitored the effect of TRIP12 depletion on the levels of IBC subunits in wild type RPE-1 or UBR5 knock-out RPE-1 cells. Depletion of TRIP12 in wild-type cells led to a mild increase in the levels of AQR and XAB2 (Figure 18A, B and C, lane 2); on the other hand, we did not observe any additive effect when depleting TRIP12 in UBR5 KO cells (lane 7). In general, TRIP12 depletion did not rescue the levels of IBC subunits as efficiently as the knock-out of UBR5; this could be explained by insufficient depletion of TRIP12, up-regulation of different isoforms, or a less crucial role of TRIP12 in targeting the subunits to degradation. In line with these results, a preliminary experiment aimed at studying the proteome-wide changes in ubiquitination caused by 48h siRNA-mediated depletion of TRIP12 did not reveal any significantly changing ubiquitination site on IBC subunits (data not shown). TRIP12 is known to modify substrates with K29-linked ubiquitin chains, and to participate in the formation of K29/K48-linked branched ubiquitin chains [180], [181]. Strikingly, we found that depletion of TRIP12 caused a statistically significant, proteome-wide decrease of K29-linked ubiquitylation, suggesting that loss of TRIP12 has the potential of causing a global change in the architecture of ubiquitin chains (Figure 18E). Notably, K29-linked ubiquitin chains often do not constitute a strong enough signal to induce substrate degradation, and it is therefore possible that blocking formation of these linkages by TRIP12 depletion would not lead to the substantial stabilization of its substrates. However, in this case TRIP12 could either boost the degradation signal laid by UBR5, or provide UBR5 with a pre-formed ubiquitin chain for addition of K48-linked monomers. Finally, we observed a clear reduction in the levels of UBR5 upon AQR and XAB2 knock-down, as well as mild increase of UBR5 levels in response to TRIP12 depletion (Figure 18D). Levels of TRIP12 also seemed reduced upon knock-down of IBC subunits, although the appearance of multiple bands complicates the interpretation of our western blot results. To our knowledge, UBR5 and TRIP12 have not been shown to regulate each other's levels, thus warranting further investigation.

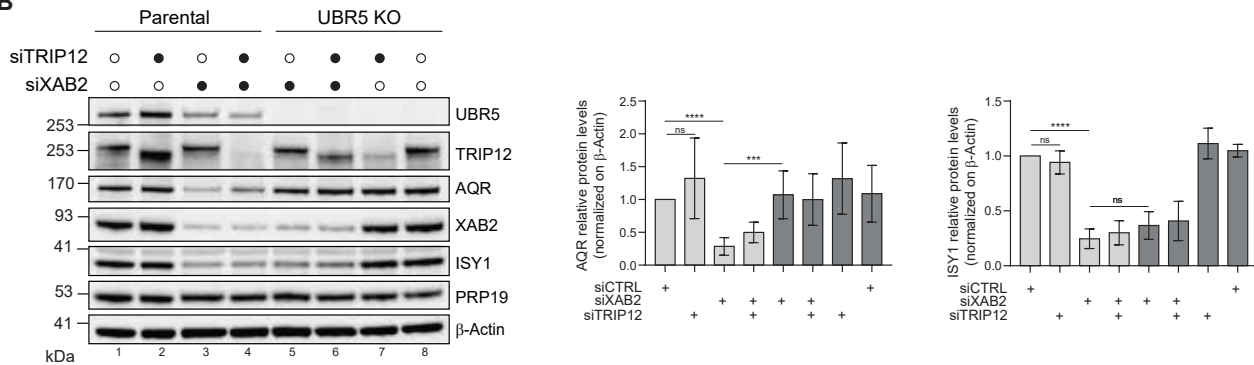
Overall, we obtained very promising results implicating UBR5 and TRIP12 in IBC regulation. Based on interaction data, we propose that these two E3 ligases specifically target a subset of subunits to degradation, namely AQR and XAB2; this leaves room for investigating the mechanisms responsible for

Results

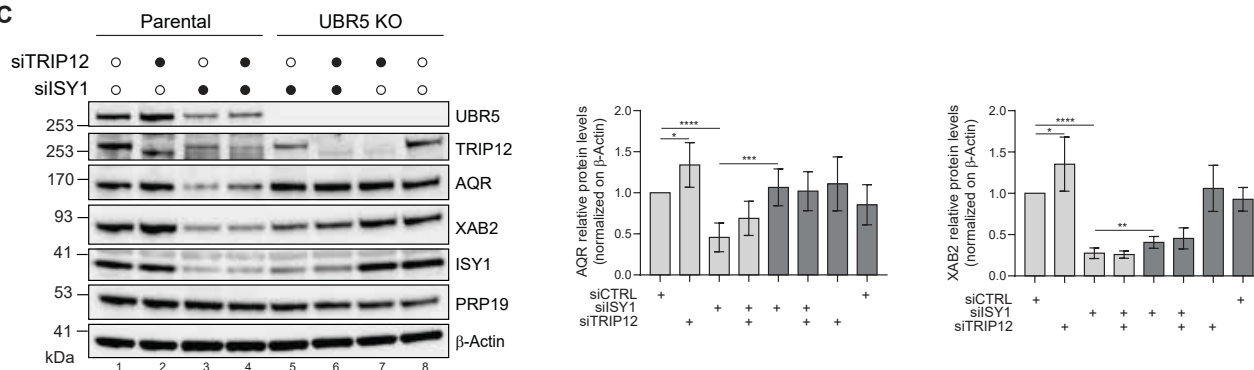
A



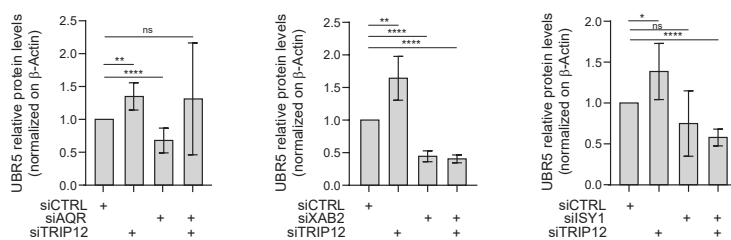
B



C



D



E

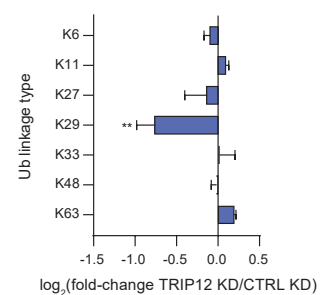


Figure 18. UBR5 knock-out opposes the loss of IBC subunits

A, B and C) Western blot analysis of IBC subunits levels in upon single or co-depletion of AQR and TRIP12 (A), XAB2 and TRIP12 (B), or ISY1 and TRIP12 (C) in WT and UBR5 KO RPE-1 cells. Quantification of AQR, XAB2 and ISY1 levels from the respective Western blot experiments (n=3) is shown on the right. Protein levels were normalized on endogenous β -Actin levels and represented as log₂ fold-change over siCTRL. Statistical significance was calculated with unpaired t-test (*P<0.05, **P<0.01, ***P<0.001, ****P<0.0001). Bar plots show mean and SD. D) Quantification of UBR5 levels from the respective Western blot experiments (n=3) is shown on the right. Protein levels were normalized on endogenous β -Actin levels and represented as log₂ fold-change over siCTRL. Statistical significance was calculated with unpaired t-test (*P<0.05, **P<0.01, ***P<0.001, ****P<0.0001). Bar plots show mean and SD. E) Quantification of ubiquitin linkage-specific peptides upon 48h depletion of TRIP12 in U2OS cells. The bar plot shows the mean and SD of log₂ SILAC ratios of ubiquitin linkage-specific peptides from three replicates.

regulating levels of the remaining subunits. We are currently working on characterizing the E3 ligase-substrate relationship involving UBR5/TRIP12 and AQR/XAB2. An outstanding question is which type ubiquitin of chains are built on these substrates: we plan to address this by using a combination of biochemical assays to gain insight into the linkage and topology of the chains. Ultimately, we hope that characterizing these mechanisms will lead to a broader understanding of the importance of IBC regulation in physiological and pathological contexts.

2.3 Cellular responses to the loss of intron-binding complex subunits

In addition to ensuring faithful splicing, a variety of cellular functions has been reported for different IBC subunits, ranging from roles in DNA repair pathways, miRNA biogenesis and R-loop regulation. However, studies that characterized IBC-depleted cells did not acknowledge that these proteins are functionally related as components of the same protein complex, and that their levels are interconnected as a result of IBC-related quality control mechanisms. In this Results section, we try to provide a more comprehensive picture of the cellular response to the loss of IBC subunits. First, we describe the global alterations at the proteome level caused by the depletion of AQR, XAB2 and ISY1, and integrate this data with transcriptomics data of AQR-depleted cells. We then focus on characterizing some of the phenotypes that emerged from this analysis: more specifically, we show that loss of IBC subunits causes irreparable DNA damage and p53 activation, which leads to cell cycle arrest and senescence. In addition, we describe the potential involvement of IBC subunits in ribosome biogenesis and stress. Lastly, we touch upon the role of IBC subunits in R-loop regulation.

2.3.1 Proteomic characterization of IBC subunits-depleted cells

To characterize the effects of IBC subunits loss on cellular processes, we employed label-free quantification (LFQ) mass spectrometry to study proteome alterations in response to 48h depletion of AQR, XAB2 or ISY1 in U2OS cells (Figure 19A). We performed the experiment in four replicates for each condition, and a Pearson correlation coefficient > 0.95 across all replicates indicated good reproducibility. Additionally, uniform manifold approximation (UMAP) clustering analysis showed that replicates for each condition clustered together and separately from other conditions (Figure 19B). After application of quality filters, 6395 proteins were quantified across all conditions. All proteins not quantified in at least three out of four replicates in at least one condition were filtered out. Missing values were imputed from a Gaussian distribution using standard settings on the Perseus software. The final dataset consisting of 4535 proteins was used in subsequent analyses. Statistically significant differences in protein abundance between conditions were determined using ANOVA test with Dunnet's multiple comparison test; an overview of the numbers of up- and down-regulated proteins in different conditions is provided in Figure 19C.

As a starting point, we assessed whether the three knock-downs elicited a similar response at the proteomic level. Overlap between significantly changing proteins upon each knock-down (fold-change ≥ 1.5 or ≤ -1.25 ; $\text{FDR} \leq 0.05$) revealed a subset of 101 up- and 121 down-regulated proteins shared between all knock-downs conditions (Figure 19D). Functional enrichment analysis revealed that shared down-regulated proteins were associated with three main biological processes: translation, ribosome

Results

biogenesis and cell cycle regulation (Figure 19E). On the other hand, the GO term “negative regulation of extrinsic apoptotic signaling pathway” was the only enriched term for shared up-regulated proteins (Figure 19F). Genes mapped to this GO term included known markers of cellular senescence, such as inhibitor of plasminogen activator SERPINE1 and anti-apoptotic factor PEA15 [182], [183], [184]. From these results, we can conclude that the proteomic landscapes of AQR-, XAB2- or ISY1-depleted cells share common features, which could hint at the existence of a core response to the loss of the subunits.

To complement this global picture, we looked for subunit-specific signatures and performed a functional enrichment analysis on proteins significantly changing in each knock-down condition (Figure 20A and B). This analysis revealed distinct responses to each knock-down, which were especially visible when looking at up-regulated proteins. XAB2 depletion, for example, lead to the highest enrichment of nuclear and chromatin-related terms for up-regulated hits across the three

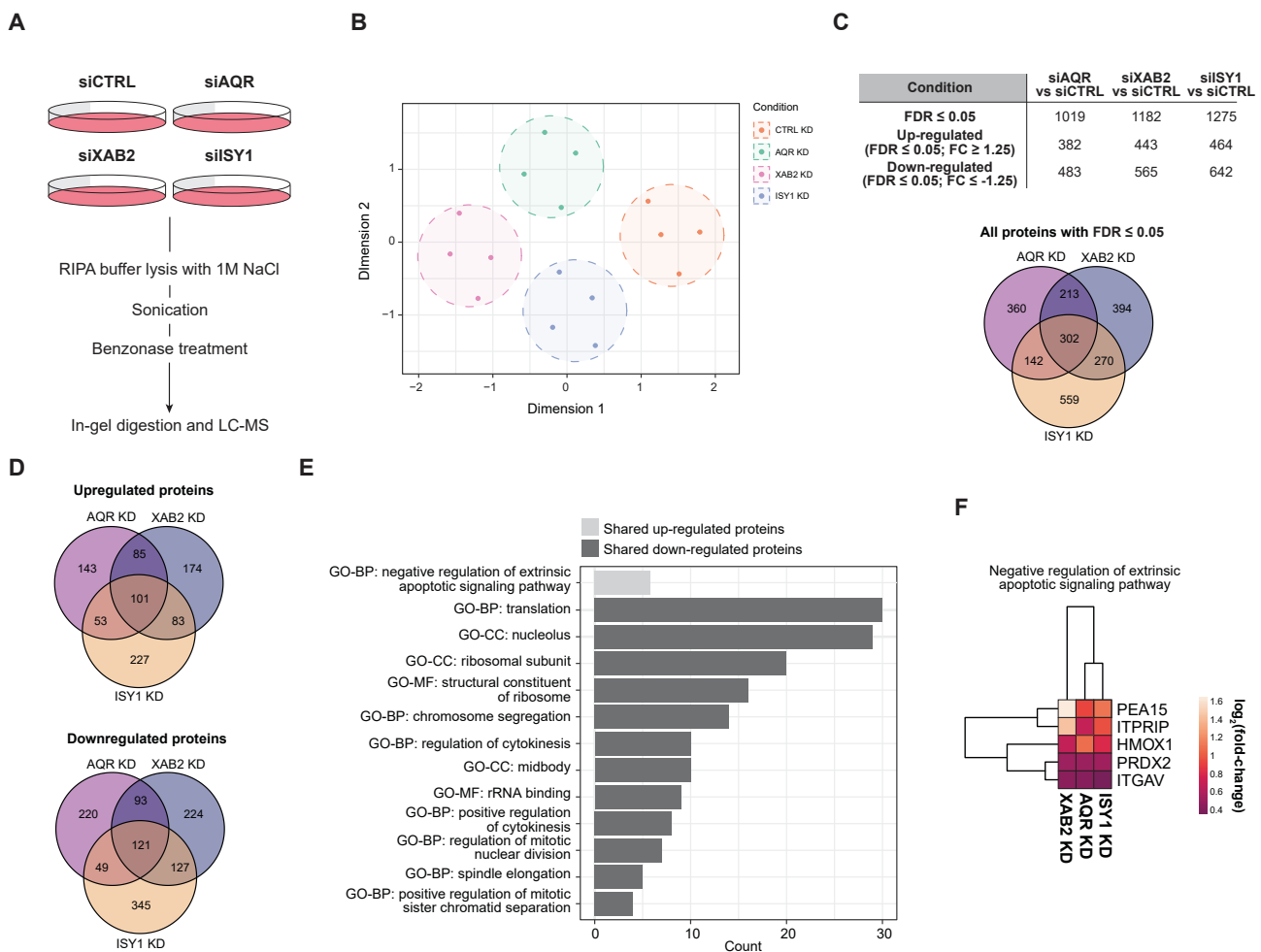


Figure 19. Proteomic characterization of IBC subunits-depleted cells

A) Workflow of label-free mass spectrometry whole-proteome analysis of AQR, XAB2- or ISY1-depleted U2OS cells. B) UMAP clustering analysis for four biological replicates. C) Table showing the number of significantly changing proteins in each condition after application of different quality filters. D) Venn diagram showing a core set of up- and down-regulated genes across all knock-down conditions (fold-change ≥ 1.25 or ≤ -1.25; FDR ≤ 0.05). E) Functional enrichment analysis performed with ClusterProfiler showing the enriched GO terms (FDR ≤ 0.05) for the common set of up- and down-regulated proteins across the three knock-downs. F) Heatmap of up-regulated proteins (FDR ≤ 0.05) shared across three knock-downs. Hierarchical clustering showing Euclidean distance was performed on both columns and rows.

knock-downs (e.g. GO terms “negative regulation of DNA-templated transcription” and “transcription regulator complex”). On the other hand, a higher percentage of the GO terms for up-regulated proteins in AQR and ISY-depleted cells associated to cytoplasmic or organellar processes (e.g. GO terms “extracellular matrix”, “secretory granule” and “regulation of protein processing”) (Figure 20A). In some cases, this signature was consistent with previously reported roles for IBC subunits. For example, depletion of ISY1 caused the up-regulation of numerous proteins involved in the response to oxidative



Figure 20. GO term analysis of differentially regulated proteins in IBC subunits-depleted cells

A and B) Functional enrichment analysis performed with ClusterProfiler showing the enriched GO terms (FDR ≤ 0.05) for up- and down-regulated proteins from full proteome analysis of AQR-, XAB2- and ISY1-depleted U2OS cells.

stress, mapping to GO terms such as “cellular detoxification” and “oxidoreductase activity”. Studies have shown that ISY1 stimulates the activity of base excision repair factor APEX1, thus preventing oxidative DNA damage [146]; we could therefore speculate that the enhanced susceptibility to oxidative stress of ISY1-depleted cells leads to up-regulation of antioxidant signaling cascades [185]. Analysis of up-regulated proteins also showed that AQR and ISY1 loss had the strongest effect on metabolic pathways. Closer examination of our datasets revealed that AQR-depleted cells up-regulated 7 out of the 10 enzymes involved in glycolysis, as well as the enzymes catalyzing the first steps of the citric acid cycle; overall, this leaves room to explore the interplay between AQR and metabolic pathways. GO terms enriched for down-regulated proteins showed a higher degree of homogeneity across the three knock-down conditions, with most terms again associating to cell cycle processes and ribosome biogenesis (Figure 20B).

To sum up, we observed extensive proteome remodeling upon depletion of single IBC subunits. Some cellular processes appeared consistently altered across the three knock-down conditions, indicating a similar response: loss of proteins involved in cell cycle and ribosome biogenesis was the key feature of this response. Additionally, up-regulation of some senescence markers was observed in all three conditions. Importantly, this global response was accompanied by subunit-specific signatures in the up-regulated proteins, which in some cases recapitulated known functions of the subunits.

2.3.2 Integration of proteome and transcriptome of AQR-depleted cells

To gain a broader view on the processes affected by the loss of IBC subunits, we collected samples for transcriptome analysis of 48h after AQR depletion in U2OS cells. We performed the experiment in five replicates, which clustered together and separately from control replicates. After filtering, we quantified 19,590 transcripts. AQR KD transcriptome and proteome data showed a good overlap, with 96% of the identified proteins represented at the transcript level (Figure 21A). Differential expression analysis revealed expression changes for 8561 genes ($FDR < 0.01$), almost equally distributed between up-regulated (4457 genes) and down-regulated (4104 genes) (Figure 21B). When introducing a cutoff of a minimal fold-change of 1.5 or -1.5, only 835 up-regulated genes passed the cutoff, against 1532 down-regulated genes (Figure 21C). Functional enrichment analysis performed on all differentially expressed genes ($FDR < 0.01$) revealed strong down-regulation of cell cycle-associated processes (GO terms “nuclear division”, “chromosome segregation” and “DNA replication”); on the other hand, up-regulated transcripts mapped to GO terms related to autophagy, ribosome biogenesis and signaling pathways, such as p53 signaling (Figure 21D).

Taking into account previous studies showing AQR’s role in ensuring splicing fidelity [96], [102], we looked whether that splicing alterations could partially explain the changes in gene and protein expression of AQR-depleted cells. Alternative splicing analysis revealed 380 significantly changing AS events in AQR-depleted cells. This number was significantly lower than what was reported in a previous study looking at AS events in HepG2 and K562 cell lines treated with a shRNA against AQR [125]. Detection of AS events from RNA-seq data can be quite challenging, due to presence of confounding variables of biological (e.g. gene expression levels) and technical origin (e.g. read length). In our case the use of short reads most likely increased the difficulty of detecting novel AS events, as short reads rarely span across one or more splice junctions. It is therefore possible that we are underestimating the effects of AQR loss on AS events. When looking at the significant AS events we detected, we found that

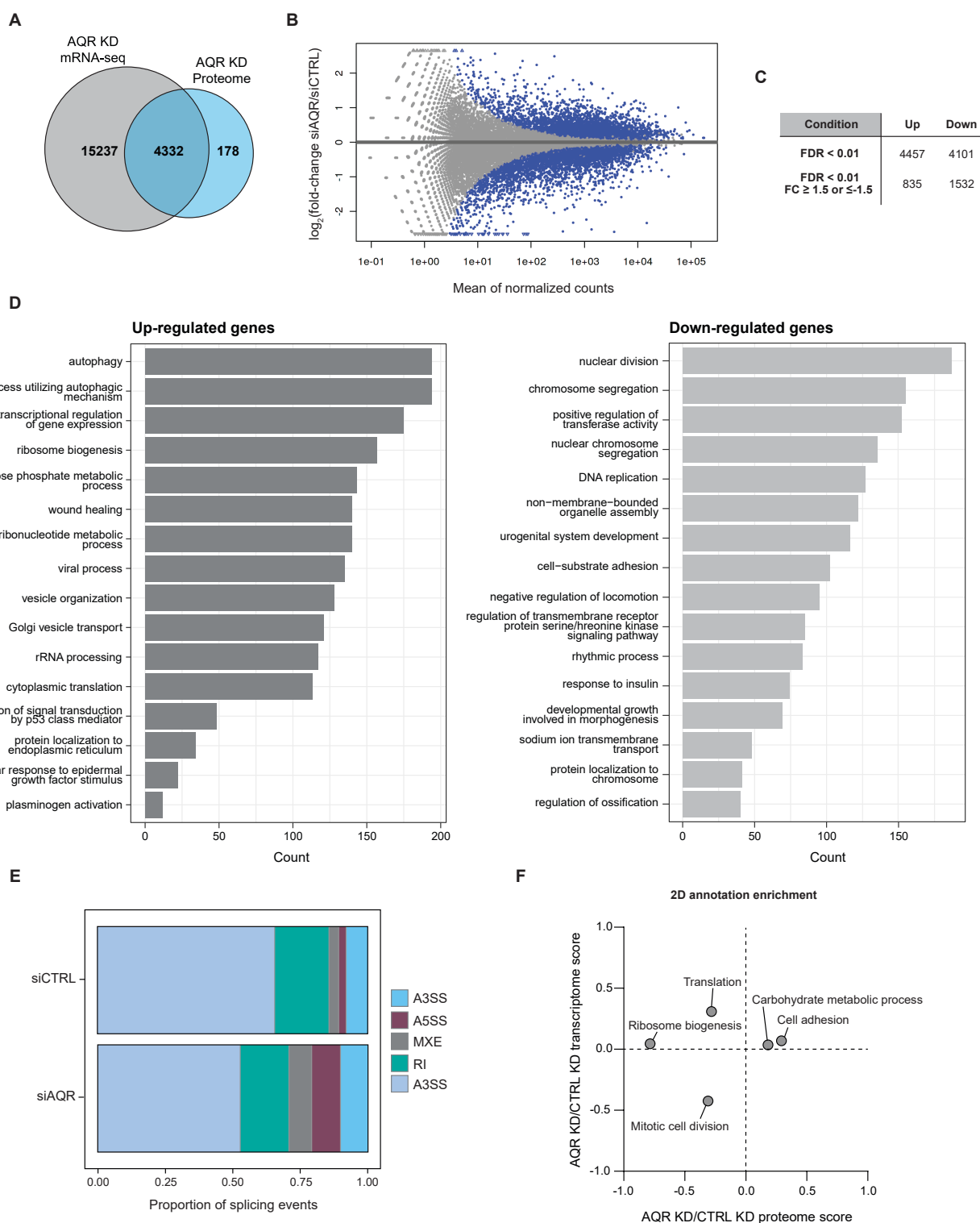


Figure 21. Integration of proteome and transcriptome of AQR-depleted cells

A) Venn diagram showing overlap between AQR KD full proteome (FDR ≤ 0.05) and mRNA-seq (FDR < 0.01). B) MA plot depicting expression fold change on log₂ scale vs. mean expression level of all annotated genes in AQR KD vs. CTRL KD cells measured by mRNA-seq. Genes with minimum fold change (FC) > 1 and adjusted-value (Benjamini-Hochberg, FDR) < 0.01 are highlighted. C) Table showing the number of significantly up- and down-regulated genes in AQR KD cells after application of different quality filters. D) Functional enrichment analysis performed with ClusterProfiler showing the enriched GO terms (FDR ≤ 0.05) for the common set of up- (left panel) and down-regulated (right panel) genes in AQR KD mRNA-seq. E) Distribution of the proportion of alternative splicing events in CTRL KD and AQR KD cells. AS events types are color-coded as indicated in the legend. F) Scatterplot of the 2D annotation enrichment of AQR KD proteome and transcriptome. Proteome and transcriptome scores are represented on the x and y axis respectively. Selected GO-BP terms are visualized.

AQR knock-down cells exhibited a higher proportion of alternative 3' splice sites (A3SS), alternative 5' splice sites (A5SS) and mutually exclusive exons (MXE) compared to control cells; on the other hand, the frequency of exon skipping (SE) and intron retention (RI) events was reduced in AQR-depleted cells (Figure 21E). Overall, 289 genes were affected by AS events: of these, 75% harbored a single AS event. The remaining 25% accumulated multiple events, with the most affected gene being ITGB1BP1 with 13 detected AS events. Further analysis of this dataset is needed to understand the significance of these AS events, and how they relate to the phenotypes we observe upon loss of AQR.

Having performed the quantitative profiling of the transcriptome and proteome of AQR-depleted cells, we went on to investigate the extent to which these two datasets correlated with each other. To achieve this, we employed a 2D annotation enrichment strategy, which allows us to identify classes of gene products showing concordant or discordant behavior between two chosen datasets [186]. In our case, we first used the results of our functional enrichment analysis to annotate genes/proteins to a subset of GO terms; we then used the annotated datasets to determine which GO terms showed a significantly deviating protein/mRNA level distribution (in other words, are processes up-regulated at the protein level also up-regulated at the mRNA level?). Finally, each GO term was assigned a transcriptome score and a protein score, which quantify its tendency to be up- or down-regulated in either dataset. Visualization of the scores for each GO term in a scatterplot revealed different correlation trends (Figure 21F): GO term “mitotic cell division” appeared in the lower left quadrant, indicating that the process is down-regulated both at the transcript and protein level. GO terms “carbohydrate metabolic process” and “cell adhesion” appeared in the upper right quadrant, showing mild up-regulation at the protein level with little to no effect at the mRNA level. The highest discrepancy between mRNA and protein levels was observed for GO terms “translation” and “ribosome biogenesis”, which localized to the upper left quadrant. For these terms, down-regulation at the protein level was matched with either no effect or mild up-regulation at the mRNA level. Overall, this approach allowed us to obtain a clear view on the processes altered by AQR loss. As previously mentioned, down-regulation of cell cycle-associated proteins is a feature of the core proteomic response to the loss of IBC subunits: here we show that this response starts at the level of transcription. Conversely, the loss of ribosomal and translation factors, another feature of the core response, is only visible at the protein level: this anti-correlating behavior between mRNA and protein suggests that the involvement of post-transcriptional regulatory mechanisms, and could reflect underlying defects in ribosome biogenesis. In the future, inclusion of a higher number of GO terms will allow us to gain a more comprehensive view on the correlation between transcriptome and proteome of AQR-depleted cells.

2.3.3 Nascent RNA sequencing sheds light on AQR-depleted cells transcriptional dynamics

mRNA sequencing of AQR-depleted cells revealed changes in the final mRNA abundance of 8561 genes, suggesting a general effect of AQR loss on transcription. However, steady-state level of any RNA in the cell is always the result of a balancing act between synthesis and degradation. We observed that loss of AQR caused a reduction in 5-Ethynyl-uridine (EU) incorporation in nascent RNA, which could be indicative of an overall decrease in RNA synthesis (Figure 22A).

To better understand the dynamic transcriptional response to AQR loss, we looked at the effect on nascent transcription specifically. To this aim, we performed TTchem-seq (Transient Transcriptome

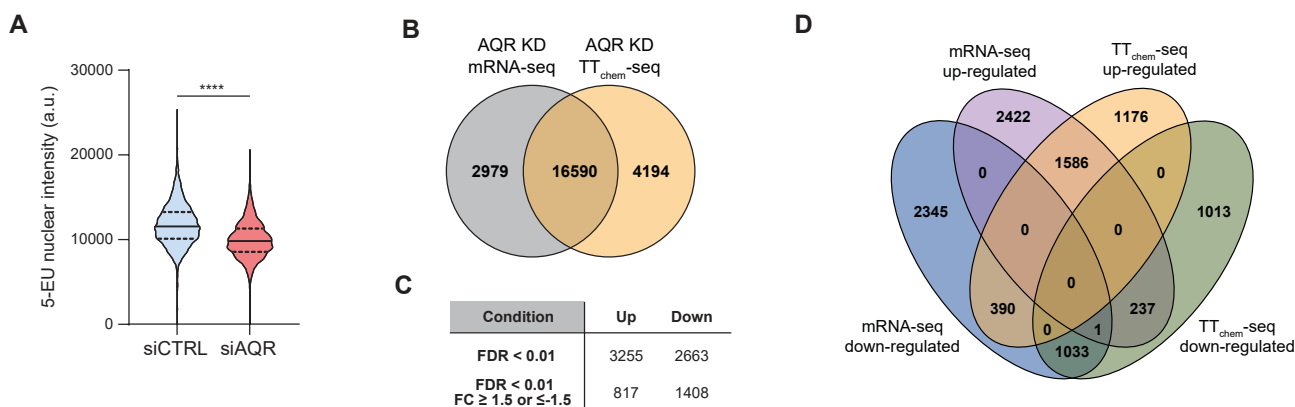


Figure 22. Nascent RNA sequencing sheds light on AQR-depleted cells transcriptional dynamics

A) Quantification of immunofluorescence analysis of 5-EU incorporation after 48h AQR knockdown in U2OS cells. Centerline shows the median and dotted lines the 25th-75th percentile. Statistical significance was calculated with unpaired t-test (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). B) Venn diagram showing overlap between AQR KD TTchem-seq and mRNA-seq ($FDR < 0.01$). C) Table showing the number of significantly up- and down-regulated nascent transcripts in AQR KD cells after application of different quality filters. D) Venn diagram showing overlap between differentially expressed genes from AQR KD TTchem-seq and mRNA-seq experiments ($FDR < 0.01$; $FC > 1$ or < -1). Only genes quantified and tested for significance in both experiments were considered.

sequencing) in unchallenged or AQR-depleted U2OS cells [187], [188]. Here, cells are metabolically labeled with the nucleoside 4-thiouridine (4-SU), which is incorporated into nascent RNAs by active polymerases. Comparison with unlabeled or total mRNA allows the identification of newly transcribed RNAs and quantification of changes in RNA synthesis. Heat shock treatment, which is known to up-regulate the transcription of specific target genes while transcriptionally repressing the majority of active genes, was used as positive control [189]. We performed the experiment in three replicates, which clustered together and separately from control replicates. After filtering, we quantified 20,782 nascent transcripts. AQR KD TTchem-seq and mRNA-seq datasets showed a good overlap, with 80% nascent transcripts represented in the transcriptome data (Figure 22B). Differential expression analysis revealed statistically significant expression changes for 5918 genes in AQR KD cells ($FDR < 0.01$), with 3255 and 2663 transcripts respectively up- and down-regulated (Figure 22C). When introducing a cutoff of a minimal fold-change of 1.5 or -1.5, only 817 up-regulated genes passed the cutoff, against 1408 down-regulated genes. Functional enrichment analysis performed on all differentially expressed genes ($FDR < 0.01$) revealed a picture similar to the one given by our mRNA-seq dataset, characterized by the up-regulation of processes involved in protein trafficking and cell adhesion, and of stress response pathways such as the MAPK and JNK cascades and p53 signaling (Appendix, Figure 31A). The similarity was more striking when considering down-regulated nascent transcripts, which mostly belonged to processes related to cell cycle regulation and DNA repair (Appendix, Figure 31B). As expected based on the similarity in their functional enrichment analysis, the overlap between differentially expressed genes in the mRNA-seq and TTchem-seq datasets was reasonably high: 1586 up- and 1033 down-regulated genes were shared between the two datasets ($FDR < 0.01$), representing roughly half and third of all differentially regulated transcripts in the TTchem-seq and mRNA-seq datasets, respectively (Figure 22D). Most of the remaining transcripts underwent significant changes in only one dataset; in addition, for a smaller subset of transcripts we observed completely opposite trends between the experiments (genes up-regulated in mRNA-seq while down-regulated in TTchem-

seq, and *vice versa*). At the moment, we are still in the preliminary stage of analyzing the TTchem-seq dataset and integrating it with other omics datasets. In the future, more in-depth analysis will allow us to better understand how the loss of AQR affects both transcription and total mRNA abundance, and ultimately to gain further insight into the cellular response to the loss of IBC subunits.

