Characterizing the spatio-temporal cellular response to transcription-blocking lesions by mass spectrometry-based proteomics

Dissertation

Zu Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

der Johannes Gutenberg-Universität Mainz

vorgelegt von

Juanjuan Wang

Geboren 21.05.1989

Shanxi, China

Mainz, April 2022

Dekan: Prof. Dr. Eckhard Thines

1. Gutachterin: Prof. Dr. Petra Beli

2. Gutachter: Dr. Julian König

Tag der mündlichen Prüfung: 18.05.2022

Preface

Summary

Exposure to ultraviolet (UV) radiation induces tumorigenesis of keratinocytes and leads to melanoma. On a molecular level, UV generates transcription-blocking DNA lesions that affect all steps of gene expression, from co-transcriptional splicing on chromatin to translation in the cytoplasm. However, a systematic investigation of protein-protein interactions and posttranslational modifications that regulate UV response remains poorly studied.

We used human keratinocytes as a model system and employed tandem mass-tag-based quantitative mass spectrometry (TMT-MS) to quantify the changes in proteome and phosphoproteome with spatial and temporal resolution. Additionally, the localization of proteins on chromatin was investigated. Our data revealed that UV stress activates the main DNA damage response kinases, ATM, ATR, and DNA-PK as well as MAPKs. After a long time of recovery post UV stress (18h), the G2/M checkpoint is activated with the activation of WEE1. The dynamic regulation of phosphorylation states results in changing protein levels and protein recruitment to and exclusion from chromatin, subsequently regulating cellular processes. At an early recovery time (0.5h), when transcription is halted, DNA repair and transcription processing are highly regulated. In comparison, when transcription recovers and restarts (2h and 6h), RNA processing including mRNA and non-coding RNA (ncRNA) processing, and translation initiation are regulated by the changes in the associated protein levels and phosphorylation. Moreover, we observed that UV stress induces R-loop accumulation when transcription restarts (6h). Also, R-loop associated proteins are regulated by phosphorylation upon UV stress. Interestingly, signatures of the inflammatory response are observed after transcription-blocking lesions are obliterated completely.

To study the role of the transcription-related kinase JNK1, we combined the chemical inhibition of JNK1 and UV light with quantitative phosphorylation profiling by mass spectrometry to identify JNK kinase substrates. We identified 206 putative JNK substrates in response to UV irradiation, with an overrepresentation of RNA binding proteins. Furthermore, we revealed that the Cleavage factor Im (CFIm) complex component (CPSF6) is phosphorylated by JNK1 on threonine 407 upon exposure to UV. The phosphorylated CPSF6 is excluded from chromatin and re-localizes into paraspeckles together with pre-mRNA upon UV stress. CPSF6 rosettes show liquid-liquid phase separation (LLPS) characteristics. The phosphorylation of CPSF6 regulates alternative polyadenylation after UV exposure and promotes cellular recovery from transcription-blocking lesions.

Zusammenfassung

Werden Hautzellen starker ultravioletter Strahlung (UV) ausgesetzt, kann dies zu Mutationen und dem Entstehen von Melanomen führen. Auf molekularer Ebene erzeugt UV-Strahlung transkriptionshemmende DNA-Läsionen, die sich auf alle Schritte der Genexpression auswirken, vom co-transkriptionalen Spleißen am Chromatin bis zur Translation im Zytoplasma. Systematische Untersuchungen der Protein-Protein-Interaktionen und posttranslationalen Modifikationen, die die UV-Reaktion regulieren, gibt es jedoch nur wenige.

In der vorliegenden Studie verwenden wir Hautzellen als Modellsystem und setzen quantitative Massenspektrometrie mit Tandem-Massentags (TMT-MS) ein, um die Veränderungen im Proteom und Phosphoproteom mit räumlicher und zeitlicher Auflösung zu quantifizieren. Zusätzlich untersuchen wir die Lokalisierung von Proteinen am Chromatin. Unsere Daten zeigen, dass UV-Stress unmittelbar die wichtigsten DNA-Schadensreaktionskinasen, ATM, ATR und DNA-PK, sowie MAPKs aktiviert. Geraume Zeit nach UV-Bestrahlung (18 Stunden) wird auch der G2/M-Kontrollpunkt durch die Aktivierung von WEE1 aktiviert. Diese dynamische Veränderung der Proteinphosphorylierung steuert die Proteinhomöostase und die Rekrutierung von Proteinen an und aus dem Chromatin. Wir beobachten, dass die DNA-Transkription unmittelbar nach der UV-Bestrahlung abgeschaltet wird (0,5 Stunden), und erst nach 2 bis 6 Stunden, einschließlich der Prozessierung von mRNA und nicht-kodierender RNA (ncRNA), sowie der Proteintranslation wiedereinsetzt. Darüber hinaus, führt UV-Stress bei Wiederaufnahme der Transkription (6 Stunden) zur Anhäufung von R-Loops, was mit der Phosphorylierung von R-loop-assoziierten Proteinen einhergeht. Interessanterweise beobachten wir Anzeichen einer Entzündungsreaktion, nachdem die transkriptionsblockierenden Läsionen vollständig verschwunden sind.

Um die Rolle der transkriptions-assoziierten Kinase JNK1 zu untersuchen, kombinieren wir die chemische Hemmung von JNK1 und UV-Bestrahlung mit der Anreicherung von Phosphopeptiden und quantitativer Massenspektrometrie. Wir identifizieren 206 potenzielle JNK-Substrate als Reaktion auf die UV-Bestrahlung, wobei RNA-bindende Proteine überrepräsentiert sind. Darüber hinaus zeigen wir, dass die Untereinheit des Cleavage-Faktor-Im (CFIm)-Komplexes (CPSF6) bei UV-Bestrahlung von JNK1 an Threonin 407 phosphoryliert wird. Das phosphorylierte CPSF6 wird bei UV-Stress aus dem Chromatin ausgeschlossen und zusammen mit der prä-mRNA in Paraspekles überführt. Die entstandenen CPSF6-Rosetten weisen Eigenschaften von Flüssig-Flüssig-Phasentrennung (LLPS) auf. Die Phosphorylierung von CPSF6 reguliert die alternative Polyadenylierung nach UV-Bestrahlung und fördert die Auflösung von transkriptionsblockierenden Läsionen.

Table of contents

1 Introduction

1.1 Principle cellular organization

Biological systems follow the rules of chemistry, physics, and biology, which is a historical science, as the forms and structures of the living world today are the results of billions of years of evolution. All organisms are related in a family tree that spreads from primitive single-cell organisms in the distant past to today's various plants, animals, and microbes, through evolution [1], [2] (**Figure 1**). The biological universe appears incredibly diverse, from single-cell bacteria to multicellular animals. Cells come in a remarkable range of sizes and shapes with rapidly changing or stationary stable structures. Some cells are killed by oxygen, while others need it to live [2]–[4]. Eukaryotic cells are typically 10–100 µm, much larger than bacteria. A typical human fibroblast cell is about 15 µm and has tens of thousands of times the volume of an *E. coli* cell [5], [6], [7]. Despite the complex variety of biological forms, it overlies a powerful uniformity. All biological systems are made up of cells that contain the same types of chemical molecules and use the same basic types of biological molecules, which are conserved over billions of years of evolution [8], [9]. The structure and function of cells and organisms rely on the interaction of chemical molecules, which are marvelously coordinated in time and space and influenced by cellular genetics and the environment.

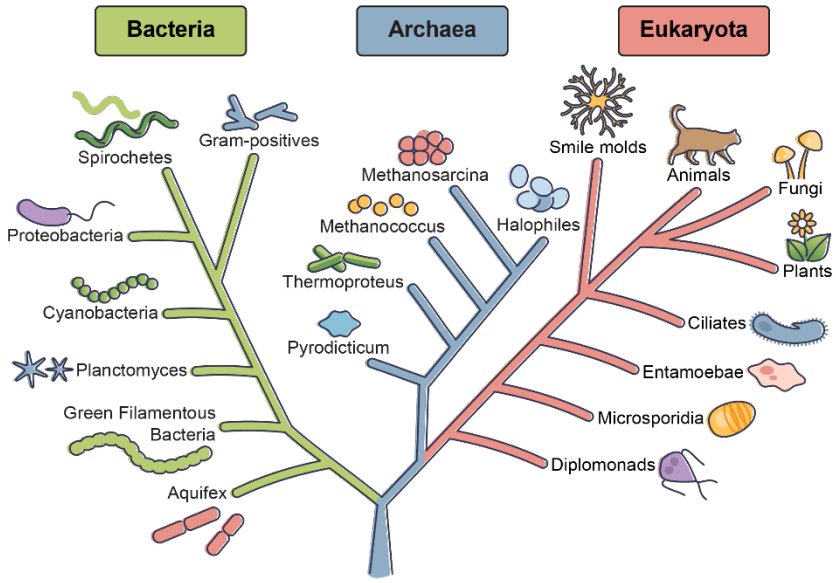

Figure 1: Phylogenetic tree is adapted from https://en.wikipedia.org/wiki/Phylogenetic_tree

Eukaryotic cells are surrounded by a plasma membrane. Most eukaryotic cells have internal membranes, which enclose and separate specific subcellular compartments from the cytoplasm, known as organelles. Many organelles and subcellular structures are also shared by all eukaryotic cells, such as the ribosome and mitochondria (**Figure 2**).

1.1.1 Membrane organelles in cells

Eukaryotic cells are incredibly brimming with deoxyribonucleic acid (DNA), ribonucleic acid (RNA), proteins, lipids, and metabolites. Total protein concentrations have been estimated to be up to 300 mg/mL, while RNA concentrations can range between 20 and 100 mg/mL in eukaryotic cells [10]. To avoid chaos and make intracellular reactions more efficient, cells have evolved several strategies to categorize and organize their content. Intracellular membranes in eukaryotic cells form specialized organelles to ensure that specific biochemical reactions and cellular functions occur in a spatially restricted manner, for example, separate transcription in the nucleus from translation in the cytosol. The cytosol, the organelle-free part of the cytoplasm, contains water, dissolved ions, small molecules, and proteins [11].

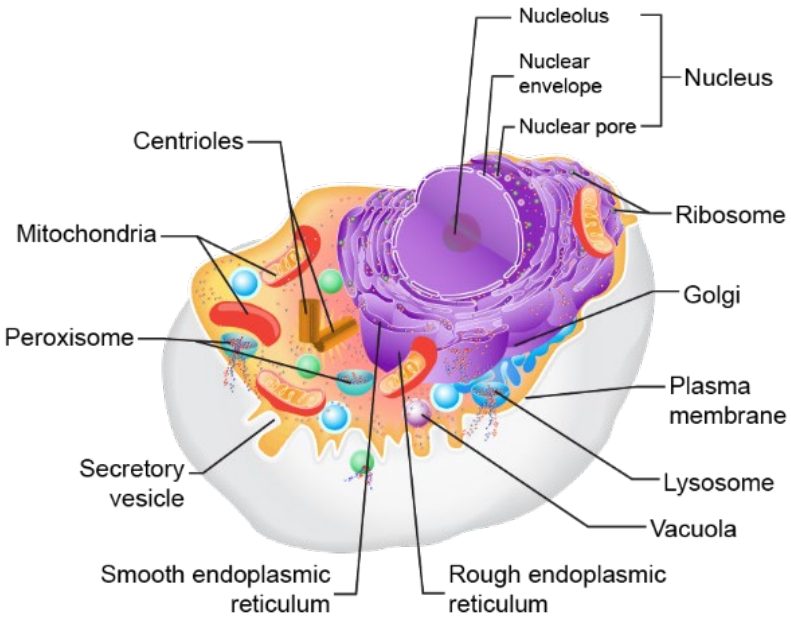

Figure 2: Simple representation of a mammalian cell adapted from [12]**.**

The majority of well-known organelles are separated from their surroundings by a membrane boundary, for example, the endomembrane system (endoplasmic reticulum (ER), Golgi, endosomes, and lysosomes) [13]. Materials are extensively transferred between organelles and their surroundings, as well as between organelles, in order to integrate various cellular activities. Diffusion or transmembrane protein-mediated transport allows these materials to enter or exit organelles (e.g., transporters and channels). Furthermore, the cytoskeleton provides 'highways' for the directed transport of RNA or vesicles to the distal end of cells and positions membrane-bounded organelles, such as ER [13]–[18].

1.1.2 Membraneless organelles in cells

There are also many membraneless organelles in cells. Example of a membraneless organelle in the nucleus are paraspeckles that are formed by the long noncoding RNA NEAT1, involved in gene expression regulation [19]. Nuclear speckles are in the interchromatin space containing high RNA-processing concentrations and some transcription factors but no DNA [20]. Cajal bodies are the centers

of assembling and modification of spliceosomal small nuclear RNPs [21]. PML (Promyelocytic leukemia) bodies are involved in multiple genome maintenance pathways, including the DNA damage response, DNA repair, and telomere homeostasis [22], [23]–[25]. Nucleoli are the sites of ribosomal RNA transcription and the centers of assembly and modification of spliceosomal small nuclear RNPs [26]. In the cytoplasm, p-bodies are condensates of enzymes involved in mRNA decay and microRNA (miRNA)-induced mRNA silencing, and stress granules (SGs) are primarily untranslated mRNA storage [27]–[29] (**Figure 3**A).

These organelles are formed in a process known as liquid-liquid phase separation (LLPS). Recent research also suggests that these organelles are supramolecular assemblies of proteins and RNA/DNA molecules, forming due to proteins LLPS, primarily governed by the interactions of multi-domain proteins or proteins containing intrinsically disordered regions (IDRs) as well as RNA-binding domains [30]–[33]. IDRs often have low-complexity domains (LCDs) with highly biased amino-acid compositions [34], [35]. LLPS is sensitive to environmental changes that affect multivalent interactions, such as composition, concentration, temperature, pH, and salt concentration. Protein posttranslational modifications (PTMs) such as phosphorylation, methylation, ubiquitination, and sumoylation change the interacting strength and valency, allowing LLPS to integrate a variety of signals, for example, by disrupting cation- π interactions via arginine methylation or by creating an SH2 domain-binding motif via tyrosine phosphorylation [36], [37], [38], [39], [40], [41] (**Figure 3**B).