In conclusion, the integration of our proteomics and transcriptomics datasets revealed that the loss of IBC subunits affect a variety of biological processes, which could lead to distinct cellular phenotypes. In the next sections we further characterize some of these phenotypes and try to draw a comprehensive picture of the cellular response to IBC subunits loss.

2.3.4 Loss of IBC subunits alters cell cycle progression

Loss of proteins associated to cell cycle regulation was among the most consistent signatures in the functional enrichment analysis of proteomes and transcriptomes of IBC-depleted cells. To confirm this phenotype, we monitored cell cycle progression in unsynchronized U2OS cells upon depletion of AQR, XAB2 or ISY1 for up to 72h with an immunofluorescence-based approach using the Opera Phenix High Content Screening System. DNA content was estimated by staining nuclei with Hoechst 33342. Additionally, in order to estimate the percentage of actively cycling cells, as well as facilitate the discrimination between different cell cycle phases, cells were labeled with 5-Ethynyl-2'-deoxyuridine (EdU), an analogue of thymidine that is incorporated in the DNA during replication and can later be visualized with click chemistry. Information obtained from Hoechst and EdU stainings was combined using the PhenoLOGIC supervised machine learning module of the Harmony software to generate a cell cycle classifier, allowing cell cycle phase determination for each cell. With this approach we observed that the knock-down of IBC subunits led to a gradual decrease of the percentage of cells in S phase across the 72 hours time-course (Figure 23A). This appeared more prominent in XAB2 and ISY1 knock-downs, with a higher percentage of cells accumulating in G2 compared to AQR knock-down; concomitantly, XAB2 and ISY1 depletion appeared to cause overall higher lethality, resulting in higher cell loss across the time-course. When looking at the percentage of EdU-positive cells, only 18% of AQR-depleted cells showed EdU staining after 72h of depletion compared to the 66% of control cells (Figure 23B). From these results, we concluded that the loss of IBC subunits causes cell cycle arrest through yet unknown mechanisms.

Considering the strong cell cycle arrest, we decided to check whether the levels of IBC subunits are affected by cell cycle phase. Madhessian and colleagues compiled a comprehensive list of cell cycle-dependent transcripts and proteins using a combination of single-cell transcriptomic and imaging of U2OS cells [190]: notably, IBC subunits were not present in this list. To corroborate these findings, we monitored levels of AQR, XAB2, ISY1 and PPIE across different cell cycle phases using the above-described cell cycle profiling approach. Aurora kinase A (AURKA), a protein whose levels are known to peak in the G2 and M phases of the cell cycle, was used as positive control [191]. As expected, we observed that only cells with a G2 DNA content were positive for Aurora kinase A. On the contrary, staining for IBC subunits appeared homogenous throughout the cell cycle, confirming that their protein levels are not cell cycle-regulated (Figure 23C).

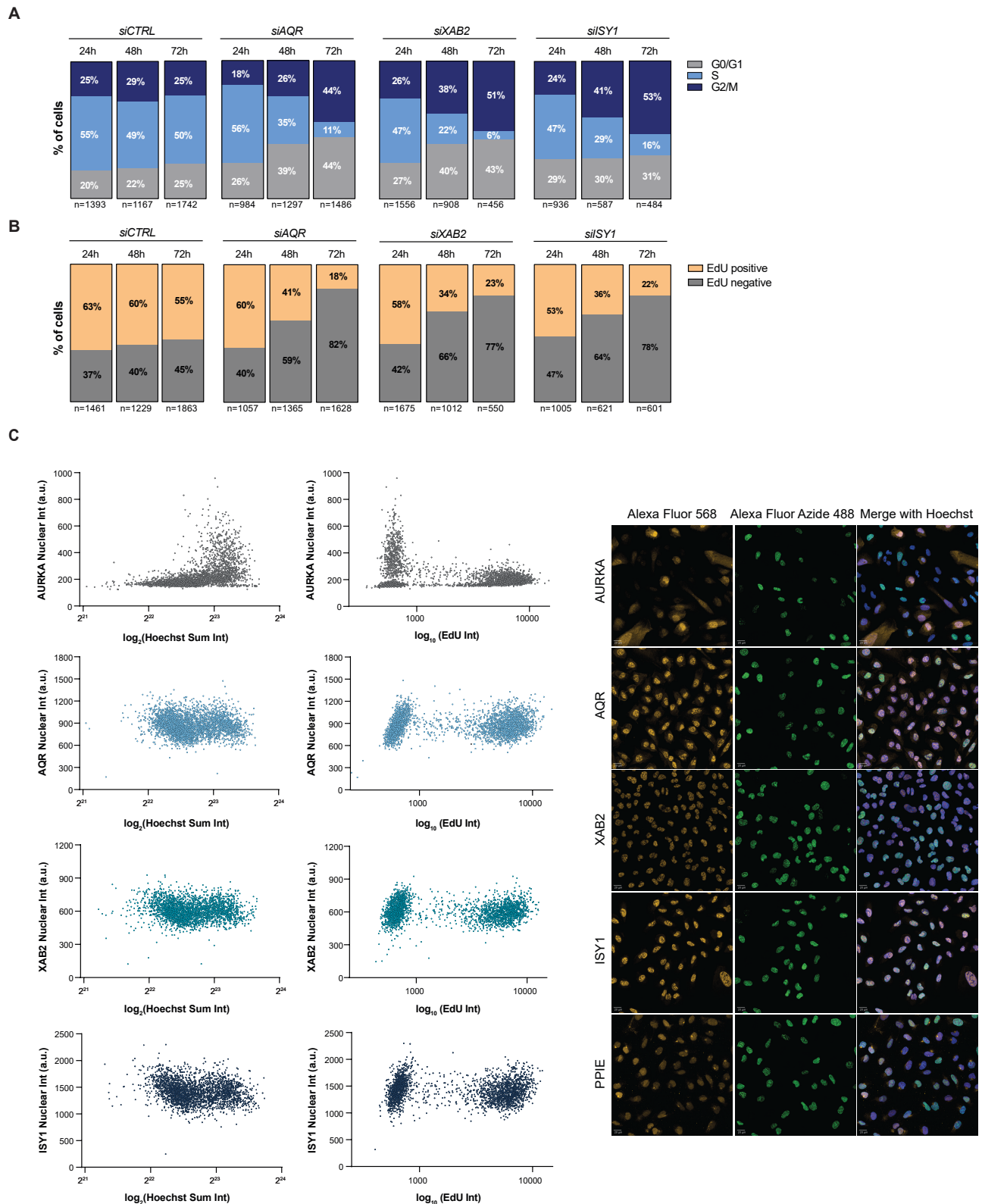


Figure 23. Loss of IBC subunits alters cell cycle progression

A) Immunofluorescence-based cell cycle profiling of U2OS cells upon AQR, XAB2 or ISY1 depletion for 24h, 48h and 72h (n=3). Numbers below bar plots indicate the number of cells profiled in each condition. B) Percentage of EdU-positive U2OS cells upon AQR, XAB2 or ISY1 depletion for 24h, 48h and 72h (n=3). Cutoff for positive cells was calculated on the 90th percentile of the intensities distribution for the siCTRL 24h condition. C) Scatterplots showing single-cell levels for a protein of interest (y axis) against Hoechst or EdU levels (x axis). Representative staining images acquired with Opera Phenix HighContent Screening System (PerkinElmer) are shown on the right.

2.3.5 Loss of IBC subunits causes p53 activation and senescence

Cell cycle arrest can occur in response to a variety of molecular stimuli and environmental insults. Prolonged cell cycle arrest can evolve into quiescence or senescence, respectively reversible and irreversible states, which are accompanied by the expression of specific markers. The strong effect on cell cycle progression observed upon loss of IBC subunits prompted us to characterize this phenotype further. Using an immunofluorescence-based approach, we found that depletion of IBC subunits from 24h up to 72h lead to increasingly higher levels of p21 (also known as CDKN1A), a well-studied mediator of growth inhibition and senescence marker, with the strongest effect observed upon loss of AQR (Figure 24A and C). Tumor suppressor p53 (also known as TP53) is the most described upstream regulator of p21; indeed, immunofluorescence analysis revealed that after 72h of depletion cells displayed higher levels of phosphorylated p53 at serine 15, a crucial phosphorylation event mediated by the DNA damage protein kinases ATM and ATR, as well as of AMPK kinase in response to metabolic stress [192], [193] (Figure 24B and D). Monitoring DNA damage markers upon depletion of IBC subunits by immunofluorescence revealed a significant increase in levels of phosphorylated histone H2AX at serine 139 (also known as γ H2AX), a marker for double-strand breaks, as well as repair factor 53BP1 after 72h depletion of AQR, XAB2 and ISY1 (Figure 24B, C, E and F). Our results echo previous studies implicating IBC subunits in the maintenance of genome stability [119], [120], [129], [130], [146].

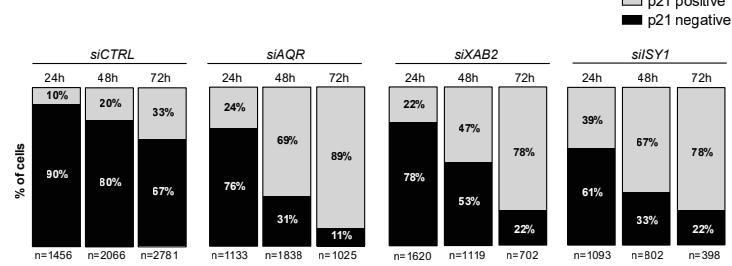
Transcription of the p21-coding locus CDKN1A can occur independently of p53 activation [194]. To control for this, we depleted AQR for 48h in wild-type, p53^{-/-} or p21^{-/-} HCT116 cells and monitored p21 activation. Consistent with our previous results, AQR depletion in wild-type HCT116 lead to a strong increase in p21 levels, as well as increased levels γ H2AX; however, the effect on p21 was almost completely abrogated in p53 knock-out HCT116 cells, suggesting that p21 activation in the context of AQR depletion is mostly p53-dependent (Figure 24G). Further strengthening these results, genes associated to the GO term “regulation of signal transduction by p53 class mediator” also resulted up-regulated in the transcriptome analysis of AQR-depleted cells (Figure 21D).

In summary, we have seen that loss of IBC subunits causes persistent genomic instability, leading to p53-mediated response and downstream activation of p21, resulting in cell cycle arrest. Interestingly, chronic DNA damage, activation of the p53-p21 axis and cell cycle arrest are considered hallmarks of

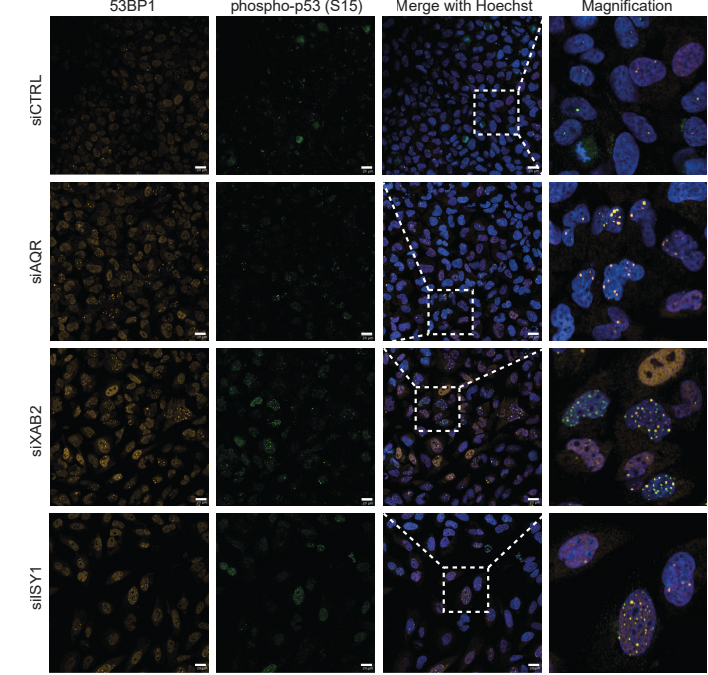
Figure 24. Loss of IBC subunits causes DNA damage-dependent p53 activation and senescence

A) Percentage of p21-positive U2OS cells upon AQR, XAB2 or ISY1 depletion for 24h, 48h and 72h. Cutoff for positive cells was calculated on the 90th percentile of the intensities distribution for the siCTRL 24h condition. Numbers below bar plots indicate the number of cells profiled in each condition. B and C) Representative immunofluorescence images of U2OS cells upon 72h depletion of the indicated proteins. Cells were fixed and stained with the indicated primary antibodies, followed by incubation with secondary antibodies Alexa Fluor 488 or 568. Nuclei were stained with Hoechst 33342. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). Scale bar is 20 μ m. D) Quantification of immunofluorescence analysis of phospho-p53(Ser15) foci after 72h knock-down with the indicated siRNAs in U2OS cells (n=3). Cells with ≥ 4 foci were considered positive. Cutoff for positive cells was calculated on the 90th percentile of the spots number distribution for the siCTRL 24h condition. Statistical significance was calculated with Fisher's exact test (*P<0.05, **P<0.01, ***P<0.001, ****P<0.0001). E and F) Quantification of immunofluorescence analysis of phospho-H2AX (S139) and 53BP1 after 72h knock-down with the indicated siRNAs in U2OS cells (n=3). Centerline of box plots represents the median, whiskers indicate the 10th-90th percentile and box limits the 25th-75th percentile. Dots represent outliers. Plus symbol indicates the mean. Statistical significance was calculated with one-way ANOVA with Tukey's post hoc test for multiple comparisons (*P<0.05, **P<0.01, ***P<0.001, ****P<0.0001). G) Western blot analysis of p21 activation in WT HCT116, p53^{-/-} HCT116 and p21^{-/-} HCT116 upon 48h AQR depletion.

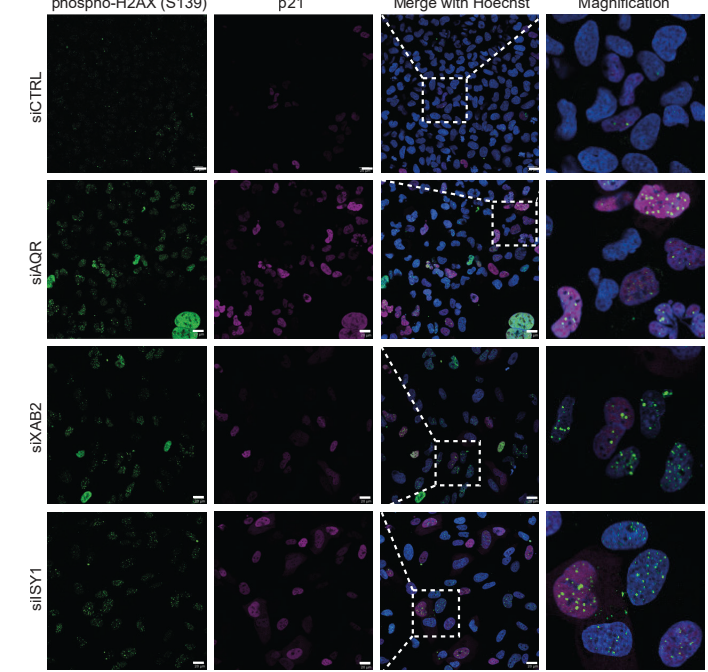
A



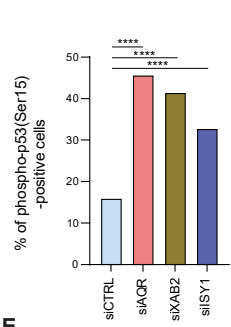
B



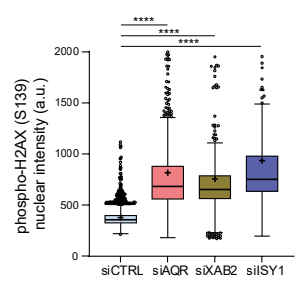
C



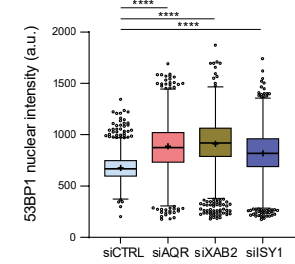
D



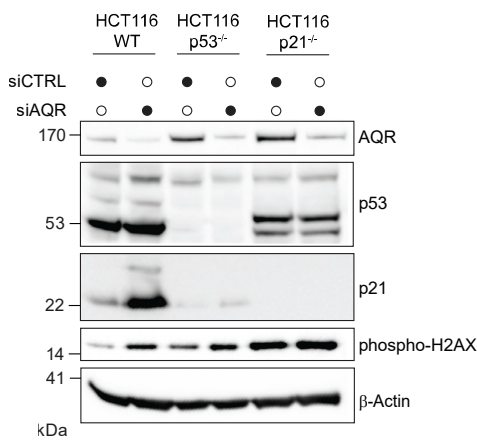
E



F



G



senescent cells [195]. Evaluation of our proteomics datasets revealed that IBC subunits-depleted cells displayed additional features of cellular senescence. For example, we observed that the up-regulation of anti-apoptotic proteins is a component of the core response to the loss of IBC subunits (Figure 19F); this could result in a general resistance to apoptosis, another hallmark of senescent cells [196]. Additionally, we observed that IBC subunits-depleted cells up-regulate various metabolic pathways, which is consistent with the metabolic alteration observed in senescent cells. Taken together, our results indicate that loss on IBC subunits is sufficient to induce senescence in U2OS cells.

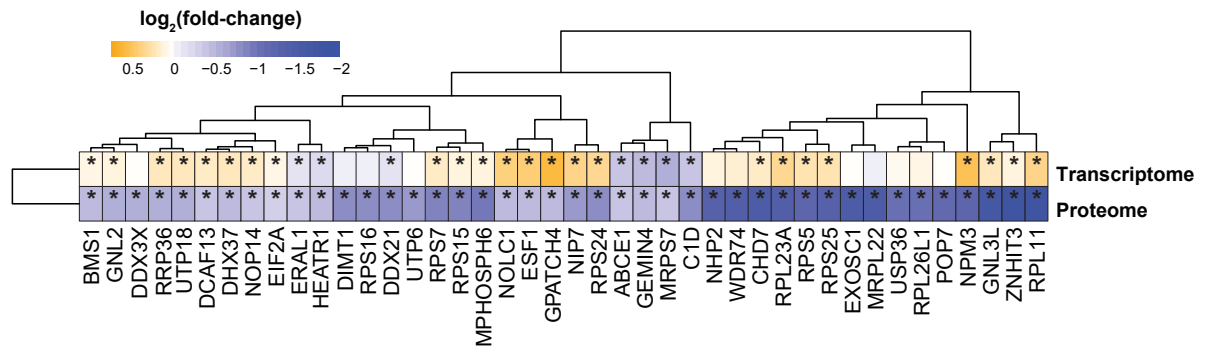
2.3.6 Loss of IBC subunits affects factors involved in ribosome biogenesis

Another feature of the core response to the loss of IBC subunits was the reduction in the levels of ribosomal proteins and translation factors. As opposed to what observed for cell cycle proteins, for which the down-regulation is part of a transcriptional response, the loss of ribosomal subunits and ribosome biogenesis factors was only seen at the protein level, which could reflect defects in the synthesis and assembly of ribosome components (Figure 25A). These results echo a recent study linking both AQR and XAB2 to the synthesis of the 40s ribosomal subunit, suggesting a role in the early steps of ribosome biogenesis [115].

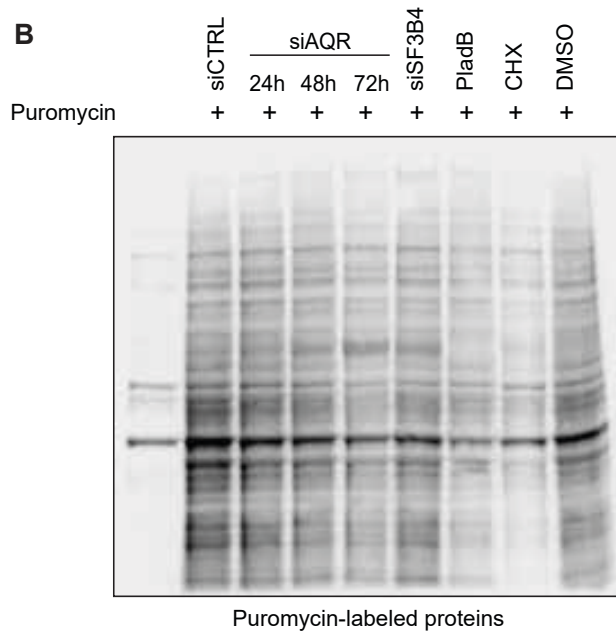
Taking this into account, we asked whether the depletion of these subunits affects global protein synthesis. For this, we monitored translation efficiency by puromycin labeling of nascent proteins in AQR-depleted U2OS cells over the course of 72 hours (Figure 25C, left panel). Translation inhibition by cyclohexamide (CHX) treatment was used as a positive control. AQR depletion resulted in a reduction of puromycin-labeled proteins, which was most evident after 72h depletion (Figure 25D). Interestingly, AQR depletion did not cause an increase in phosphorylation of the α subunit of the translation initiation factor eIF2 α at serine 51 (p-eIF2 α (S51)), a key event in the cellular stress response usually leading to shutdown of translation initiation (Figure 25C, right panel) [197]. Considering that AQR is a spliceosome-associated protein, we checked if the observed effect on translation could be the result of a more general stress caused by faulty splicing. Both 48h depletion of the SF3B complex subunit SF3B4 and splicing inhibition by Pladienolide B treatment resulted in a reduction in protein synthesis. Interestingly, and in contrast with AQR depletion, splicing inhibition caused increased levels of p-eIF2 α (S51), indicating that the observed translation attenuation could be coordinated by different and stressor-dependent upstream signaling cascades.

Importantly, ribosome biogenesis and p53 regulation are tightly linked processes. Depletion of ribosomal proteins and ribosome biogenesis factors is known to cause p53 induction and cell cycle arrest through the activation of the Impaired Ribosome Biogenesis Checkpoint (IRBC). This mechanism is mediated by the 5S RNP, an assembly intermediate of the large ribosomal subunit composed of 5S rRNA and ribosomal proteins RPL11 and RPL5, which sequesters the E3 ubiquitin ligase MDM2 leading to p53 stabilization [198], [199], [200], [201]. Surprisingly, RPL11 was the most down-regulated protein across all the ribosomal subunits found in our AQR KD proteome, which would speak against the activation of the IRBC (Figure 25A). However, as ribosomal stress can lead to p53 activation through IRBC-independent mechanisms, we cannot exclude that something similar occurs in AQR-depleted cells. Shifting our perspective, p53 activation itself could be the root cause of ribosomal stress [202]; for example, it was shown that DNA damage inhibits the trafficking of ribosomal proteins to the

A



B



C

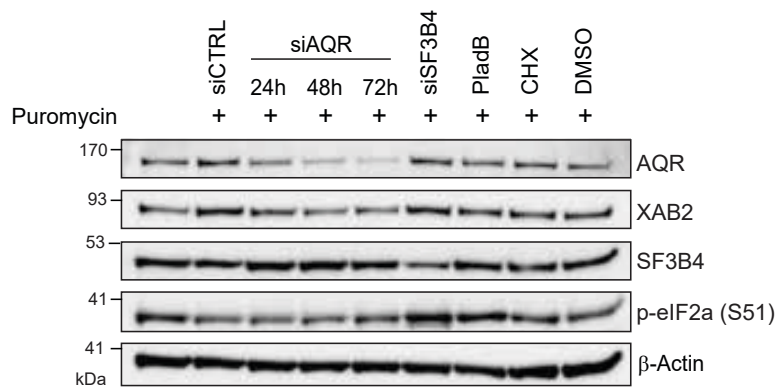
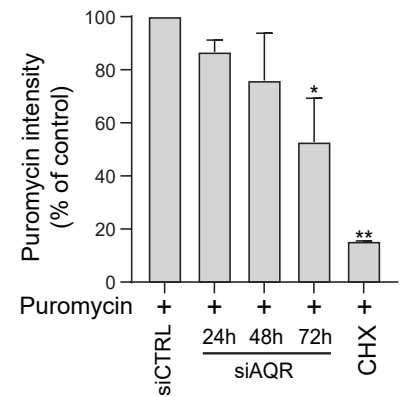


Figure 25. Loss of AQR affects factors involved in ribosomal biogenesis

A) Heatmap showing levels of transcripts and proteins involved in ribosomal biogenesis (from term GO:0042254 “ribosome biogenesis”) from transcriptome and full proteome analysis of AQR-depleted cells. Color bar indicates the log₂ fold-change of AQR KD over control condition. Asterisk indicates FDR ≤ 0.05. Hierarchical clustering showing Euclidean distance was performed on both columns and rows. B) Western blot analysis of puromycin-labeled nascent proteins upon AQR depletion (24h, 48h, 72h), SF3B4 depletion (48h), and treatment with 100μg/mL CHX (3h) or 1μM Padienolide B (3h). Phosphorylation of eIF2α was monitored. C) Quantification of puromycin intensity from the respective Western blot experiments (n=2). Puromycin signal was normalized on endogenous β-Actin levels and represented as fold-change over control condition. Statistical significance was calculated with one-way ANOVA with Tukey’s post hoc test for multiple comparisons (*P<0.05, **P<0.01, ***P<0.001, ****P<0.0001).

nucleolus in a p53-dependent way, which could lead to the degradation of unassembled subunits [203]. Overall, understanding the interplay between ribosome biogenesis and p53 activation in IBC-depleted cells requires additional work, and provides an exciting opportunity for future research.

2.3.7 Loss of IBC subunits causes irreparable genomic instability

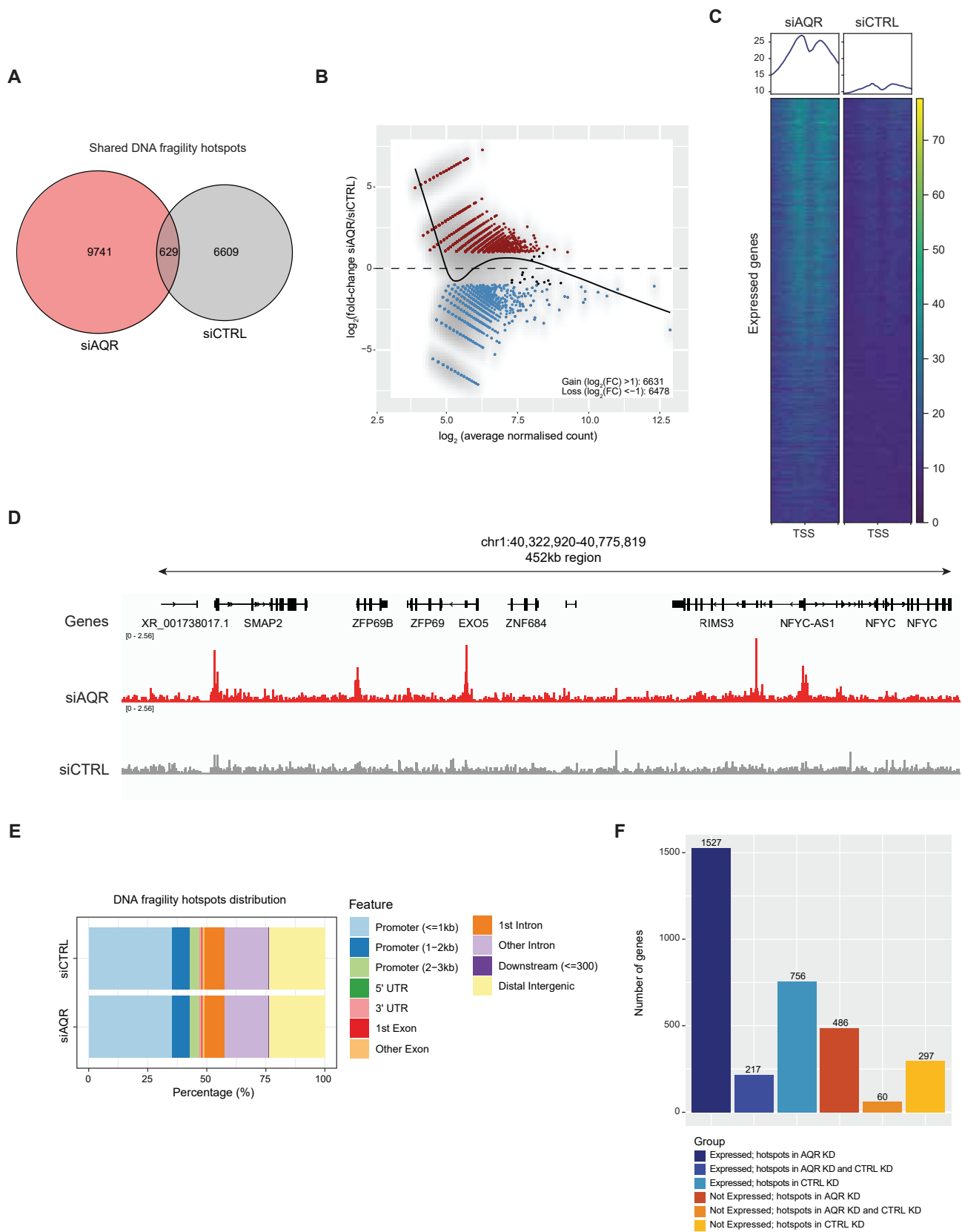
As previously shown, depletion of IBC subunits caused increased genomic instability. To gain further insight in this phenotype, we went on to map double-strand breaks (from now on referred to as DSBs) in unchallenged or AQR-depleted U2OS cells using sBLISS (in-suspension Break Labeling In Situ and Sequencing) [204]. This method allows the profiling of recurrent DSBs at nucleotide resolution and genome-wide, thus providing insight into the DSB landscape.

The experiment was performed in three technical replicates; the DlvA (DSB Inducible via AsiSI) cell line, in which AsiSI-mediated DSBs at known genomic locations can be induced by 4-hydroxytamoxifen (4-OHT), was used as a positive control [205]. We identified 10,370 and 7,238 DSBs hotspots in AQR knock-down and CTRL knock-down, respectively: of these, 629 hotspots were identified in both conditions (Figure 26A). We detected a similar amount of DSB gains and losses (6631 and 6478 respectively; $\log_2$ FC > 1 or < -1) in AQR-depleted cells (Figure 26B). Around 50% of the identified hotspots mapped to promoter regions (Figure 26C, D and E), in line with previous studies [121], [206]. Notably, loss of AQR did not lead to any change in hotspot distribution across different genomic regions (Figure 26E). sBLISS signal mainly accumulated around the transcription start site (TSS), which is expected as transcription itself sensitizes genomic DNA to different sources of DNA damage (Figure 26C) [207]. In line with this, we found that the propensity of a gene to accumulate DSBs correlated well with its expression level, obtained from our AQR KD mRNA-seq analysis: expressed genes from our mRNA-seq dataset generally accumulated more DSBs than non-expressed genes and, similarly, highly expressed genes accumulated more DSBs than lowly expressed genes (Figure 26C and F). Notably, 486 non-expressed genes (defined as the genes whose expression is lower than the mean expression level across the whole dataset) still accumulated DSBs in an AQR-dependent manner. This could suggest that gene expression is not the only determinant of genome instability in AQR-depleted cells (Figure 26F).