Some membraneless organelles are constitutive, such as nucleoli, whereas others are transient and must be resolved quickly to avoid pathological stabilization, such as stress granules in neurodegeneration [42]–[44]. In addition, some membraneless organelles are formed only when cells are under stress, for example, paraspeckles. Paraspeckles are initiated by the specific long noncoding RNA (lncRNA) of NEAT1 (nuclear paraspeckle assembly transcript 1). NEAT1 is composed of two isoforms transcribed by RNA polymerase II from a single exon: the polyadenylated short isoform NEAT1_1 (3.7 kb in humans, 3.0 kb in mice) and the non-polyadenylated long isoform NEAT1_2 (22.7 kb in humans, 20.7 kb in mice) [45]. When lncRNA NEAT1_2 is produced in cells, it will be captured by the mammalian DBHS (*Drosophila melanogaster* behavior, human splicing) protein family, consisting of PSPC1, NONO, and SFPQ to start assembling of paraspeckles and elongation of paraspeckles with other RNA binding proteins, such as FUS (**Figure 3**A) [46]–[48]. The increase of NEAT1_2 proportion can be induced by proteasome inhibition, mitochondrial defects, or modification of NEAT1_1 polyadenylation sites. These will affect the number and size of paraspeckles and the number of elongated paraspeckles [49]–[54]. DBHS proteins' RNA recognition motif and the coiled-coil domain at C-terminal, which mediates dimerization, are required for the formation of paraspeckles [55]–[57]. Under normal conditions, DBHS proteins are highly dynamic and cycle between the nucleoplasm and the nucleolus. However, when RNA polymerase II (RNA Pol II)-driven transcription is inhibited, they accumulate within the perinucleolar cap structures, paraspeckles [58]. Paraspeckles are restricted to mammalian nuclei and size in around 0.5–1.0 μm, and their numbers differ depending on

cell populations and cell types [59], [60]. Paraspeckles have been widely reported to play a role in cell apoptosis and gene expression regulation by retaining RNA in the nucleus [54], [61]–[63]. In addition, paraspeckles and aberrant expression of NEAT1 have been found in breast cancer and gynecologic cancers (e.g., ovarian, cervical, endometrial, and vulvar) [63], [64], [65], [66], [67], [68].

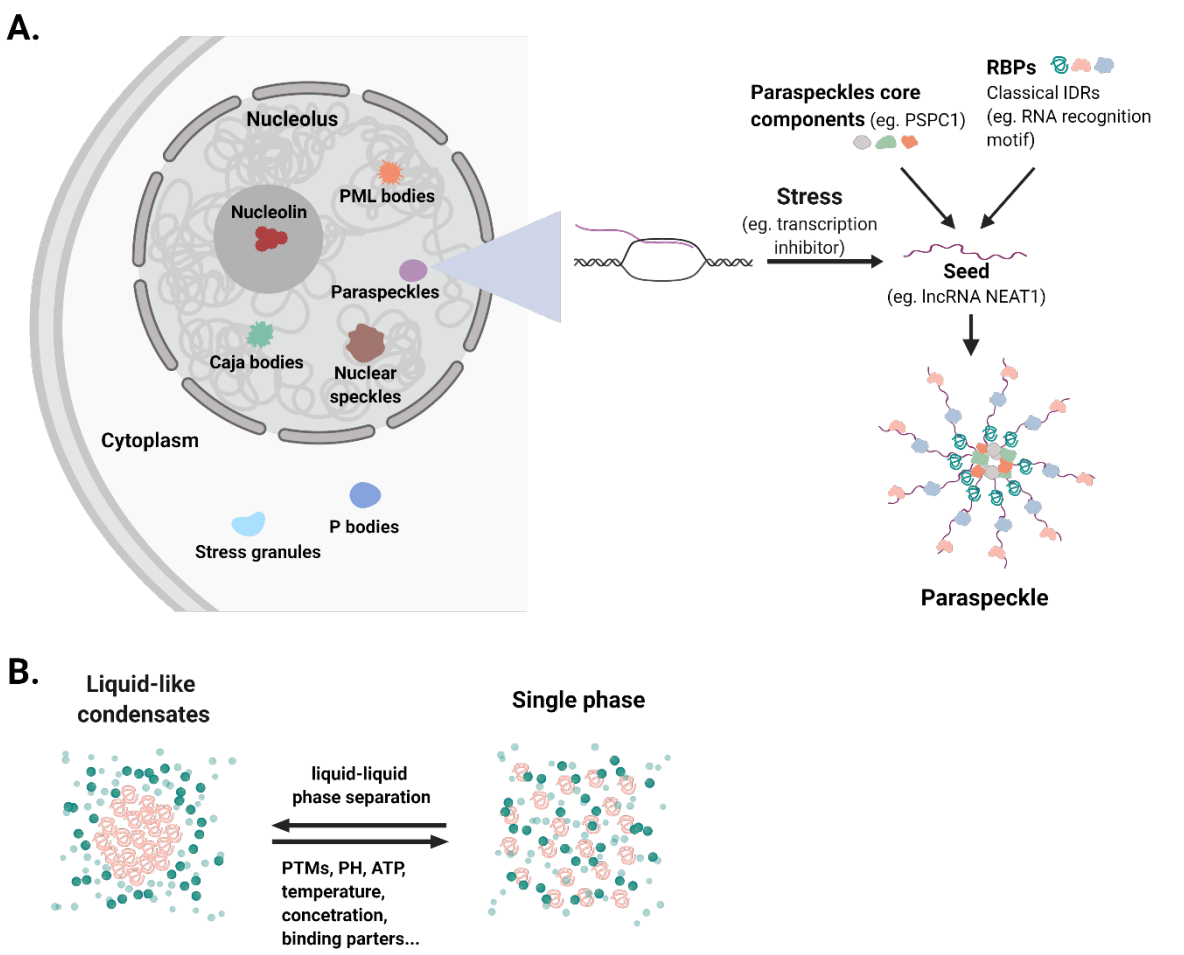

Figure 3: Membraneless organelles in cells and the formation of paraspeckles. A. lncRNA NEAT1 is induced by cellular stress, such as transcription inhibition. It will function as a scaffold or seed and recruit paraspeckles core components (e.g., PSPC1) and other RNA binding proteins (RBPs) with intrinsically disordered regions (IDRs) to form paraspeckles. **A.** LLPS is regulated by post-translational modifications, pH, temperature, or the concentration of proteins, RNA or ATP [33], [69], [70]. The images are created with BioRender.com.

1.1.3 DNA is packed in chromosomes in the nucleus

In human cells, the largest organelle is the nucleus which is surrounded by two membranes (**Figure 2**). The two nuclear membranes fuse at nuclear pore complexes, composed of specific membrane proteins that allow molecules and signals to transport between the nucleus and the cytosol [71]–[77].

Chromosome

Many molecules and subcellular structures play essential roles in cellular functions and responding to cellular stress in the nucleus, such as chromosomes, DNA (deoxyribonucleic acid), RNA (Ribonucleic acid) and proteins. DNA is wrapped around spools of histone proteins to form nucleosomes, which package together to produce chromatin that contains genes, collectively called

the genome [78], [79]. The haploid human genome contains about 3 billion base pairs of DNA organized into 23 chromosomes. However, most human body cells, except female ova and male sperm, are diploid, with 6 billion base pairs of DNA in 23 pairs of chromosomes [80]–[83]. Chromatin is a fiber with a diameter of about 30 nm. Nucleosomes are the basic unit of the eukaryotic chromosome. Every individual nucleosome is composed of a histone octamer with eight histone proteins, two molecules each of histones H2A, H2B, H3, and H4, and double-stranded DNA that is 146 nucleotide pairs long [84], [85]. In addition to the histones found in the nucleosome core, H1 histone protein binds externally to both the nucleosome and the "linker" DNA between nucleosomes, supporting DNA compacting (**Figure 6**) [86], [87].

Chromosome organization

Although its high level of compaction helps DNA transport during cell division, it will reduce DNA accessibility for other cellular functions, such as DNA replication and transcription. Therefore, chromosome structure changes in how tight DNA is packaged depending on the stage of the cell cycle and the level of gene activity required. To ensure that the chromosomes are evenly aligned and separated into daughter cells during cell division, the DNA must be thoroughly condensed [88]–[96]. The local chromosomes must be relaxed or opened during other cell cycle stages to allow gene expression [97], [98], [99], [100]. The dynamic organization of DNA into chromatin regulates gene expression and other cellular processes, such as DNA replication, DNA repair, recombination, and chromosome segregation.

ATP (adenosine triphosphate)-driven chromatin remodeling complexes and modifications on histone tails regulate the mobilization of nucleosomes via the alteration of histone–DNA interactions [101], [102], [103]. Chromatin remodeling complexes have been reported to regulate gene activation or inactivation [104]–[107]. The N-terminal and C-terminal tails of histones are reported to contain post-translational modifications (PTMs), such as acetylation, methylation, phosphorylation, ubiquitination, ADP-ribosylation, and sumoylation [102], [108]–[114]. Histone modifications have been widely linked to gene expression at multiple levels [115]–[117]. Together with DNA methylation, these modifications control the nucleosome positioning in chromosomes and mediate cellular signaling processes in response to genotoxic stimuli [101], [113], [118]–[124]. For example, H3 trimethylated at lysine 4 (H3K4me3) is defined at transcription promoters and transcription start sites (TSS) of active genes, while mono-methylated H3K4 is specifically at gene enhancers [125]–[128], [129]–[131]. Therefore, they are often considered as marks for "open" or "relaxed" chromatin. However, methylated H3K9, H3K27me3, and H4K20me3 spread widely in chromatin which is defined as "closed" chromatin or silent regions (**Figure 4**) [132]–[135], [136], [137], [138]–[142].

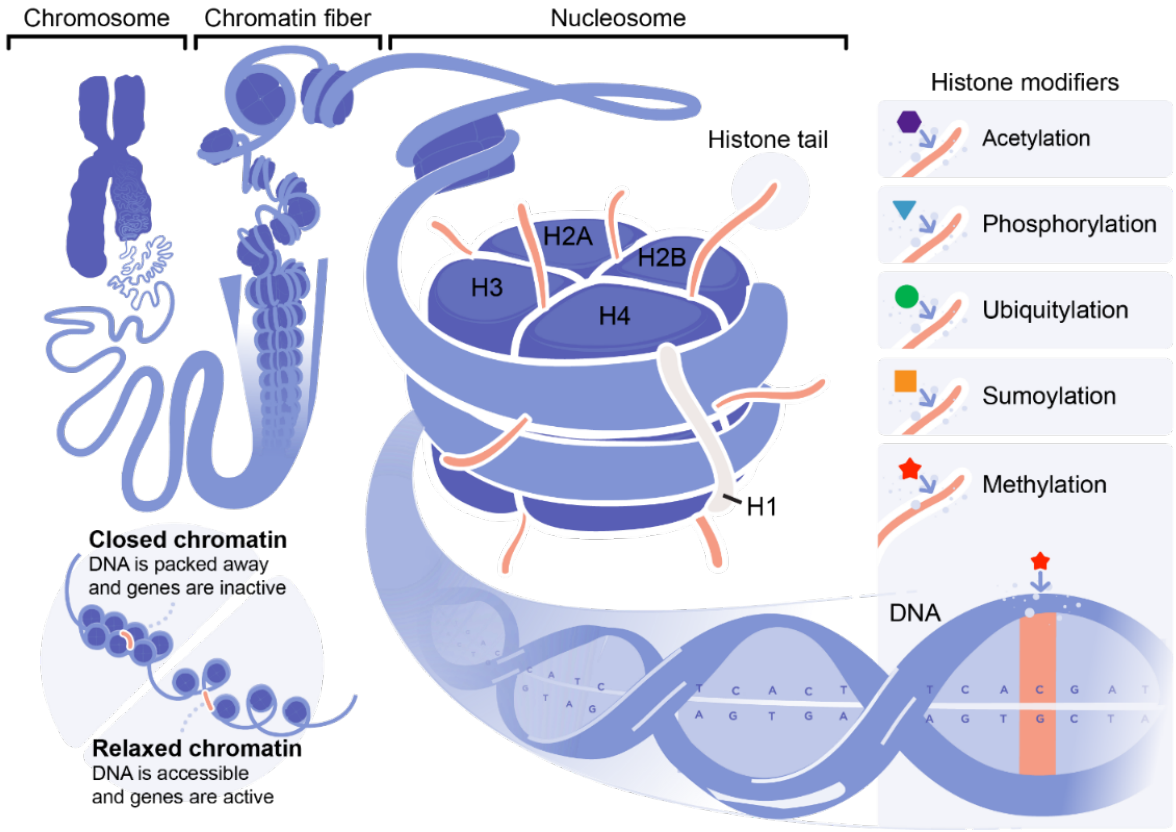

Figure 4: Simplified scheme of the eukaryotic chromosome. Adapted from the National Institutes of Health (NIH) and Annunziato, A.(2008) [80].

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)

DNA on chromatin contains heredity information in the nucleotide sequence within a genetic code which is a triplet code [143]. This information is required to produce RNA serving as the blueprints for proteins in cells, which will support cellular processes. Messenger RNA (mRNA) is very similar to DNA, primarily informational molecules, carrying information in the exact sequence of their nucleotides [143], [144]. All organisms have only four different nucleotides for DNA or RNA (**Figure 5**). All nucleotides consist of an organic base linked to a 5'-carbon sugar with a phosphate group attached to the 5'-carbon. In DNA, the sugar is deoxyribose, while in RNA, it is ribose. The nucleotides used in the synthesis of DNA and RNA contain five different nucleobases. The nucleobases adenine (A) and guanine (G) are purines, which have a pair of fused rings, and the bases cytosine (C), thymine (T), and uracil (U) are pyrimidines, which contain a single ring. A, C and G are shared in both DNA and RNA. T is only in DNA and U is in RNA (**Figure 5**) [144], [145]. In 1953, James D. Watson and Francis H. C. Crick first proposed the DNA three-dimensional (3D) structure, which consists of two long helical strands that are coiled around a common axis to form a double helix. One DNA strand is formed by joining four different nucleotides (A, T, C, and G), with the base parts presenting inward from the strand's backbone. DNA double helix is formed when two DNA strands bind together via their bases and twist through hydrogen bonds. Hydrogen bonds allow bases to interact. G~C, A~T (in DNA) and G~C, A~U (in RNA) are the Watson-Crick base pairs (**Figure 5**). The native 3D structures of DNA and

RNA are stabilized by base pairing. The double-helix structure of DNA ensures the phenomenon of heredity, which is the transport of genetically determined characteristics from one generation to the next [146]–[149].

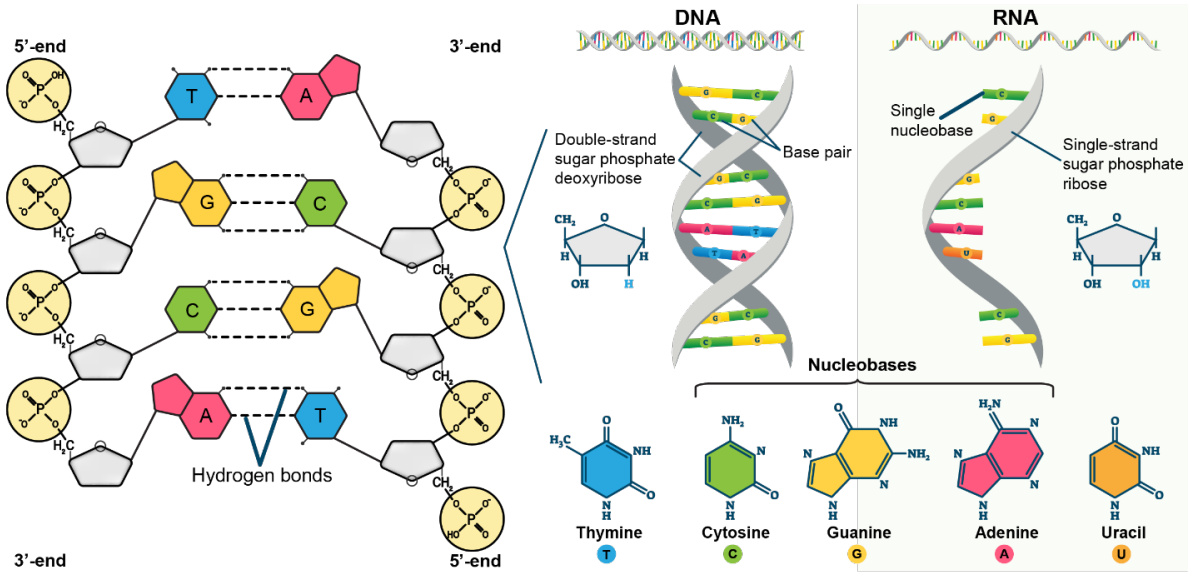

Figure 5: Simple scheme of DNA and RNA structure. DNA double helix (left and middle). This schematic shows the two sugar-phosphate backbones and hydrogen bonding between the Watson-Crick base pairs. **RNA single strand (right).** Nucleotides and nucleobases for DNA and RNA are shown in different colors, adapted from https://commons.wikimedia.org/wiki/File:Difference_DNA_RNA-EN.svg, https://www.technologynetworks.com/genomics/lists/what-are-the-key-differences-between-dna-and-rna-296719, and https://www.thoughtco.com/dna-versus-rna-608191.

1.2 Gene expression is controlled by transcription

1.2.1 The central dogma of molecular biology

Specific segments of DNA, termed genes, carry instructions for making specific proteins. Commonly, genes contain two parts: the coding region specifies the amino acid sequence of a protein, and the regulatory region binds specific proteins and controls when and in which cells the gene's protein is made. The nucleus is metabolically active in a growing or differentiating cell as it is the site of DNA replication and the synthesis of ribosomal RNA (rRNA), messenger RNA (mRNA), and a broad group of non-coding RNAs.

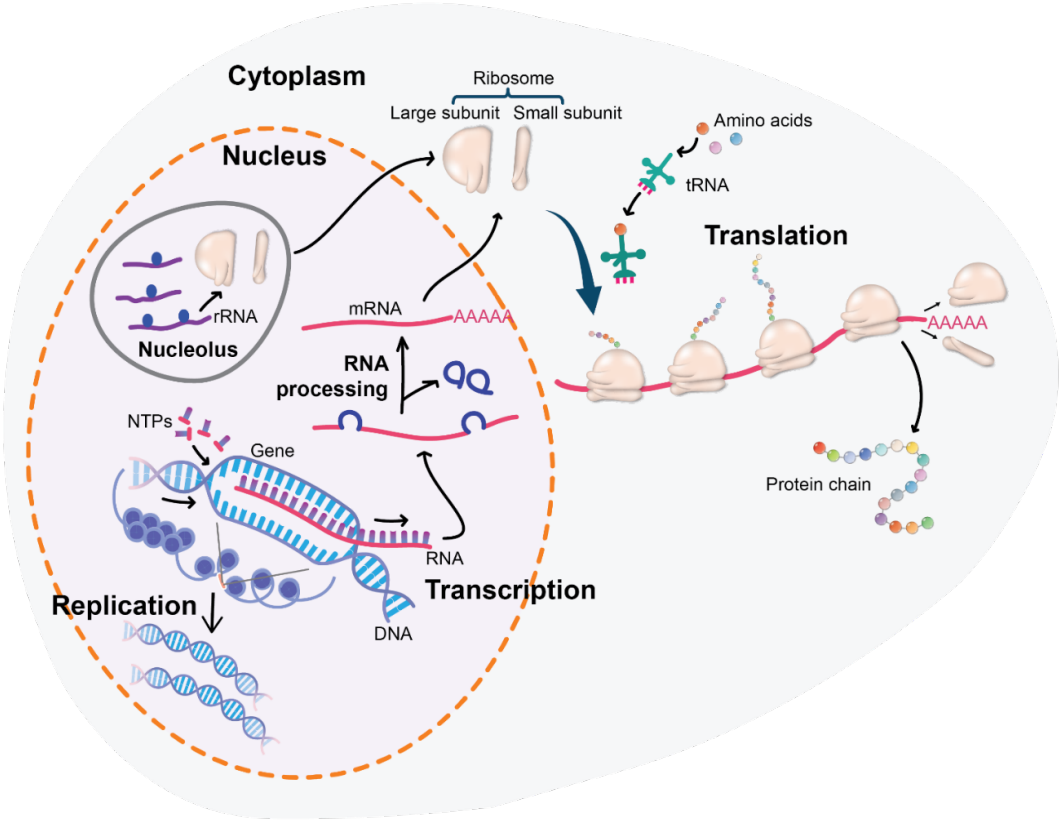

Figure 6: The simplified scheme of an overview of four basic molecular genetic processes. The production of proteins (transcription and translation**)** and DNA replication, adapted from Biology2e by OpenStax and https://www.cancer.gov/publications/dictionaries/genetics-dictionary/def/translation [150]**.**

Human cells convert the coded information in DNA into specific proteins by two processes: transcription and translation (**Figure 6**). Thus genes ultimately define the biological structure and maintain the integration of cellular function. In the first transcription process, the gene coding region is copied into RNA whose sequence is identical to one of the two in the double-stranded DNA. RNA polymerase II (RNA Pol II), which is a large enzyme, catalyzes the linkage of nucleotides into an RNA chain using DNA as a template. One strand of DNA acts as a template for RNA polymerase, which is commonly named the template strand. As it "reads" this template one base at a time, the polymerase builds an RNA molecule out of complementary nucleotides, producing a chain from 5' to 3'. The RNA transcript carries the same information as the non-template (non-coding) strand of DNA, but it contains

the base uracil (U) instead of thymine (T). The initial RNA product in eukaryotic cells is processed into a messenger RNA (mRNA) via post-transcriptional modifications, for instance, RNA polyadenylation, splicing, capping with associated machineries, such as the spliceosome. Then matured mRNA is transported to the cytoplasm. The nucleotide sequence of an mRNA molecule contains accurate information that specifies the correct order of amino acids during the synthesis of a protein (**Figure 6**). The second process, known as translation, is carried out here by the ribosome, an enormously complex molecular machine made up of ribosomal RNA (rRNA) and proteins. rRNA is synthesized, and ribosomes are assembled in a dense nuclear subcompartment, termed the nucleolus. During translation, the ribosome stepwise assembles and connects amino acids in the accurate order dictated by the mRNA sequence, as dictated by the gene code in DNA. In this stepwise process, the nucleotide sequence of an mRNA molecule is "read" by transfer RNA (tRNA) with the aid of ribosome. Then the correct amino acids are brought into sequence by tRNAs; peptide bonds link them to make protein chains, known as polymerization (**Figure 6**) [151], [152].

1.2.2 The phosphorylation of RNA Pol II regulates transcription

The simplified representation of proteins' production is named the central dogma as DNA → RNA → protein [152]. However, it is more complicated in cells, especially cells facing internal and external stimuli all the time.

Faithful transcription processes and gene expression regulation are essential to the cellular ability to respond and adapt to changes in its environment. The control of transcription rates is directly regulated by transcription factors (TFs) and indirectly by chromatin state, cell signaling, and other regulatory elements. The phosphorylation or dephosphorylation often orchestrates the modulation of TF activity by protein kinases (PKs) or phosphatases (PPs), especially RNA polymerase II (RNA Pol II) [153]–[155].

RNA Pol II is a large complex (550 kDa) with 12 highly conserved subunits (Rpb1-12) in eukaryotic cells [156]. The carboxyl-terminal domain (CTD) of the largest RNA Pol II subunit, Rpb1 (200kD), is an important recipient of regulatory signals during all steps of transcription. This is critical for nascent RNA synthesis and cotranscriptional processing [157], [158]. The CTD contains a series of heptapeptide repeats that range from 52 in humans to 26 in yeast, with the consensus sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 [159]–[164]. CTD heptapeptide repeats were subjected to a variety of modifications, including methylation, acetylation, and phosphorylation [156], [165], [166]. The phosphorylation on Ser2, Ser5, Ser7, Thr4, and Tyr1 of the RNA Pol II CTD is a major regulation model during transcription in eukaryotes [167]–[171]. Many enzymes are involved in the modification of RNA Pol II CTD during transcription. Among them, CDKs (cyclin-dependent protein kinases) play critical roles in this regulation of RNA Pol II phosphorylation. CDKs are serine/threonine-protein kinases and are essential for many cellular processes, such as cell cycle, cell growth, proliferation, and transcriptional

regulation in response to extracellular and intracellular signals. CDKs are constitutively expressed and perform their functions only after binding with the appropriate cyclin. To date, at least 21 CDKs and 30 cyclins have been confirmed based on structural and functional differences. CDKs are classified into two types based on their functions: cell-cycle CDKs (e.g., CDK1/2/4/6) and transcriptional CDKs (e.g., CDK7/8/9/12/13/19) [172]–[175].

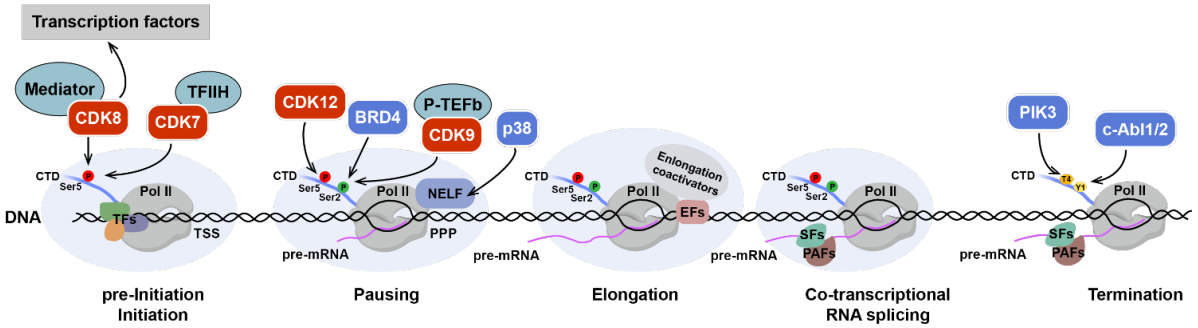

Figure 7: The brief scheme of the transcription process regulated by kinases and transcriptional condensates. Light blue bubbles indicate the transcriptional condensates.

In eukaryotic cells, transcription initiation starts with the recognition of the region upstream of the promoter by general transcription factors and positioning of RNA polymerase, RNA Pol II to a specific DNA sequence referred to as a "promoter", such as TATA box. RNA Pol II recruitment initiates the stepwise transcription process to form a pre-initiation complex (PIC) with TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH at the promoter site, followed by transcription initiation, elongation, and termination. The interval between initiation and elongation is generally known as pausing of gene transcription [176]. RNA Pol II unwinds 14 base pairs of DNA to form an RNA polymerase-promoter "open complex". The promoter DNA is partially unwound and single-stranded in the "open complex". The exposed, single-stranded DNA is referred to as "transcription bubble". RNA Pol II then selects a transcription start site in the transcription bubble, binds to an initiating NTP and an extending NTP (or a short RNA primer and an extending NTP) complementary to the transcription start site sequence, and catalyzes bond formation to yield an initial RNA product with the help of one or more general transcription factors. Transcription initiation is regulated by additional proteins known as activators and repressors and associated coactivators or corepressors in some cases, which modulate the formation and function of the transcription initiation complex. After the first bond is synthesized, the RNA polymerase must escape the promoter. It occurs mechanistically through DNA scrunching, which provides the energy required to break interactions between RNA Pol II holoenzyme and the promoter. RNA Pol II will be forced to pause during the early stages of transcript elongation. It will be released by the regulation of CTD of RNA Pol II and pausing complex, such as the Negative Elongation Factor (NELF) complex (see below for more details). At the end of elongation, the 3' end of the gene, RNA Pol II will encounter the specific DNA sequence, termed as terminators that signal that RNA transcript is complete. Once they are transcribed, they will cause the transcript, pre-mRNA, to be released from the RNA Pol

II. During all steps of transcription, phosphorylation plays a key role, especially the phosphorylation of RNA Pol II.

At the pre-initiation of gene transcription, CDK8/19 and CDK7 are major kinases that phosphorylate CTD. CDK8 and its paralog CDK19 play critical roles in transcription regulation by interacting with mediator complexes or phosphorylated transcription factors (**Figure 7**) [176], [177]. After pre-initiation, CDK7 primarily regulates the transcriptional process by phosphorylating Ser5 and Ser7 of the RNA Pol II CTD, thus promoting promoter clearance and transcription initiation. CDK7 can also phosphorylate various CDK(1, 2, 4, and 6) kinases, as the catalytic core of the CDK-activated kinase (CAK) complex, which is involved in cell cycle regulation (**Figure 7**) [178]. During elongation, BRD4 (bromodomain-containing protein 4) and CDK9 are critical kinases that ensure the smooth progression of gene transcription. CDK9, a P-TEFb (Positive Transcription Elongation Factor b) subunit, phosphorylates Ser2 of CTD, eliminating pausing and shifting RNA pol II into an elongation mode. However, the arrival of CDK9 to promoters is inextricably linked to the phosphorylation of Ser2 during pausing, which is carried out either directly or indirectly by BRD4 [179]. CDK12 is also an important kinase in the transition from transcriptional initiation to transcriptional elongation mode of RNA Pol II. According to research conducted in Drosophila and yeast, CDK9 phosphorylates Ser2 during the early stages of transcription before CDK12 takes over, which phosphorylates Ser2 of CTD at the elongation stage (**Figure 7**) [180], [181]. At the 3'-end of the active gene, RNA Pol II encounters the poly-A signal on mRNA then terminates the transcription. During transcription termination, Plk3 (Polo-like kinases 3) and c-Abl1/2 (Abelson murine leukemia viral oncogene 1 and 2) phosphorylate Thr4 and Tyr1 of CTD (**Figure 7**) [182], [183]. However, the specific effects of Plk3 and c-Abl1/2 on the regulation of transcription still remain debatable.

Under normal conditions, the phosphorylation levels of CTD at the corresponding amino acids are certain, ensuring RNA Pol II performs its respective functions during the various stages of the transcription process. The enhancement or overexpression of CTD kinases would influence the RNA Pol II functions, resulting in abnormal gene transcription. However, it is rapidly changed and regulated when cells are under stress, especially transcription stress, as cells are naturally exposed to internal and external stimuli.