To gain further insight into the relationship between transcription and DNA damage in AQR-depleted cells, we analyzed the overlap between differentially expressed genes from our TTchem-seq dataset and genes displaying increased DSB formation in AQR-depleted cells. We found that the majority of AQR-dependent DSBs occurred in genes whose levels were unaffected by AQR loss (Figure 27A).

Figure 26. Loss of AQR causes irreparable genomic instability

A) Venn diagram showing the overlap of blacklist-filtered peaks based on merged replicates mapped by sBLISS in AQR and CTRL knockdown U2OS cells. B) MA plot depicting $\log_2$ average normalized counts against the $\log_2$ fold-change for the peaks detected in AQR and CTRL knockdown cells measured by sBLISS. Fold change (FC) > 2 was taken to determine genomic regions with DSB gains and DSB loss. C) Heatmaps showing the distribution of the median sBLISS signal of all expressed genes after 48 h AQR or CTRL knock-down in U2OS cells. sBLISS signal was sorted based on average gene expression level obtained from AQR KD mRNA-seq dataset from high to low. The region -2kb/+ 2kb around the TSS is displayed. Metagene profiles are shown above the associated heatmaps. D) Representative snapshot of a genomic region showing DNA fragility profiled by sBLISS in CTRL and AQR knockdown U2OS cells. E) Distribution of DNA fragility hotspots mapped by sBLISS in CTRL and AQR knockdown U2OS cells across genomic features. Features are color-coded as indicated in the legend. F) Bar plot displaying the overlap between hotspot-containing genes in CTRL and AQR knockdown U2OS identified by sBLISS, and expressed genes in U2OS cells identified by mRNA-seq. Not expressed genes were defined as the genes whose expression is lower than the mean expression level across the whole dataset.



Results

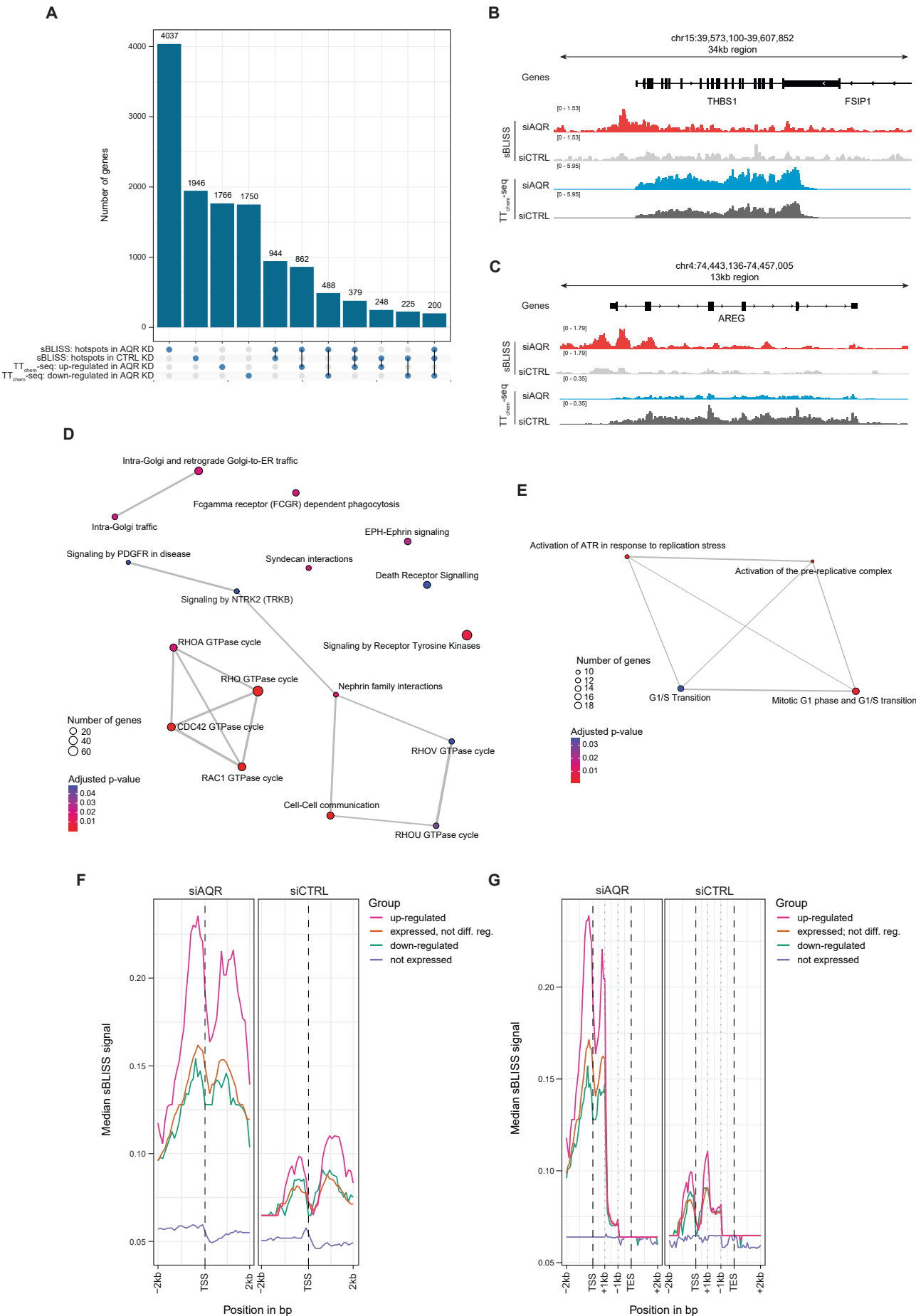


Figure 27. AQR-dependent transcriptional changes drive genomic instability

A) Upset plot displaying the intersections between hotspot-containing genes in CTRL and AQR knockdown U2OS identified by sBLISS, and differentially expressed nascent transcripts in AQR-depleted U2OS cells measured by TTchem-seq. B and C) Representative snapshots of genomic regions highlighting the presence of DSBs in up-regulated (B) or down-regulated (C) nascent transcripts in AQR-depleted cells. D and E) Reactome pathway analysis of genes showing increased DNA fragility in AQR KD cells and that appear up (D) or down-regulated (E) in AQR KD TTchem-seq. F and G) Metagene profiles of sBLISS signal after 48 h AQR or CTRL knock-down in U2OS cells. The region ranging from -2kb/+2kb around the transcription start site (TSS) (F), or ranging from -2kb before the TSS to +2kb after the transcription end site (TES) (G), is shown. Genes are grouped according to their transcriptional status obtained from AQR KD TTchem-seq: up-regulated in AQR KD (pink), down-regulated in AQR KD (green), expressed but not differentially regulated in AQR KD (orange), not expressed (violet). The profiles display the mean of signal within a window (100bp) and the median for the genes in each group.

Additionally, 862 up-regulated and 488 down-regulated genes accumulated DSBs in an AQR-dependent manner (Figure 27A, B and C). Functional enrichment analysis revealed that the set of 862 up-regulated genes mainly associated to signal transduction pathways typically up-regulated in senescent cells, such as signaling by Rho GTPases, Receptor Tyrosine Kinases and the MAPK pathway (Figure 27D) [208], [209]. As the increased sensitivity of highly transcribed genes to DNA damage is a hallmark of transcription-associated genomic instability, we can conclude that these genes accumulate DSBs as the result of AQR-dependent transcription-associated genomic instability [207], [210]. Down-regulated genes bearing AQR-dependent DSBs associated with cell cycle regulation and response to replication stress (Figure 27E).

The reason why down-regulated genes still accumulate DSBs upon AQR loss is more elusive, as we would intuitively expect less transcribed genes to be less prone to DNA damage. One hypothesis could be that, despite the down-regulation observed upon loss of AQR, these genes retain a high enough basal expression to drive genomic instability. On this note, metagene profiles of sBLISS signal for up-regulated and down-regulated genes revealed a very similar picture in terms of the regions affected by DSBs formation: in both cases, most of the signal localized around the transcription start site (TSS), and decreased along the gene body until the transcription end site (TES) (Figure 27F and G).

To sum up, we could show that loss of AQR leads to persistent genomic instability. In some cases, this is due to AQR-dependent transcription-associated genomic instability. However, is likely that other factors contribute to the genomic instability phenotype, which seems too widespread to be depending only on transcriptional changes.

2.3.8 IBC subunits contribute to R-loops homeostasis

Previous studies have linked AQR and XAB2 to the processing and resolution of R-loops, co-transcriptionally-formed structures composed of an RNA:DNA hybrid and displaced single strand of DNA [116], [117]; specifically, loss of AQR and XAB2 was shown to cause the accumulation of these structures, resulting in R-loop-dependent genomic instability [119], [141]. To the present day, the molecular mechanisms behind the role of IBC subunits in R-loop regulation are unknown. It is also unclear to what extent these proteins carry out their function as R-loop regulators alone or as part of a complex.

To shed light on the contribution of different IBC subunits to R-loop homeostasis, we depleted AQR, XAB2 or ISY1 for 24h, 48h or 72h in U2OS cells and monitored R-loop levels using an immunofluorescence-based approach. Known R-loop regulators DDX41, SETX and RNaseH2A were

Results

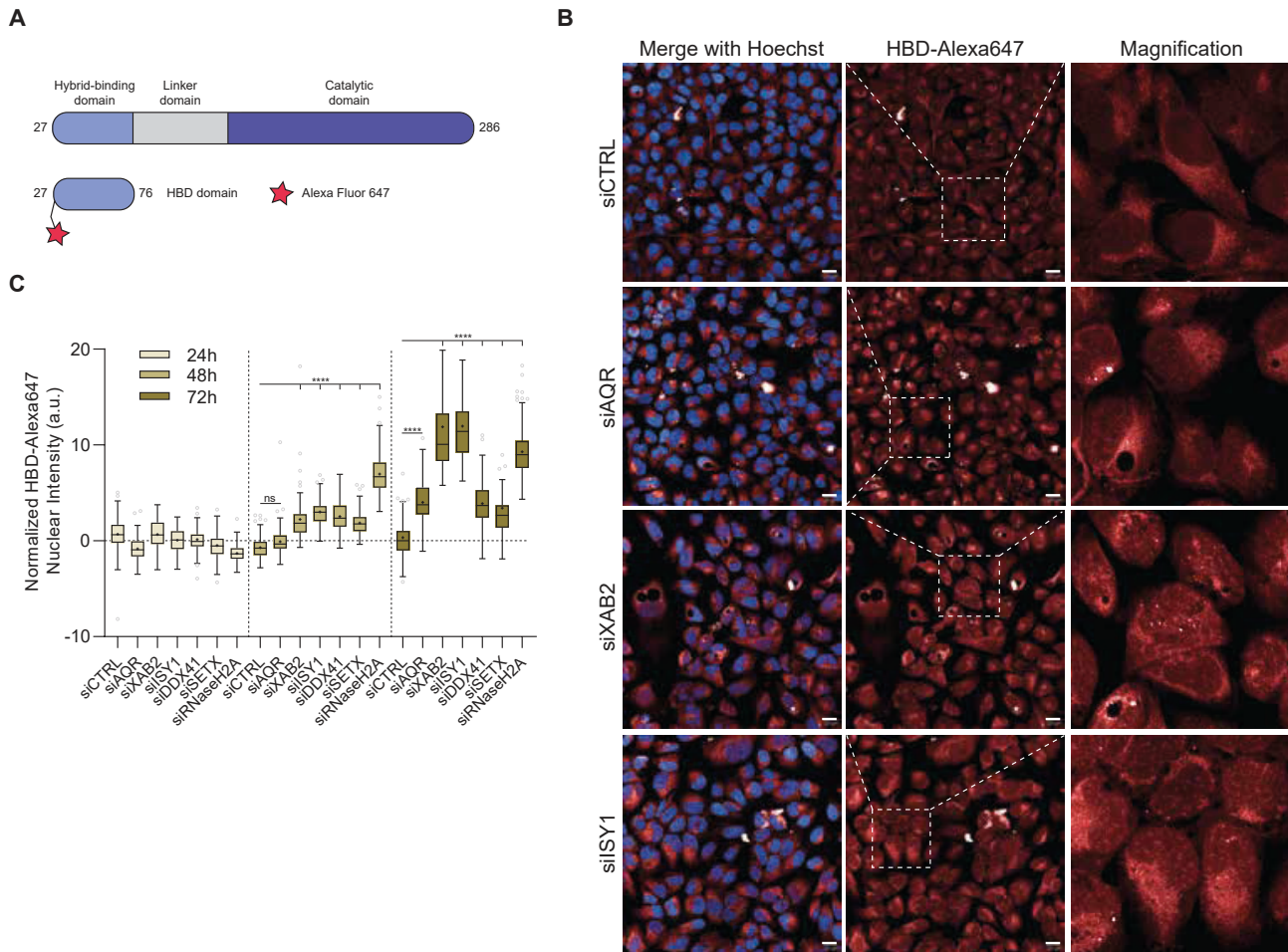


Figure 28. IBC subunits contribute to R-loops homeostasis

A) Schematic representation of the hybrid-binding domain of RNaseH1 (HBD) coupled to Alexa Fluor 647 dye. B) Representative immunofluorescence images of U2OS cells upon 72h depletion of the indicated proteins stained with the HBD-Alexa647 probe. Nuclei were stained with Hoechst 33342. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). Scale bar is 20µm. C) Quantification of HBD-Alexa647 signal after 24h, 48h or 72h knock-down with the indicated siRNAs in U2OS cells. Centerline of box plots represents the median, whiskers indicate the 10th-90th percentile and box limits the 25th-75th percentile. Dots represent outliers. Plus symbol indicates the mean. Statistical significance was calculated with one-way ANOVA with Tukey's post hoc test for multiple comparisons (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$).

included as positive controls [121], [124], [211], [212], [213]. To visualize and quantify R-loops, we employed an RNaseH1-based approach relying on the intrinsic ability of the hybrid-binding domain of RNaseH1 to bind to RNA:DNA hybrids. Briefly, the HBD domain of the nuclear isoform of RNaseH1 (RNaseH1 M27; aa 27-76) was purified and N-terminally coupled to an Alexa Fluor 647 dye, allowing for use in direct immunofluorescence applications (Figure 28A). The development of this tool by the Beli lab, in collaboration with IMB's Protein Production Core Facility, has been a multi-year effort, and this project is currently led by Christian Blum. The details of the development, testing and optimization of this tool are beyond the scope of this thesis and will not be discussed. After inter-plate normalization, we observed a trend of consistent increase in RNA:DNA hybrids levels upon prolonged depletion of IBC subunits (Figure 28B). Interestingly, AQR depletion had the weakest effect among the three IBC subunits included in this analysis, while the effect of XAB2 and ISY1 loss was comparable to RNaseH2A depletion at the 72h time-point (Figure 28C).

At this point, we envision two possible scenarios: as AQR is the only IBC subunit that, in virtue of its helicase activity, could have the ability to unwind R-loops, it is possible that the regulation of R-loop levels by the IBC is mediated by AQR as an R-loop resolving enzyme. In this case, the increase in R-loop levels observed in IBC-depleted cells could be the consequence of inefficient removal of these structures. Alternatively, the IBC could regulate R-loop levels indirectly, by modulating transcription or splicing; alterations in any of these processes could lead to increased R-loop levels. Importantly, R-loop accumulation is a known source of genomic instability: this leaves us with the intriguing prospect that these structures could contribute to the genomic instability phenotype observed in AQR-depleted cells.

3. Discussion

In this thesis, we sought to gain novel insight into the biology of the poorly characterized intron-binding complex (IBC). To this aim, two questions were addressed: on one hand, we investigated the upstream molecular mechanisms regulating the stability of intron-binding complex subunits. Secondly, we looked downstream of the complex formation and studied how cells respond to the chronic absence of IBC subunits.

3.1 Regulation of the stability of intron-binding complex subunits

3.1.1 Expression and stability of IBC subunits

In recent years, increasing evidence has highlighted how the total abundance of protein complex subunits in the cell is regulated at the post-transcriptional and post-translational levels, rather than through modulation of gene expression [49], [50], [52], [214]. As a result, proteins belonging to the same complex often have similar expression patterns, which can differ from the ones expected based on mRNA expression or gene copy-number [95]. We decided to test if this held true for the IBC by studying the co-expression patterns of its subunits in cancer cells, drawing on the published datasets from the Cancer Cell Line Encyclopedia (CCLE) as a source for protein and mRNA levels [160], [161]. Calculation and visualization of pairwise Pearson correlation coefficients for each possible pair of subunits revealed that specific subunits correlated very positively with each other across all the cancer lineages examined in the CCLE, while for other pairs the correlation varied greatly from positive to negative values. This pattern was not visible at the mRNA level; additionally, protein levels of the five subunits showed no correlation with the respective gene copy-number (DepMap, data not shown), which could have been a contributing factor in light of the deep genomic rearrangements often occurring in cancer cells. Based on the differences in correlation of protein levels, the five IBC subunits clustered into two subgroups: AQR and XAB2 belonged to one subgroup as they consistently correlated well with each other, as did ISY1, PPIE and ZNF830 in the second subgroup. The assembly of many heteromeric protein complexes was shown to occur in an ordered fashion and through multiple parallel pathways; as a result, this can lead to the formation of temporary sub-complexes [1], [215]. Considering that very little is known about the IBC assembly dynamics *in vivo*, it is intriguing to speculate that these subgroups of subunits could represent biologically distinct sub-complexes, which would then assemble in the final IBC in a stepwise manner. We can hypothesize that these sub-complexes are the target of different quality control mechanisms. Interestingly, we observed that one of the putative regulators of IBC stability emerging from this study, the E3 ubiquitin ligase TRIP12, physically interacted with AQR and XAB2 but not with the other members of the complex, which could support the hypothesis of different regulatory mechanisms targeting specific subunits.

Co-expression patterns are not exclusively observed at the steady-state. Reduction of the levels of a subunit of a protein complex, for example as a result of hemizygous deletion of its encoding gene, can lead to the widespread loss of the other subunits of the same complex, in a phenomenon known as collateral loss [95], [216]. We found that the single depletion of the subunit AQR indeed caused collateral loss of subunits XAB2 and ISY1. A previous study has observed a similar effect caused by depletion of XAB2 [133]; however, no hypothesis was formulated to explain this phenomenon. We found

that this observation held true independently of which subunit we depleted and across different cell lines, both tumorous and non-transformed, thereby pointing to a conserved regulatory mechanism. We found that the decrease in XAB2 and ISY1 levels in AQR-depleted cells was not matched by a decrease at the transcript level, suggesting the existence of regulatory mechanisms acting post-translationally (data not shown). Until now, the IBC has mostly been studied in relation to its role as a component of the spliceosome [102]. Recent studies nominated this complex as a pre-spliceosome building block, which assembles before its integration in the spliceosome [96]. Interestingly, loss of IBC subunits did not affect protein levels of spliceosomal factors with which the complex physically interacts, such as the E3 ubiquitin ligase PRP19. We therefore speculated that the observed loss of IBC subunits is the result of quality control mechanisms acting on the complex before it becomes a part of the spliceosome.

The post-translational control of protein complexes requires the involvement of protein degradation pathways. Considering the compensatory loss of IBC subunits observed when actively depleting AQR, XAB2 or ISY1, we hypothesized that the inhibition of proteasomal degradation or autophagy would be sufficient to reverse this phenotype. Surprisingly, we found that this was not the case. None of the conditions we tested (different inhibitors, different treatment lengths, different IBC subunits knock-downs) appeared to significantly affect the levels of IBC subunits in AQR-depleted cell or in control cells. The fact that the subunits steady-state levels seemed, for the most part, not responsive to inhibition of degradation pathways, prompted us to look into their degradation kinetic in unchallenged cells. Inhibition of protein synthesis for up to 8h caused only a slight decrease in the levels of IBC subunits, suggesting these are stable proteins with a slow turnover. We found these results to be in agreement with a 2016 study which, by performing pulse-chase experiments on a proteome-wide scale, highlighted the existence of exponential (ED) and non-exponential (NED) degradation kinetics [61]. Here AQR, XAB2 and PPIE were categorized as exponentially degraded proteins and their abundance was quantified to slowly decrease over the course of an 8h pulse-chase experiment, thus mirroring quite closely our results. Taking into account the slow turnover of these proteins, we cannot exclude that extending the length of the proteasome or autophagy inhibitor treatment would yield a more significant increase of their levels in unchallenged cells. However, the reason why the inhibition of protein degradation pathways is not sufficient to rescue the compensatory loss of IBC subunits in AQR-depleted cells is still unclear. Plausibly, this is due to a mixture of biological and technical reasons. In this sense, one of the main drawbacks of our approach is the use of a siRNA-mediated knock-down strategy. This has a couple of consequences: first, it complicates the overall picture, introducing new biological variables derived from the cellular response to the prolonged loss of the subunits. Second, by forcing us to work on a longer time frame, it prevents us from catching events that would occur quickly after we destabilized the system. In other words, it is possible that after 48h depletion the system has already adapted to match the availability of the less abundant subunit (as we are not working with a knock-out system, we always have residual levels of the subunit targeted by siRNA-mediated depletion), thus returning to homeostasis. We saw that the levels of IBC subunits in unchallenged cells were mostly unaffected by the inhibition of protein degradation pathways: if homeostasis has been restored after 48h of depletion, we could expect to observe such a refractory behavior also in this situation.

Taking this into account, this study would benefit from using complementary strategies to study IBC dynamics. One straightforward way to address the time frame issue would be to switch to an acute depletion strategy. Degron systems or dTAG systems have been successfully employed to study the

cellular response to acute depletion of BAF complex subunits or nucleoporins [217], [218], [219], [220]; in our case, this would enable us to monitor IBC subunits levels from the moment the degradation of one subunit is induced. From the practical point of view, it is important to keep in mind that the very nature of these systems would restrict, to an extent, the type of scientific questions we can ask: for example, experiments involving the use of proteasomal inhibitors would not be possible. Another way we could tackle this problem is by directly interfering with specific steps in the complex assembly. IBC formation *in vitro* is dependent on the presence of AQR, and more specifically of its structural N-terminal domain [96]. The inducible overexpression of an N-terminally deleted dominant negative AQR variant in cells could disrupt complex formation, and therefore serve as an additional tool to study IBC dynamics.

The above-mentioned strategies will prove helpful in better understanding the role of protein degradation pathways in IBC homeostasis. In addition to this, we decided to tackle this scientific question from a different perspective by investigating proteins that could be involved in regulating the levels of IBC subunits.

3.1.2 The global interaction network of AQR

With the goal of identifying potential IBC regulators, we employed mass spectrometry-based approaches to define the interactome and the proximal proteome of the IBC subunit AQR. The two proteomic strategies we used complement each other: affinity purification mass spectrometry (AP-MS) allows the study of stable protein-protein interactions (PPIs), while proximity-dependent biotinylation mass spectrometry (PDB-MS) catches transient PPIs *in vivo* [221], [222]. Importantly, the labeling radius of the enzymes catalyzing the biotinylation reaction, in our case the ascorbate peroxidase APEX2 (labeling radius of ~ 20 nm), results in the enrichment of all proteins in close proximity to the bait, without the need for a physical interaction [223], [224]. Workflows combining AP-MS and PDB-MS usually involve the use of overexpression systems [225], [226]. In our case, to avoid potentially altering the homeostasis of the IBC by overexpressing one of its subunits, we favored working with the endogenous levels of our bait: for this, we used a CRISPR-based knock-in strategy to tag the endogenous *AQR* locus with an AQR-APEX2-3×FLAG cassette in U2OS cells.

By performing AP-MS and PDB-MS from the same cell line, we successfully outlined the global interaction network of AQR. In line with its function in the spliceosome, we identified 100 and 14 known splicing factors in the AQR proximal proteome and interactome respectively. These factors included IBC subunits XAB2, ISY1 and PPIE, confirming that the presence of the APEX2-3×FLAG tag does not impair the formation of the complex. Additionally, we confirmed that depletion of AQR in the AQR-APEX2-3×FLAG cell line caused the same collateral loss of XAB2 and ISY1 observed in wild-type U2OS cells (data not shown). We found the majority of AQR-proximal proteins to be involved in mRNA processing, transcription and chromatin organization, thus giving a comprehensive picture of the transcriptional process (Figure 5G and 8B). Interestingly, ATPase core subunits of the SWI/SNF complex SMARCA2 and SMARCA4, as well as subunits ARID2, PBRM1 and SMARCC2, showed up as AQR interactors in our AP-MS dataset. A number of studies demonstrated that SMARCA2 and SMARCA4 can influence splicing outcome through the interaction with the splicing machinery and factors involved in other steps of RNA processing [227], [228], [229], [230], [231]. To our knowledge, this study is the first one reporting the interaction of the SWI/SNF complex with AQR, and potentially with the IBC.

Another biological process we found enriched across both datasets was rRNA processing and, more generally, ribosome biogenesis. This was particularly evident in the AQR AP-MS dataset, where ~20% of the identified hits associated to the GO term “ribosome biogenesis”. This finding is in line with the previously described role of AQR as a regulator of the maturation and assembly of box C/D snoRNPs, small nucleolar ribonucleoproteins that mediate post-transcriptional modifications of precursor rRNAs [111], [232]. Indeed, we found AQR to physically interact with NOP56, one of the four core box C/D snoRNP proteins, as well as with DKC1, a pseudouridine synthase and core subunit of H/ACA snoRNPs (Figure 7D). Importantly, other IBC subunits have not been described to play a role in this pathway; we could therefore speculate that a subset of AQR interactors might not be shared with the IBC, but is rather the reflection of functions specific to AQR.

3.1.3 Looking for IBC regulators

Among the pathways that stood out from the functional enrichment analysis of the AQR proximal proteome were “ubiquitin-like protein transferase activity” and “ubiquitin mediated proteolysis”. Given our goal of understanding the mechanisms regulating IBC stability, we decided to follow up more closely on this set of proteins. Our AQR PDB-MS dataset revealed 32 proteins involved in ubiquitin signaling enriched in AQR proximity, comprising of two E1 enzymes (UBA1 and UBA6), one E2 enzyme (UBE2R2), 21 E3 enzymes, and 8 deubiquitinating enzymes. 14 subunits of E3 ubiquitin ligase complexes were also identified: these included members of the cell cycle-regulating anaphase-promoting complex (APC) and Skp1-Cullin-F-box protein complex (SCF), and subunits of the DDB1-CUL4-X-box (DCX), C-terminal to LisH complex (CLTH), and of the COP signalosome (CSN). In addition to members of the ubiquitin system, enzymes involved in signaling cascades of ubiquitin-like molecules, such as SUMO and NEDD8, were present in proximity of AQR. Last but not least, the ubiquitin-binding ATPase VCP/p97 and the substrate-shuttling factor RAD23B, which select and deliver ubiquitylated substrates to the proteasome, were also identified. Recent studies have shown that VCP and RAD23B, together with ubiquitin ligase UBE3A, are necessary for the formation of nucleoplasmic proteasome condensates upon hyperosmotic stress and nutrient starvation, in which the degradation of orphan ribosomal proteins takes place [233], [234]. As we identify these three proteins in proximity of AQR, it is tempting to hypothesize that unassembled IBC subunits could also be targeted by a similar mechanism in specific stress conditions. In addition to UBE3A, HECT-type E3 ubiquitin ligases UBR5 and TRIP12 also showed up as AQR proximal proteins. Interestingly, both proteins have been shown to play a role in the regulation of protein complexes: UBR5 exerts orphan quality control on unpaired subunits of multiple transcription factors, while TRIP12 is involved in the regulation of the SWI/SNF complex stoichiometry [173], [235]. Interestingly, UBR5 was reported to interact in embryonic stem cells with H/ACA snoRNPs box proteins such as DKC1 and NOP10, which we also encountered in the AQR AP-MS [236]: this places UBR5 in the functional and physical vicinity of AQR. The known roles of TRIP12 and UBR5 in the regulation of protein complex homeostasis prompted us to further investigate their association with AQR and the IBC. TRIP12 showed up as an interactor of AQR in our AP-MS dataset; immunoprecipitation of endogenous TRIP12 complemented this data, confirming the binary interaction between AQR and TRIP12. Among other IBC subunits, only XAB2 co-immunoprecipitated with TRIP12 in addition to AQR: we therefore hypothesized that TRIP12 is involved in regulating the stability of a potential sub-complex formed by AQR and XAB2. TRIP12 and UBR5 are known to work together, for example in the context of double-strand break repair, where they control the cellular pool of RNF168 thereby regulating the spread of ubiquitylation on chromatin [179]. In addition, TRIP12 homolog in *S.cerevisiae*, Ufd4, was shown to

cooperate with the RING-type Ubr1 E3 ubiquitin ligase in the degradation of Arg/N-degron substrates [85]. As UBR5 is one of the E3 ubiquitin ligase mediating the Arg/N-degron pathway in humans, this raises the possibility that IBC subunits could be targeted to degradation by such a mechanism [178].

We reasoned that, if UBR5 plays a role in regulating the stability of the IBC, loss of UBR5 might rescue the collateral loss effect caused by depletion of one of the subunits. We tested this by depleting AQR, XAB2 or ISY1 in a RPE-1 UBR5 KO cell line. We found that UBR5 knock-out affected different subunits to a different extent, and these effects were also dependent on which subunits we depleted. In particular, three observations stood out: 1) UBR5 knock-out always opposes loss of AQR levels independently of which subunit is depleted, 2) UBR5 knock-out rescues the collateral loss of XAB2 and ISY1 caused by AQR depletion, and 3) UBR5 knock-out is not sufficient to rescue the reduction in levels of ISY1 and XAB2 observed XAB2- and ISY-depleted cells, respectively. Bringing TRIP12 into the picture adds another layer of complexity. The effects of TRIP12 depletion in wild-type cells were not as strong and reproducible as the ones caused by absence of UBR5: however, TRIP12-depleted cells in general showed higher levels of AQR and XAB2, while ISY1 levels were not affected. TRIP12 depletion in UBR5 KO cells had no additive effect, which could indicate that UBR5 has a more prominent role in regulating the levels of IBC subunits. However, we noticed that TRIP12 depletion led to an increase in UBR5 levels, which could hint towards a cross-regulation between the two E3 ubiquitin ligases.

Overall, this data substantiates a role for UBR5 and TRIP12 in regulating the protein levels of a subset of IBC subunits. However, how and when UBR5 and TRIP12 targets XAB2 and other IBC subunits for degradation is still unclear.

3.1.4 The role of TRIP12 and UBR5 in the regulation of IBC subunits

In E3 ligase-substrate relationships, absence of the E3 ligase can result in reduced ubiquitination and subsequent stabilization of the substrate. Indeed, mapping changes in ubiquitination proteome-wide by ubiquitin remnant profiling revealed that specific lysine residues on XAB2 are less ubiquitylated in UBR5 knock-out cells, thus making XAB2 a potential substrate of UBR5 (data not shown, experiment performed by Ekaterina Isaakova). However, which type of ubiquitin chains are built on XAB2 and other subunits remains an open question. In fact, while ubiquitin remnant profiling is an extremely powerful technique for identifying novel substrates of E3 ubiquitin ligases, it does not allow us to obtain information on the nature of ubiquitin chains, which ultimately dictates the fate of the substrate. E3 ubiquitin ligases show specificity in the type of ubiquitin chains they can build on substrates. UBR5, for example, is able to synthesize homotypic K48-linked chains, and heterotypic K11/K48-linked or K48/K63-linked branched chains, all of which mark substrates for proteasomal degradation [237], [238], [239], [240]. TRIP12 on the other hand mediates the formation of homotypic K29-linked chains and heterotypic K29/K48-linked branched ubiquitin chains [180], [241], [242]. Strikingly, results from our ubiquitin remnant profiling analysis of TRIP12-depleted cells showed a global decrease in K29-linked ubiquitylation, which could have profound effects on the architecture of ubiquitin chains and influence the fate of the substrates. We found that depletion of TRIP12 led to a mild increase in AQR and XAB2 levels, which could indicate that these proteins are less targeted to degradation. A recent study by Yu and colleagues showed that a synthetic antigen-binding fragment specific for K29-linked ubiquitin efficiently enriched AQR and XAB2 from protein lysates, suggesting that these proteins are indeed modified with K29-linked ubiquitin chains in cells [243]. Therefore, from our current knowledge we

can speculate that UBR5 and TRIP12 cooperate in building K29/K48-linked branched ubiquitin chains on AQR and XAB2, thus targeting them for degradation. Importantly, recent studies showed that UBR5 works as a chain-elongating E3 ligase, meaning that it is able to modify pre-ubiquitylated substrates [239], [240]. In this scenario, TRIP12 and UBR5 could act in sequence (Figure 29). In the first step, IBC subunits would be modified by TRIP12 with K29-linked ubiquitin chains: this first modification could already work as a weak degradation signal. Branching of these chains with K48 linkages by UBR5 would potentiate this signal and target the substrates to the proteasome for degradation. Notably, a similar mechanism was observed in the yeast UFD pathway: in this case, K29-linked chains built by Ufd4 (the yeast homolog of TRIP12) are preferentially recognized and branched with K48 linkages by E4 ubiquitin elongation enzyme Ufd2 [181]. Different scenarios are of course possible: for example, TRIP12 and UBR5 could be joined in this ubiquitination cascade by a third enzyme, which could prime the substrates for ubiquitin chain formation.

Determining when, in the sequence of events from subunit synthesis to complex assembly, these ubiquitination events take place will be important for understanding how the IBC is regulated. UBR5 is a known player in the orphan quality control pathway, and was found to target for degradation free subunits exposing degron signals that would normally be buried inside the assembled complex [235]. Therefore, it is possible that unassembled IBC subunits are degraded through a similar mechanism. In line with this, it would be interesting to investigate whether IBC subunits are substrates of the Arg/N-degron pathway, in which UBR5 plays a role. In addition, it is possible that these proteins are degraded during translation itself; in this process, known as co-translational degradation, nascent polypeptides can be targeted for degradation before their synthesis is completed [244], [245], [246], [247]. To our

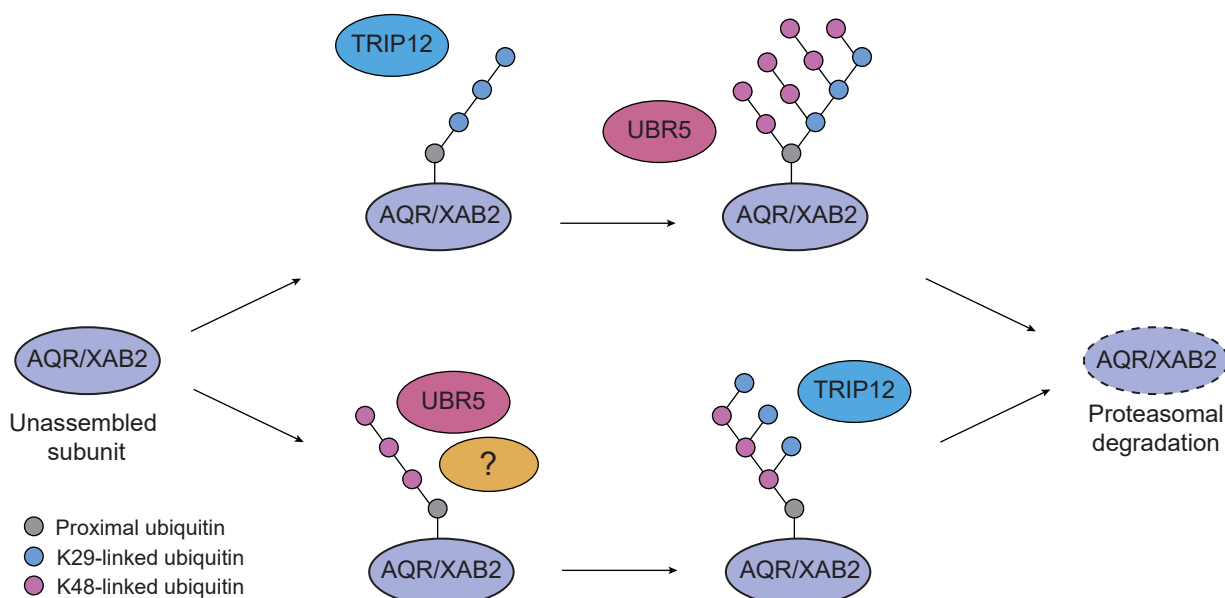


Figure 29. Possible model of IBC subunits degradation by TRIP12 and UBR5

Unassembled AQR or XAB2 subunits are first modified by TRIP12 with K29-linked ubiquitin chains (upper part); UBR5 initiates the branching of the pre-existent chains adding K48-linked ubiquitin monomers, thus targeting the substrates to degradation. Alternatively, substrates are first modified by a third unidentified E3 ubiquitin ligase (lower part); this first modification acts as a seed for UBR5, which builds K48-linked chains. TRIP12 cooperates by adding K29-linked ubiquitin monomers. The resulting K29/K48-linked branched ubiquitin chain targets substrates for degradation.

knowledge, the ubiquitylome of nascent polypeptides has been profiled by two separate studies in yeast and human cells, neither of which identified IBC subunits [248], [249]. Nevertheless, we cannot exclude that at least a fraction of these proteins is degraded before reaching its fully folded conformation in a condition when IBC formation is impaired. Lastly, it is also possible that unassembled IBC subunits misfold and form protein aggregates; however, using a denaturing lysis buffer to recover insoluble proteins did not improve the recovery of XAB2 levels in AQR-depleted cells, indicating that aggregation is potentially a less likely fate for unassembled IBC subunits.

In summary, we have outlined different ways through which free IBC subunits could be removed from the cell. Our data strongly suggest a role for UBR5 and TRIP12 in regulating the levels of IBC subunits AQR and XAB2: this supports the hypothesis that these two proteins form a sub-complex before joining the other three subunits in the fully assembled IBC. Additional work will allow us to gain further insight into how and when UBR5 and TRIP12 targets IBC subunits and other IBC subunits for degradation. Additionally, the mechanisms responsible for regulating levels of the remaining subunits remain completely unexplored. Ultimately, we hope that characterizing these mechanisms will lead to a broader understanding of the importance of IBC regulation in physiological and pathological contexts.

3.2 Cellular responses to the loss of intron-binding complex subunits

IBC subunits have been associated with various biological processes in cells, ranging from splicing, to maintenance of genomic stability, to miRNA biogenesis and cell differentiation [128], [130], [145]. Some of these functions are carried out through interactions with different binding partners: for example, ISY1 interacts with base excision repair factor APEX1 [146], and XAB2 was first characterized as binder of transcription-coupled and global genome nucleotide excision repair pathway (TC- and GG-NER) factor XPA [128]. How and in response to which stimuli IBC subunits navigate between different interaction partners is unknown. In the previous section we have touched upon different aspects of the regulation of the IBC, such as co-expression and collateral loss. So far, studies on IBC members have focused on single subunits, without taking into account the interconnection between levels of different complex members; however, a phenomenon like the compensatory loss of subunits could have wide repercussions on cellular processes. Here, we want to acknowledge that the picture is more complex, and that multiple factors need to be taken into account when studying the role of IBC subunits in biological processes.

3.2.1 “Omics” approaches to characterize the response to loss of IBC subunits

To disentangle the cellular response to the loss of IBC subunits, we used proteomics and transcriptomics to gain a broad perspective on which biological processes are affected by loss of specific members of the complex.

For the proteomic profiling of IBC-depleted cells, we chose to compare the effects of the 48h depletion of three subunits, AQR, XAB2 and ISY1 (the same subunits and conditions we used in the first part of this study). Currently, we can complement this data with transcriptomic data of 48h AQR-depleted cells; in the future, generation of mRNA-seq datasets for XAB2- and ISY1-depleted cells will allow to paint a comprehensive picture of the different cellular processes affected by the loss of IBC subunits.

The datasets at hand allowed us to draw some interesting conclusions. We observed that loss of AQR, XAB2 and ISY1 affected the proteome to a similar extent, influencing the levels of roughly a fourth of all the quantified proteins; here, ISY1 caused the most widespread effect, affecting levels of 1275 proteins.

Overlap of knock-down-specific changes identified a set of 2240 proteins responsive to the depletion of at least one IBC subunit; of these, 302 proteins were shared across the three conditions, suggesting the existence of a core response to the loss of IBC subunits. This core response was accompanied by IBC subunit-specific signatures, which were visible in our functional enrichment analysis. We first focused our attention on the core response; here, key features were the loss of cell cycle-associated proteins and the up-regulation of proteins involved in anti-apoptotic pathways, among which were known senescence markers. Integration of proteomic and transcriptomic data for AQR-depleted cells revealed that this response begins at the level of transcription, indicating the activation of specific transcriptional programs. On the other hand, reduction in the levels of ribosome biogenesis factors, another feature of the core response, was observed only at the proteome level: this suggests that the stability of these proteins in response to loss of AQR (and others IBC subunits) is regulated through post-transcriptional mechanisms.

The presence of this conserved response upon depletion of AQR, XAB2 and ISY1 prompted us to investigate if the changes at the proteome levels were mirrored in cellular phenotypes.

3.2.2 DNA damage, p53 activation and cell cycle arrest: the core response to the loss of IBC subunits

Our proteomic profiling analysis revealed the down-regulation of many cell cycle regulators. These were, for the most part, mitotic factors such as mitotic kinase Aurora kinase B (AURKB), components of the chromosomal passenger complex INCENP and Borealin (CDCA8), as well as a number of kinesin-like proteins required for mitotic spindle formation and cytokinesis (KIF18B, KIF20A, KIF23). The reduction in protein levels extended to factors regulating G1/S phase transition, such as cyclin-dependent kinase CDK6, and to replication-associated proteins, such as subunits of the GINS and MCM complex and DNA topoisomerase binding protein TOPBP1. Monitoring cell cycle progression revealed that these proteomic alterations were accompanied by a strong phenotype: prolonged depletion of IBC subunits led to a gradual reduction of actively cycling cells, as seen by a stark reduction in the percentage of EdU-incorporating cells, with concurrent accumulation in G1/G0 and G2. In addition to the reduction in proliferation, we could observe high toxicity and cell loss, particularly for XAB2 and ISY1-depleted cells.

Cell cycle regulation is orchestrated by changes in the activity, chromatin association and overall abundance of many proteins. Regulation of abundance of cell cycle factors is a multilayered process: post-translationally, the E3 ubiquitin ligases complexes anaphase-promoting complex (APC/C) and Skp1-Cul1-F-box (SCF) target a plethora of cell cycle proteins for degradation, allowing progression through cell cycle phases and regulation of cell cycle checkpoints [250]. In addition, specific transcriptional programs coordinate the levels of most proteins necessary for cell cycle progression: this regulatory network includes different families of transcription factors, such as the E2F/RB family and the DREAM complex [251]. Importantly, absence of mitogenic factors or presence of stress conditions leads to the down-regulation of cell cycle gene expression, thus halting proliferation. Stress signals typically converge on the activation of the p53–p21 pathway, which through multiple mechanisms coordinates the transcriptional

repression response [252], [253]. This mechanism was perfectly represented in our transcriptomics data of AQR-depleted cells, which showed both activation of the p53-p21 pathway and down-regulation of cell cycle-related genes. Consistent with this data, prolonged depletion of IBC subunits caused the gradual increase of p21 expression, which was strictly dependent on the presence of p53.

An extensive crosstalk exists between cell cycle regulation, p53-dependent pathways and DNA damage repair (DDR) signaling. Activation of DDR upstream kinases ATM-CHK2 and ATR-CHK1 in response to different types of genotoxic stressors induces the phosphorylation and stabilization of p53, as well as modulating cell cycle checkpoints [254], [255], [256], [257]. In line with previous reports, our data showed that loss of IBC subunits caused genomic instability, as seen by the spontaneous increase in levels of double-strand breaks (DSBs) marker γ H2AX and of DSB repair pathway regulator 53BP1.

Overall, our results show that the prolonged loss of IBC subunits causes spontaneous DNA damage and activates the p53-p21 axis, eliciting a transcriptional response that ultimately results in cell cycle arrest. Importantly, this seems to be a core response induced by the loss of all three IBC subunits.

3.2.3 Senescence is the ultimate consequence of the loss of IBC subunits

Activation of the p53-p21 signaling pathway is typically observed in senescent cells. Senescence is a multifaceted phenomenon with profound implications in physiological and pathological contexts. Senescent phenotypes are very heterogeneous, and they usually reflect differences in cell type and upstream stress stimulus: for example, pathways activated in by oncogene-induced stress differ from the ones activated in response to overproduction of reactive oxygen species, but both stimuli result in a senescent phenotype. Despite this variability, some hallmarks of cellular senescence have been defined throughout the years [195]: these include chronic DNA damage, irreversible cell cycle arrest, resistance to apoptosis, rewiring of metabolic pathways, activation of the unfolded protein response, a specific secretory phenotype, and recognizable morphological alterations. Interestingly, we found that cells depleted of IBC subunits display most of these hallmarks. As highlighted in the previous section, genome instability and cell cycle arrest are key features of the cellular response to the loss of IBC subunits. Additionally, our proteomic analysis revealed that the up-regulation of a subset of anti-apoptotic proteins is part of the core proteomic response to the loss of the subunits.

As mentioned before, senescent cells usually undergo metabolic rewiring, which includes the alteration of several pathways involved in glucose and lipid metabolism [258], [259]. In line with studies showing increased glycolysis rate in senescent cells, our proteomic analysis revealed that loss of IBC subunits leads to the up-regulation of pathways involved in carbon metabolism, such as glycolysis and the pentose phosphate pathway. Interestingly, this effect is seen strongest in AQR- and ISY1-depleted cells, suggesting underlying differences between the responses to the loss of specific subunits.

An additional hallmark of senescent cells is the ability of releasing cytokines, chemokines, growth factors and proteases in the extracellular milieu. This mixture of proteins is collectively known as senescent-associated secretory phenotype (SASP) and it mediates paracrine signaling, favoring the onset of senescence in nearby cells [260]. Despite SASP composition being heterogeneous and dependent on the stimulus that triggered senescence, proteomic studies have highlighted subsets of proteins that are shared across secretomes of different senescence conditions [184], [261], [262]. We were able to

identify several of these components amongst proteins up-regulated upon loss of IBC subunits, such as subunits of type VI collagen (COL6A1 and COL6A2), danger associated molecular pattern molecules (HMGB1) and members of the Serpin superfamily (SERPINE1) (data not shown). This suggests that, upon loss of IBC subunits, cells are able to engage in SASP production, which would in turn facilitate the spreading of senescence to neighboring cells.

Lastly, senescent cells typically experience defects in ribosome biogenesis and global reduction of protein synthesis [263]. We saw that the loss of IBC subunits led to the strong down-regulation of numerous ribosomal subunits and ribosome biogenesis factors, and that prolonged depletion of AQR caused a reduction in protein synthesis. Additionally, our transcriptomic data revealed the up-regulation of genes involved in MAPK kinase signaling pathways, which play a key role in coordinating the cellular response to a variety of stressors, including ribotoxic stress [264]. Taken together, these results suggest that ribosomal stress is another feature of IBC-depleted cells. Interestingly, ribosomal stress can both be cause and consequence of p53 activation, raising the question of which phenotype comes first when IBC subunits are lost. Considering the role of AQR in the biogenesis of snoRNPs, which regulate the maturation of rRNA, it is tempting to speculate that the defects in ribosome biogenesis come first, and are the source of p53 activation and senescence. In the future, additional work will hopefully allow us to find the answer to this chicken-and-egg question.

In conclusion, we could show that the acquisition of a multifaceted senescent phenotype is the ultimate consequence of the loss of IBC subunits. Given the complexity of this phenotype, we set out to better characterize some of its features. As DNA damage is one the driving factors of senescence, this prompted us to investigate which could be the underlying causes of the chronic genomic instability phenotype observed upon loss of IBC subunits. For simplicity, we decided to focus our efforts on the characterization of one subunit, AQR.

3.2.4 Interplay of transcription and DNA damage in AQR-depleted cells

As mentioned in the previous section, chronic DNA damage is one of the main features of senescent cells. Which mechanisms drive the genomic instability phenotype observed upon loss of IBC subunits remains an open question. Considering the involvement of the IBC in splicing and transcription-associated processes, we could speculate that genomic instability results, to an extent, from misregulation of these processes. We started tackling these questions by exploring the interplay between AQR-driven transcriptional changes and genomic instability, and looked at the genome-wide effects of AQR loss on nascent transcription and DSBs formation.

Mapping the sites of recurrent DSB formation genome-wide with sBLISS revealed that AQR depletion caused alterations in the DSB landscape of U2OS cells. In line with the observed genomic instability phenotype, loss of AQR increased the number of persistent DSBs; notably, AQR-induced DSBs only minorly overlapped with endogenous DSBs (identified in the control condition). One explanation for this observation could be that the high induction of DSBs in AQR-depleted cells leads to detecting fewer spontaneous endogenous breaks, which would instead be more easily detected in the control condition. Nevertheless, the distribution of DSBs across different genomic features was identical in AQR-depleted and control cells, indicating that the overall fragility of different genomic regions (e.g. promoter regions) is unchanged between the two conditions.

Endogenous DSBs are typically enriched around the transcription start site (TSS) of transcribed protein-coding genes [204], [206], [265], [266], [267], [268]. The relationship between transcription rate and increased propensity to accumulate mutations and recombination events has been subject of studies since the early '90s [207], [269], [270], [271]. This phenomenon, known as transcription-associated genomic instability, follows the principle by which elevated transcription rate increases the sensitivity of genomic DNA to different sources of DNA damage, resulting in the formation of DSBs [210]. Nowadays the picture appears to be more complex, with recent studies showing that other factors, such as the release of RNA polymerase II pausing, are also contributing to the formation of endogenous DSBs [272]. The correlation between transcriptional status and DNA fragility was visible in our datasets: expressed genes from our mRNA-seq dataset generally accumulated more DSBs than non-expressed genes and, similarly, highly expressed genes accumulating more DSBs than lowly expressed genes.

Up-regulation of specific transcriptional programs can be accompanied by an increase in the occurrence of endogenous DSBs; this was seen, for example, in the context of neuronal differentiation, where highly expressed genes involved in neural cell fate determination were prone to spontaneous DNA damage [268]. We therefore investigated if AQR-dependent transcriptional changes, obtained from our transient transcriptome dataset, correlated with increased DNA fragility. This overlap allowed us to define three groups of AQR-dependent DSBs. Group 1 included DSBs occurring at genes that were transcriptionally up-regulated in AQR-depleted cells. DSBs in Group 2 occurred at genes that did not respond to AQR loss: the vast majority of DSBs belonged to this group. Lastly, Group 3 included DSBs that accumulated at genes whose transcription is repressed by AQR depletion.

The presence of these three groups gives us some insight into how the loss of AQR could affect genomic stability. For Group 1 DSBs, a mechanistic explanation is quite straightforward. As these DSBs are found at genes that are more transcribed in response to AQR depletion, they could be the direct result of transcription-associated genomic instability. Importantly, we found that not all up-regulated genes accumulated AQR-dependent DSBs, indicating that transcription-associated genomic instability is likely not a general feature of the transcriptional response to AQR depletion. These DSBs are especially affecting genes involved in signaling pathways usually up-regulated in senescent cells, such as the MAPK pathway or TGF- β signaling pathway [208], [273].

As mentioned before, Group 2 includes AQR-dependent DSBs occurring at genes for which we detect no change in transcription rate. Here we see the apparent uncoupling of AQR-dependent effects on genome stability and transcription: AQR loss simply increases DSBs formation at these genes, without affecting their transcriptional status. The mechanism behind AQR-induced genomic instability at these genes is, as a result, more difficult to pinpoint. In the future, checking the extent of the correlation between DNA fragility with expression level and with factors influencing RNA Polymerase II pausing, such as gene length and bidirectional transcription, will allow us to gain more insight on the origin of these DSBs.

As it is somewhat counterintuitive that less transcribed genes would accumulate more DNA damage, Group 3 DSBs are more difficult to explain. Here we envision two possible explanations: on one hand, it is possible that genes, despite being down-regulated, retain a level of expression high enough to drive genomic instability. Alternatively, it is possible that the persistence of these DSBs is not an indication of higher incidence, but rather of less efficient repair. Active genes are known to be repaired faster than

inactive genes through transcription-coupled double-strand break repair (TC-DSBR) [274], [275], [276]. In this context, transcriptional repression caused by AQR depletion could reduce DNA repair efficiency of spontaneous DSBs, which would therefore be more persistent.

Finally, we want to acknowledge that, despite its complexity, the picture we have so far presented is still a simplified one, and many additional factors could be contributing to DSBs formation. Some factors could still be linked to transcriptional processes, such as the formation of unscheduled R-loops; however, DSBs might arise for reasons completely independent from transcription, such as other endogenous DNA damage sources.

3.2.5 The IBC in R-loop regulation

In recent years, R-loops have emerged as major players in transcription-associated DNA damage [277]. The relationship between R-loops and transcription is multifaceted. First of all, transcription provides the perfect conditions for R-loop formation: the negative supercoiling of the DNA stretch behind the elongating RNA polymerase favors the re-invasion and annealing of the nascent transcript to the complementary DNA strand, leading to the formation of an RNA:DNA hybrid and displacing the single strand of DNA. As a result, transcription rate typically correlates well with R-loop levels [278]. R-loops presence itself affects transcription; in recent years, many studies shed light on various mechanisms through which R-loops can regulate gene expression, for example by modulating chromatin accessibility or promoting transcription termination [116]. Timely removal of R-loops is critical, and accumulation of these structures contributes to genomic instability. The displaced single strand of DNA is particularly vulnerable to DNA damage, giving rise to single-strand breaks (SSBs) which can be later processed into DSBs [119], [279], [280], [281]. Furthermore, R-loop accumulation can contribute to transcription-replication conflicts, leading to collapse of the replication fork and generation of DSBs [282]. Despite the causal relationship between high R-loop levels and genomic instability being very well established, it is worth pointing out that both R-loops and RNA:DNA hybrids can also contribute to the maintenance of genome stability [283]. For example, different studies showed that the transient formation of RNA:DNA hybrids at subsets of DSBs facilitates DNA repair [284], [285], [286]. All in all, the interplay between transcription, R-loops and DNA repair requires the dynamic regulation of all processes involved.

Proteins acting in transcription and pre-mRNA processing are great candidates for roles as R-loop regulators. Mutations in general splicing factors were shown to cause an increase in R-loops formation and genome instability in myelodysplastic syndromes and chronic lymphocytic leukemia [287], [288], [289], [290]. Additionally, the displacement of the whole spliceosome from chromatin in response to transcription-blocking DNA lesions could also provide an opportunity for the formation of unscheduled R-loops [291]. In the case of the IBC, subunits AQR and XAB2 have been previously implicated in R-loop regulation [119], [141]. R-loops were shown to accumulate upon depletion of either protein, resulting in R-loop-dependent genomic instability; additionally, XAB2 was found to interact with nucleotide excision repair (NER) factors XPF and XPG and mediate their localization at RNA:DNA hybrids, thus regulating R-loop processing. Despite these reports, our current understanding of why R-loops accumulate in absence of IBC subunits is limited.

Our data showed that loss of AQR, XAB2 and ISY1 leads to the progressive increase in R-loops levels. It is possible that each one of these proteins causes R-loop accumulation through separate mechanisms;

the helicase activity of AQR, for example, makes it an attractive candidate as an R-loop resolving factor. In the future, purification of wild-type AQR will allow to test its ability to unwind R-loops *in vitro*. Alternatively, R-loop accumulation upon depletion of single subunits could reflect, at least to an extent, a general response to reduced levels of the whole IBC. In other words, R-loop formation could be another aspect of the core response to the loss of IBC subunits. In the future, mapping R-loops genome-wide upon loss of different complex members could give us valuable insight into where this structure form, and how the loss of each subunits shapes the R-loop landscape.

As the IBC is part of the spliceosome, R-loop accumulation could be the result of faulty splicing mechanisms, which could in turn affect transcription. In this sense, understanding where R-loops form in AQR-depleted cells, and how they could affect genomic instability, will provide a new perspective on the relationship between nascent transcription and AQR-dependent DSB formation. By looking at this global picture, we hope to unravel the complex crosstalk between transcription and genome instability, and understand the role of the IBC plays in it.

3.3 Concluding remarks

In this thesis, I have presented two stories sharing the same protagonist, the intron-binding complex. This work is a good example of how a serendipitous observation can lead to new ideas and projects; in this case, the collateral loss of IBC subunits is the finding we were not looking for, that spurred a lot of the research presented here. Without prior knowledge, the interpretation of experimental data, and particularly of “unexpected” data that is not directly linked to the primary research hypothesis, is very challenging. Similarly, analyzing specific phenotypes without acknowledging the bigger picture can lead to an only partial understanding of the underlying biology. In the case of the IBC, previous studies have investigated how cells react to the loss of individual subunits; however, we can expect that the biology and regulation of the IBC play a role in such cellular response. If the levels of IBC subunits are interconnected, and the absence of one subunit is matched by a reduction in the levels of the others, this will have an impact on downstream phenotypes. Of course, acknowledging this regulation increases the complexity of the research question, and requires taking into account many new biological variables. With this thesis I have tried, to the best of my abilities, to bridge the gap between IBC regulation and its involvement in cellular processes. In the process, this research uncovered novel quality control mechanisms for the IBC and highlighted new functions for its subunits. Most of the work presented here is part of projects that are still ongoing, and there are many exciting future directions and open questions. One of the goals for future research will be to explore the relevance of these mechanisms in physiological and pathological contexts, such as aging and cancer.

4. Materials and Methods

4.1 List of all consumables, software and machines

Table 1: Buffers/Solutions/Consumables

Cell culture	Composition / vendor
4-thiouridine (4SU)	Jena Bioscience
5-Ethynyl-2-deoxyuridine (EdU)	Jena Bioscience
5-Ethynyl uridine (EU)	Jena Bioscience
Bafilomycin A (BafA)	Sigma Aldrich
Biotin-phenol	Iris Biochem
Bortezomib	Selleckchem
Cyclohexamide (CHX)	Merck
Dialyzed FBS (10,000 molecular weight cut-off)	Thermo Fisher Scientific
Dimethyl sulfoxide (DMSO)	Sigma Aldrich
D-MEM for SILAC without lysine and arginine	Sigma Aldrich
Dulbecco's Modified Eagle Medium (D-MEM)	Gibco
Dulbecco's Phosphate-Buffered Saline (D-PBS)	Gibco
Fetal bovine serum (FBS)	Gibco
Human bone osteosarcoma epithelial cells (U2OS)	ATCC
Human colorectal carcinoma epithelial cells (HCT116)	ATCC
Human embryonic kidney cells (HEK293T)	ATCC
Human hTERT-immortalized retinal pigment epithelial cells (RPE-1)	ATCC
Hydrogen Peroxide	Sigma Aldrich
L-arginine (Arg0)	Cambridge Isotope Laboratories
L-lysine (Lys0)	Cambridge Isotope Laboratories
L-arginine-U-13C6 99% (Arg6)	Cambridge Isotope Laboratories
L-lysine-4,4,5,5-D4 96–98% (Lys4)	Cambridge Isotope Laboratories
L-arginine-U-13C6-15N4 99% (Arg10)	Cambridge Isotope Laboratories
L-lysine-U-13C6-15N2 99% (Lys8)	Cambridge Isotope Laboratories

Cell culture	Composition / vendor
L-glutamine	Gibco
MG132	Sigma Aldrich
Penicillin/Streptomycin	Gibco
Pladienolide B (PladB)	Santa Cruz Biotechnologies
Puromycin	InvivoGen
Quenching solution	1 mM Sodium azide, 10mM Sodium ascorbate and 5mM Trolox in D-PBS
Roswell Park Memorial Institute 1640 Medium (RPMI)	Gibco
Trypsin-EDTA (0.05%)	Gibco
Transfection	Composition / vendor
Lipofectamine RNAiMAX	Life Technologies
Opti-MEM with GlutaMAX	Gibco
Cell lysis	Composition / vendor
Complete protease inhibitor cocktail tablets	Roche Diagnostics
Modified RIPA buffer	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% Sodium-deoxycholate
No-salt modified RIPA buffer	50 mM Tris-HCl pH 7.5, 1 mM EDTA, 1% NP-40, 0.1% Sodium-deoxycholate
N-ethylmaleimide (NEM)	Sigma Aldrich
NuPAGE LDS Sample Buffer (4×) (LDS SB)	Thermo Fisher Scientific
Phosphatase inhibitors:	
1 mM sodium orthovanadate	Sigma Aldrich
5 mM β-glycerophosphate	Sigma Aldrich
5 mM sodium fluoride	Sigma Aldrich
QuickStart Bradford 1x Dye Protein Reagent	BioRad
RIPA buffer	50 mM Tris-HCl, pH 7.4, 1% Triton X-100, 0.5% Sodium Deoxycholate, 0.1% SDS, 150mM NaCl
Sm Nuclease	IMB PPCF
Pull-downs	Composition / vendor
ANTI-FLAG M2 Affinity Gel	Merck Millipore

Pull-downs	Composition / vendor
Dynabeads Protein G beads	Thermo Fisher Scientific
Neutravidin beads	Thermo Fisher Scientific
Urea	Sigma
SDS-PAGE and Western blotting	Composition / vendor
0.45 µm nitrocellulose	Sigma Aldrich
Blocking buffer	10% skimmed milk solution in PBS-T or 5% BSA in PBS-T or 3% BSA in PBS-T
Bovine serum albumin (BSA)	Sigma Aldrich
Dithiothreitol (DTT)	Sigma Aldrich
NuPAG MOPS SDS Running Buffer (20X)	Thermo Fisher Scientific
NuPAGE Bis-Tris gels 4-12%	Thermo Fisher Scientific
PBS-T	1x PBS, 0.1% Tween-20
Ponceau S	Sigma Aldrich
Ponceau S solution	0.1% (w/v) Ponceau S, 5% acetic acid
SuperSignal West Pico Chemiluminescent Substrate	Thermo Fisher Scientific
Transfer buffer	25 mM Tris, 192 mM Glycine, 20% (v/v) ethanol, pH 8.3
Microscopy and flow cytometry	Composition / vendor
4% paraformaldehyde (PFA) in PBS	Affymetrix
Alexa Fluor 647 azide	Thermo Fisher Scientific
Copper (II) sulfate pentahydrate in PBS	Sigma Aldrich
Blocking buffer	5% BSA in PBS-T
Hoechst 33342	Thermo Fisher Scientific
Permeabilization buffer	0.3% Triton X-100 in PBS
Sodium L-ascorbate	Sigma Aldrich
Mass spectrometry	Composition / vendor
In-gel digestion	
Buffer B	80% ACN, 0.5% acetic acid
Chloroacetamide (CAA)	Sigma Aldrich

Mass spectrometry	Composition / vendor
Colloidal Blue Staining Kit	Life Technologies
Destaining solution	50% Ethanol, 50 mM ABC buffer pH 8.0
Digestion buffer	25 mM ABC buffer pH 8.0
Peptide extraction buffer	30% ACN, 3% TFA
Sequencing grade Trypsin (0.5 µg/µl in 50 mM acetic acid)	Serva
In-solution digestion	
Denaturation buffer	6 M urea, 2 M thiourea in 10 mM HE-PES-NaOH pH 8.0
SepPAK C18 cartridges	Waters
Stage tipping	
Buffer A	0.1% formic acid
Buffer B	80% ACN, 0.1% formic acid
C18 elution buffer	50% ACN, 0.1% formic acid
C18 Empore 47 mm extraction disks	CDS Analytical
Micro-SCX fractionation	
SCX elution buffers	40 mM acetic acid, 40 mM boric acid, 40 mM phosphoric acid. (Adjust pH to the indicated pH values with sodium hydroxide. Add 40% ACN before use)
SCX Empore Cation 47 mm extraction disks	CDS Analytical
SCX wash buffer	40% ACN, 0.1% TFA
DiGly remnant enrichment	Composition / vendor
DiGly immunoaffinity purification buffer (IAP, 10x)	500 mM MOPS pH 7.2, 100 mM Na ₃ PO ₄ , 500 mM NaCl
DiGly wash buffer 1	1x IAP buffer, 150 mM NaCl, 0.5% NP-40
PTMScan Ubiquitin Remnant Motif beads	Cell signaling
mRNA sequencing	Composition / vendor
DNase	New England Biolabs
RNeasy Plus Mini Kit	Qiagen
RNase-free water	Sigma Aldrich
RNAaseZAP	Sigma Aldrich
TruSeq Stranded mRNA LT Sample Prep Kit	Illumina

TTchem-seq	Composition / vendor
KAPA RNA HyperPrep kit	Roche
μMACS™ Streptavidin Kit	Miltenyi Biotec
MTSEA-Biotin-XX	Biotium
RNA clean-up beads	Beckman Coulter
TRIzol	Invitrogen

sBLISS	Composition / vendor
Klenow fragment	New England Biolabs
NEBNext dA-tailing	New England Biolabs
CUTSmart buffer	New England Biolabs

Table 2: List of primary antibodies used in this study.