1.2.3 The cellular response to transcription-blocking lesions (TBLs)

Various toxic agents will induce DNA lesions during gene transcription, affecting faithful transcription. Environmental agents such as UV (ultraviolet), X-rays, and several genotoxic chemicals cause alterations in DNA structure [184], [185]. The normal cellular metabolites will also induce oxidative lesions, and spontaneous depurination will generate abasic sites [186]. Despite these oxidative lesions transiently hampering the forward translocation of RNA Pol II, transcription machinery could bypass the damage in cooperation with Cockayne Syndrome protein B (CSB), an ATP-dependent 3′-to-5′ single-strand DNA translocase from the SWI2/SNF2-family. In contrast, some bulky DNA lesions induced by UV irradiation or other chemical agents (Benzo[α]pyrene (BAP) and 4-Nitroquinoline-1-oxide (4NQO)) can not be bypassed by transcriptional machinery [187], [188], [189]. These lesions will induce RNA Pol II stalling and thus induce prolonged transcriptional arrest, referred as "transcription-blocking lesions (TBLs)" [190], [191].

Briefly, TBLs can be classified into four species: Bulky base adducts, DNA-protein crosslinks, inter/intra-strand crosslinks, and double-strand breaks (DSBs).

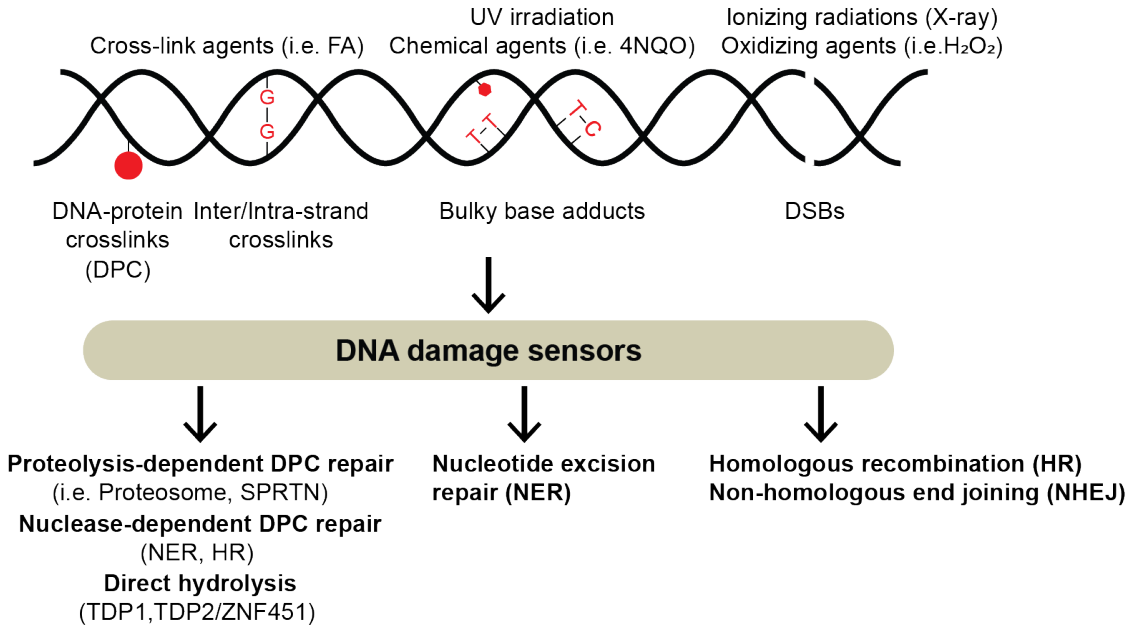

Figure 8: Transcription-blocking DNA lesions, their origins, and the corresponding repair pathways. TDP stands for. NER and HR pathways can directly cleave DNA molecules in a nuclease-dependent manner. Tyrosyl-DNA phosphodiesterase1 (TDP1) and TDP2/ZNF451 are able to directly hydrolyze the chemical bond between topoisomerases 1/2 (TOP1/TOP2) and DNA. SPRTN is a DNA-dependent metalloprotease. Proteasome and SPRTN are merged to the proteolysis-dependent removal of covalently bound proteins. The details of the pathways (NER, HR and NHEJ) are described below.

Bulky base adducts

Some chemical agents are known to induce TBLs in cells, such as Benzo[a]pyrene (BAP) and 4-Nitroquinoline-1-oxide (4NQO). Benzo[a]pyrene (BAP), present in tobacco smoke, is metabolized by cytochrome P-450 (CYP450) to form a number of active intermediates, including (+)- and (-)-anti-benzo[a]pyrene diol epoxide (BPDE), (+)- and (-)-3,4-dihydroxy-1,2-epoxy-1,2,3,4-tetrahy-dro-

benzo[c]phenanthrene (BPhDE). BPDE and BPhDE covalently bind and form adducts at the N2 position of guanine and adenine bases, resulting in the formation of bulky lesions that block RNA Pol II elongation. 4-Nitroquinoline-1-oxide (4NQO) is a quinoline derivative and a tumorigenic compound that could form UV damage-mimetic bulky adducts primarily with guanine bases. In response to 4NQO-induced DNA damage, RNA Pol II stalls and is known to be degraded [192].

UV irradiation has been well known to induce bulky DNA lesions in cells and inhibit transcription [144]. UV light is classified based on its wavelength range. UV-C (190-290 nm) from the sun is primarily blocked by the stratospheric ozone layer. UV-B (290-320 nm) radiation is the most dangerous spectrum of UV light that reaches the surface and contributes to the development of skin cancer, causing the same types of DNA damage as UVC. UV-B is mostly absorbed by the ozone layer. Only lower energetic UV-A radiation (320-400 nm) reaches the planet's surface almost completely unfiltered (95%) [193]–[195]. Nonetheless, UV-C is widely used in laboratories to study UV-induced DNA damage. This is due to the high energy content of UV-C and the overlap with the energy absorption maximum of DNA at 260 nm, which results in the strong formation of photoproducts [196]–[198]. UV-C light induces dimerization of adjacent pyrimidines in DNA strands, resulting in photolesions, such as 6-4 photoproducts (6,4-PPs) and cyclobutane pyrimidine dimers (CPDs) (**Figure 9**) [199], [200]. RNA Pol II can incorporate nucleotides opposite CPDs and 6,4-PPs, but it stalls at the damage, forming a stable ternary complex of RNA Pol II, the damaged template, and a nascent transcript. These RNA Pol II conformation changes render the lesion sites inaccessible to DNA repair machinery [201].

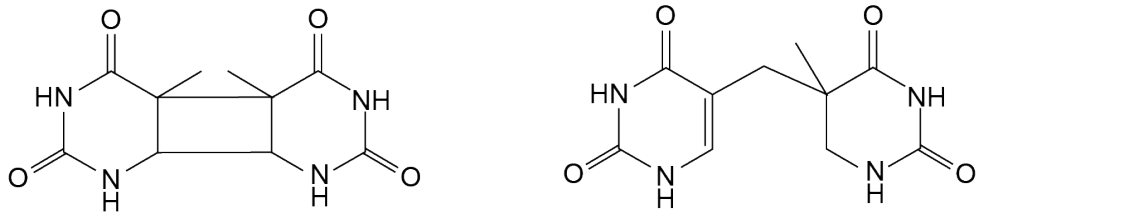

Figure 9: UV-induced photolesions.

Nucleotide-excision repair (NER) pathway

The stalled RNA Pol II serves as a detection machinery and sensor of TBLs in actively transcribed genes, which will signal the transcription-coupled NER (TC-NER) to start. CSB will be the first protein recruited to the stalled RNA Pol II. Under normal conditions, the interaction between CSB and RNA Pol II is highly dynamic. Therefore, it can be overcome easily by a known process, bypass [202]. However, when this complex stalls at the bulky DNA lesion, CSB stably interacts with RNA Pol II. One possible explanation is that the C-terminus of CSB latches onto RNA Pol II or binds to DNA upstream of RNA Pol II to mediate the initial interaction, then disassociates from the ATPase domain, thus allowing the access to form a more stable interaction with RNA Pol II [203], [204].

Globally, when UV photolesions happen at the active and inactive gene regions, they are repaired in a pathway known as the global-genome NER (GG-NER). The DNA lesion will be recognized by two complexes, including XPC-RAD23B and XPC-DNA-damage binding (DDB) or UV-DDB complex. The XPC-Rad23B complex is in charge of distortion detection, composed of XPC, either RAD23A or RAD23B and CETN2 [205], [206]. The UV-DDB complex is a ubiquitin ligase composed of DDB1, CUL4A or CUL4B, RBX1, and a GG-NER specific protein DDB2 [163]. It can also recognize DNA damage induced by UV light. *In vitro*, the UV-DDB complex is only necessary for GG-NER mediated DNA repair, but not TC-NER [207], [208].

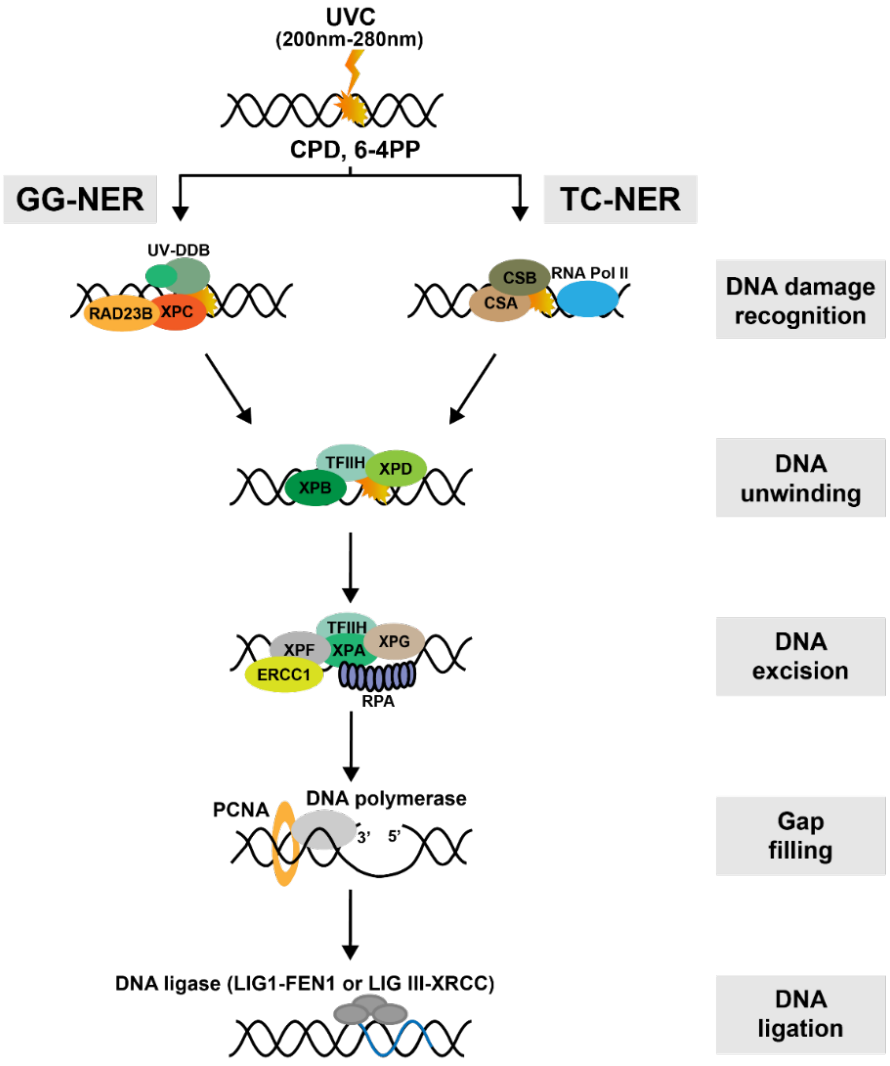

Figure 10: The simplified scheme of nucleotide excision repair (NER) pathway. The main steps of the global genome NER (GG-NER) and transcription-coupled NER (TC-NER) repair mechanisms are shown.

Following these DNA lesion recognition and TC-NER initiation steps, the stalled RNA Pol II will be dissociated through two known modulations: degradation and backtracking [209]. The removal of the stalled RNA Pol II will provide access to the DNA lesion for the multi transcription factor II H (TFIIH) complex subunits. Then XPA is recruited to TFIIH, followed by a dissociation of CDK-activating kinase (CAK) complex that activates the helicase activity of XPD [210]–[212]. Two TFIIH

helicase subunits, XPD and XPB, in collaboration with XPA, verify DNA damage and unwind the DNA helix around the lesions, thereby creating an open "bubble" platform for the recruitment of later repair machinery. Then XPG and XPF–ERCC1 complex is recruited to the TFIIH complex at DNA lesions stabilizing TFIIH. XPG also has endonuclease activity, and it cuts DNA damage on the 3'-end, whereas the XPF–ERCC1 complex cuts on the 5'-end. The dual incision leads to the removal of a piece of ssDNA with 25~30 nucleotides from the strand containing lesions and produces exposed ssDNA and a gap around the DNA lesion [213]–[215]. This ssDNA will be covered by RPA (replication protein A) [216]. On the other hand, the XPA-RPA complex also helps orient the endonuclease process [217], [218]. The gap will be filled in the replication process. The proliferating cell nuclear antigen (PCNA) is loaded by replication factor C (RFC) onto the DNA strand. This enables DNA polymerases involved in repair (δ, ε and/or κ) to translocate and copy the undamaged strand. To complete NER, DNA ligase I (LIG1) and Flap endonuclease 1 (FEN1) or the Ligase-III-XRCC1 complex will seal the nicks [219]–[221].

DNA-protein crosslinks and inter/intra-strand crosslinks

In eukaryotic cells, DNA is coated with proteins and forms a highly compact and dynamic chromatin structure. Dynamic interactions of DNA and proteins are essential for various cellular processes in a timely and spatially regulated manner, such as transcription and DNA replication. However, under certain conditions, proteins can covalently bind to DNA, resulting in the formation of bulky DNA lesions, such as DNA–protein crosslinks (DPC). Based on the properties of crosslinked proteins, DPCs can be classified as enzymatic or nonenzymatic. Enzymatic DPC formation can be induced when many DNA-related enzymatic reaction intermediates are trapped to DNA, such as DNA topoisomerases, DNA methyltransferases, and AP lyases [222]–[224]. Of note, Topoisomerase I (TOPI)-DNA crosslinks can be generated when TOP1 is trapped at the 3′-end of a DNA single-strand nick in cells with topoisomerase poisons like camptothecin (CPT) [225]. Non-enzymatic DPCs are very common in cells, as they can be induced by endogenous factors, such as aldehyde production (e.g., formaldehyde) during cell metabolism. Some exogenous sources like ionizing radiation, UV light, and chemotherapeutic agents will also lead to DPC formation [223], [226], [227]. Depending on the properties of DPCs, cells employ distinct repair pathways to deal with them. In mammalian cells and bacteria, the nucleotide excision repair (NER) pathway repairs formaldehyde-induced DPCs with crosslinked proteins of less than 16 kDa [228], [229], [230], [231]. TOPI-DNA crosslinks are resolved by tyrosyl-DNA phosphodiesterases (TDP1) and NER [223].