Unless otherwise stated, indicated dilutions refer to western blotting.

Antibody	Product number	Manufacturer	Dilution	Origin
53BP1	MAB3802	Millipore	1:400 (IF)	mouse
ACTB	A2228	Sigma-Aldrich	1:5000	mouse
AURKA	14475	Cell Signaling	1:500 (IF)	rabbit
c-Myc	ab32072	Abcam	1:1000	rabbit
DDX41	15076	Cell Signaling	1:1000	rabbit
FLAG M2	F1804	Sigma-Aldrich	1:1000	mouse
GFP	sc-9996	Santa Cruz Biotechnology	1:500	mouse
γH2AX (S139)	05-636	Millipore	1:1000	mouse
1:500 (IF)				
HCNP (XAB2)	sc-271037	Santa Cruz Biotechnology	1:500	mouse
1:200 (IF)				
IBP160 (AQR)	A302-547A	Bethyl Laboratories	1:1000	rabbit
1:250 (IF)				
ISY1	HPA016995	Sigma (Atlas Antibodies)	1:1000 (IF)	rabbit
MAP1LC3B	NB100-2220	Novus Biologicals	1:500	rabbit
p21 Waf1/Cip1	2947	Cell Signaling	1:1000	rabbit

Antibody	Product number	Manufacturer	Dilution	Origin
p-eIF2 α (S51)	3398	Cell Signaling	1:1000	rabbit
p-p53 (S15)	284	Cell Signaling	1:1000	rabbit
PP1E	sc-100700	Santa Cruz Biotechnology	1:500 (IF)	mouse
PRP19	sc-514338	Santa Cruz Biotechnology	1:500	mouse
Puromycin	MABE343	Sigma Aldrich	1:1000	mouse
RNaseH2A	ab92876	Abcam	1:1000	mouse
SETX	A301-104A-T	Bethyl Laboratories	1:1000	rabbit
SF3B4	ab183483	Abcam	1:1000	mouse
SQSTM1/p62	88588	Cell Signaling	1:1000	mouse
TRIP12	A301-814A	Bethyl Laboratories	1:1000	rabbit
Ubiquitin	sc-8017	Santa Cruz Biotechnology	1:1000	mouse
UBR5	65344	Cell Signalling	1:1000	rabbit

Table 3: List of secondary antibodies used in this study.

Indicated dilutions refer to western blotting, unless otherwise stated.

Antibody	Product number	Manufacturer	Dilution	Origin
Alexa Fluor 488, 568	A11001, A11004	Life Technologies	1:1000 (IF)	mouse
Alexa Fluor 488, 568	A1108, A11011	Life Technologies	1:1000 (IF)	rabbit
Secondary anti-bodies coupled to horseradish peroxidase		Jackson ImmunoResearch Laboratories	1:5000	mouse rabbit
Streptavidin-HRP	21130	Thermo Scientific	1:5000	

Table 4: List of siRNAs used in this study.

siRNAs	Sequence 5'-3'
AQR	CUGAAUAUGGCGGUGUAGU-TT
CTRL	UGGUUUACAUGUUGUGUGA-TT
CTRL pool	D-001820-10 – Horizon Discovery

siRNAs	Sequence 5'-3'
DDX41 pool	L-010394-00-0005 – Horizon Discovery
ISY1	GAGCCGAGUUAGUGGAAAA-TT
RNaseH2A	GGGCGGGUCAAAUACAACCUGAACU-TT
SETX	CCAUCUAACUCUGUACAACUUGCUU-TT
SF3B4	TGCCTTCAAGAAGGACTCCAA-TT
TRIP2	s17810 - ThermoFisher
XAB2	ACGCAGCACUCUCGAAUUU-TT

Table Table 5. List of software used in this study.

Software	
Adobe Illustrator CC2021	Harmony High-Content Imaging and Analysis Software (PerkinElmer) (version 4.4)
Bioconductor (version 3.14)	Integrated Genome Browser (version 2.12.3)
Blc2fastq (version 1.19)	MaxQuant (version 1.5.2.8 or 2.3.0.1)
BWA (version 0.7.17-r1188)	Perseus (version 1.6.15.0)
Cytoscape (version 3.8.2)	R studio ("Cherry Blossom" Release (3c53477a, 2023-03-09) / R (version 4.2.3)
DESeq2 (version 1.34)	rMATs (version 4.1.2)
featureCounts (version 2.0.0)	STAR (version 2.7.10)
GraphPad Prism (version 7.04)	Thermo Xcalibur (Thermo Fisher Scientific)

Table 6: Machines.

Machines	Vendor
3311 Forma Steri-Cult CO2Incubator	Eppendorf
Bioruptor NGS	Diagenode
ChemiDoc imaging system	BioRad
Easy-LC-1000	ThermoFisher Scientific
Easy-LC-1200	ThermoFisher Scientific
NanoDrop 2000	ThermoFisher Scientific
NextSeq550	Illumina

Machines	Vendor
NuPage Novex Gel System	ThermoFisher Scientific
Opera Phenix High Content Screening System	PerkinElmer
Orbitrap Exploris 480 mass spectrometer	ThermoFisher Scientific
Q Exactive Plus	ThermoFisher Scientific
Qubit 2.0 Fluorometer	Life Technologies
Sonifier 450	Branson
Thermoshaker	Eppendorf
VacuFuge Plus	Eppendorf
Xcell II Blot-Modul	ThermoFisher Scientific

4.2 Cell culture

4.2.1 Cell cultivation and passaging

Human osteosarcoma cells (U2OS), human colorectal carcinoma cells (HCT116), hTERT-immortalized retinal pigment epithelial cells (RPE-1) and human embryonic kidney 293 cells (HEK293T) cells were obtained from the ATCC and cultured in DMEM supplemented with 10% fetal bovine serum, L-glutamine, penicillin and streptomycin. p53^{-/-}, p21^{-/-} and isogenic wild-type HCT-116 cells were a gift from Holger Bastians (University Medical Center Göttingen), and were cultured in RPMI supplemented with 10% fetal bovine serum, L-glutamine, penicillin and streptomycin. Cells were routinely tested for mycoplasma infection using a PCR-based method. For SILAC labeling, cells were cultured in media containing either L-arginine and L-lysine, (13C6), L-arginine and (2H4)L-lysine, or (13C6,15N4) L-arginine and (13C6,15N2)L-lysine (Cambridge Isotope Laboratories) as described previously [292]. All cells were cultured at 37°C in a humidified incubator containing 5% CO₂.

4.2.2 siRNA transfection and treatments

Cells were transfected with siRNAs using Lipofectamine RNAiMAX (Life Technologies) according to the manufacturer's instructions. For proteasomal inhibition, cells were treated with either 10 μM MG-132 or 500 nM Bortezomib for indicated the indicated time-points. To inhibit autophagy, cells were treated with 200 nM Bafilomycin A1 for 4 hours. To inhibit translation, cells were treated with 100 μg/mL cyclohexamide for the indicated time-points. For splicing inhibition, cells were treated with 1 μM Pladienolide B for 3 hours. For labeling of nascent proteins, cells were incubated with 10 μg/ml puromycin at 37°C for 10 minutes.

4.3 Cell-based methods

4.3.1 Generation of whole cell lysates

Whole cell lysates were obtained by lysing cells in RIPA buffer supplemented with inhibitors (complete protease inhibitor cocktail, 1 mM sodium orthovanadate, 5 mM β -glycerophosphate, 5 mM sodium fluoride, 10 mM N-ethylmaleimide). Alternatively, a denaturing lysis was performed using 6M Urea/2M Thiourea supplemented with 0.1% SDS. Lysates were supplemented with Sm Nuclease (IMB core facility) to a 1:500 dilution, 2mM MgCl₂ and NaCl up to 1M. To release proteins from chromatin, lysates were sonicated with Bioruptor from Diagenode (30s ON/30s OFF, 10 cycles, intensity high). Lysates were clarified by centrifugation 16 000 $\times$ g for 15 min at 4 degrees and protein concentration was measured using QuickStart Bradford Protein assay (BioRad), followed by addition of NuPAGE LDS Sample Buffer (Life Technologies) supplemented with 1 mM DTT and boiling.

4.3.2 SDS-PAGE and Western Blotting

Proteins were resolved on 4%–12% gradient SDS-PAGE gels (NuPAGE Bis-Tris Precast Gels, Life Technologies) and transferred onto nitrocellulose membranes. Membranes were blocked using either 5% (w/v) skim milk, 5% (w/v) skim milk supplemented with 5% FBS, or 5% (w/v) BSA solutions in PBS supplemented with 0.1% Tween-20 (PBS-T). For proximity-dependent biotinylation experiments membranes were blocked with 3% (w/v) BSA solutions in PBST-T. For a complete list of antibodies, see Table 1. Secondary antibodies coupled to horseradish peroxidase (Jackson ImmunoResearch Laboratories) were used for immunodetection. For detection of biotinylated proteins in proximity-dependent biotinylation experiments, membranes were incubated with Streptavidin-HRP avoiding light exposure, followed by washes with three washes with H₂O, PBS and PBS-T each. Detection was performed with the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

4.3.3 Cell lysis for diGly-remnant enrichment experiments

Cells were washed twice with ice-cold PBS and lysed by scraping on ice with 1mL modified RIPA (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% sodium deoxycholate) supplemented with inhibitors (complete protease inhibitor cocktail, 1 mM sodium orthovanadate, 5 mM β -glycerophosphate, 5 mM sodium fluoride, and 10 mM N-ethylmaleimide). Lysates were supplemented with 1/10 volume of 5M NaCl and sonicated using a Branson Sonifier sonicator (3x30s, output = 1, on ice). Lysates were clarified by centrifugation 16 000 $\times$ g for 15 min at 4 degrees and protein concentration was estimated with QuickStart Bradford Protein assay (BioRad). SILAC-labeled extracts were then combined in a 1:1 ratio.

4.3.4 Proximity-dependent biotinylation for mass spectrometry (PDB-MS)

For the endogenous AQR proximal proteome, light and heavy-labeled SILAC U2OS cells were incubated with 10mM biotin-phenol in DMEM for 2h and subsequently treated with 1mM H₂O₂ for 2 minutes at room temperature. Cells were rapidly washed three times with quenching solution (10mM sodium azide, 10mM sodium ascorbate, 5mM Trolox in PBS) to stop the labeling reaction, followed by three washes with ice-cold PBS. Cells were scraped in ice-cold PBS supplemented with inhibitors (complete protease inhibitor cocktail, 1 mM sodium orthovanadate, 5 mM β -glycerophosphate, 5 mM sodium fluoride, and 10 mM N-ethylmaleimide) and lysed in RIPA buffer supplemented with inhibitors. NaCl concentration in RIPA buffer was brought up to 1M. To release proteins from chromatin, lysates were

sonicated and treated with Sm nuclease and 2mM MgCl₂ for 20 minutes on rotating wheel at 4 degrees. Lysates were clarified by centrifugation 16 000 × g for 15 min at 4 degrees and diluted 1:4 with no-salt modified RIPA buffer supplemented with inhibitors to bring down NaCl concentration to 200mM. After estimation of protein concentration QuickStart Bradford Protein assay (BioRad), equal amounts of SILAC-labeled extracts were combined (9 milligrams of lysate/SILAC condition were used) and incubated with pre-equilibrated NeutrAvidin agarose beads for 2h on rotating wheel at 4 °C. After the pulldown, beads were washed once with modified RIPA and three times with 2 M Urea. Bound proteins were eluted in NuPAGE LDS Sample Buffer supplemented with 1 mM DTT and 10mM biotin and boiled at 95 °C for 15 min.

4.3.5 Affinity purification for mass spectrometry (AP-MS)

For affinity purification of endogenous proteins, cells were lysed in modified RIPA buffer supplemented with inhibitors. NaCl concentration in RIPA buffer was brought up to 1M during lysis. Cells were lysed for 30 minutes on rotating wheel at 4 degrees. Lysates were clarified by centrifugation 16 000 × g for 15 min at 4 °C and diluted 1:3 with no-salt modified RIPA buffer to bring down NaCl concentration to 150mM. For pull-down of untagged endogenous baits, Dynabeads Protein G beads were washed with cold mRIPA buffer and incubated with the suggested amount of primary antibody (typically 2µg for each pull-down) or equal amounts of IgG for 1h on rotating wheel at 4 °C. For pull-down of endogenous FLAG-tagged AQR, anti-FLAG M2 affinity gel A2220 beads was washed with cold mRIPA. After estimation of protein concentration QuickStart Bradford Protein assay (BioRad), lysates (typically around 2.5 mg) were loaded on the beads and incubated for 2h on rotating wheel at 4 degrees. After the pulldown, beads were washed three times with modified RIPA; different SILAC conditions were combined on the last wash. Proteins were eluted in 2X NuPAGE LDS Sample Buffer supplemented with 1 mM DTT and boiled at 70°C for 15 min.

4.3.6 Immunofluorescence and high-throughput confocal microscopy

For high-throughput microscopy experiments, 10,000 or 20,000 cells were seeded into 96-well cell carrier plates depending on the experimental setup. For metabolic labeling of RNA, cells were incubated with 5-Ethynyluridine (5-EU) diluted at a final concentration of 10 µM in growth medium for 2h at 37°C prior to fixation. For immunofluorescence-based cell cycle profiling, cells were incubated with 5-Ethynyl-2-deoxyuridine (EdU) diluted at a final concentration of 10 µM in growth medium for 1 hour at 37°C prior to fixation. On the day of harvest, cells were washed twice with PBS and fixed with 4% PFA solution for 20 min at RT. After PFA removal, cells were washed 3 times with PBS containing 0.1% Tween-20 (PBS-T) and nuclei were permeabilized by incubation with a solution of 0.3% Triton X-100 in PBS for 10 minutes at RT. Cells were subsequently washed twice with PBS-T and incubated in blocking buffer (5% BSA in PBS-T) for 1 hour at RT. Primary antibodies were diluted according to the manufacturer's instructions in blocking buffer and added to the cells overnight at 4 °C. Cells were washed 3 times with PBS-T and then incubated with a solution of the secondary antibody (1:1000 in blocking buffer) for 1h at RT avoiding light exposure. After three washes with PBS-T, nuclei were stained with a solution of Hoechst 33342 (Sigma-Aldrich) in PBS (final concentration 2 µg/µl) for 20 minutes at RT avoiding light exposure. Cells were washed two more times with PBS-T, once with PBS and stored in PBS at 4°C until the measurement. Images were acquired on the Opera Phenix High Content Screening System using a 40X water objective lens. Per well about 30 fields with 5-10 z-planes separated by 0.5 µm were obtained. Image analysis was performed in the Harmony High-Content Imaging and Analysis

Software. Cell cycle profiling analysis was performed using the PhenoLOGIC supervised machine learning module of the Harmony software. Significance values were calculated in GraphPad Prism (one-way ANOVA followed by Dunnett's test). When needed, interplate normalization was performed as described in [293]: each cell was standardized by subtracting the median of control cells and dividing by the MAD (median absolute deviation) of controls per plate.

4.4 Mass spectrometry-based proteomics

4.4.1 In-gel digestion

Samples were alkylated with 5.5 mM chloroacetamide (CAA) for 30 minutes in the dark before being loaded onto 4-12% gradient SDS-PAGE gels. Proteins were stained with the Life Technologies Colloidal Blue Staining Kit. For subsequent processing, each lane was divided into 2-3 pieces or 5-6 pieces depending on the application. Each piece was then divided into 1 mm squares and incubated 4x with destaining buffer for 20 minutes. Gel pieces were dehydrated twice with 1 mL 100% EtOH for 10 minutes. Protein digestion was started by adding 50 μ L of trypsin solution (125 ng/ μ L trypsin in digestion buffer) to the pieces followed by incubation for 10~30 minutes at 500 rpm. The pieces were then covered with digestion buffer and incubated at 37 °C overnight. Digestion was stopped by adding 50 μ L peptide extraction buffer. The supernatant was transferred to a new tube after 20 minutes of incubation at 500 rpm. Peptide extraction was then achieved in a volume large enough (~100 μ L) to cover the gel pieces by sequential incubation with extraction buffer, in-gel buffer B, and acetonitrile (ACN). Supernatants from the same gel piece were combined, vacuum concentrated at 45 °C to a volume of 100 l, and subjected to desalting with StageTip purification (see 4.5.6).

4.4.2 In-solution digestion

For ubiquitin remnant profiling, SILAC-labeled extracts were combined in a 1:1 ratio and precipitated in a fourfold excess of ice-cold acetone overnight at -20 °C. After centrifugation at 1000 x g for 5 minutes, pellets were re-solubilized in 6 M urea/2 M thiourea buffered in 10 mM HEPES (pH 8.0). Cysteine residues were reduced with 1 mM DTT for 45 minutes and subsequently alkylated with 5.5 mM CAA for 30 minutes in the dark. Samples were diluted 1:4 with H₂O prior to digestion with sequencing-grade modified trypsin (Serva) in a 1:150 ratio overnight at RT in dark. Digestion was terminated by acidification with 0.5% trifluoroacetic acid (0.5% TFA) for 30 minutes at 4 °C. Precipitates were removed by centrifugation at 4000 x g for 10 minutes. Supernatants were purified and concentrated using reverse-phase Sep-Pak C18 cartridges (Waters). Cartridges were prepared by washing three times with 0.1% TFA and once with 100% ACN. After loading with peptide mixture, columns were washed three times with H₂O, dried, and either stored at 4 °C or eluted immediately.

4.4.3 DiGly-remnant enrichment

Peptides were eluted from reverse-phase Sep-Pak C18 cartridges in 2 ml of 50% aqueous ACN and peptide concentration was measured with NanoDrop. 100 μ g of peptides were put aside for full proteome analysis; the rest was vacuum concentrated until reaching 25% of the initial volume and combined with 100 μ l of immunoprecipitation buffer. Precipitates were removed by centrifugation for 10 minutes at 16.000 x g. PTMScan Ubiquitin Remnant Motif beads were equilibrated by washing three times in

immunoprecipitation buffer. Peptide mixtures were loaded on diGly beads and incubated overnight at 4 °C with rotation. Beads were washed twice with diGly wash buffer 1, twice with immunoprecipitation buffer, and twice with H₂O. Peptides were eluted by incubating three times with 100 µl 0.15% TFA at 1000 rpm for 5 minutes, and further processed by micro-SCX.

4.4.4 Micro-tip based strong cation exchange chromatography (Micro-SCX)

To build micro-SCX tips, six disks from 47 mm cation exchange extraction disks were cut with a 17-gauge Hamilton syringe and placed into a 200 µL pipette tip [294]. Micro-SCX tips were equilibrated with 50 µl of methanol, SCX elution buffer pH 4 (diGLY-remnant enrichment) or pH 3 (full proteome), SCX elution buffer pH 11, and 100 µl of SCX wash buffer by centrifugation at 500 x g. Acidified samples (pH < 2) were loaded on micro-SCX tips by centrifugation at 500 x g and fractionated by eluting with 100 µl of SCX elution buffers at 500 x g from lowest to highest pH. ACN was removed by vacuum centrifugation for 20 minutes at 45 °C and eluates were subsequently desalted by C18 StageTipping.

pH of SCX-buffers:

Full proteome: 3.0 / 3.6 / 3.9 / 4.25 / 4.8 / 5.5 / 7.0 / 11.0

DiGly-remnant peptides: 4.0 / 5.0 / 6.0 / 7.0 / 8.5 / 11.0

4.4.5 Desalting and concentration of peptides

Peptide purification was carried out with self-made stop-and-go extraction tips (StageTips) [294]. To build StageTips, two disks from 47 mm C18 extraction disks were cut with a 17-gauge Hamilton syringe and placed into a 200 µL pipette tip. StageTips were equilibrated with 25 µl of methanol, 25 µl of stage tip buffer B and 2 x 25 µl of stage tip buffer A by centrifugation at 500 x g. Acidified samples (pH < 2) were loaded on StageTips by centrifugation at 500 x g. StageTips were washed with 50 µl of stage tip buffer A and stored at RT. Peptides were eluted from StageTips with 50 µl elution buffer into 96-well sample plates by centrifugation at 500 x g. Samples were vacuum concentrated until reaching a volume of 4 µl and combined with 2 µl of 0.1% FA. For full proteome, samples were vacuum concentrated until reaching a volume of 7 µl and combined with 7 µl of 0.1% FA.

4.4.6 MS analysis

Peptides were analyzed on a quadrupole Orbitrap mass spectrometer (Q Exactive Plus or Exploris 480, Thermo Scientific) equipped with a UHPLC system (EASY-nLC 1000 or EASY-nLC 1200, Thermo Scientific) as described [295], [296], [297]. Peptide samples were loaded onto C18 reversed-phase columns and eluted with a linear 8 to 40% acetonitrile (ACN) gradient containing 0.1% formic acid. The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS₂ acquisition. Survey full scan MS spectra (m/z 300–1,700) were acquired in the Orbitrap mass analyzer. The 10 or 15 most intense ions were sequentially isolated and fragmented by higher energy C-trap dissociation (HCD) [298]. An ion selection threshold of 5,000 was used. Peptides with unassigned charge states, as well as with charge states < +2, were excluded from fragmentation. Fragment spectra were acquired in the Orbitrap mass analyzer.

4.4.7 MS peptide quantification

Raw data files were analyzed using MaxQuant software (development version 1.5.2.8 or 2.3.0.1) [299]. Parent ion and MS2 spectra were searched against a database containing 98,566 or 96,788 human protein sequences obtained from the UniProtKB (April 2018 or February 2020 release) using Andromeda search engine [300]. Spectra were searched with a mass tolerance of 6 ppm in MS mode, 20 ppm in HCD MS2 mode, strict trypsin specificity and allowing up to two miscleavages. Cysteine carbamidomethylation was set as a fixed modification, while cysteine modification with NEM, protein N-terminal acetylation and methionine oxidation were set as variable modifications. For diGLY-remnant profiling GlyGly modification of lysine was set as a variable modification. The dataset was filtered based on posterior error probability (PEP) to arrive at a FDR < 1% estimated using a target-decoy approach [301]. For SILAC-based quantification, the match between runs and re-quantify features were switched on. LFQ quantification was performed in the MaxQuant software using at least two LFQ ratio counts; the fast LFQ and the match between run options were switched on.

4.4.8 Data processing and visualization

Processed data from MaxQuant was analyzed in RStudio (R version 4.2.3) or Perseus (version 1.6.15.0) [302], [303]. Proteins or peptides flagged as “reverse”, “only identified by site” or “potential contaminant” were excluded from downstream analysis. Only proteins identified by no less than two peptides and at least one unique peptide were used in downstream analysis. Only GlyGly peptides with a localization probability > 0.9 were considered for analysis. For the AQR-proximal proteome non-normalized ratios were used. For label-free data, missing values were imputed using random values from a Gaussian distribution using the default parameters in Perseus. P-values were calculated using the LIMMA package [304] or with a standard two-sided t-test. Benjamini-Hochberg correction was used for FDR calculation. Unsupervised hierarchical clustering and heatmaps were created with the pheatmap package. UMAP clustering was performed with the umap package [305]. GO term analysis, KEGG pathways analysis and Reactome pathways analysis were performed with the ClusterProfiler package [306]. Protein interaction networks were analyzed with the STRING database and visualized on Cytoscape (version 3.8.2) [307], [308]. 2D annotation enrichment was performed on Perseus [186]. Protein expression data (CCLE project [160]) and mRNA expression data (CCLE project [161]) for IBC subunits were downloaded from the DepMap portal. Pairwise Pearson correlation coefficients (PCCs) were calculated for expression values for different pairs of genes across the CCLE cell line panel. Only lineages comprising at least 3 cell lines were used for PCC computation.

4.5 Genomics methods

4.5.1 mRNA-sequencing

RNA was extracted using the RNeasy Plus Mini Kit (Qiagen). In brief, cells were lysed and genomic DNA was depleted. Samples were treated with DNase to remove residual DNA. After purification using spin columns, RNA was eluted in RNase-free water and stored at -80 °C until library preparation. NGS library prep was performed with Illumina’s TruSeq stranded mRNA LT Sample Prep Kit following TruSeq Stranded mRNA Reference Guide (Oct.2017) (Document # 1000000040498v00). Libraries were prepared with a starting amount of 1000 ng and amplified in 10 PCR cycles. Libraries were profiled in

a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent Technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life Technologies). Samples were pooled in equimolar ratio and sequenced on a NextSeq 500 Highoutput FC, SR for 1 × 84 cycles plus 7 cycles for the index read.

4.5.2 Nascent RNA quantification by TTChem-sequencing

TTChem-seq was used for nascent RNA analysis in U2OS cells as described elsewhere (Gregersen et al., 2020). Heat-shock treatment (42°C for 30') was used as a positive control. Briefly, cells were pulsed with 1mM of 4-thiouridine (4sU, Jena Bioscience N-1090-100) for 30 minutes before being lysed with TRIzol (Invitrogen 15596026) directly on the plate to quench 4sU incorporation. Total RNA was extracted following manufacturer's instructions. The total RNA was chemically fragmented with 166mM of NaOH for 20 minutes on ice. The reaction was quenched by adding 400mM of Tris pH 6.8 followed by 2 consecutive clean-up reactions on Micro Bio-Spin P-30 columns (Bio-Rad 7326250). The thiol groups in the labeled RNA were biotinylated with MTSEA Biotincapcap; 2-((6-((6-((biotinoyl)amino)hexanoyl)amino)hexanoyl)amino)ethylmethanethiosulfonate (MTSEA-Biotin-XX, Biotium 90066) before RNA precipitation using phenol/chloroform/isoamyl alcohol (25:24:1 (vol/vol/vol)). Biotin-streptavidin pull-down of labeled RNA was performed with the μ MACS™ Streptavidin Kit (Miltenyi Biotec 130-074-101) and the 4sU-labeled RNA was purified using 2x volumes of RNA Clean-up beads (Beckman Coulter A63987). Finally, 66ng of 4sU-labeled RNA was used for stranded RNA library preparation using the KAPA RNA HyperPrep kit (Roche KK8540) following the manufacturer's instructions. Libraries were sequenced in a NextSeq550 illumina sequencer in 1x75bp mode.

4.5.3 Data analysis for mRNA-sequencing and TTChem-sequencing

The libraries for both the mRNA sequencing and the TTSeq data were sequenced on an Illumina NextSeq 500 instrument. The mRNA with a read length of 84 bp and TTSeq libraries were sequenced with a read length of 75 bp in paired end mode. Demultiplexed was carried out using bcl2fastq v1.19. The samples were mapped against hg38 (Gencode primary release (v26) [309] using STAR (v2.7.10) [310]. To improve splice site detection the mapping second pass mapping was performed using the detected splice junctions in the first pass as an additional input for STAR. Reads per gene were quantified using featureCounts (v. 2.0.0) [311]. For the mRNA analysis exons were used as a reference to quantify the amount of reads per gene. In the TTSeq analysis the gene as a whole was used as a reference for quantification. The differential expression analysis was performed using Bioconductor (v3.14) [312] / DESeq2 (v1.34) [313]. Genes were deemed sign. diff. regulated with an FDR below 1%. In the mRNA-Seq splicing event analysis was conducted using rMATS (v4.1.2) [314]. Splicing events with an FDR smaller than 10% were considered significant.

4.5.4 BLISS and data analysis

sBLISS in U2OS cells was performed as previously described^{48,49} with the following modifications: after blunting of DNA DSB ends in fixed nuclei, samples were 3' adenylated using Klenow Fragment (3'-->5' exo-) (NEB M0212) at a final amount of 60U per reaction in 1× NEBNext® dA-Tailing Reaction Buffer (NEB B6059). The A-tailing reaction was incubated at 37 °C for 1 h in a thermo-shaker at 300 rpm. Fixed nuclei were then washed 2× with CST buffer (CutSmart buffer B7204 supplemented with 0.1% Triton) in order to washout the Klenow enzyme. DNA DSB-end labeling was performed as described⁴⁸ with the following modification: sBLISS linkers containing one thymine overhang at the

3' end of the reverse oligo were used. Downstream sample processing steps were carried out with 150 ng of DNA template input from each sample for in vitro transcription reaction. Sequenced sBLISS libraries were processed as previously described [204]. The process is implemented in the NGSpice2go sBLISS and BreakTag pipeline (<https://gitlab.rlp.net/imbforge/NGSpice2go/-/tree/master/pipelines/breaktag>). The pipeline first scans reads to keep those containing the expected 8 nucleotide-long UMI followed by the 8 nucleotide-long sample barcode in the 5' end. Valid reads were aligned to the human reference genome version hg38 downloaded from UCSC with timestamp 2014-01-15 21:14, using the mem command in BWA (version 0.7.17-r1188) [315] with a seed length of 19 and default scoring/penalty values for mismatches, gaps, and read clipping. Uniquely mapping reads which mapped with a minimum quality score Q=60 were retained. A final deduplication step was performed in which spatial consecutive reads mapping within a window of 30 nt and their UMIs differing by up to two mismatches were considered close PCR duplicates, and only one read was kept. The resulting reads were aggregated per position and reported as double-stranded breaks in a BED file. For further processing the hotspots, called on merged replicates per group, were filtered using blacklisted regions based on published Encode data provided by the Boyle lab on github (<https://github.com/Boyle-Lab/Blacklist/blob/61a04d2c5e49341d76735d485c61f0d1177d08a8/lists/hg38-blacklist.v2.bed.gz>) [316]. The scatterplot was generated using the union of blacklist-filtered 1 hotspots for both conditions. Additionally, the Venn diagram contrasting the called hotspots is based on hotspots called on merged replicates as well.

5. List of abbreviations

Table 7: Abbreviations and respective full forms

Abbreviation	Full form
4-OHT	4-hydroxytamoxifen
5'TOP	5' terminal oligo-pyrimidine
53BP1	TP53-binding protein 1
A3SS	Alternative 3' splice site
A5SS	Alternative 5' splice site
aa	Aminoacid
ACN	Acentonitrile
ACTB	Actin beta
AMPK	5'-AMP-activated protein kinase
AP	Affinity purification
APC	Anaphase promoting complex
APEX1	DNA-(apurinic or apyrimidinic site) endonuclease
APEX2	L-ascorbate peroxidase 2
AQR	Aquarius intron-binding spliceosomal factor
ARID2	AT-rich interaction domain 2
ARM	Armadillo
AS	Alternative splicing
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATPase	ATP hydrolase
ATR	Ataxia telangiectasia and Rad3 related
AURKA	Aurora kinase A
AURKB	Aurora kinase B
BafA	Bafilomycin A1
BAG6	BAG family molecular chaperone regulator 6
BCAS2	Breast carcinoma-amplified sequence 2
BER	Base excision repair
BP	Biological process
CAA	Chloroacetamide

Abbreviation	Full form
Cas9	CRISPR associated protein 9
CC	Cellular compartment
CCLE	Cancer cell line encyclopedia
CDCA8	Cell division cycle associated 8
CDC40	Cell division cycle 40
CDK6	Cyclin-dependent kinase 6
CDKN1A	Cyclin dependent kinase inhibitor 1A
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
CHX	Cyclohexamide
CLTH	C-terminal to LisH complex
COL6A1	Collagen type VI alpha 1 chain
COL6A2	Collagen type VI alpha 2 chain
CPSF3	Cleavage and polyadenylation specificity factor subunit 3
CRISPR	Clustered regularly interspaced short palindromic repeats
CRNKL1	Crooked neck pre-mRNA splicing factor 1
CSA	Cockayne syndrome protein A
CSB	Cockayne syndrome protein B
CSN	COP signalosome
CTRL	Control
CWC22	CWC22 Spliceosome Associated Protein Homolog
Cwf3	Pre-mRNA-splicing factor cwf3
DCX	DDB1-CUL4-X-box complex
DDX41	Probable ATP-dependent RNA helicase DDX41
DDR	DNA damage response
Degron	Degradation signal
DiGly	Lys- ϵ -Gly-Gly
DlvA	DSB Inducible via AsiSI
DKC1	Dyskerin pseudouridine synthase 1
DNA	Deoxyribonucleic acid
DREAM	Dimerization partner, RB-like, E2F and multi-vulval class B complex
DSB	Double-strand break
DTT	Dithiothreitol
E1	Ubiquitin-activating enzyme

Abbreviation	Full form
E2F	E2 transcription factor
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin-ligating enzyme
EBP50	Ezrin-binding protein 50
ED	Exponentially degraded
EdU	5-Ethynyl-2-deoxyuridine
eIF2 α	Eukaryotic translation initiation factor 2 subunit alpha
EJC	Exon junction complex
ESC	Embryonic stem cell
EU	5-Ethynyl uridine
FA	Formic acid
FC	Fold change
FDR	False discovery rate
GIN5	Go-ichi-ni-san complex
GO	Gene ontology
GG-NER	Global genome NER
GTP	Guanosine triphosphate
GTPase	GTP hydrolysis
H2AX	H2A.X variant histone
HAT	Half a TPR repeat
HBD	Hybrind binding domain
HCD	higher energy C-trap dissociation
HCT116	Human colorectal carcinoma epithelial cells
HEAT	Huntington, Elongation Factor 3, PR65/A, TOR
HECT	Homologous to the E6AP carboxyl terminus
HEK293T	Human embryonic kidney cells
HMGB1	High mobility group box 1
HR	Homologous recombination
HUWE1	HECT, UBA and WWE domain-containing protein 1
IBC	Intron-binding complex
INCENP	Inner centromere protein
IP	Immunoprecipitation
IRBC	Impaired ribosome biogenesis checkpoint
ISY1	ISY1 splicing factor homolog

Abbreviation	Full form
ITGB1BP1	Integrin subunit beta 1 binding protein 1
LC-MS/MS	High pressure liquid chromatography and tandem mass spectrometry
LFQ	Label-free quantification
KD	Knock-down
KDa	1,000 Dalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knock-out
KIF18B	Kinesin family member 18B
KIF20A	Kinesin family member 20A
KIF23	Kinesin family member 23
KMT2A	Histone-lysine N-methyltransferase 2A
m/z	Mass to charge
MAP1LC3	Microtubule-associated proteins 1A/1B light chain 3
MAPK	Mitogen-activated protein kinase
MCM	Minichromosome maintenance protein complex
MDM2	E3 ubiquitin-protein ligase Mdm2
MF	Molecular function
Micro-SCX	Micro-tip based strong cation exchange chromatography
MLL1	Mixed lineage leukemia protein 1
miRNA	Micro-RNA
mRNA	Messenger RNA
MS	Mass spectrometry
MS/MS	Tandem MS
MXE	Mutually exclusive exons
MYC	Myc proto-oncogene protein
NAC	Nascent polypeptide-associated complex
NED	Non-exponentially degraded
NEDD8	Neural precursor cell expressed, developmentally down-regulated 8
NER	Nucleotide excision repair
NMD	Nonsense-mediated decay
NOP10	Nucleolar protein 10
NOP56	Nucleolar protein 56
nt	Nucleotide

Abbreviation	Full form
NTC	NineTeen complex
PEA15	Proliferation And Apoptosis Adaptor Protein 15
PBRM1	Polybromo 1
PCC	Pearson correlation coefficient
PDB	Proximity-dependent biotinylation
PEP	Posterior error probability
PFA	Paraformaldehyde
PFD	Prefoldin
Plad B	Pladienolide B
PPI	Protein-protein interaction
PPIE	Peptidylprolyl isomerase E
PPIase	Peptidyl-prolyl cis-trans isomerase
pro-miRNA	Progenitor miRNA
PRP19	Pre-mRNA processing factor 19
RAC	Ribosome-associated complex
RAD23B	UV excision repair protein RAD23 homolog B
RANBP2	RAN inbding protein 2
RB	Retinoblastoma protein
RBP	RNA binding protein
RBR	RING-between-RING
RI	Intron retention
RING	Really interesting new gene
RIPA	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RNaseH1	Ribonuclease H1
RNaseH2A	Ribonuclease H2A
RNF168	Ring Finger Protein 168
RNP	Ribonucleoprotein
RPE-1	Human hTERT-immortalized retinal pigment epithelial cells
RPL5	Ribosomal protein L5
RPL11	Ribosomal protein L11
RRM	RNA recognition motif
SAGA	Spt-Ada-Gcn5 acetyltransferase

Abbreviation	Full form
SASP	Senescent-associated secretory phenotype
sBLISS	In-suspension Break Labeling In Situ and Sequencing
SCF	Skp1-Cullin-F-box protein complex
SE	Skipped exon
SERPINE1	Serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
SETX	Senataxin
SF1	Superfamily 1
SF2	Superfamily 2
SF3B1	Splicing factor 3B subunit 1
SF3B3	Splicing factor 3B subunit 3
SF3B4	Splicing factor 3B subunit 4
SILAC	Stable isotope labeling with amino acids in cell culture
siRNA	Short interfering RNA
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 2
SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, Member 4
SMARCC2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily C Member 2
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein
SQSTM1	Sequestosome-1
SRRM2	Serine/Arginine repetitive matrix 2
SSB	Single-strand break
SUMO	Small ubiquitin-related modifier
SWI/SNF	SWItch/Sucrose Non-Fermentable
Syf1	Pre-mRNA-splicing factor SYF1
TC-DSBR	Transcription-coupled double-strand break repair
TC-NER	Transcription-coupled NER
TES	Transcription end site
TFA	Trifluoroacetic acid
TFIID	Transcription factor II D
TGF- β	Transforming growth factor beta
Tom1	Target of Myb protein 1
TOPBP1	Topoisomerase 2 binding protein 1

Abbreviation	Full form
TP53	Tumor protein p53
TPR	Tetratricopeptide repeat
TRiC	T-complex protein ring complex
TRIP12	Thyroid hormone receptor interacting protein 12
tRNA	Transfer RNA
TREX-2	Three prime repair exonuclease 2
TSS	Transcription start site
TTchem-seq	Transient transcriptome sequencing
U2OS	Human bone osteosarcoma epithelial cells
Ub	Ubiquitin
UBA1	Ubiquitin-activating enzyme E1 1
UBA6	Ubiquitin-like modifier-activating enzyme 6
UBE2O	Ubiquitin conjugating enzyme E2 O
UBE2R2	Ubiquitin conjugating enzyme E2 R2
UBE3A	HECT-type ubiquitin transferase E3A
UBR1	Ubiquitin protein ligase E3 component N-recognin 1
UBR2	Ubiquitin protein ligase E3 component N-recognin 2
UBR4	Ubiquitin protein ligase E3 component N-recognin 4
UBR5	Ubiquitin protein ligase E3 component N-recognin 5
UFD	Ub fusion degradation
Ufd2	Ubiquitin fusion degradation protein 2
Ufd4	Ubiquitin fusion degradation protein 4
UMAP	Uniform manifold approximation
UPF1	UPF1 RNA helicase and ATPase
UPS	Ubiquitin-proteasome system
USP28	Ubiquitin carboxyl-terminal hydrolase 28
VCP	Valosin-containing protein
XAB2	XPA binding protein 2
XPA	Xeroderma pigmentosum group A-complementing protein
XPF	Xeroderma pigmentosum, complementation group F
XPG	Xeroderma pigmentosum, complementation group G
ZNF830	Zinc finger protein 830

6. Appendix

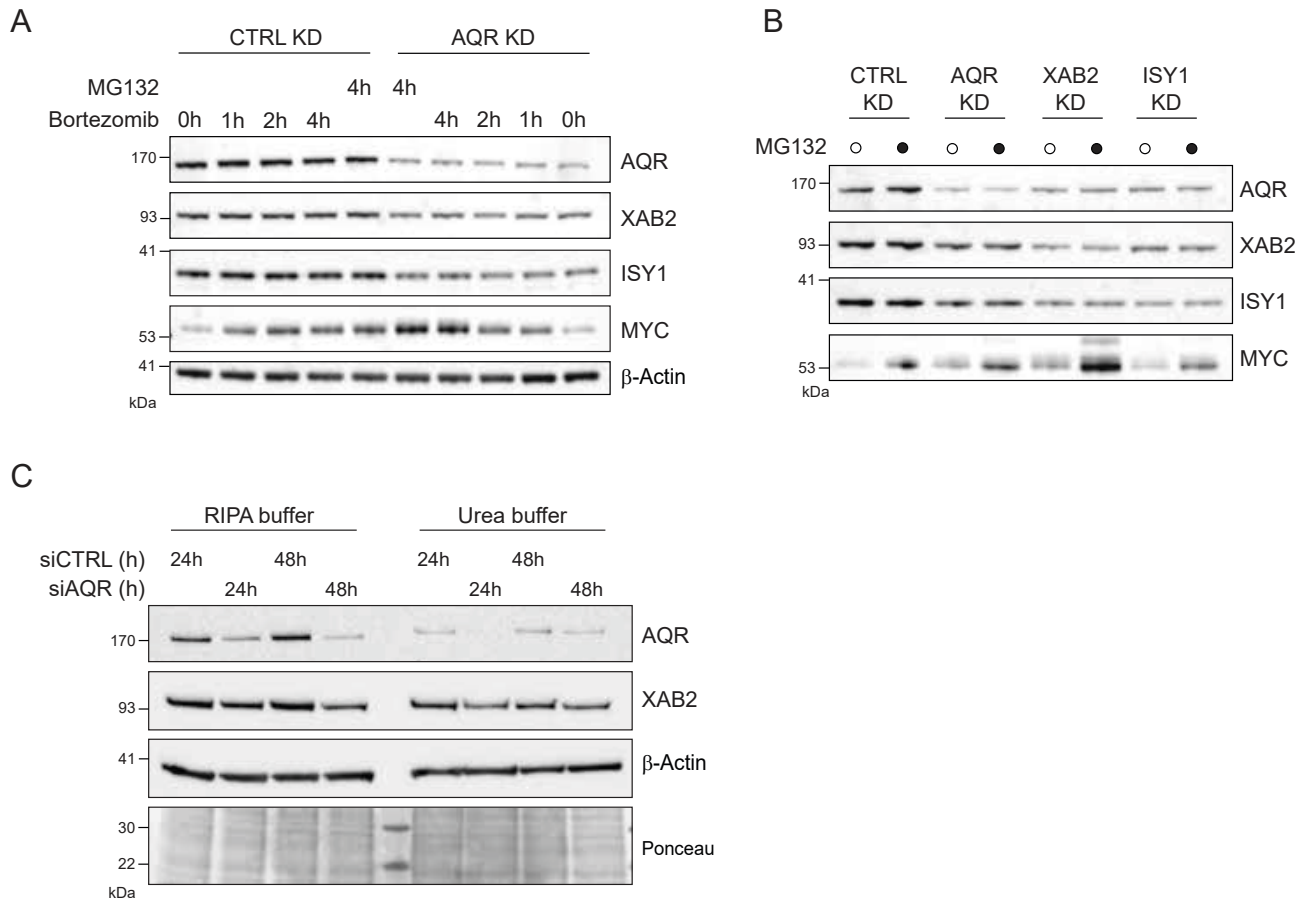
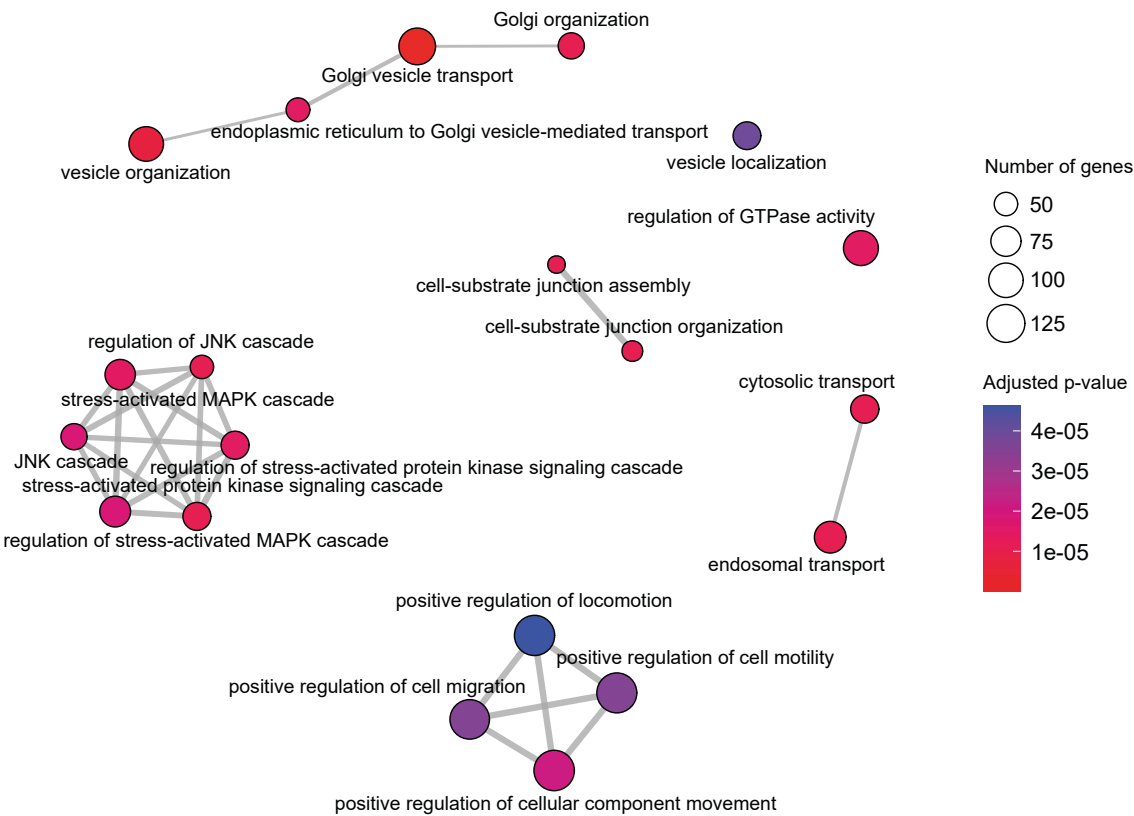


Figure 30. Additional attempts to rescue the collateral loss phenotype upon IBC subunits depletion

A) Western blot analysis of levels of indicated proteins in CTRL or AQR-depleted U2OS cells. Cells were treated with 10 μ M MG132 or 500nM Bortezomib for the indicated time-points. Levels were monitored after 48h of knock-down. B) Western blot analysis of IBC subunits levels after 48h AQR, XAB2 or ISY1 knock-down in RPE-1 cells. Cells were treated with 10 μ M MG132 for 8h. C) Western blot images showing the side-by-side comparison of different cell lysis protocol. Cells were lysed with either high-salt RIPA buffer or with a denaturing buffer containing 6M Urea/2M Thiourea and 0.1% SDS. Levels of AQR and XAB2 were monitored after 24h or 48h of AQR knock-down in U2OS cells.

A



B

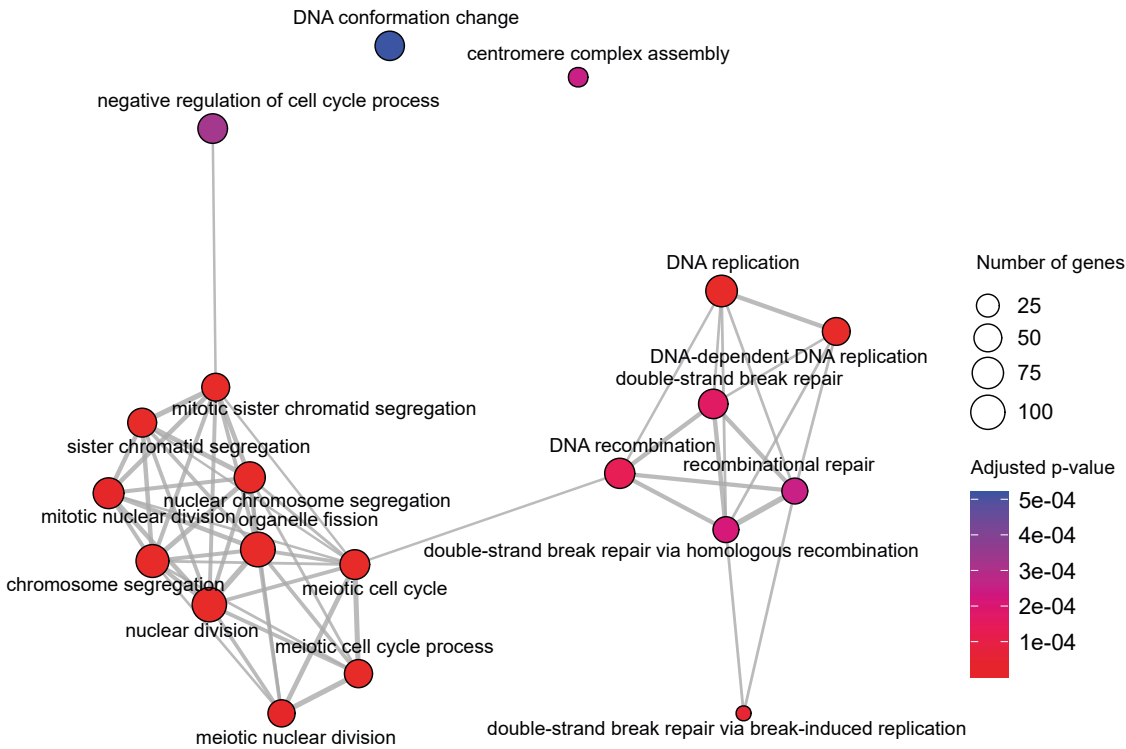


Figure 31. GO term analysis of AQR KD TTchem-seq

A and B) Functional enrichment analysis performed with ClusterProfiler showing the enriched GO terms (FDR ≤ 0.05) for the up-regulated (A) or down-regulated genes (B) in AQR KD TTchem-seq (FDR < 0.01).

7. References

- [1] J. A. Marsh and S. A. Teichmann, 'Structure, dynamics, assembly, and evolution of protein complexes', *Annual Review of Biochemistry*, vol. 84, 2015. doi: 10.1146/annurev-biochem-060614-034142.
- [2] N. J. Krogan et al., 'Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*', *Nature*, vol. 440, no. 7084, 2006, doi: 10.1038/nature04670.
- [3] M. Y. Hein et al., 'A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances', *Cell*, vol. 163, no. 3, 2015, doi: 10.1016/j.cell.2015.09.053.
- [4] E. L. Huttlin et al., 'The BioPlex Network: A Systematic Exploration of the Human Interactome', *Cell*, vol. 162, no. 2, 2015, doi: 10.1016/j.cell.2015.06.043.
- [5] M. Giurgiu et al., 'CORUM: The comprehensive resource of mammalian protein complexes - 2019', *Nucleic Acids Res*, vol. 47, no. D1, 2019, doi: 10.1093/nar/gky973.
- [6] K. Luck et al., 'A reference map of the human binary protein interactome', *Nature*, vol. 580, no. 7803, 2020, doi: 10.1038/s41586-020-2188-x.
- [7] P. C. Havugimana et al., 'A census of human soluble protein complexes', *Cell*, vol. 150, no. 5, 2012, doi: 10.1016/j.cell.2012.08.011.
- [8] E. D. Levy, E. B. Erba, C. V. Robinson, and S. A. Teichmann, 'Assembly reflects evolution of protein complexes', *Nature*, vol. 453, no. 7199, 2008, doi: 10.1038/nature06942.
- [9] J. A. Marsh et al., 'Protein complexes are under evolutionary selection to assemble via ordered pathways', *Cell*, vol. 153, no. 2, 2013, doi: 10.1016/j.cell.2013.02.044.
- [10] S. E. Ahnert, J. A. Marsh, H. Hernández, C. V. Robinson, and S. A. Teichmann, 'Principles of assembly reveal a periodic table of protein complexes', *Science* (1979), vol. 350, no. 6266, 2015, doi: 10.1126/science.aaa2245.
- [11] V. A. Raker, G. Plessel, and R. Lührmann, 'The snRNP core assembly pathway: Identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro', *EMBO Journal*, vol. 15, no. 9, 1996, doi: 10.1002/j.1460-2075.1996.tb00579.x.
- [12] M. Lutzmann, R. Kunze, A. Buerer, U. Aebi, and E. Hurt, 'Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins', *EMBO Journal*, vol. 21, no. 3, 2002, doi: 10.1093/emboj/21.3.387.
- [13] N. Gallastegui and M. Groll, 'The 26S proteasome: assembly and function of a destructive machine',

- Trends in Biochemical Sciences, vol. 35, no. 11. 2010. doi: 10.1016/j.tibs.2010.05.005.
- [14] A. G. Matera and Z. Wang, 'A day in the life of the spliceosome', *Nature Reviews Molecular Cell Biology*, vol. 15, no. 2. 2014. doi: 10.1038/nrm3742.
 - [15] C. A. Woolhead, P. J. McCormick, and A. E. Johnson, 'Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins', *Cell*, vol. 116, no. 5, 2004, doi: 10.1016/S0092-8674(04)00169-2.
 - [16] A. A. Komar, 'A pause for thought along the co-translational folding pathway', *Trends in Biochemical Sciences*, vol. 34, no. 1. 2009. doi: 10.1016/j.tibs.2008.10.002.
 - [17] O. B. Nilsson et al., 'Cotranslational Protein Folding inside the Ribosome Exit Tunnel', *Cell Rep*, vol. 12, no. 10, 2015, doi: 10.1016/j.celrep.2015.07.065.
 - [18] V. Albanèse, A. Y. W. Yam, J. Baughman, C. Parnot, and J. Frydman, 'Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells', *Cell*, vol. 124, no. 1, 2006, doi: 10.1016/j.cell.2005.11.039.
 - [19] F. Morales-Polanco, J. H. Lee, N. M. Barbosa, and J. Frydman, 'Cotranslational Mechanisms of Protein Biogenesis and Complex Assembly in Eukaryotes', *Annual Review of Biomedical Data Science*, vol. 5. 2022. doi: 10.1146/annurev-biodatasci-121721-095858.
 - [20] E. Deuerling, M. Gamerding, and S. G. Kreft, 'Chaperone interactions at the ribosome', *Cold Spring Harb Perspect Biol*, vol. 11, no. 11, 2019, doi: 10.1101/cshperspect.a033977.
 - [21] E. Natan, J. N. Wells, S. A. Teichmann, and J. A. Marsh, 'Regulation, evolution and consequences of cotranslational protein complex assembly', *Current Opinion in Structural Biology*, vol. 42. 2017. doi: 10.1016/j.sbi.2016.11.023.
 - [22] A. Schwarz and M. Beck, 'The Benefits of Cotranslational Assembly: A Structural Perspective', *Trends in Cell Biology*, vol. 29, no. 10. 2019. doi: 10.1016/j.tcb.2019.07.006.
 - [23] K. Khan and P. L. Fox, 'Benefits of co-translational complex assembly for cellular fitness', *BioEssays*, vol. 45, no. 5. 2023. doi: 10.1002/bies.202300024.
 - [24] J. N. Wells, L. T. Bergendahl, and J. A. Marsh, 'Co-translational assembly of protein complexes', *Biochem Soc Trans*, vol. 43, 2015, doi: 10.1042/BST20150159.
 - [25] M. Bertolini et al., 'Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly', *Science* (1979), vol. 371, no. 6524, 2021, doi: 10.1126/science.abc7151.
 - [26] C. D. Nicholls, K. G. McLure, M. A. Shields, and P. W. K. Lee, 'Biogenesis of p53 involves cotranslational dimerization of monomers and posttranslational dimerization of dimers. Implications on the

- dominant negative effect', *Journal of Biological Chemistry*, vol. 277, no. 15, 2002, doi: 10.1074/jbc.M108815200.
- [27] K. Khan, B. Long, V. Gogonea, G. M. Deshpande, K. Vasu, and P. L. Fox, 'Multimodal cotranslational interactions direct assembly of the human multi-tRNA synthetase complex', *Proc Natl Acad Sci U S A*, vol. 119, no. 36, 2022, doi: 10.1073/pnas.2205669119.
 - [28] F. Liu, D. K. Jones, W. J. De Lange, and G. A. Robertson, 'Cotranslational association of mRNA encoding subunits of heteromeric ion channels', *Proc Natl Acad Sci U S A*, vol. 113, no. 17, 2016, doi: 10.1073/pnas.1521577113.
 - [29] O. O. Panasenko et al., 'Co-translational assembly of proteasome subunits in NOT1-containing assemblysomes', *Nat Struct Mol Biol*, vol. 26, no. 2, 2019, doi: 10.1038/s41594-018-0179-5.
 - [30] F. Morales-Polanco et al., 'Core Fermentation (CoFe) granules focus coordinated glycolytic mRNA localization and translation to fuel glucose fermentation', *iScience*, vol. 24, no. 2, 2021, doi: 10.1016/j.isci.2021.102069.
 - [31] C. D. S. Duncan and J. Mata, 'Widespread cotranslational formation of protein complexes', *PLoS Genet*, vol. 7, no. 12, 2011, doi: 10.1371/journal.pgen.1002398.
 - [32] A. Shiber et al., 'Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling', *Nature*, vol. 561, no. 7722, 2018, doi: 10.1038/s41586-018-0462-y.
 - [33] I. Kamenova et al., 'Co-translational assembly of mammalian nuclear multisubunit complexes', *Nat Commun*, vol. 10, no. 1, 2019, doi: 10.1038/s41467-019-09749-y.
 - [34] O. Lautier et al., 'Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis', *Mol Cell*, vol. 81, no. 11, 2021, doi: 10.1016/j.molcel.2021.03.030.
 - [35] M. Seidel et al., 'Co-translational assembly orchestrates competing biogenesis pathways', *Nat Commun*, vol. 13, no. 1, 2022, doi: 10.1038/s41467-022-28878-5.
 - [36] M. Badonyi and J. A. Marsh, 'Large protein complex interfaces have evolved to promote cotranslational assembly', *Elife*, vol. 2022, no. 11, 2022, doi: 10.7554/ELIFE.79602.
 - [37] A. Halbach et al., 'Cotranslational assembly of the yeast SET1C histone methyltransferase complex', *EMBO Journal*, vol. 28, no. 19, 2009, doi: 10.1038/emboj.2009.240.
 - [38] K. Khan et al., 'Cotranslational interaction of human EBP50 and ezrin overcomes masked binding site during complex assembly', *Proc Natl Acad Sci U S A*, vol. 119, no. 7, 2022, doi: 10.1073/pnas.2115799119.
 - [39] H. M. Schnell et al., 'Structures of chaperone-associated assembly intermediates reveal coordinated

References

- mechanisms of proteasome biogenesis', *Nat Struct Mol Biol*, vol. 28, no. 5, 2021, doi: 10.1038/s41594-021-00583-9.
- [40] O. Q. H. Zinani, K. Keseroğlu, and E. M. Özbudak, 'Regulatory mechanisms ensuring coordinated expression of functionally related genes', *Trends in Genetics*, vol. 38, no. 1. 2022. doi: 10.1016/j.tig.2021.07.008.
- [41] M. Dávila López, J. J. Martíne Guerra, and T. Samuelsson, 'Analysis of gene order conservation in eukaryotes identifies transcriptionally and functionally linked genes', *PLoS One*, vol. 5, no. 5, 2010, doi: 10.1371/journal.pone.0010654.
- [42] O. Meyuhas and T. Kahan, 'The race to decipher the top secrets of TOP mRNAs', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, vol. 1849, no. 7. 2015. doi: 10.1016/j.bbagr.2014.08.015.
- [43] G. W. Li, D. Burkhardt, C. Gross, and J. S. Weissman, 'Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources', *Cell*, vol. 157, no. 3, 2014, doi: 10.1016/j.cell.2014.02.033.
- [44] J. C. Taggart and G.-W. Li, 'Production of Protein-Complex Components Is Stoichiometric and Lacks General Feedback Regulation in Eukaryotes', *Cell Syst*, vol. 7, no. 6, pp. 580-589.e4, Dec. 2018, doi: 10.1016/j.cels.2018.11.003.
- [45] B. Dahlmann, T. Ruppert, L. Kuehn, S. Merforth, and P. M. Kloetzel, 'Different proteasome subtypes in a single tissue exhibit different enzymatic properties', *J Mol Biol*, vol. 303, no. 5, 2000, doi: 10.1006/jmbi.2000.4185.
- [46] E. Emmott, M. Jovanovic, and N. Slavov, 'Ribosome Stoichiometry: From Form to Function', *Trends in Biochemical Sciences*, vol. 44, no. 2. 2019. doi: 10.1016/j.tibs.2018.10.009.
- [47] Y. Liu, A. Beyer, and R. Aebersold, 'On the Dependency of Cellular Protein Levels on mRNA Abundance', *Cell*, vol. 165, no. 3. 2016. doi: 10.1016/j.cell.2016.03.014.
- [48] E. M. Torres et al., 'Effects of aneuploidy on cellular physiology and cell division in haploid yeast', *Science (1979)*, vol. 317, no. 5840, 2007, doi: 10.1126/science.1142210.
- [49] T. Geiger, J. Cox, and M. Mann, 'Proteomic changes resulting from gene copy number variations in cancer cells', *PLoS Genet*, vol. 6, no. 9, 2010, doi: 10.1371/journal.pgen.1001090.
- [50] S. Stingele, G. Stoehr, K. Peplowska, J. Cox, M. Mann, and Z. Storchova, 'Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells', *Mol Syst Biol*, vol. 8, 2012, doi: 10.1038/msb.2012.40.
- [51] N. Dephoure et al., 'Quantitative proteomic analysis reveals posttranslational responses to

- aneuploidy in yeast', *Elife*, vol. 3, no. July2014, 2014, doi: 10.7554/eLife.03023.
- [52] E. Gonçalves, A. Fragoulis, L. Garcia-Alonso, T. Cramer, J. Saez-Rodriguez, and P. Beltrao, 'Widespread Post-transcriptional Attenuation of Genomic Copy-Number Variation in Cancer', *Cell Syst*, vol. 5, no. 4, 2017, doi: 10.1016/j.cels.2017.08.013.
 - [53] K. Ishikawa, K. Makanae, S. Iwasaki, N. T. Ingolia, and H. Moriya, 'Post-Translational Dosage Compensation Buffers Genetic Perturbations to Stoichiometry of Protein Complexes', *PLoS Genet*, vol. 13, no. 1, 2017, doi: 10.1371/journal.pgen.1006554.
 - [54] W. R. Blevins et al., 'Extensive post-transcriptional buffering of gene expression in the response to severe oxidative stress in baker's yeast', *Sci Rep*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-47424-w.
 - [55] E. M. Teyssonniere, Y. Shichino, M. Mito, A. Friedrich, S. Iwasaki, and J. Schacherer, 'Translation variation across genetic backgrounds reveals a post-transcriptional buffering signature in yeast', *Nucleic Acids Res*, Jan. 2024, doi: 10.1093/nar/gkae030.
 - [56] E. P. Kusnadi, C. Timpone, I. Topisirovic, O. Larsson, and L. Furic, 'Regulation of gene expression via translational buffering', *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1869, no. 1. 2022. doi: 10.1016/j.bbamcr.2021.119140.
 - [57] J. Lorent et al., 'Translational offsetting as a mode of estrogen receptor α -dependent regulation of gene expression', *EMBO J*, vol. 38, no. 23, 2019, doi: 10.15252/embj.2018101323.
 - [58] J. C. Taggart, H. Zaubert, M. Selbach, G. W. Li, and E. McShane, 'Keeping the Proportions of Protein Complex Components in Check', *Cell Systems*, vol. 10, no. 2. 2020. doi: 10.1016/j.cels.2020.01.004.
 - [59] R. T. Schimke and D. Doyle, 'Control of enzyme levels in animal tissues.', *Annual review of biochemistry*, vol. 39. 1970. doi: 10.1146/annurev.bi.39.070170.004433.
 - [60] U. Schubert, L. C. Antón, J. Gibbs, C. C. Norbury, J. W. Yewdell, and J. R. Bennink, 'Rapid degradation of a large fraction of newly synthesized proteins by proteasomes', *Nature*, vol. 404, no. 6779, 2000, doi: 10.1038/35008096.
 - [61] E. McShane et al., 'Kinetic Analysis of Protein Stability Reveals Age-Dependent Degradation', *Cell*, vol. 167, no. 3, pp. 803-815.e21, Oct. 2016, doi: 10.1016/j.cell.2016.09.015.
 - [62] M. K. Sung, J. M. Reitsma, M. J. Sweredoski, S. Hess, and R. J. Deshaies, 'Ribosomal proteins produced in excess are degraded by the ubiquitin-proteasome system', *Mol Biol Cell*, vol. 27, no. 17, 2016, doi: 10.1091/mbc.E16-05-0290.
 - [63] K. L. Hsu, H. C. S. Yen, and C. H. Yeang, 'Cooperative stability renders protein complex formation more robust and controllable', *Sci Rep*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-14362-z.