Platinum-containing chemical agents, such as cisplatin, are broadly used for anti-cancer treatments, such as ovarian cancer and breast cancer [232]–[235]. They will induce intra-strand crosslinks and inter-strand crosslinks in cancer cells [236], [237], [238]. RNA Pol II will be stalled at intra-strand or inter-strand crosslink lesions and function as a DNA damage sensor to trigger the TC-NER pathway, and thus remove these TBLs [239]. In addition, inter-strand crosslinks can also be repaired by HR or translesion synthesis (TLS).

DNA double-strand breaks (DSBs)

Double strand breaks (DSBs) have been reported as a serious type of DNA damage. DSBs can be induced by X-ray or some chemical agents during replication or transcription [240]–[242]. It will affect the transcription initiation and elongation when DSBs are near or at transcription sites. However, DSB-induced transcription blocking is not a physical blockage because it can be run off by RNA Pol II. This transcriptional blockage induced by DSBs is regulated by kinases ATM and DNA-PK signaling pathways, as well as some proteins, such as transcriptional factor NELF (Negative elongation factor) and chromatin remodeling and organizing factor cohesin [243], [240], [244], [245]. ATM regulates the transcription suppression near DSB sites at transcription start sites (TSS) by recruiting RNF8/RNF168 E3 ubiquitin ligases to force RNA Pol II to pause and thus decrease the phosphorylation of RNA Pol II at Ser2 [246]. Of note, ATM also phosphorylates PBAF (Polymorphic BRG-/BRM-associated factor), which belongs to the SWI/SNF chromatin remodeling complexes. These phosphorylations will induce chromatin structure changes and negatively regulate transcription. ATM-dependent transcription repression has been proposed to activate the NHEJ pathway to repair the DSB. However, the mechanisms remain unclear [247], [248]. In contrast, DNA-PKcs is reported to regulate DSB-induced transcriptional repression within gene bodies by removing RNA Pol II. The stalled RNA Pol II will be degraded by DNA-PKcs signaling with the involvement of the HECT E3 ubiquitin ligase WWP2 (WW Domain Containing E3 Ubiquitin Protein Ligase 2) to promote the repair of DSBs by the NHEJ pathway [249], [250]. WWP2 is recruited by DNA-PKcs in response to DSBs and then ubiquitylates eight lysine residues of the CTD of RNA Pol II subunit, RPB1. This will result in the removal of the stalled RNA Pol II for proteasomal degradation from the DSB sites in the gene bodies [249]. NELF promotes RNA Pol II to pause at the transcription start sites (TSS) with the aid of DSIF (DRB sensitivity inducing factor) to inhibit transcriptional elongation [251]. NELF-E, a subunit of NELF, can be phosphorylated by ATM and recruited to DSB sites to inhibit transcription [244]. Furthermore, poly (ADP-ribose) polymerase 1 (PARP1) has been shown to regulate DSB-induced transcription repression by modifying RNA Pol II to enhance its interaction with the NELF complex [244], [245]. The association of NELF with RNA Pol II is reported to promote the DSB repair by both NHEJ and HR pathways [244].

DNA double-strand break repair

Double-strand break repair can be accomplished through different pathways, which are commonly classified into two broad categories based on whether or not a homologous DNA sequence is used as a template, like Homologous recombination (HR) and non-homologous end-joining (NHEJ). DNA double-strand breaks rapidly recruit and activate PARP1, as one of the earliest events [252]. PARP1 synthesizes a structurally-complex polymer composed of ADP-ribose units upon activation, which aids in local chromatin relaxation and the recruitment of DNA repair factors by inducing the ADP-Ribosylation of histones and other associated proteins. Both HR and NHEJ pathways include similar steps to the NER pathway: DNA damage recognition, DNA end processing, and ligation. However, the involved proteins differ in HR and NHEJ.

Non-homologous end joining (NHEJ)

Non-homologous end joining (NHEJ) is the direct resealing of two broken ends that occur regardless of sequence homology. Arguably, the NHEJ pathway is more straightforward, which uses minimal processing of the break ends before rejoining them. NHEJ is initiated by the detection of the DSBs by the Ku heterodimer (Ku70-Ku80, or XRCC6-XRCC5) within seconds. It forms the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) (DNA-PKcs) holoenzyme with Ku70/Ku80 [253]. Upon DNA-PKcs recruitment to DSBs sites, DNA-PKcs is activated to phosphorylate itself and its targets, including Protein Artemis (DCLRE1C), which trims DSB ends via its nuclease activity to remove the single nucleotides from DNA ends [254], [255], [256]. In addition, PARP1 also facilitates the Poly(ADP-ribosyl)ation (PARylation) of DNA-PKcs to increase its activity [257], [258]. Further, this DNA-PKcs/Ku70/Ku80 complex promotes the recruitment of downstream NHEJ factors XRCC4, XRCC4-like factor (XPF), DNA ligase 4 (LIG4), which mediate the rejoining process. The LIG4-XRCC4-XPF complex and the newly identified PAXX (paralog of XRCC4 and XLF) re-ligates the blunt DNA ends [259]. Since repair by NHEJ can include end-processing steps and does not rely on the sequence homology of a sister chromatid, it is inherently error-prone and can contribute to genome instability through chromosomal translocations and deletions [260]. An alternative but less characterized NHEJ pathway is Ku-independent. Fittingly, it is termed alternative end-joining (aEJ) or microhomology-mediated end-joining (MMEJ), as it relies on the presence of small stretches of homology (5-25 nts) between the sequences flanking the DSB [261]. Since overlapping DNA flaps are excised, aEJ always results in sequence deletions [262], [263]. Although being active throughout the cell cycle, NHEJ is relatively more important during G1 (**Figure 11**) [264], [265].

Homologous recombination (HR)

In contrast to NHEJ, homologous recombination (HR) necessitates the use of a homologous DNA sequence as a template to recover sequence information lost at the break site and involves extensive DNA-end processing as well [263]. As expected, HR is extremely accurate, resulting in error-free repair. HR predominantly uses the homologous DNA sequence from the sister chromatid as a template rather than the homologous chromosome [266]. This tight regulation is ensured through a strong inhibition of HR during G1 when a sister chromatid is absent and the nature of the newly replicated chromatin, which favors HR. Therefore, HR is restricted to late S and G2 of the cell cycle [267], [268]. DSBs are recognized by the MRN complex, which is formed by MRE11, RAD50, and NBS1 and recruited by PARP1 [269], [270], [271]. The MRN complex, with the aid of CtBP-interacting protein (CtIP), generates short 3' overhangs, known as "DNA end resection" [272]–[274]. Similar to Ku

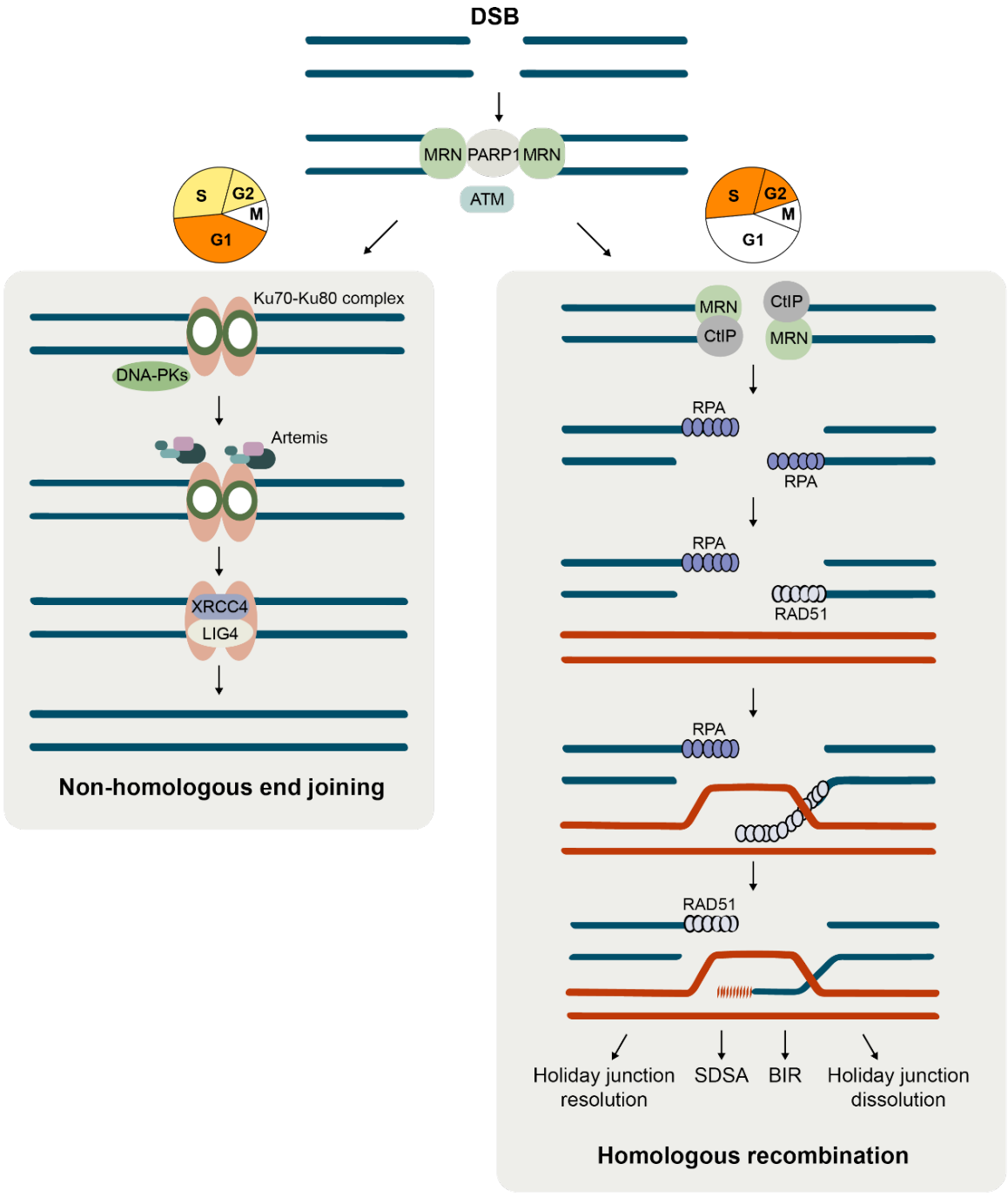

Figure 11: DSB repair pathways. The main steps of the non-homologous end joining (NHEJ) and homologous recombination (HR) repair mechanisms are shown. Homologous recombination can occur via multiple pathways, all of which share the same initial steps (Holliday junction resolution, synthesis-dependent strand annealing (SDSA), break-induced replication (BIR), and holiday junction dissolution). The cell cycle is a major determinant of which DSB repair mechanisms are used. NHEJ occurs throughout the cell cycle, but mainly in the G1 phase. Homologous recombination occurs only during the S/G2 phases [275].

and DNA-PKcs in NHEJ, MRN is important for the recruitment and activation of ATM (Ataxia Telangiectasia Mutated) kinase. ATM mediates a signaling cascade on chromatin surrounding the break site through a series of PTMs and protein recruitment and amplifies the damage signal [276]–[278]. The resections are extended by Exonucleases 1 (EXO1), Bloom syndrome protein (BLM), DNA replication ATP-dependent helicase DNA2 (DNA2), which generates extensive tracks of ssDNA that are rapidly covered by the heterotrimer RPA (RPA1-2-3). The binding of RPA stabilizes the ssDNA by preventing

its degradation, the formation of secondary structures, and the spontaneous recombination between regions of microhomology [279]–[281]. Then recombinase RAD51 replaces RPA, and directs homology search and strand invasion of the homologous template, resulting in the formation of a displacement loop (D-loop) structure [282]–[285]. After strand invasion, POLδ or translesion polymerases use the invading strand in the D-loop as a primer for strand elongation [286].

Most of the extended D-loops are disrupted and subsequently repaired by synthesis-dependent strand annealing (SDSA). SDSA always results in a non-crossover [287]. The extended D-loop can also be captured or invaded on the other end to form a double Holliday junction (dHJ). As a result of the BLM-TOP3A-RMI (BTR) complex, which drives the two junctions towards each other before they are cleaved, dHJ dissolution can result in non-crossover products [287], [288], [289]. Crossover products are formed when the GEN1 resolvase or the MUS81-EME1/SLX1-SLX4 nuclease complex is used. Alternatively, non-error-free pathways, such as break-induced replication (BIR) and single-strand annealing (SSA), can also take place (**Figure 11**) [290].

1.3 TBLs affect transcription and RNA metabolism

1.3.1 Transcription is shut down in response to TBLs

TBLs in transcribed genes cause a global transcriptional shutdown, for example, in response to UV-induced damage, while some individual genes are highly upregulated as a result of DNA damage. RNA-seq analysis for the level of newly synthesized RNA (by RNA Pol II) following UV irradiation revealed an immediate inhibition of transcription elongation within 1h after UV and subsequent inhibition of transcription initiation at the transcription start sites (TSS) in both damaged and un-damaged genes within 3h after UV irradiation [291], [292]. This is accompanied by the release of RNA Pol II from promoter-proximal pause (PPP) sites and the global loss of the hypophosphorylated form of RNA Pol II, decreasing its hypophosphorylated form globally [293], [191]. The normal transcription level is fully restored after 24-48 hours post UV stress. However, transcriptional initiation recovers faster than elongation as the latter remains slow for many hours [291], [294]. Furthermore, transcription appears to be 'spatially restricted' for long periods of time, with the promoter-proximal 20–25 kb of genes showing much more activity than areas further downstream in response to UV-induced DNA damage (**Figure 12**) [292], [293], [295]. This is also one of the reasons that almost all genes induced by DNA damage are short, and more short mRNAs are produced upon UV irradiation.

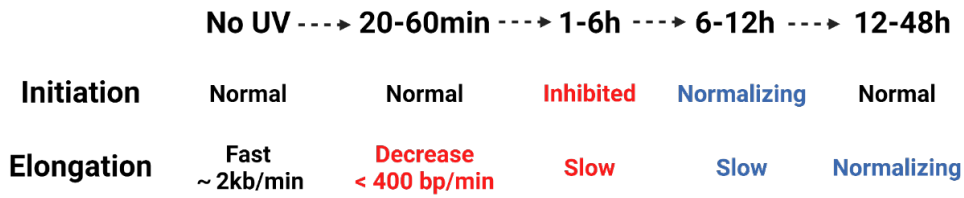

Figure 12: UV induces transcription inhibition. Recovery dynamics is summarized.