References

- [64] K. Husnjak and I. Dikic, 'Ubiquitin-binding proteins: Decoders of ubiquitin-mediated cellular functions', *Annu Rev Biochem*, vol. 81, 2012, doi: 10.1146/annurev-biochem-051810-094654.
- [65] N. Zheng and N. Shabek, 'Ubiquitin ligases: Structure, function, and regulation', *Annual Review of Biochemistry*, vol. 86, 2017. doi: 10.1146/annurev-biochem-060815-014922.
- [66] J. Peng et al., 'A proteomics approach to understanding protein ubiquitination', *Nat Biotechnol*, vol. 21, no. 8, 2003, doi: 10.1038/nbt849.
- [67] D. Komander and M. Rape, 'The ubiquitin code', *Annu Rev Biochem*, vol. 81, 2012, doi: 10.1146/annurev-biochem-060310-170328.
- [68] E. B. Dammer et al., 'Polyubiquitin linkage profiles in three models of proteolytic stress suggest the etiology of alzheimer disease', *Journal of Biological Chemistry*, vol. 286, no. 12, 2011, doi: 10.1074/jbc.M110.149633.
- [69] D. Finley, 'Recognition and processing of ubiquitin-protein conjugates by the proteasome', *Annual Review of Biochemistry*, vol. 78, 2009. doi: 10.1146/annurev.biochem.78.081507.101607.
- [70] D. L. Haakonsen and M. Rape, 'Branching Out: Improved Signaling by Heterotypic Ubiquitin Chains', *Trends in Cell Biology*, vol. 29, no. 9, 2019. doi: 10.1016/j.tcb.2019.06.003.
- [71] M. E. French, C. F. Koehler, and T. Hunter, 'Emerging functions of branched ubiquitin chains', *Cell Discovery*, vol. 7, no. 1, 2021. doi: 10.1038/s41421-020-00237-y.
- [72] W. Li et al., 'Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling', *PLoS One*, vol. 3, no. 1, 2008, doi: 10.1371/journal.pone.0001487.
- [73] A. Varshavsky, 'N-degron and C-degron pathways of protein degradation', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 2, 2019. doi: 10.1073/pnas.1816596116.
- [74] K. Y. E. Kong, J. P. L. Coelho, M. J. Feige, and A. Khmelinskii, 'Quality control of mislocalized and orphan proteins', *Exp Cell Res*, vol. 403, no. 2, 2021, doi: 10.1016/j.yexcr.2021.112617.
- [75] C. Padovani, P. Jevtić, and M. Rapé, 'Quality control of protein complex composition', *Molecular Cell*, vol. 82, no. 8, 2022. doi: 10.1016/j.molcel.2022.02.029.
- [76] C. Pla-Prats and N. H. Thomä, 'Quality control of protein complex assembly by the ubiquitin-proteasome system', *Trends in Cell Biology*, vol. 32, no. 8, 2022. doi: 10.1016/j.tcb.2022.02.005.
- [77] T. Gilon, O. Chomsky, and R. G. Kulka, 'Degradation signals for ubiquitin system proteolysis in *Saccharomyces cerevisiae*', *EMBO Journal*, vol. 17, no. 10, 1998, doi: 10.1093/emboj/17.10.2759.

- [78] T. Hessa, A. Sharma, M. Mariappan, H. D. Eshleman, E. Gutierrez, and R. S. Hegde, 'Protein targeting and degradation are coupled for elimination of mislocalized proteins', *Nature*, vol. 475, no. 7356, 2011, doi: 10.1038/nature10181.
- [79] M. C. Rodrigo-Brenni, E. Gutierrez, and R. S. Hegde, 'Cytosolic Quality Control of Mislocalized Proteins Requires RNF126 Recruitment to Bag6', *Mol Cell*, vol. 55, no. 2, 2014, doi: 10.1016/j.molcel.2014.05.025.
- [80] A. T. Nguyen et al., 'UBE2O remodels the proteome during terminal erythroid differentiation', *Science* (1979), vol. 357, no. 6350, 2017, doi: 10.1126/science.aan0218.
- [81] K. Yanagitani, S. Juszkiwicz, and R. S. Hegde, 'UBE2O is a quality control factor for orphans of multiprotein complexes', *Science* (1979), vol. 357, no. 6350, 2017, doi: 10.1126/science.aan0178.
- [82] D. B. Grabarczyk et al., 'HUWE1 employs a giant substrate-binding ring to feed and regulate its HECT E3 domain', *Nat Chem Biol*, vol. 17, no. 10, 2021, doi: 10.1038/s41589-021-00831-5.
- [83] T. Arnesen et al., 'Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans', *Proc Natl Acad Sci U S A*, vol. 106, no. 20, 2009, doi: 10.1073/pnas.0901931106.
- [84] H. Aksnes, R. Ree, and T. Arnesen, 'Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases', *Molecular Cell*, vol. 73, no. 6, 2019, doi: 10.1016/j.molcel.2019.02.007.
- [85] C. S. Hwang, A. Shemorry, and A. Varshavsky, 'N-terminal acetylation of cellular proteins creates specific degradation signals', *Science* (1979), vol. 327, no. 5968, 2010, doi: 10.1126/science.1183147.
- [86] A. Shemorry, C. S. Hwang, and A. Varshavsky, 'Control of Protein Quality and Stoichiometries by N-Terminal Acetylation and the N-End Rule Pathway', *Mol Cell*, vol. 50, no. 4, 2013, doi: 10.1016/j.molcel.2013.03.018.
- [87] A. Ori et al., 'Integrated Transcriptome and Proteome Analyses Reveal Organ-Specific Proteome Deterioration in Old Rats', *Cell Syst*, vol. 1, no. 3, 2015, doi: 10.1016/j.cels.2015.08.012.
- [88] D. M. Walther et al., 'Widespread proteome remodeling and aggregation in aging *C. elegans*', *Cell*, vol. 161, no. 4, 2015, doi: 10.1016/j.cell.2015.03.032.
- [89] G. R. Keele et al., 'Global and tissue-specific aging effects on murine proteomes', *Cell Rep*, vol. 42, no. 7, 2023, doi: 10.1016/j.celrep.2023.112715.
- [90] M. Leutert et al., 'Multidimensional proteomics identifies molecular trajectories of cellular aging and rejuvenation', *bioRxiv*, 2023.

References

- [91] M. A. D'Angelo, M. Raices, S. H. Panowski, and M. W. Hetzer, 'Age-Dependent Deterioration of Nuclear Pore Complexes Causes a Loss of Nuclear Integrity in Postmitotic Cells', *Cell*, vol. 136, no. 2, 2009, doi: 10.1016/j.cell.2008.11.037.
- [92] B. H. Toyama et al., 'Identification of long-lived proteins reveals exceptional stability of essential cellular structures', *Cell*, vol. 154, no. 5, 2013, doi: 10.1016/j.cell.2013.07.037.
- [93] G. E. Janssens et al., 'Protein biogenesis machinery is a driver of replicative aging in yeast', *Elife*, vol. 4, no. September 2015, 2015, doi: 10.7554/eLife.08527.
- [94] M. S. Hipp, P. Kasturi, and F. U. Hartl, 'The proteostasis network and its decline in ageing', *Nature Reviews Molecular Cell Biology*, vol. 20, no. 7. 2019. doi: 10.1038/s41580-019-0101-y.
- [95] C. J. Ryan, S. Kennedy, I. Bajrami, D. Matallanas, and C. J. Lord, 'A Compendium of Co-regulated Protein Complexes in Breast Cancer Reveals Collateral Loss Events', *Cell Syst*, vol. 5, no. 4, 2017, doi: 10.1016/j.cels.2017.09.011.
- [96] I. De et al., 'The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes', *Nat Struct Mol Biol*, vol. 22, no. 2, 2015, doi: 10.1038/nsmb.2951.
- [97] R. Yonemasu et al., 'Disruption of mouse XAB2 gene involved in pre-mRNA splicing, transcription and transcription-coupled DNA repair results in preimplantation lethality', *DNA Repair (Amst)*, vol. 4, no. 4, 2005, doi: 10.1016/j.dnarep.2004.12.004.
- [98] S. P. Chan, D. I. Kao, W. Y. Tsai, and S. C. Cheng, 'The Prp19p-associated complex in spliceosome activation', *Science (1979)*, vol. 302, no. 5643, 2003, doi: 10.1126/science.1086602.
- [99] O. V Makarova et al., 'A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing', *EMBO J*, vol. 23, no. 12, pp. 2381–2391, Jun. 2004, doi: 10.1038/sj.emboj.7600241.
- [100] C. H. Chen et al., 'Functional and physical interactions between components of the Prp19p-associated complex', *Nucleic Acids Research*, vol. 30, no. 4. 2002. doi: 10.1093/nar/30.4.1029.
- [101] M. D. Ohi, A. J. Link, L. Ren, J. L. Jennings, W. H. McDonald, and K. L. Gould, 'Proteomics Analysis Reveals Stable Multiprotein Complexes in Both Fission and Budding Yeasts Containing Myb-Related Cdc5p/Cef1p, Novel Pre-mRNA Splicing Factors, and snRNAs', *Mol Cell Biol*, vol. 22, no. 7, 2002, doi: 10.1128/mcb.22.7.2011-2024.2002.
- [102] J. Schmitzová, C. Cretu, C. Dienemann, H. Urlaub, and V. Pena, 'Structural basis of catalytic activation in human splicing', *Nature*, vol. 617, no. 7962, pp. 842–850, May 2023, doi: 10.1038/s41586-023-06049-w.
- [103] M. E. Fairman-Williams, U.-P. Guenther, and E. Jankowsky, 'SF1 and SF2 helicases: family matters',

- Curr Opin Struct Biol, vol. 20, no. 3, pp. 313–324, Jun. 2010, doi: 10.1016/j.sbi.2010.03.011.
- [104] M. A. Andrade, C. Petosa, S. I. O'Donoghue, C. W. Müller, and P. Bork, 'Comparison of ARM and HEAT protein repeats', J Mol Biol, vol. 309, no. 1, pp. 1–18, May 2001, doi: 10.1006/jmbi.2001.4624.
 - [105] A. H. Huber, W. J. Nelson, and W. I. Weis, 'Three-Dimensional Structure of the Armadillo Repeat Region of β -Catenin', Cell, vol. 90, no. 5, pp. 871–882, Sep. 1997, doi: 10.1016/S0092-8674(00)80352-9.
 - [106] T. Hirose, T. Ideue, M. Nagai, M. Hagiwara, M.-D. Shu, and J. A. Steitz, 'A Spliceosomal Intron Binding Protein, IBP160, Links Position-Dependent Assembly of Intron-Encoded Box C/D snoRNP to Pre-mRNA Splicing', Mol Cell, vol. 23, no. 5, pp. 673–684, Sep. 2006, doi: 10.1016/j.molcel.2006.07.011.
 - [107] H. Le Hir, 'The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions', EMBO J, vol. 19, no. 24, pp. 6860–6869, Dec. 2000, doi: 10.1093/emboj/19.24.6860.
 - [108] H. Le Hir, J. Saulière, and Z. Wang, 'The exon junction complex as a node of post-transcriptional networks', Nat Rev Mol Cell Biol, vol. 17, no. 1, pp. 41–54, Jan. 2016, doi: 10.1038/nrm.2015.7.
 - [109] T. Ideue, Y. T. F. Sasaki, M. Hagiwara, and T. Hirose, 'Introns play an essential role in splicing-dependent formation of the exon junction complex', Genes Dev, vol. 21, no. 16, pp. 1993–1998, Aug. 2007, doi: 10.1101/gad.1557907.
 - [110] M. Shiimori, K. Inoue, and H. Sakamoto, 'A Specific Set of Exon Junction Complex Subunits Is Required for the Nuclear Retention of Unspliced RNAs in *Caenorhabditis elegans*', Mol Cell Biol, vol. 33, no. 2, pp. 444–456, Jan. 2013, doi: 10.1128/MCB.01298-12.
 - [111] T. Hirose, M.-D. Shu, and J. A. Steitz, 'Splicing-Dependent and -Independent Modes of Assembly for Intron-Encoded Box C/D snoRNPs in Mammalian Cells', Mol Cell, vol. 12, no. 1, pp. 113–123, Jul. 2003, doi: 10.1016/S1097-2765(03)00267-3.
 - [112] G. Dieci, M. Preti, and B. Montanini, 'Eukaryotic snoRNAs: A paradigm for gene expression flexibility', Genomics, vol. 94, no. 2, pp. 83–88, Aug. 2009, doi: 10.1016/j.ygeno.2009.05.002.
 - [113] J. Eralles et al., 'Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes', Proceedings of the National Academy of Sciences, vol. 114, no. 49, pp. 12934–12939, Dec. 2017, doi: 10.1073/pnas.1707674114.
 - [114] S. Khoshnevis, R. E. Dreggors-Walker, V. Marchand, Y. Motorin, and H. Ghalei, 'Ribosomal RNA 2'-O-methylations regulate translation by impacting ribosome dynamics', Proceedings of the National Academy of Sciences, vol. 119, no. 12, Mar. 2022, doi: 10.1073/pnas.2117334119.
 - [115] L. Badertscher et al., 'Genome-wide RNAi Screening Identifies Protein Modules Required for 40S Subunit Synthesis in Human Cells', Cell Rep, vol. 13, no. 12, pp. 2879–2891, Dec. 2015, doi:

10.1016/j.celrep.2015.11.061.

- [116] C. Niehrs and B. Luke, 'Regulatory R-loops as facilitators of gene expression and genome stability', *Nature Reviews Molecular Cell Biology*, vol. 21, no. 3. 2020. doi: 10.1038/s41580-019-0206-3.
- [117] A. Aguilera and T. García-Muse, 'R Loops: From Transcription Byproducts to Threats to Genome Stability', *Molecular Cell*, vol. 46, no. 2. 2012. doi: 10.1016/j.molcel.2012.04.009.
- [118] R. D. Paulsen et al., 'A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability', *Mol Cell*, vol. 35, no. 2, pp. 228–239, Jul. 2009, doi: 10.1016/j.molcel.2009.06.021.
- [119] J. Sollier, C. T. Stork, M. L. García-Rubio, R. D. Paulsen, A. Aguilera, and K. A. Cimprich, 'Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability', *Mol Cell*, vol. 56, no. 6, pp. 777–785, Dec. 2014, doi: 10.1016/j.molcel.2014.10.020.
- [120] R. Sakasai et al., 'Aquarius is required for proper CtIP expression and homologous recombination repair', *Sci Rep*, vol. 7, no. 1, p. 13808, Oct. 2017, doi: 10.1038/s41598-017-13695-4.
- [121] T. Mosler et al., 'R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability', *Nat Commun*, vol. 12, no. 1, p. 7314, Dec. 2021, doi: 10.1038/s41467-021-27530-y.
- [122] S. Dehghani-Tafti and C. M. Sanders, 'DNA substrate recognition and processing by the full-length human UPF1 helicase', *Nucleic Acids Res*, vol. 45, no. 12, pp. 7354–7366, Jul. 2017, doi: 10.1093/nar/gkx478.
- [123] G. H. P. Ngo, J. W. Grimstead, and D. M. Baird, 'UPF1 promotes the formation of R loops to stimulate DNA double-strand break repair', *Nat Commun*, vol. 12, no. 1, p. 3849, Jun. 2021, doi: 10.1038/s41467-021-24201-w.
- [124] Z. Hasanova, V. Klapstova, O. Porrua, R. Stefl, and M. Sebesta, 'Human senataxin is a bona fide R-loop resolving enzyme and transcription termination factor', *Nucleic Acids Res*, vol. 51, no. 6, pp. 2818–2837, Apr. 2023, doi: 10.1093/nar/gkad092.
- [125] E. L. Van Nostrand et al., 'A large-scale binding and functional map of human RNA-binding proteins', *Nature*, vol. 583, no. 7818, pp. 711–719, Jul. 2020, doi: 10.1038/s41586-020-2077-3.
- [126] A. K. Das, 'The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions', *EMBO J*, vol. 17, no. 5, pp. 1192–1199, Mar. 1998, doi: 10.1093/emboj/17.5.1192.
- [127] P. J. Preker and W. Keller, 'The HAT helix, a repetitive motif implicated in RNA processing', *Trends in Biochemical Sciences*, vol. 23, no. 1. 1998. doi: 10.1016/S0968-0004(97)01156-0.

- [128] Y. Nakatsu et al., 'XAB2, a novel tetratricopeptide repeat protein involved in transcription-coupled DNA repair and transcription', *Journal of Biological Chemistry*, vol. 275, no. 45, 2000, doi: 10.1074/jbc.M004936200.
- [129] D. O. Onyango, S. M. Howard, K. Neherin, D. A. Yanez, and J. M. Stark, 'Tetratricopeptide repeat factor XAB2 mediates the end resection step of homologous recombination.', *Nucleic Acids Res*, vol. 44, no. 12, pp. 5702–16, Jul. 2016, doi: 10.1093/nar/gkw275.
- [130] A. B. Sharma et al., 'XAB2 promotes Ku eviction from single-ended DNA double-strand breaks independently of the ATM kinase.', *Nucleic Acids Res*, vol. 49, no. 17, pp. 9906–9925, Sep. 2021, doi: 10.1093/nar/gkab785.
- [131] W. H. McDonald, R. Ohi, N. Smelkova, D. Frendewey, and K. L. Gould, 'Myb-Related Fission Yeast cdc5p Is a Component of a 40S snRNP-Containing Complex and Is Essential for Pre-mRNA Splicing', *Mol Cell Biol*, vol. 19, no. 8, pp. 5352–5362, Aug. 1999, doi: 10.1128/MCB.19.8.5352.
- [132] P. Fabrizio et al., 'The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome', *Mol Cell*, vol. 36, no. 4, pp. 593–608, Nov. 2009, doi: 10.1016/j.molcel.2009.09.040.
- [133] I. Kuraoka et al., 'Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair', *Journal of Biological Chemistry*, vol. 283, no. 2, 2008, doi: 10.1074/jbc.M706647200.
- [134] J. Deckert et al., 'Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions.', *Mol Cell Biol*, vol. 26, no. 14, pp. 5528–43, Jul. 2006, doi: 10.1128/MCB.00582-06.
- [135] S. Bessonov et al., 'Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis', *RNA*, vol. 16, no. 12, 2010, doi: 10.1261/rna.2456210.
- [136] D. Haselbach et al., 'Structure and Conformational Dynamics of the Human Spliceosomal Bact Complex', *Cell*, vol. 172, no. 3, 2018, doi: 10.1016/j.cell.2018.01.010.
- [137] X. Zhang, C. Yan, X. Zhan, L. Li, J. Lei, and Y. Shi, 'Structure of the human activated spliceosome in three conformational states.', *Cell Res*, vol. 28, no. 3, pp. 307–322, Mar. 2018, doi: 10.1038/cr.2018.14.
- [138] X. Zhan, C. Yan, X. Zhang, J. Lei, and Y. Shi, 'Structure of a human catalytic step I spliceosome.', *Science*, vol. 359, no. 6375, pp. 537–545, Feb. 2018, doi: 10.1126/science.aar6401.
- [139] X. Zhang, C. Yan, J. Hang, L. I. Finci, J. Lei, and Y. Shi, 'An Atomic Structure of the Human Spliceosome.', *Cell*, vol. 169, no. 5, pp. 918–929.e14, May 2017, doi: 10.1016/j.cell.2017.04.033.

References

- [140] S. Hou et al., 'XAB2 depletion induces intron retention in POLR2A to impair global transcription and promote cellular senescence,' *Nucleic Acids Res*, vol. 47, no. 15, pp. 8239–8254, Sep. 2019, doi: 10.1093/nar/gkz532.
- [141] E. Goulielmaki et al., 'The splicing factor XAB2 interacts with ERCC1-XPF and XPG for R-loop processing,' *Nat Commun*, vol. 12, no. 1, p. 3153, May 2021, doi: 10.1038/s41467-021-23505-1.
- [142] I. Dix, C. Russell, S. B. Yehuda, M. Kupiec, and J. D. Beggs, 'The identification and characterization of a novel splicing protein, Isy1p, of *Saccharomyces cerevisiae*,' *RNA*, vol. 5, no. 3, pp. 360–8, Mar. 1999, doi: 10.1017/s1355838299981396.
- [143] T. Villa and C. Guthrie, 'The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing,' *Genes Dev*, vol. 19, no. 16, pp. 1894–904, Aug. 2005, doi: 10.1101/gad.1336305.
- [144] P. Du, L. Wang, P. Sliz, and R. I. Gregory, 'A Biogenesis Step Upstream of Microprocessor Controls miR-17~92 Expression,' *Cell*, vol. 162, no. 4, pp. 885–99, Aug. 2015, doi: 10.1016/j.cell.2015.07.008.
- [145] P. Du et al., 'An Intermediate Pluripotent State Controlled by MicroRNAs Is Required for the Naive-to-Primed Stem Cell Transition,' *Cell Stem Cell*, vol. 22, no. 6, pp. 851–864.e5, Jun. 2018, doi: 10.1016/j.stem.2018.04.021.
- [146] A. S. Jaiswal et al., 'The splicing component ISY1 regulates APE1 in base excision repair,' *DNA Repair (Amst)*, vol. 86, p. 102769, Feb. 2020, doi: 10.1016/j.dnarep.2019.102769.
- [147] C.-L. Yuan et al., 'APE1 overexpression is associated with poor survival in patients with solid tumors: a meta-analysis,' *Oncotarget*, vol. 8, no. 35, pp. 59720–59728, Aug. 2017, doi: 10.18632/oncotarget.19814.
- [148] M. Li et al., 'APE1 deficiency promotes cellular senescence and premature aging features,' *Nucleic Acids Res*, vol. 46, no. 11, pp. 5664–5677, Jun. 2018, doi: 10.1093/nar/gky326.
- [149] G. Fischer, H. Bang, and C. Mech, '[Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides],' *Biomed Biochim Acta*, vol. 43, no. 10, pp. 1101–11, 1984.
- [150] A. Matouschek, S. Rospert, K. Schmid, B. S. Glick, and G. Schatz, 'Cyclophilin catalyzes protein folding in yeast mitochondria,' *Proc Natl Acad Sci U S A*, vol. 92, no. 14, pp. 6319–6323, Jul. 1995, doi: 10.1073/pnas.92.14.6319.
- [151] S. F. Göthel and M. A. Marahiel, 'Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts,' *Cell Mol Life Sci*, vol. 55, no. 3, pp. 423–436, Mar. 1999, doi: 10.1007/s000180050299.

- [152] H. Mi, O. Kops, E. Zimmermann, A. Jäschke, and M. Tropschug, 'A nuclear RNA-binding cyclophilin in human T cells', *FEBS Lett*, vol. 398, no. 2–3, pp. 201–205, Dec. 1996, doi: 10.1016/S0014-5793(96)01248-3.
- [153] Y. Wang et al., 'Human CyP33 binds specifically to mRNA and binding stimulates PPIase activity of hCyP33', *FEBS Lett*, vol. 582, no. 5, pp. 835–839, Mar. 2008, doi: 10.1016/j.febslet.2008.01.055.
- [154] N. R. Lloyd and D. S. Wuttke, 'Cyp33 binds AU-rich RNA motifs via an extended interface that competitively disrupts the gene repressive Cyp33-MLL1 interaction in vitro.', *PLoS One*, vol. 16, no. 2, p. e0237956, 2021, doi: 10.1371/journal.pone.0237956.
- [155] M. Blatter et al., 'RNA binding induces an allosteric switch in Cyp33 to repress MLL1-mediated transcription.', *Sci Adv*, vol. 9, no. 16, p. eadf5330, Apr. 2023, doi: 10.1126/sciadv.adf5330.
- [156] J. Artus, S. Vandormael-Pournin, M. Frödin, K. Nacerddine, C. Babinet, and M. Cohen-Tannoudji, 'Impaired mitotic progression and preimplantation lethality in mice lacking OMCG1, a new evolutionarily conserved nuclear protein.', *Mol Cell Biol*, vol. 25, no. 14, pp. 6289–302, Jul. 2005, doi: 10.1128/MCB.25.14.6289-6302.2005.
- [157] T. Léguillier et al., 'Omcg1 is critically required for mitosis in rapidly dividing mouse intestinal progenitors and embryonic stem cells.', *Biol Open*, vol. 1, no. 7, pp. 648–57, Jul. 2012, doi: 10.1242/bio.20121248.
- [158] G. Chen et al., 'ZNF830 mediates cancer chemoresistance through promoting homologous-recombination repair', *Nucleic Acids Res*, vol. 46, no. 3, pp. 1266–1279, Feb. 2018, doi: 10.1093/nar/gkx1258.
- [159] Z. Qiu et al., 'PLK1-mediated phosphorylation of PPIL2 regulates HR via CtIP', *Front Cell Dev Biol*, vol. 10, p. 902403, 2022, doi: 10.3389/fcell.2022.902403.
- [160] D. P. Nusinow et al., 'Quantitative Proteomics of the Cancer Cell Line Encyclopedia', *Cell*, vol. 180, no. 2, 2020, doi: 10.1016/j.cell.2019.12.023.
- [161] M. Ghandi et al., 'Next-generation characterization of the Cancer Cell Line Encyclopedia', *Nature*, vol. 569, no. 7757, 2019, doi: 10.1038/s41586-019-1186-3.
- [162] G. Senger, S. Santaguida, and M. H. Schaefer, 'Regulation of protein complex partners as a compensatory mechanism in aneuploid tumors', *Elife*, vol. 11, 2022, doi: 10.7554/eLife.75526.
- [163] A. L. Goldberg, 'Development of proteasome inhibitors as research tools and cancer drugs', *Journal of Cell Biology*, vol. 199, no. 4, 2012, doi: 10.1083/jcb.201210077.
- [164] Y. Chen, X. X. Sun, R. C. Sears, and M. S. Dai, 'Writing and erasing MYC ubiquitination and SUMOylation', *Genes and Diseases*, vol. 6, no. 4, 2019, doi: 10.1016/j.gendis.2019.05.006.

References

- [165] C. Mauvezin and T. P. Neufeld, 'Bafilomycin A1 disrupts autophagic flux by inhibiting both V-ATPase-dependent acidification and Ca-P60A/SERCA-dependent autophagosome-lysosome fusion', *Autophagy*, vol. 11, no. 8, 2015, doi: 10.1080/15548627.2015.1066957.
- [166] A. Yamamoto, Y. Tagawa, T. Yoshimori, Y. Moriyama, R. Masaki, and Y. Tashiro, 'Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells', *Cell Struct Funct*, vol. 23, no. 1, 1998, doi: 10.1247/csf.23.33.
- [167] R. Huang et al., 'Deacetylation of nuclear LC3 drives autophagy initiation under starvation', *Mol Cell*, vol. 57, no. 3, pp. 456–466, Feb. 2015, doi: 10.1016/j.molcel.2014.12.013.
- [168] N. Mizushima et al., 'A protein conjugation system essential for autophagy', *Nature*, vol. 395, no. 6700, 1998, doi: 10.1038/26506.
- [169] N. Mizushima, T. Yoshimori, and B. Levine, 'Methods in Mammalian Autophagy Research', *Cell*, vol. 140, no. 3, 2010. doi: 10.1016/j.cell.2010.01.028.
- [170] V. Kirkin and V. V. Rogov, 'A Diversity of Selective Autophagy Receptors Determines the Specificity of the Autophagy Pathway', *Molecular Cell*, vol. 76, no. 2, 2019. doi: 10.1016/j.molcel.2019.09.005.
- [171] T. Schneider-Poetsch et al., 'Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin', *Nat Chem Biol*, vol. 6, no. 3, 2010, doi: 10.1038/nchembio.304.
- [172] P. V. Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham, and E. Skrzypek, 'PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations', *Nucleic Acids Res*, vol. 43, no. D1, 2015, doi: 10.1093/nar/gku1267.
- [173] B. R. Keppler and T. K. Archer, 'Ubiquitin-dependent and Ubiquitin-independent Control of Subunit Stoichiometry in the SWI/SNF Complex', *Journal of Biological Chemistry*, vol. 285, no. 46, pp. 35665–35674, Nov. 2010, doi: 10.1074/jbc.M110.173997.
- [174] Y. Park, S. K. Yoon, and J. B. Yoon, 'TRIP12 functions as an E3 ubiquitin ligase of APP-BP1', *Biochem Biophys Res Commun*, vol. 374, no. 2, 2008, doi: 10.1016/j.bbrc.2008.07.019.
- [175] I. I. C. Chio et al., 'TRADD contributes to tumour suppression by regulating ULF-dependent p19Arf ubiquitylation', *Nat Cell Biol*, vol. 14, no. 6, pp. 625–633, Jun. 2012, doi: 10.1038/ncb2496.
- [176] D. Chen, J. Shan, W. G. Zhu, J. Qin, and W. Gu, 'Transcription-independent ARF regulation in oncogenic stress-mediated p53 responses', *Nature*, vol. 464, no. 7288, 2010, doi: 10.1038/nature08820.
- [177] N. Hah et al., 'A role for BAF57 in cell cycle-dependent transcriptional regulation by the SWI/SNF chromatin remodeling complex', *Cancer Res*, vol. 70, no. 11, 2010, doi: 10.1158/0008-5472.CAN-09-2767.