In addition, RNA Pol II shifts from promoter-proximal pause sites to the slow elongating mode of RNA Pol II, thus increasing TBL detection, providing space for repair proteins, and preventing collisions between the DNA damage repair and transcription machinery [295]. *De novo* transcription of short genes is temporarily increased, boosting TBL detection chances by RNA Pol II, which is regulated by phosphorylation events, such as those that take place at the p38-NELFE axis [292], [296]. After UV irradiation, p38 MAPK signaling is activated and promotes the releasing of the RNA-binding protein RBM7 from the 7SK small nuclear ribonucleoprotein (snRNP), which sequesters active p-TEFb into inactive 7SK snRNP, resulting in chromatin localization and activation of p-TEFb [297]. The p38 MAPK also phosphorylates NELFE and promotes its dissociation from RNA Pol II on chromatin. Subsequently, these regulations will induce changes in the phosphorylation of RNA Pol II, stimulating a wave-like release of the promoter-proximal paused RNA Pol II into productive elongation [296]. These actions, taken together, induce the continuous release of promoter-proximal paused RNA Pol II into productive elongation, possibly accelerating TBL detection, recognition, and subsequent TC-NER initiation.

1.3.2 Splicing of pre-mRNA in response to UV irradiation

Most pre-mRNAs mature to mRNAs co-transcriptionally by the addition of a 7-methyl guanosine cap to the 5' end (capping), the removal of intronic sequences via splicing catalyzed by the spliceosome, and the endonucleolytic cleavage at the 3' end followed by the synthesis of a polyadenosine (poly(A)) tail, known as polyadenylation [298]–[301].

Alternative splicing requires the use of various splice site combinations, resulting in a variety of splicing patterns, including alternative 5′ splice site selection (Alt5), alternative 3′ splice site selection (Alt3), the skipping of complete exons (SE), and the retention of introns (RI) (**Figure 13**) [302], [303]. Alternative splicing of pre-mRNA is an important mechanism for increasing the complexity of gene expression, cellular differentiation, and organism development [304]–[308].

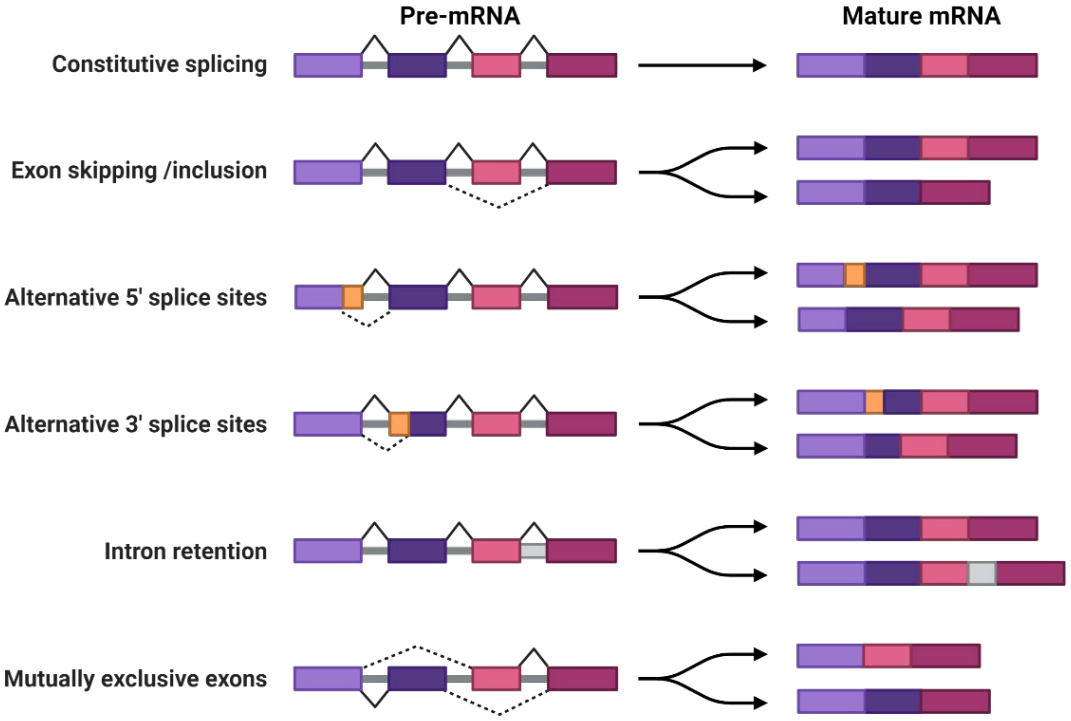

Figure 13: The simplified scheme of splicing types. It was created with BioRender.com.

TBLs appear to induce alternative mRNA splicing. UV irradiation affects splicing efficiency and decisions by changing the transcription elongation rate to induce spliceosome mobilization, such as the removal of spliceosome subunit small nuclear ribonucleoproteins (snRNP) from chromatin [309]. In response to UV irradiation, alternative last exon (ALE) splicing occurs preferentially at the 5′ end of genes, such as the pro-apoptotic mRNA isoforms of Bcl-x [310], [311]. TBL-induced alternative splicing is known to be mediated by two major stress kinases, ATR and ATM. ATR is activated after UV damage and mediates hyperphosphorylation of RNA Pol II, which causes transcription to slow and alternative splicing to occur. Nonetheless, it appears that the formation of UV-induced DNA dimers (e.g., CPD), rather than stalled RNA Pol II, is responsible for this process [312]. UV-induced TBLs also stimulate noncanonical ATM activation. In response to UV, late-stage spliceosomes are rapidly

excluded from RNA Pol II stalled sites, increasing R-loop formation, which activates ATM and further regulates spliceosome dissociation and global alternative splicing events [309], [313].

1.3.3 Polyadenylation of pre-mRNA

Polyadenylation (poly(A)) is the process by which pre-mRNA is cleaved by endonucleases, and the poly(A) tail is added at the cleavage/polyadenylation sites (polyA sites or pA sites) [314]. Alternative polyadenylation (APA) produces multiple mRNA polyadenylation isoforms through coordinative actions of several factors. The 3′-processing factors are the major targets of APA regulation. The interaction of specific sequence elements within the pre-mRNA with the 3' end processing complex determines the 3' end cleavage site. This complex is assembled by the poly(A) polymerase and four multisubunit complexes: cleavage polyadenylation specificity factor (CPSF), cleavage factors Im and IIm (CFIm and CFIIm), and cleavage stimulation factor (CSTF) [315]–[318]. Normaly, APA processing includes the following steps: (1) CFIm (cleavage factor I) binds to the UGUA field of pre-mRNA upstream of the pA site and attracts CPSF and CSTF to assemble at the end of RNA Pol II; (2) CPSF binds to the specific pA signal sequence (e.g., AAUAAA) and CSTF is transferred to the new mRNA precursor to bind to the GU or U-rich sequence [319]–[322]; (3) Then CPSF and CSTF initiate the cleavage after the pA signal sequence for around ~ 35 nucleosides, and polyadenylation binding protein (PABPN1) will bind to the polyadenylation tail sequence to start the APA process [323], [324]; (4) When poly(A) polymerase-mediated polyadenylation is going, adenosine tails of 50–250 nucleotides (nt) are prepared, and CPSF dissociates from the binding sequence; (5) PABPN1 determines when the polyadenylation process stops and then Poly(A) polymerase starts dissociating from the RNA [325], [326], [327]. The combination of the preceding 5 steps and the 5′-capping process promotes mRNA maturation. Eventually mRNA is exported from the nucleus to the cytoplasm for translation [316], [318], [322], [328]–[332].

APA is a dynamic and spatio-temporally coordinated process involving a number of key factors. Approximately 50-80% of mammalian pre-mRNA typically contains multiple polyA sites (PAS) [333], [334], [335]. Current studies show that the distance between the proximal and distal PASs, the RNA Pol II elongation rate, and the efficiency of PAS recognition at both proximal and distal sites determine the use of the proximal PAS [336], [337], [338]. Normally, the usage of the proximal PAS generates short isoforms. This may result in the suppression of translation, which is regulated by many factors in cells (reference). Among them, the necessary mRNA 3′ processing factor CFIm appears to be especially important as CFIm-mediated APA regulation has been linked to tumor suppression, hepatocellular carcinoma as well as neurological disorders [318], [339], [340]. It is regulated by the CFIm complex. The CFIm complex is made up of CFIm25 (also known as NUDT21 or CPSF5), CFIm59 (CPSF7), and CFIm68 (CPSF6), which are members of the SR superfamily proteins. CFIm25 binds specifically to a UGUA motif and forms a dimer. CFIm68/59 binds to the CFIm25 dimer via their RNA recognition motif (RRM) domains to form a tetrameric CFIm complex [341]–[345]. The depletion of CPSF5 or CPSF6, but not CPSF7, results in the selection of proximal PAS and a shortening of the 3′ UTR

(untranslated region) in HEK293 cells [346], [328]. Especially, CPSF6 plays an important role in activating mRNA 3′ end processing by binding to the RE/D domain of the CPSF subunit FIP1L1, thus promoting the recruitment of core processing machinery in yeast and mammalian cells. More recently, the function of CPSF6 in regulating APA has been linked to the progression of hepatocellular carcinoma (HCC) through the short isoform of NAD(P)H Quinone Dehydrogenase 1 (NQO1). The high levels of CPSF6 were linked to a poor prognosis in HCC patients [347]. In addition, the phosphorylated CPSF6 at the arginine/serine (RS)-rich domain binds less efficiently to FIP1L1, resulting in more short mRNA production [318], [340]. This suggests that the hyperphosphorylation of CPSF6 seems to inhibit their functions (**Figure 14**).

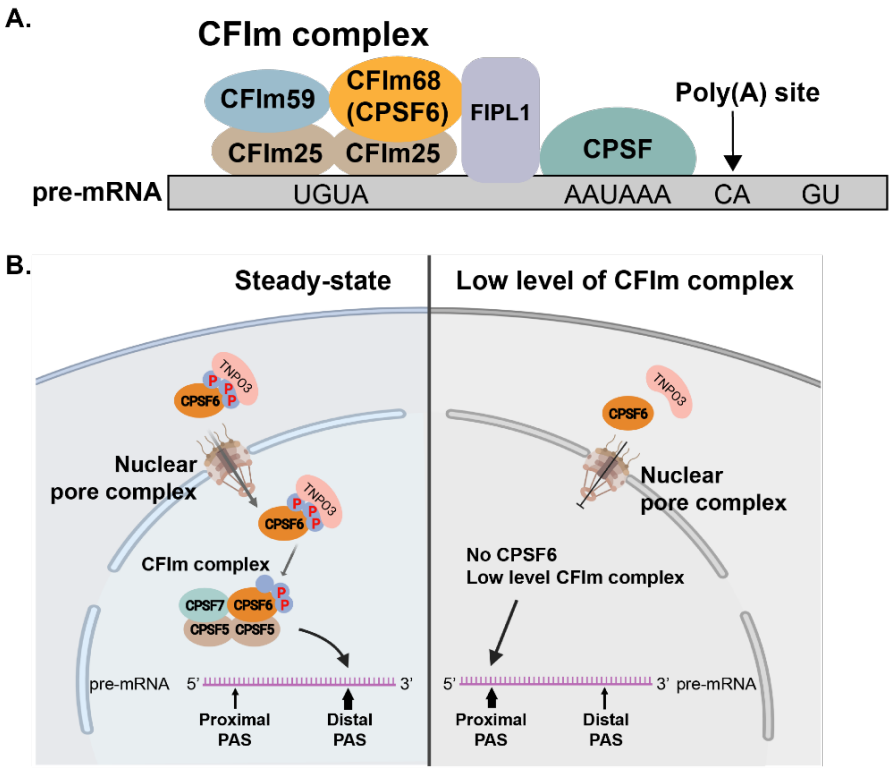

Figure 14: The CFIm complex and the regulation of alternative polyadenylation activity by CPSF6. A. The CFIm complex. The CFIm complex binds to the conserved upstream UGUA region to mediate the cleavage reaction and recruit other proteins. After forming a complex with PAP, this complex moves through the pre-mRNA in a 5′ to 3′ direction. When the adenosine acidification signal CPSF arrives at the AAUAAA region, it recognizes the polyadenylation signal AAUAAA and cleaves the mRNA. After that, CSTF binds to the GU- or U-rich sequence. Between the polyadenylation signal AAUAAA and the cleavage site is a U-rich region bound to the FIP1L1 subunit of the CPSF. PAPs catalyze the addition of untemplated adenosines, while Symplekin acts as a scaffold protein. **B. The phosphorylation of CPSF6 regulates APA.** In the steady-state, CPSF6 is a nuclear protein. Following the formation of the CFIm complex in the nucleus, RSLD phosphorylation governs the choice of pre-mRNA PAS. Nuclear import proceeds efficiently in the absence of RSLD phosphorylation of CPSF6, but RSLD hypophosphorylation results in a significant number of abnormally distal PASs. When CPSF6 is missing from the nucleus due to gene knockout or the expression of a mutant protein that is unable to bind TNPO3, such as hyperphosphorylated CPSF6, PASs shift dramatically to the proximal position. PAS stands for polyadenylation site [348]. The figure is created with BioRender.com.

Moreover, CFIm subunits have been found in purified human spliceosomes [349]. It implies that CFIm is involved in splicing regulation. Recently, it was discovered that CPSF binds to spliceosome subunit U1-70K and regulates global alternative splicing [350]. It has been demonstrated that spliceosome subunit U2AF65 interacts with CFIm59 to stimulate mRNA 3′ processing [351]. These interactions between polyadenylation factors, CFIm complex and splicing factors at terminal exons of pre-mRNA possibly provide a common binding platform or regulatory machinery for cross-regulation

and coordination of APA and alternative splicing processing. Last intron removal is known to be influenced by alternative polyadenylation in conjunction with alternative splicing [352], [353]. However, whether APA events influence alternative splicing decisions at upstream exons is not very well characterized. Recent studies proposed that the mechanistic link between APA and pre-mRNA splicing is restricted to terminal exon definition [354], [355]. The generation of different 3′-UTRs is primarily responsible for the regulation of mRNA isoform stability and translatability. Therefore, the coordination of APA and upstream AS could result in alternatively spliced mRNA isoforms with distinct 3'-UTRs that govern mRNA half-lives and protein product functions [356].

1.3.4 Production of short mRNA is increased in response to UV irradiation

As mentioned above, UV irradiation induced RNA Pol II stalling at transcription blocking lesions. RNA Pol II molecules are dynamically and synchronously released from promoter-proximal regions into elongation in response to UV irradiation, promoting the detection of TBLs and accelerating the surveillance by global transcription, evidenced by RNA Pol II progressive accumulation at TBLs and a general transcription inhibition in gene bodies [291]. Although UV irradiation induced transcription inhibition globally, RNA-seq and ChIP-seq analyses have revealed that more short genes

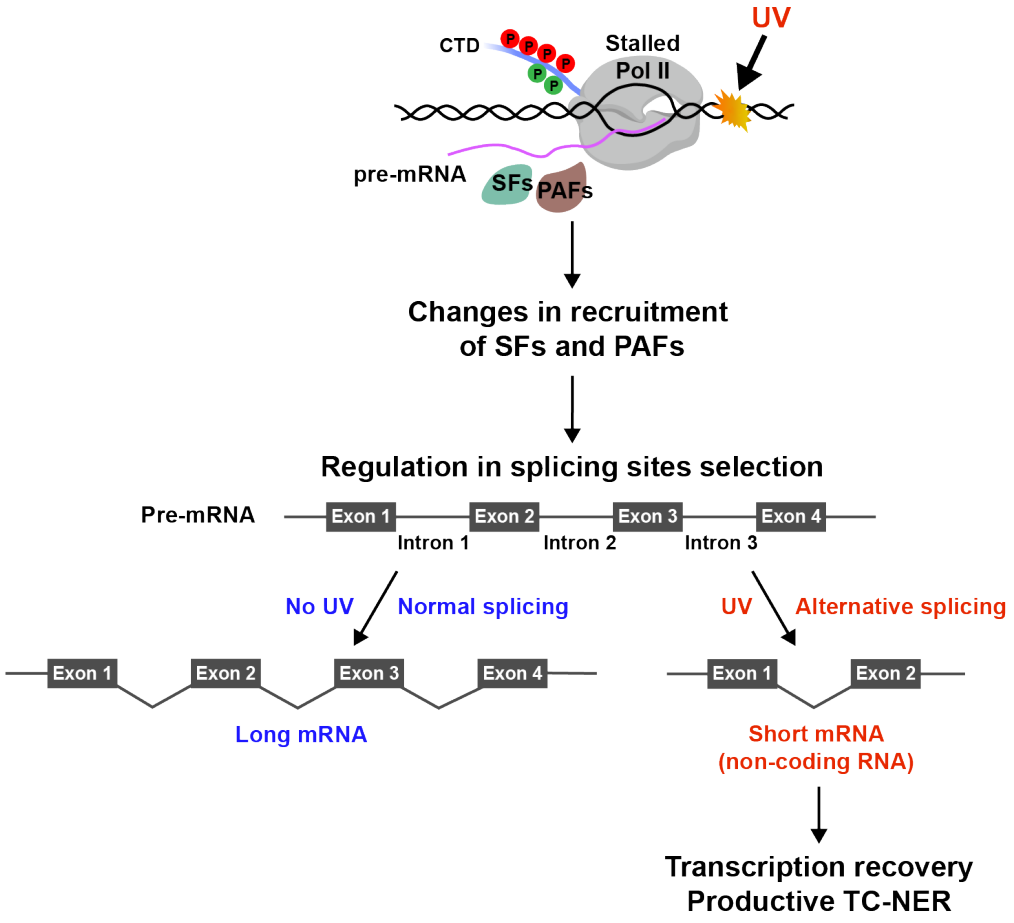

Figure 15: UV induces more short mRNA production in cells. Stalled RNA Pol II induced by UV irradiation will affect the co-transcriptional recruitment of splicing factors (SFs) and polyadenylation factors (PAFs). Consequently, it will induce alternative splicing or alternative polyadenylation, thus producing more short mRNA in cells to promote a productive TC-NER and efficient transcription recovery (as was shown for the short isoform of ASSC3 [292]).

are expressed because of RNA Pol II elongating wave-like in active genes following UV irradiation, resulting in more short mRNA products in cells after DNA damage.

Moreover, in response to UV-induced TBLs, the inhibition of RNA Pol II forward translocation also affects the co-transcriptional mRNA maturation processes, such as pre-RNA splicing. Studies established that the E3 ligase MDM2 is alternatively spliced into an inactive protein isoform, stabilizing p53 following UV irradiation [357], [358], [359]. More recently, the Svejstrup group carried out RNA-seq and revealed that UV light induced alternative splicing response, in which genes are expressed into a shorter mRNA incorporating alternative last exons (ALEs) that are more proximal to the transcription start site. Remarkably, some of these shorter transcripts appear to play additional roles in the cellular stress response. For example, upon the induction of TBLs, a short isoform (25 kb, 4 exons) of ASCC3 is produced rather than a long isoform (>370 kb, 42 exons), which is the largest subunit of activating signal co-integrator complex (ASCC) and was characterized as a DNA helicase that unwinds DNA by translocating on one strand in 3'-to-5' direction during DNA alkylation repair [360]. This ASCC3 ALE short isoform does not encode a protein but rather functions as a non-coding RNA required for efficient transcription recovery after UV-induced DNA damage, evidenced by the fact that deficiency of the short ASCC3 isoform inhibited transcription recovery after UV irradiation (**Figure 15**) [292].

However, more details on the exact mechanism of more short mRNA production following UV irradiation need to be characterized.

1.3.5 R-loop formation interplays with TBLs

R-loops are three-stranded structures consisting of a displaced single-stranded DNA (ssDNA) and DNA:RNA hybrids. During normal transcription, the nascent RNA binds strongly to the template DNA strand, resulting in the formation of a strange DNA:RNA hybrid structure that displaces the non-template ssDNA [361]. R-loops form behind the elongating RNA Pol II co-transcriptionally. Genome-wide mapping studies indicate R-loops are abundant at gene promoters [362], [363]. R-loops aid some gene transcription at their promoters by preventing the binding of transcriptional repressors or DNA-methylating enzymes (DNMTs) or acting as transcription factor binding sites [364]. R-loops as Janus-faced modulators could also block transcription by enforcing RNA Pol II stalling. It has been shown that R-loops inhibit transcription *in vitro* [365]. R-loops could also inhibit transcription by blocking the transcription factors' binding at the gene promoter [366]. During transcription termination, R-loops promote RNA Pol II pausing and the cleavage of the pre-mRNA from its template, either by recruiting exonucleases for releasing the pre-mRNA (e.g., XRN2) or R-loop resolution factors (e.g., DHX9 and SETX (Senataxin)) or by recruiting the RNAi-silencing machinery. It has been very well established that the loss of DHX9, SETX, and XRN2 will lead to R-loop accumulation and induce a defective termination [367], [368], [369].

On the other hand, DSB- and UV-induced persistent stalling of RNA Pol II appears to accumulate the formation of R-loops from nascent RNA [370]. Recent studies showed that UV-induced

TBLs induce the exclusion of late-stage spliceosomes from DNA damage sites which are composed of U2, U5, and U6 small nuclear ribonucleoproteins [371]. This displacement of co-transcriptional spliceosomes from arrested RNA Pol II increases R-loop formation. This R-loop formation in TBLs causes non-canonical activation of the ataxia-telangiectasia mutated (ATM) kinase, which signals the further mobilization of spliceosomes from RNA Pol II located distal to DNA lesions [371]. Additionally, in response to UV-induced TBLs, ATM signaling culminates in wide-spread alternative splicing and gene expression changes [309].

Persisting R-loops are genotoxic by interfering with DNA replication, transcription, and DNA repair through increasing the likelihood of replication fork collapse after colliding with stalled transcription complexes and promoting unscheduled replication via transcription-associated recombination. For example, TBL-induced arrest or backtrack of RNA Pol II will hinder replication fork progression and result in transcription-replication collisions when a replication fork comes in from the opposite direction in a head-on orientation relative to transcription, which is most commonly seen at transcription start sites [372]–[378]. Such head-on collisions result in widespread checkpoint activation, as evidenced by gammaH2AX spreading and ATR activation [379].

Based on *in vitro* and *in vivo* studies, cells are equipped with specialized RNA hydrolases (RNaseH1 and H2) or RNA/DNA helicases (e.g., Aquarius, DHX9, SETX, Bloom Syndrome RecQ Like Helicase (BLM)) that unwind the DNA:RNA hybrids to counteract R-loop toxicity. In addition, other factors are also reported to prevent or resolve R-loop accumulation, which are involved in TC-NER (e.g., XPF and XPG), RNA processing (e.g., THO-TREX complex), transcription, and chromatin remodeling [380], [381].

1.3.6 The consequence of TBLs

Unresolved TBLs eventually induce a global transcription arrest and cell cycle arrest, allowing cells to process stalled RNA Pol II and the TC-NER pathway. Failure of TBL removal will result in: (1) an accumulation of unrepaired DNA damage; (2) a failure to resume transcription; (3) a dysregulation of cell cycle progression; and (4) increased cell death. Persistently stalled RNA Pol II will also increase the formation of R-loops, which will promote DNA breakage and subsequent transcriptional inhibition, impairing cellular functions further. All of these biological consequences manifest clinically as photosensitivity, premature aging, neurodegeneration, cancer, and immunodeficiency. Mutations in NER proteins lead to hereditary disorders of high skin photosensitivity. For example, the most well-studied TCR deficiency diseases are Cockayne Syndrome (CS) and UV-sensitive syndrome (UVSS). A mutation in any of the XP factors (XPA-G) causes Xeroderma Pigmentosum (XP) [382]–[384]. The deficiency or mutation in CSA/CSB will cause Cockayne Syndrome [385], [386]. Patients with CS or XP frequently exhibit accelerated aging, progressive neurological degeneration, and severe photosensitivity [387], [388].

1.4 Translation in response to DNA damage

Almost all protein-coding mRNAs (except histones) have 3′ ends polyadenylated. The poly(A) tail protects mRNAs against degradation and its length influences translation initiation. RBPs binding to the 3′-UTR upon export to the cytoplasm may recruit polyadenylation or deadenylation proteins that may enhance or repress mRNA transcript translation.

Global translation control is achieved by regulating the availability and/or phosphorylation status of canonical eukaryotic initiation factors (eIFs) and their binding partners, whereas specific subsets of mRNAs are regulated by sequence/structured elements within the mRNAs 5' and 3'-UTRs and the proteins that interact with these elements [389]. The mechanism used by most mRNAs to initiate translation is known as cap-dependent scanning, which involves binding of the eIF4F complex (which includes eIF4E, eIF4G, and eIF4A) to the mRNA's 7-methyl G cap. The 40S ribosomal subunit is recruited through an eIF4G-eIF3-40S interaction with the ternary complex, which includes eIF2, GTP, and the initiator $tRNA_{imet,}$ and eIF2 is required for initiator $tRNA^{Met}$ delivery to the translation machinery [389], [390], [391].

The response to DNA damage is heavily reliant on signaling to gene expression. UV and ionizing radiation both increase the transcription of some genes, including those involved in NER, such as XPC, DDB2, PCNA, and GADD45A [392], [393]. It has also been reported that UV irradiation significantly inhibits cellular protein synthesis. This paradox could result from the differential regulation of mRNA translation, which has been observed in other cell stress conditions, such as heat shock [394]. Studies for mRNA translation efficiency profiling during cell stress revealed that certain mRNAs avoid the global inhibition of protein synthesis [395]. Furthermore, many of these mRNAs are translated via alternative translation initiation mechanisms, such as internal ribosome entry (IRE). These mRNAs, in general, encode proteins that are required for the stress response. For example, selective mRNA translation increases the synthesis of chromatin remodeling proteins during apoptosis, like Histone acetyltransferase 1 (HAT1) [396], [397]. However, during the response to hypoxic stress, mediators of the unfolded protein response are preferentially translated, like activating transcription factor 4 (ATF4) [398], [399]. These changes are mainly regulated by translation initiation factors, such as eIF2. UV light has been shown to induce eIF2 phosphorylation that is DNA-PKcs or GCN2 (eIF-2-alpha kinase GCN2) dependent and subsequently regulate translation [400], [401], [402].

Furthermore, ribosomes play a major role in translation besides translation initiation factors. In eukaryotic cells, ribosome biogenesis is a complex process in which the 35S pre-rRNA is processed to mature 18S, 5.8S, and 25S rRNAs, which are then assembled with the large number of ribosomal proteins that comprise the small (40S) or large (60S) ribosomal subunits [403], [404]. The majority of these steps occur in the nucleolus, but the final ones appear in the nucleoplasm or after nascent subunits are exported to the cytoplasm. The complexity of the processing and export pathway and the large number of factors involved raise the possibility of errors and the generation of defective preribosomal

subunits [405]–[409]. Ribosome biogenesis is a tightly controlled process that is closely linked to cell growth and division [403], [409], [410]. Some studies have revealed that cellular stress (such as oxidation and alkylation), induces the ubiquitylation of ribosome subunits (the 40S and 60S subunit) depending on the activity of E3 ligase ZNF598 (Zinc Finger Protein 598) [411], [412], [413]–[416]. ZNF598 is in a complex with the translation repressor proteins EIF4E2/4EHP and GIGYF2 [417]. ZNF598 is also required for ribosome stalling at polyA sequences, and its E3 ubiquitin ligase activity is linked to the inhibition of ribosome function and translation arrest by ubiquitylating the 40S subunit ribosomal proteins RPS10 and RPS20 [418], [419]. When ribosomes stall, the ribosome-associated quality control (RQC) pathway is activated, which targets collided ribosomes and causes subunit dissociation, followed by the proteasomal degradation of the nascent peptide [414], [420].

1.5 Post-translational modifications (PTM) in NER

1.5.1 Post-translational modifications

Studies have discovered that the human proteome is far more complex than the human genome in the last decades. While the human genome is thought to contain between 20,000 and 25,000 genes, the total number of proteins in the human proteome is estimated to be over one million [421]. These estimates show that a single gene can encode multiple proteins. Mechanisms that generate different mRNA transcripts from a single gene include genomic recombination, transcription initiation at alternative promoters, differential transcription termination, and alternative splicing or polyadenylation of pre-mRNA, as well as translation regulation, which were explained above. During or shortly after polymerization of the translation process, a linear chain of amino acids folds into a complex shape giving the protein a distinct three-dimensional structure and further forming complexes with other proteins to apply specific functions [422]–[427] (**Figure 16**).