- [178] E. Matta-Camacho, G. Kozlov, F. F. Li, and K. Gehring, 'Structural basis of substrate recognition and specificity in the N-end rule pathway', *Nat Struct Mol Biol*, vol. 17, no. 10, pp. 1182–1187, Oct. 2010, doi: 10.1038/nsmb.1894.
- [179] T. Gudjonsson et al., 'TRIP12 and UBR5 Suppress Spreading of Chromatin Ubiquitylation at Damaged Chromosomes', *Cell*, vol. 150, no. 4, pp. 697–709, Aug. 2012, doi: 10.1016/j.cell.2012.06.039.
- [180] A. Kaiho-Soma et al., 'TRIP12 promotes small-molecule-induced degradation through K29/K48-branched ubiquitin chains', *Mol Cell*, vol. 81, no. 7, pp. 1411–1424.e7, Apr. 2021, doi: 10.1016/j.molcel.2021.01.023.
- [181] C. Liu, W. Liu, Y. Ye, and W. Li, 'Ufd2p synthesizes branched ubiquitin chains to promote the degradation of substrates modified with atypical chains', *Nat Commun*, vol. 8, 2017, doi: 10.1038/ncomms14274.
- [182] R. M. Kortlever, P. J. Higgins, and R. Bernards, 'Plasminogen activator inhibitor-1 is a critical downstream target of p53 in the induction of replicative senescence', *Nat Cell Biol*, vol. 8, no. 8, 2006, doi: 10.1038/ncb1448.
- [183] Y. Y. Lee, H. S. Kim, and I. K. Lim, 'Downregulation of PEA-15 reverses G1 arrest, and nuclear and chromatin changes of senescence phenotype via pErk1/2 translocation to nuclei', *Cell Signal*, vol. 27, no. 6, 2015, doi: 10.1016/j.cellsig.2015.02.014.
- [184] N. Basisty et al., 'A proteomic atlas of senescence-associated secretomes for aging biomarker development', *PLoS Biol*, vol. 18, no. 1, 2020, doi: 10.1371/journal.pbio.3000599.
- [185] A. Moretton and J. I. Loizou, 'Interplay between cellular metabolism and the dna damage response in cancer', *Cancers*, vol. 12, no. 8. 2020. doi: 10.3390/cancers12082051.
- [186] J. Cox and M. Mann, '1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data.', *BMC Bioinformatics*, vol. 13 Suppl 16, 2012, doi: 10.1186/1471-2105-13-S16-S12.
- [187] B. Schwalb et al., 'TT-seq maps the human transient transcriptome', *Science* (1979), vol. 352, no. 6290, pp. 1225–1228, Jun. 2016, doi: 10.1126/science.aad9841.
- [188] L. H. Gregersen, R. Mitter, and J. Q. Svejstrup, 'Using TTchem-seq for profiling nascent transcription and measuring transcript elongation', *Nat Protoc*, vol. 15, no. 2, pp. 604–627, Feb. 2020, doi: 10.1038/s41596-019-0262-3.
- [189] S. Cugusi et al., 'Heat shock induces premature transcript termination and reconfigures the human transcriptome', *Mol Cell*, vol. 82, no. 8, 2022, doi: 10.1016/j.molcel.2022.01.007.
- [190] D. Mahdessian et al., 'Spatiotemporal dissection of the cell cycle with single-cell proteogenomics',

References

- Nature, vol. 590, no. 7847, 2021, doi: 10.1038/s41586-021-03232-9.
- [191] M. Tanaka et al., 'Cell-cycle-dependent Regulation of Human aurora A Transcription Is Mediated by Periodic Repression of E4TF1', *Journal of Biological Chemistry*, vol. 277, no. 12, pp. 10719–10726, Mar. 2002, doi: 10.1074/jbc.M108252200.
- [192] J. Loughery, M. Cox, L. M. Smith, and D. W. Meek, 'Critical role for p53-serine 15 phosphorylation in stimulating transactivation at p53-responsive promoters', *Nucleic Acids Res*, vol. 42, no. 12, 2014, doi: 10.1093/nar/gku501.
- [193] N. J. MacLaine and T. R. Hupp, 'The regulation of p53 by phosphorylation: a model for how distinct signals integrate into the p53 pathway', *Aging*, vol. 1, no. 5, 2009, doi: 10.18632/aging.100047.
- [194] T. Abbas and A. Dutta, 'P21 in cancer: Intricate networks and multiple activities', *Nature Reviews Cancer*, vol. 9, no. 6. 2009. doi: 10.1038/nrc2657.
- [195] A. Hernandez-Segura, J. Nehme, and M. Demaria, 'Hallmarks of Cellular Senescence', *Trends in Cell Biology*, vol. 28, no. 6. 2018. doi: 10.1016/j.tcb.2018.02.001.
- [196] B. G. Childs, D. J. Baker, J. L. Kirkland, J. Campisi, and J. M. van Deursen, 'Senescence and apoptosis: dueling or complementary cell fates?', *EMBO Rep*, vol. 15, no. 11, 2014, doi: 10.15252/embr.201439245.
- [197] E. Boye and B. Grallert, 'eIF2 α phosphorylation and the regulation of translation', *Curr Genet*, vol. 66, no. 2, pp. 293–297, Apr. 2020, doi: 10.1007/s00294-019-01026-1.
- [198] E. Nicolas, P. Parisot, C. Pinto-Monteiro, R. De Walque, C. De Vleeschouwer, and D. L. J. Lafontaine, 'Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress', *Nat Commun*, vol. 7, 2016, doi: 10.1038/ncomms11390.
- [199] P. Moudry, K. Chroma, S. Bursac, S. Volarevic, and J. Bartek, 'RNA-interference screen for p53 regulators unveils a role of WDR75 in ribosome biogenesis', *Cell Death Differ*, vol. 29, no. 3, 2022, doi: 10.1038/s41418-021-00882-0.
- [200] Z. Turi, M. Senkyrikova, M. Mistrik, J. Bartek, and P. Moudry, 'Perturbation of RNA Polymerase I transcription machinery by ablation of HEATR1 triggers the RPL5/RPL11-MDM2-p53 ribosome biogenesis stress checkpoint pathway in human cells', *Cell Cycle*, vol. 17, no. 1, 2018, doi: 10.1080/15384101.2017.1403685.
- [201] M. S. Lindström, J. Bartek, and A. Maya-Mendoza, 'p53 at the crossroad of DNA replication and ribosome biogenesis stress pathways', *Cell Death and Differentiation*, vol. 29, no. 5. 2022. doi: 10.1038/s41418-022-00999-w.
- [202] L. Golomb, S. Volarevic, and M. Oren, 'P53 and ribosome biogenesis stress: The essentials', *FEBS*

- Letters, vol. 588, no. 16. 2014. doi: 10.1016/j.febslet.2014.04.014.
- [203] F. M. Boisvert and A. I. Lamond, 'p53-Dependent subcellular proteome localization following DNA damage', *Proteomics*, vol. 10, no. 22, 2010, doi: 10.1002/pmic.201000213.
- [204] B. A. M. Bouwman et al., 'Genome-wide detection of DNA double-strand breaks by in-suspension BLISS', *Nat Protoc*, vol. 15, no. 12, 2020, doi: 10.1038/s41596-020-0397-2.
- [205] J. S. Iacovoni et al., 'High-resolution profiling of γ H2AX around DNA double strand breaks in the mammalian genome', *EMBO J*, vol. 29, no. 8, pp. 1446–1457, Apr. 2010, doi: 10.1038/emboj.2010.38.
- [206] H. J. Gothe et al., 'Spatial Chromosome Folding and Active Transcription Drive DNA Fragility and Formation of Oncogenic MLL Translocations', *Mol Cell*, vol. 75, no. 2, 2019, doi: 10.1016/j.molcel.2019.05.015.
- [207] H. Gaillard and A. Aguilera, 'Transcription as a Threat to Genome Integrity', *Annu Rev Biochem*, vol. 85, 2016, doi: 10.1146/annurev-biochem-060815-014908.
- [208] C. Aterillas, K. Abdelmohsen, and M. Gorospe, 'Regulation of senescence traits by MAPKs', *GeroScience*, vol. 42, no. 2. 2020. doi: 10.1007/s11357-020-00183-3.
- [209] M. A. Lemmon and J. Schlessinger, 'Cell signaling by receptor tyrosine kinases', *Cell*, vol. 141, no. 7. 2010. doi: 10.1016/j.cell.2010.06.011.
- [210] B. Boulianne and N. Feldhahn, 'Transcribing malignancy: Transcription-associated genomic instability in cancer', *Oncogene*, vol. 37, no. 8. 2018. doi: 10.1038/onc.2017.402.
- [211] S. Cohen et al., 'Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations', *Nat Commun*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-018-02894-w.
- [212] A. Lockhart et al., 'RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids', *Cell Rep*, vol. 29, no. 9, 2019, doi: 10.1016/j.celrep.2019.10.108.
- [213] A. Cristini et al., 'RNase H2, mutated in Aicardi-Goutières syndrome, resolves co-transcriptional R-loops to prevent DNA breaks and inflammation', *Nat Commun*, vol. 13, no. 1, 2022, doi: 10.1038/s41467-022-30604-0.
- [214] Y. Liu et al., 'Systematic proteome and proteostasis profiling in human Trisomy 21 fibroblast cells', *Nat Commun*, vol. 8, no. 1, 2017, doi: 10.1038/s41467-017-01422-6.
- [215] M. W. T. Talkington, G. Siuzdak, and J. R. Williamson, 'An assembly landscape for the 30S ribosomal subunit', *Nature*, vol. 438, no. 7068, pp. 628–632, Dec. 2005, doi: 10.1038/nature04261.

References

- [216] L. F. Pemberton, M. P. Rout, and G. Blobel, 'Disruption of the nucleoporin gene NUP133 results in clustering of nuclear pore complexes', *Proc Natl Acad Sci U S A*, vol. 92, no. 4, 1995, doi: 10.1073/pnas.92.4.1187.
- [217] B. Nabet et al., 'The dTAG system for immediate and target-specific protein degradation', *Nat Chem Biol*, vol. 14, no. 5, 2018, doi: 10.1038/s41589-018-0021-8.
- [218] A. Yesbolatova et al., 'The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice', *Nat Commun*, vol. 11, no. 1, 2020, doi: 10.1038/s41467-020-19532-z.
- [219] S. Schick et al., 'Acute BAF perturbation causes immediate changes in chromatin accessibility', *Nat Genet*, vol. 53, no. 3, 2021, doi: 10.1038/s41588-021-00777-3.
- [220] V. Aksenova, A. Arnaoutov, and M. Dasso, 'Analysis of Nucleoporin Function Using Inducible Degron Techniques', in *Methods in Molecular Biology*, vol. 2502, 2022. doi: 10.1007/978-1-0716-2337-4_9.
- [221] A. C. Gingras, K. T. Abe, and B. Raught, 'Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles', *Current Opinion in Chemical Biology*, vol. 48. 2019. doi: 10.1016/j.cbpa.2018.10.017.
- [222] I. Bludau and R. Aebersold, 'Proteomic and interactomic insights into the molecular basis of cell functional diversity', *Nature Reviews Molecular Cell Biology*, vol. 21, no. 6. 2020. doi: 10.1038/s41580-020-0231-2.
- [223] J. D. Martell et al., 'Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy', *Nat Biotechnol*, vol. 30, no. 11, 2012, doi: 10.1038/nbt.2375.
- [224] S. S. Lam et al., 'Directed evolution of APEX2 for electron microscopy and proximity labeling', *Nat Methods*, vol. 12, no. 1, 2014, doi: 10.1038/nmeth.3179.
- [225] L. Yadav et al., 'Systematic Analysis of Human Protein Phosphatase Interactions and Dynamics', *Cell Syst*, vol. 4, no. 4, 2017, doi: 10.1016/j.cels.2017.02.011.
- [226] X. Liu, K. Salokas, R. G. Weldatsadik, L. Gawriyski, and M. Varjosalo, 'Combined proximity labeling and affinity purification–mass spectrometry workflow for mapping and visualizing protein interaction networks', *Nat Protoc*, vol. 15, no. 10, 2020, doi: 10.1038/s41596-020-0365-x.
- [227] E. Batsché, M. Yaniv, and C. Muchardt, 'The human SWI/SNF subunit Brm is a regulator of alternative splicing', *Nat Struct Mol Biol*, vol. 13, no. 1, 2006, doi: 10.1038/nsmb1030.
- [228] T. Ito et al., 'Brm transactivates the telomerase reverse transcriptase (TERT) gene and modulates the splicing patterns of its transcripts in concert with p54nrb', *Biochemical Journal*, vol. 411, no. 1, pp. 201–209, Apr. 2008, doi: 10.1042/BJ20071075.

- [229] E. Allemand, M. P. Myers, J. Garcia-Bernardo, A. Harel-Bellan, A. R. Krainer, and C. Muchardt, 'A Broad Set of Chromatin Factors Influences Splicing', *PLoS Genet*, vol. 12, no. 9, 2016, doi: 10.1371/journal.pgen.1006318.
- [230] S. Yu et al., 'SWI/SNF interacts with cleavage and polyadenylation factors and facilitates pre-mRNA 3' end processing', *Nucleic Acids Res*, vol. 46, no. 16, pp. 8557–8573, Sep. 2018, doi: 10.1093/nar/gky438.
- [231] A. Gañez-Zapater et al., 'The SWI/SNF subunit BRG1 affects alternative splicing by changing RNA binding factor interactions with nascent RNA', *Molecular Genetics and Genomics*, vol. 297, no. 2, 2022, doi: 10.1007/s00438-022-01863-9.
- [232] S. Massenet, E. Bertrand, and C. Verheggen, 'Assembly and trafficking of box C/D and H/ACA snoRNPs', *RNA Biology*, vol. 14, no. 6. 2017. doi: 10.1080/15476286.2016.1243646.
- [233] S. Yasuda et al., 'Stress- and ubiquitylation-dependent phase separation of the proteasome', *Nature*, vol. 578, no. 7794, 2020, doi: 10.1038/s41586-020-1982-9.
- [234] M. Uriarte et al., 'Starvation-induced proteasome assemblies in the nucleus link amino acid supply to apoptosis', *Nat Commun*, vol. 12, no. 1, 2021, doi: 10.1038/s41467-021-27306-4.
- [235] K. G. Mark et al., 'Orphan quality control shapes network dynamics and gene expression', *Cell*, vol. 186, no. 16, 2023, doi: 10.1016/j.cell.2023.06.015.
- [236] I. Saez et al., 'The E3 ubiquitin ligase UBR5 interacts with the H/ACA ribonucleoprotein complex and regulates ribosomal RNA biogenesis in embryonic stem cells', *FEBS Lett*, vol. 594, no. 1, 2020, doi: 10.1002/1873-3468.13559.
- [237] R. G. Yau et al., 'Assembly and Function of Heterotypic Ubiquitin Chains in Cell-Cycle and Protein Quality Control', *Cell*, vol. 171, no. 4, 2017, doi: 10.1016/j.cell.2017.09.040.
- [238] F. Ohtake, H. Tsuchiya, Y. Saeki, and K. Tanaka, 'K63 ubiquitylation triggers proteasomal degradation by seeding branched ubiquitin chains', *Proc Natl Acad Sci U S A*, vol. 115, no. 7, 2018, doi: 10.1073/pnas.1716673115.
- [239] Z. Hodáková et al., 'Cryo-EM structure of the chain-elongating E3 ubiquitin ligase UBR5', *EMBO J*, vol. 42, no. 16, 2023, doi: 10.15252/embj.2022113348.
- [240] L. A. Hehl et al., 'Structural snapshots along K48-linked ubiquitin chain formation by the HECT E3 UBR5', *Nat Chem Biol*, 2023, doi: 10.1038/s41589-023-01414-2.
- [241] E. S. Johnson, P. C. M. Ma, I. M. Ota, and A. Varshavsky, 'A proteolytic pathway that recognizes ubiquitin as a degradation signal', *Journal of Biological Chemistry*, vol. 270, no. 29, 1995, doi: 10.1074/jbc.270.29.17442.

References

- [242] M. Koegl, T. Hoppe, S. Schlenker, H. D. Ulrich, T. U. Mayer, and S. Jentsch, 'A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly', *Cell*, vol. 96, no. 5, 1999, doi: 10.1016/S0092-8674(00)80574-7.
- [243] Y. Yu et al., 'K29-linked ubiquitin signaling regulates proteotoxic stress response and cell cycle', *Nat Chem Biol*, vol. 17, no. 8, pp. 896–905, Aug. 2021, doi: 10.1038/s41589-021-00823-5.
- [244] S. Sato, C. L. Ward, and R. R. Kopito, 'Cotranslational ubiquitination of cystic fibrosis transmembrane conductance regulator in vitro', *Journal of Biological Chemistry*, vol. 273, no. 13, 1998, doi: 10.1074/jbc.273.13.7189.
- [245] M. Zhou, E. A. Fisher, and H. N. Ginsberg, 'Regulated co-translational ubiquitination of apolipoprotein B100: A new paradigm for proteasomal degradation of secretory protein', *Journal of Biological Chemistry*, vol. 273, no. 38, 1998, doi: 10.1074/jbc.273.38.24649.
- [246] G. C. Turner and A. Varshavsky, 'Detecting and measuring cotranslational protein degradation in vivo', *Science* (1979), vol. 289, no. 5487, 2000, doi: 10.1126/science.289.5487.2117.
- [247] T. J. Eisenack and D. B. Trentini, 'Ending a bad start: Triggers and mechanisms of co-translational protein degradation', *Frontiers in Molecular Biosciences*, vol. 9. 2023. doi: 10.3389/fmolb.2022.1089825.
- [248] S. Duttler, S. Pechmann, and J. Frydman, 'Principles of cotranslational ubiquitination and quality control at the ribosome', *Mol Cell*, vol. 50, no. 3, 2013, doi: 10.1016/j.molcel.2013.03.010.
- [249] F. Wang, L. A. Durfee, and J. M. Huibregtse, 'A cotranslational ubiquitination pathway for quality control of misfolded proteins', *Mol Cell*, vol. 50, no. 3, 2013, doi: 10.1016/j.molcel.2013.03.009.
- [250] F. Dang, L. Nie, and W. Wei, 'Ubiquitin signaling in cell cycle control and tumorigenesis', *Cell Death and Differentiation*, vol. 28, no. 2. 2021. doi: 10.1038/s41418-020-00648-0.
- [251] M. Fischer, A. E. Schade, T. B. Branigan, G. A. Müller, and J. A. DeCaprio, 'Coordinating gene expression during the cell cycle', *Trends in Biochemical Sciences*, vol. 47, no. 12. 2022. doi: 10.1016/j.tibs.2022.06.007.
- [252] A. E. Schade, M. Fischer, and J. A. DeCaprio, 'RB, p130 and p107 differentially repress G1/S and G2/M genes after p53 activation', *Nucleic Acids Res*, vol. 47, no. 21, 2019, doi: 10.1093/nar/gkz961.
- [253] S. Uxa et al., 'DREAM and RB cooperate to induce gene repression and cell-cycle arrest in response to p53 activation', *Nucleic Acids Res*, vol. 47, no. 17, 2019, doi: 10.1093/NAR/GKZ635.
- [254] S. Banin et al., 'Enhanced phosphorylation of p53 by ATM in response to DNA damage', *Science* (1979), vol. 281, no. 5383, 1998, doi: 10.1126/science.281.5383.1674.

- [255] R. Maya et al., 'ATM-dependent phosphorylation of Mdm2 on serine 395: Role in p53 activation by DNA damage', *Genes Dev*, vol. 15, no. 9, 2001, doi: 10.1101/gad.886901.
- [256] J. Bartek and J. Lukas, 'Chk1 and Chk2 kinases in checkpoint control and cancer', *Cancer Cell*, vol. 3, no. 5. 2003. doi: 10.1016/S1535-6108(03)00110-7.
- [257] Y. Abuetabh et al., 'DNA damage response revisited: the p53 family and its regulators provide endless cancer therapy opportunities', *Experimental and Molecular Medicine*, vol. 54, no. 10. 2022. doi: 10.1038/s12276-022-00863-4.
- [258] C. D. Wiley and J. Campisi, 'The metabolic roots of senescence: mechanisms and opportunities for intervention', *Nature Metabolism*, vol. 3, no. 10. 2021. doi: 10.1038/s42255-021-00483-8.
- [259] E. L. James et al., 'Senescent human fibroblasts show increased glycolysis and redox homeostasis with extracellular metabolomes that overlap with those of irreparable DNA damage, aging, and disease', *J Proteome Res*, vol. 14, no. 4, 2015, doi: 10.1021/pr501221g.
- [260] M. Xu et al., 'Senolytics improve physical function and increase lifespan in old age', *Nat Med*, vol. 24, no. 8, 2018, doi: 10.1038/s41591-018-0092-9.
- [261] J. P. Coppé, P. Y. Desprez, A. Krtolica, and J. Campisi, 'The senescence-associated secretory phenotype: The dark side of tumor suppression', *Annual Review of Pathology: Mechanisms of Disease*, vol. 5. 2010. doi: 10.1146/annurev-pathol-121808-102144.
- [262] S. Özcan et al., 'Unbiased analysis of senescence associated secretory phenotype (SASP) to identify common components following different genotoxic stresses', *Aging*, vol. 8, no. 7, 2016, doi: 10.18632/aging.100971.
- [263] M. J. Payea, C. Anerillas, R. Tharakan, and M. Gorospe, 'Translational Control during Cellular Senescence', *Mol Cell Biol*, vol. 41, no. 2, 2021, doi: 10.1128/mcb.00512-20.
- [264] A. C. Vind, A. V. Genzor, and S. Bekker-Jensen, 'Ribosomal stress-surveillance: Three pathways is a magic number', *Nucleic Acids Research*, vol. 48, no. 19. 2020. doi: 10.1093/nar/gkaa757.
- [265] S. V. Lensing, G. Marsico, R. Hänsel-Hertsch, E. Y. Lam, D. Tannahill, and S. Balasubramanian, 'DSBCapture: In situ capture and sequencing of DNA breaks', *Nat Methods*, vol. 13, no. 10, 2016, doi: 10.1038/nmeth.3960.
- [266] B. Schwer et al., 'Transcription-associated processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells', *Proc Natl Acad Sci U S A*, vol. 113, no. 8, 2016, doi: 10.1073/pnas.1525564113.
- [267] P. C. Wei et al., 'Three classes of recurrent DNA break clusters in brain progenitors identified by 3D proximity-based break joining assay', *Proc Natl Acad Sci U S A*, vol. 115, no. 8, 2018, doi: 10.1073/

References

- pnas.1719907115.
- [268] R. Ballarino et al., 'An atlas of endogenous DNA double-strand breaks arising during human neural cell fate determination', *Sci Data*, vol. 9, no. 1, 2022, doi: 10.1038/s41597-022-01508-x.
- [269] J. A. Nickoloff, 'Transcription enhances intrachromosomal homologous recombination in mammalian cells.', *Mol Cell Biol*, vol. 12, no. 12, 1992, doi: 10.1128/mcb.12.12.5311.
- [270] A. Datta and S. Jinks-Robertson, 'Association of increased spontaneous mutation rates with high levels of transcription in yeast', *Science* (1979), vol. 268, no. 5217, 1995, doi: 10.1126/science.7777859.
- [271] N. Kim, A. L. Abdulovic, R. Gealy, M. J. Lippert, and S. Jinks-Robertson, 'Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication', *DNA Repair (Amst)*, vol. 6, no. 9, 2007, doi: 10.1016/j.dnarep.2007.02.023.
- [272] G. I. Dellino et al., 'Release of paused RNA polymerase II at specific loci favors DNA double-strand-break formation and promotes cancer translocations', *Nat Genet*, vol. 51, no. 6, 2019, doi: 10.1038/s41588-019-0421-z.
- [273] S. Ueda et al., 'TGF- β 1 is involved in senescence-related pathways in glomerular endothelial cells via p16 translocation and p21 induction', *Sci Rep*, vol. 11, no. 1, p. 21643, Nov. 2021, doi: 10.1038/s41598-021-01150-4.
- [274] A. Chakraborty et al., 'Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes', *Nat Commun*, vol. 7, 2016, doi: 10.1038/ncomms13049.
- [275] S. Guha and S. R. Bhaumik, 'Transcription-coupled DNA double-strand break repair', *DNA Repair (Amst)*, vol. 109, 2022, doi: 10.1016/j.dnarep.2021.103211.
- [276] P. Chaurasia, R. Sen, T. K. Pandita, and S. R. Bhaumik, 'Preferential repair of DNA double-strand break at the active gene in vivo', *Journal of Biological Chemistry*, vol. 287, no. 43, 2012, doi: 10.1074/jbc.M112.364661.
- [277] S. Hamperl and K. A. Cimprich, 'The contribution of co-transcriptional RNA: DNA hybrid structures to DNA damage and genome instability', *DNA Repair (Amst)*, vol. 19, 2014, doi: 10.1016/j.dnarep.2014.03.023.
- [278] L. Wahba, L. Costantino, F. J. Tan, A. Zimmer, and D. Koshland, 'S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation', *Genes Dev*, vol. 30, no. 11, 2016, doi: 10.1101/gad.280834.116.

- [279] M. P. Crossley, M. Bocek, and K. A. Cimprich, 'R-Loops as Cellular Regulators and Genomic Threats', *Molecular Cell*, vol. 73, no. 3. 2019. doi: 10.1016/j.molcel.2019.01.024.
- [280] A. Cristini et al., 'Dual Processing of R-Loops and Topoisomerase I Induces Transcription-Dependent DNA Double-Strand Breaks', *Cell Rep*, vol. 28, no. 12, 2019, doi: 10.1016/j.celrep.2019.08.041.
- [281] V. Marabitti et al., 'ATM pathway activation limits R-loop-associated genomic instability in Werner syndrome cells', *Nucleic Acids Res*, vol. 47, no. 7, 2019, doi: 10.1093/nar/gkz025.
- [282] T. García-Muse and A. Aguilera, 'Transcription-replication conflicts: How they occur and how they are resolved', *Nature Reviews Molecular Cell Biology*, vol. 17, no. 9. 2016. doi: 10.1038/nrm.2016.88.
- [283] A. Marnef and G. Legube, 'R-loops as Janus-faced modulators of DNA repair', *Nature Cell Biology*, vol. 23, no. 4. 2021. doi: 10.1038/s41556-021-00663-4.
- [284] C. Ohle, R. Tesorero, G. Schermann, N. Dobrev, I. Sinning, and T. Fischer, 'Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair', *Cell*, vol. 167, no. 4, 2016, doi: 10.1016/j.cell.2016.10.001.
- [285] F. Michelini et al., 'Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks', *Nat Cell Biol*, vol. 19, no. 12, 2017, doi: 10.1038/ncb3643.
- [286] G. D'Alessandro et al., 'BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment', *Nat Commun*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-018-07799-2.
- [287] X. Li and J. L. Manley, 'Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability', *Cell*, vol. 122, no. 3, 2005, doi: 10.1016/j.cell.2005.06.008.
- [288] L. Chen et al., 'The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations', *Mol Cell*, vol. 69, no. 3, 2018, doi: 10.1016/j.molcel.2017.12.029.
- [289] H. D. Nguyen et al., 'Spliceosome mutations induce R loop-associated sensitivity to ATR inhibition in myelodysplastic syndromes', *Cancer Res*, vol. 78, no. 18, 2018, doi: 10.1158/0008-5472.CAN-17-3970.
- [290] M. Cusan et al., 'SF3B1 mutation and ATM deletion codrive leukemogenesis via centromeric R-loop dysregulation', *Journal of Clinical Investigation*, vol. 133, no. 17, 2023, doi: 10.1172/JCI163325.
- [291] M. Tresini et al., 'The core spliceosome as target and effector of non-canonical ATM signalling', *Nature*, vol. 523, no. 7558, 2015, doi: 10.1038/nature14512.

References

- [292] S.-E. Ong et al., 'Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics,' *Mol Cell Proteomics*, vol. 1, no. 5, pp. 376–386, 2002, doi: 10.1074/mcp.M200025-MCP200.
- [293] Y. E. Pearson, S. Kremb, G. L. Butterfoss, X. Xie, H. Fahs, and K. C. Gunsalus, 'A statistical framework for high-content phenotypic profiling using cellular feature distributions,' *Commun Biol*, vol. 5, no. 1, 2022, doi: 10.1038/s42003-022-04343-3.
- [294] J. Rappsilber, M. Mann, and Y. Ishihama, 'Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips,' *Nat Protoc*, vol. 2, no. 8, 2007, doi: 10.1038/nprot.2007.261.
- [295] A. Michalski et al., 'Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer,' *Molecular and Cellular Proteomics*, vol. 10, no. 9, 2011, doi: 10.1074/mcp.M111.011015.
- [296] C. D. Kelstrup, C. Young, R. Lavalley, M. L. Nielsen, and J. V. Olsen, 'Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer,' *J Proteome Res*, vol. 11, no. 6, 2012, doi: 10.1021/pr3000249.
- [297] D. B. Bekker-Jensen et al., 'A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients,' *Molecular & Cellular Proteomics*, vol. 19, no. 4, pp. 716–729, Apr. 2020, doi: 10.1074/mcp.TIR119.001906.
- [298] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann, 'Higher-energy C-trap dissociation for peptide modification analysis,' *Nat Methods*, vol. 4, no. 9, 2007, doi: 10.1038/nmeth1060.
- [299] J. Cox and M. Mann, 'MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification,' *Nat Biotechnol*, vol. 26, no. 12, pp. 1367–1372, Dec. 2008, doi: 10.1038/nbt.1511.
- [300] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, and M. Mann, 'Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment,' *J Proteome Res*, vol. 10, no. 4, pp. 1794–1805, Apr. 2011, doi: 10.1021/pr101065j.
- [301] J. E. Elias and S. P. Gygi, 'Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry,' *Nat Methods*, vol. 4, no. 3, 2007, doi: 10.1038/nmeth1019.
- [302] S. Tyanova and J. Cox, 'Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research,' in *Methods in Molecular Biology*, vol. 1711, 2018. doi: 10.1007/978-1-4939-7493-1_7.

- [303] R Core Team, 'R Core Team (2014). R: A language and environment for statistical computing,' R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>, 2014.
- [304] G. K. Smyth, 'Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments', *Stat Appl Genet Mol Biol*, vol. 3, no. 1, pp. 1–25, Jan. 2004, doi: 10.2202/1544-6115.1027.
- [305] L. McInnes, J. Healy, N. Saul, and L. Großberger, 'UMAP: Uniform Manifold Approximation and Projection', *J Open Source Softw*, vol. 3, no. 29, 2018, doi: 10.21105/joss.00861.
- [306] T. Wu et al., 'clusterProfiler 4.0: A universal enrichment tool for interpreting omics data', *Innovation*, vol. 2, no. 3, 2021, doi: 10.1016/j.xinn.2021.100141.
- [307] P. Shannon et al., 'Cytoscape: A software Environment for integrated models of biomolecular interaction networks', *Genome Res*, vol. 13, no. 11, 2003, doi: 10.1101/gr.1239303.
- [308] D. Szklarczyk et al., 'The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest', *Nucleic Acids Res*, vol. 51, no. 1 D, 2023, doi: 10.1093/nar/gkac1000.
- [309] A. Frankish et al., 'GENCODE reference annotation for the human and mouse genomes', *Nucleic Acids Res*, vol. 47, no. D1, 2019, doi: 10.1093/nar/gky955.
- [310] A. Dobin et al., 'STAR: Ultrafast universal RNA-seq aligner', *Bioinformatics*, vol. 29, no. 1, 2013, doi: 10.1093/bioinformatics/bts635.
- [311] Y. Liao, G. K. Smyth, and W. Shi, 'FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features', *Bioinformatics*, vol. 30, no. 7, 2014, doi: 10.1093/bioinformatics/btt656.
- [312] W. Huber et al., 'Orchestrating high-throughput genomic analysis with Bioconductor', *Nat Methods*, vol. 12, no. 2, 2015, doi: 10.1038/nmeth.3252.
- [313] M. I. Love, W. Huber, and S. Anders, 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, vol. 15, no. 12, 2014, doi: 10.1186/s13059-014-0550-8.
- [314] S. Shen et al., 'rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data', *Proc Natl Acad Sci U S A*, vol. 111, no. 51, 2014, doi: 10.1073/pnas.1419161111.
- [315] H. Li, '[Heng Li - Compares BWA to other long read aligners like CUSHAW2] Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM', arXiv preprint arXiv, 2013.
- [316] H. M. Amemiya, A. Kundaje, and A. P. Boyle, 'The ENCODE Blacklist: Identification of Problematic Regions of the Genome', *Sci Rep*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-45839-z.

8. Acknowledgments

9. Curriculum vitae