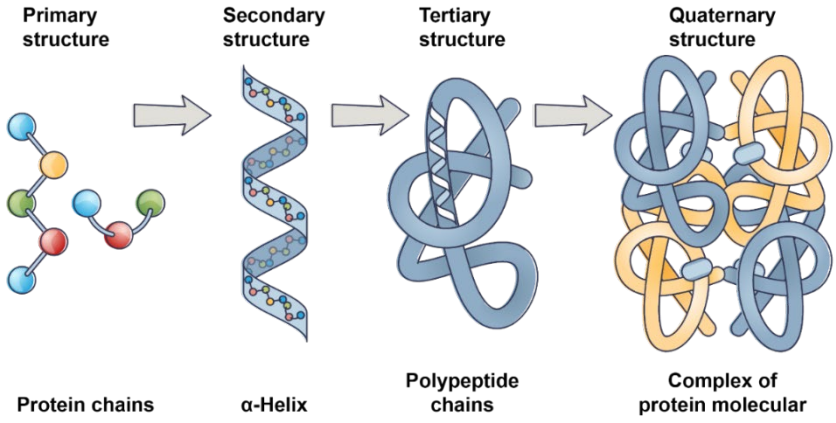

Figure 16: The simplified scheme of protein structures adapted from [428].

Moreover, proteins have a variety of functions in cells, with the contribution of post-translational modifications (PTM) to increase in complexity from the genome to the proteome. PTMs are chemical modifications that regulate proteins' activity, localization, and interaction with other cellular molecules such as proteins, DNA, RNA, and lipids. PTMs occur at specific amino acid side chains or peptide linkages mediated by enzymatic activity. It is estimated that 5% of the proteome is made up of enzymes that perform over 200 different types of PTMs, including kinases, phosphatases, transferases, and ligases, which add or remove functional groups, proteins, lipids, or sugars from the amino acid side chains. In addition, proteases also modify proteins by cleaving peptide bonds to remove specific sequences or regulatory subunits.

Furthermore, the human proteome is dynamic and changes in response to a variety of stimuli. Moreover, post-translational modifications are frequently used to regulate cellular activity. PTMs have been reported to be important in DNA damage repair (DDR) through altering protein activity without requiring *de novo* protein synthesis by attaching small molecules to substrate proteins. Protein PTMs play a key role in the first phase of DDR because they mediate protein-protein interactions and regulate

protein trafficking, localization, activity, and stability, assuring accurate and timely removal of DNA lesions. The most prominent PTMs are phosphorylation, ubiquitination, sumoylation, acetylation, methylation, and poly(ADP-ribosyl)ation (**Figure 17**).

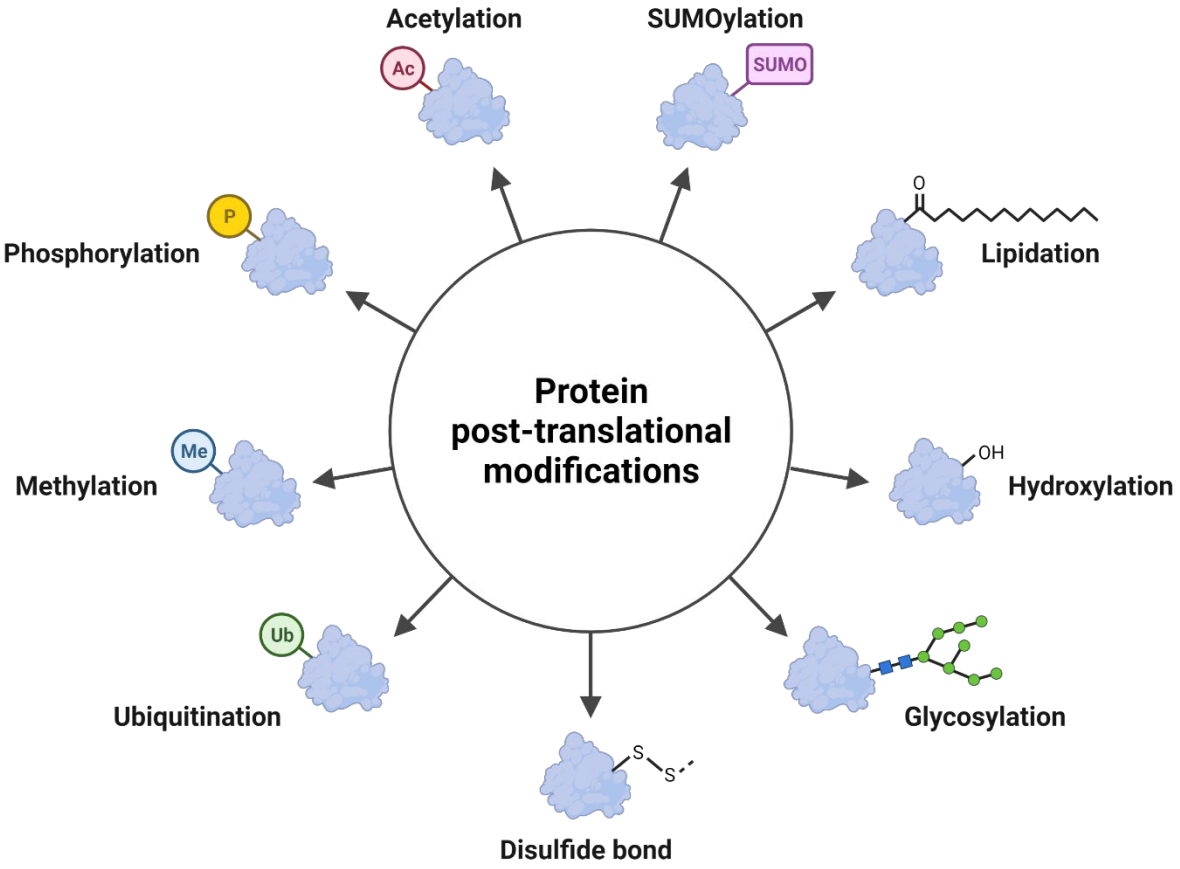

Figure 17: Several examples of protein post-translational modifications. This figure was created with BioRender.com.

1.5.2 Protein ubiquitylation

Ubiquitylation is a complex and reversible covalent modification catalyzed sequentially by a series of enzymes. Ubiquitin is a highly conserved 8-kDa protein made up of 76 amino acids that is covalently attached to the εNH2 of lysine in target proteins via ubiquitin's C-terminal glycine [429], [430], [431]. Ubiquitylation occurs in three steps in mammalian cells. First, ubiquitin is activated by one of the two activation enzymes (E1), with the aid of ATP. The activated ubiquitin is then transferred to one of the ~40 ubiquitin-conjugation enzymes (E2). Finally, ubiquitin is attached to the target protein via one of the ~600 ubiquitin ligases (E3) via an isopeptide bond between the C-terminus of ubiquitin and the [epsilon]-amino group of the lysine residue in the target protein [431], [432] (**Figure 18**A). Ubiquitin contains seven lysines which can be linked together to form a polyubiquitin chain. Polyubiquitylated proteins are then recognized by the 26S proteasome, which catalyzes ubiquitin degradation and ubiquitin recycling (**Figure 18**B). This process may be repeated until multiple lysine residues of the target protein are ubiquitinated, or ubiquitin chains are formed, connected through

specific isopeptide bonds (K6, K11, K27, K29, K33, K48, K63, N-terminal methionine) [432]–[436]. These various ubiquitin modifications take on distinct conformations and result in specific outcomes for proteins in cells. For instance, Lys48 is linked to another ubiquitin on target proteins to form the most common polyubiquitin chain. This chain facilitates protein recognition and degradation by the proteasome complexes or lysosomes [437]–[439]. However, target proteins are not degraded by the proteasome after Lys63-linked polyubiquitylation. Lys63-linked polyubiquitylation has been reported to be involved in DNA repair, protein kinase activation, and other biological processes [440]–[443].

Furthermore, ubiquitination is a reversible process that can be reversed by over 100 deubiquitinating enzymes (DUBs) [444]. DUBs are classified into two groups based on their catalytic sites: Cysteine peptidases and metalloproteases. DUBs are further classified into seven protein families based on sequence and structure similarity: Cysteine peptidases, ubiquitin carboxyl-terminal hydrolases (UCH), ubiquitin-specific proteases (USP), ovarian tumor proteases (OTU), Machado-Josephin domain-containing proteases (MJD), motif-interacting with ubiquitin-containing novel DUB family (MINDY), and zinc finger with U (ZUFSP). Only the JAB1, MPN, and MOV34 families (JAMM) are metalloproteases [445]–[448]. DUBs, like ubiquitin ligases and conjugases, typically bind specific ubiquitin linkage types as substrates to cleave single ubiquitin monomers from the distal end of a chain or by breaking the bond between the proximal ubiquitin and the substrate [449]–[451]. DUBs regulate ubiquitin signaling by removing ubiquitin and thus disassembling chains and thereby signals while recycling ubiquitin for further conjugation [452].

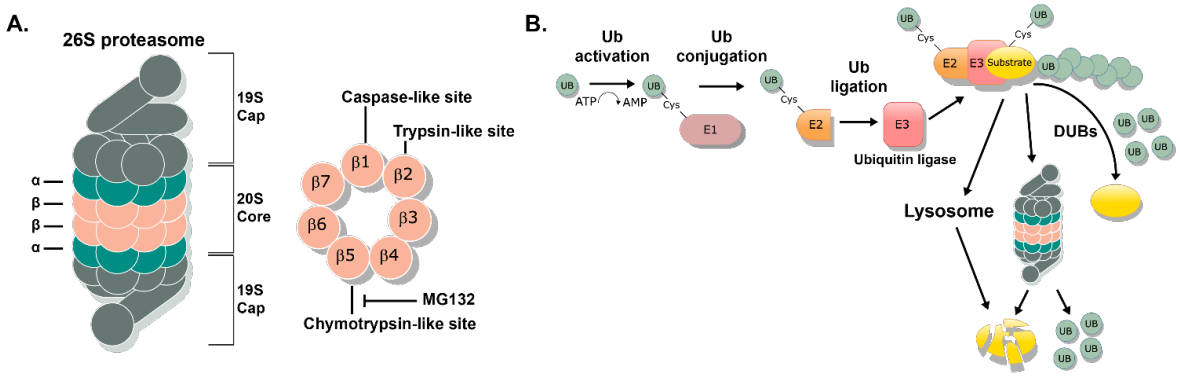

Figure 18: The simplified scheme of ubiquitylation and the proteasome. A. The diagram of the 26S proteasome. Left panel: Schematic diagram of the 26S proteasome. Positions of subunits are indicated accordingly. Right panel: Subunit structure of core particle (20S proteasome) is depicted schematically. Caspase-like, trypsin-like, and chymotrypsin-like activities are associated with the 1, 2, and 5 subunits, respectively. **B. The ubiquitylation steps and recycling of ubiquitin.** Ubiquitin is activated by a ubiquitin-activating enzyme (E1) in an ATP-dependent manner. It is transferred to a ubiquitin-conjugating enzyme (E2) and can be specifically attached to the substrate with the assistance of an E3 ligase. Ubiquitin chains are formed through the repeated attachment of ubiquitin. Deubiquitination enzymes are responsible for ubiquitin removal (DUBs). Proteasomes and lysosomes will degrade ubiquitylated proteins [453], [454].

Ubiquitylation in NER

RNA polymerase II (RNA Pol II) stalling acts as a DNA damage sensor for the TC-NER pathway. The stalled RNA Pol II at TBLs undergoes proteolytic clearance via a ubiquitin-proteasome

system with the participation of multiple E3 ubiquitin ligases for RNA Pol II ubiquitination and the involvement of CSB and CSA. CSB contains a ubiquitin-binding domain (UBD) in the C-terminal region as well as in the central region. A highly conserved CSA-interacting motif (CIM) was recently discovered in the C-terminus of CSB, upstream of the UBD domain. RNA Pol II-CSB will recruit CSA via the CIM motif [455]. Then the $CRL4^{CSA}$ (CSA-DDB1/ Cul4A/Roc1) E3 ubiquitin ligase complex is recruited by interacting with the C-terminal CSA-interacting motif (CIM) of CSB. The CSB-$CRL4^{CSA}$ complex is subsequently activated by dissociating from the COP9 signalosome, which plays a central role in the ubiquitination of RNA Pol II and CSB in human cells [456]–[459]. Recently, studies have shown a single DNA damage-induced ubiquitination site (K1268) in the largest subunit of RNA Pol II, RPB1, plays a key role in TC-NER. The loss of ubiquitination at RPB1-K1268 results in a failure of RNA Pol II removal from TBLs on transcribed strands and transcription recovery in response to UV irradiation is impaired [460]–[462]. Recent research has provided more insight into this process by identifying the transcription elongation factor ELOF1 as a new TC-NER factor in the regulation of RNA Pol II ubiquitination after UV irradiation. In normal cells, ELOF1 binds to elongating RNA Pol II which is close to the K1268 site and thus stimulates transcription elongation. When cells are treated with UV light, ELOF1 interacts with $CRL4^{CSA}$ and directs the E3 ligase catalytic domain to the proximity of the K1268 site, subsequently promoting RNA Pol II ubiquitination without affecting the association of the CSB-$CRL4^{CSA}$ complex with stalled RNA Pol II. Following this step, the Valosin-containing protein (VCP)/p97 complex co-extracts ubiquitinated CSB and RNA Pol II from DNA lesion sites, allowing for the sequential incorporation of arriving core NER factors into the pre-incision complex. VCP/p97 eventually presents ubiquitinated RNA Pol II to the proteasome for degradation, whereas deubiquitination rescues ubiquitinated CSB [463], [460]. In addition, UVSSA (UV Stimulated Scaffold Protein A) is recruited to the stalled and ubiquitinated RNA Pol II complex via the interaction with the N-terminal CSA-interacting region (CIR) of CSA, USP7 (ubiquitin-specific proteases 7), and ubiquitin chains [461], [464], [465]. USP7 is involved in the deubiquitination of RNA Pol II and CSB to prevent their degradation [466].

1.5.3 Protein phosphorylation

Phosphorylation is a common and widely used posttranslational modification that is mediated by kinases attaching phosphate groups to a substrate protein's serine, threonine, histidine, or tyrosine residues [467], [468], [469], [470]. Signal transduction from upstream sensors to downstream effectors occurs in a kinase signaling cascade which is composed of the phosphorylation of a series of proteins, such as a first kinase phosphorylating and activating a second kinase, which then phosphorylates and activates the third kinase to phosphorylate the target proteins [471]–[473]. Phosphorylation is a reversible process as well. Dephosphorylation is the separation of a phosphate group from a substrate protein [474]–[476].

Many kinases play a role in promoting and regulating the DNA damage response. ATM, ATR, and DNA-PKcs are the most prominent examples, with the former two considered the primary

transducers or master regulators of the DDR response [248], [477], [478]. All three kinases are members of the phosphatidylinositol 3-kinase-related kinase (PIKKs) family, have a similar structure, and phosphorylate serine or threonine residues frequently followed by glutamine (S/T-Q motif) [479], [480]. When activated, these kinases phosphorylate hundreds of substrates, controlling a wide range of cellular processes. Some substrates (e.g., H2AX) can be phosphorylated by all three kinases, whereas others are specifically targeted by one but not the others [481], [482]. Despite their similarities, these kinases serve different functions in the DDR [248], [483]. The function of ATM and ATR in DNA repair has been described above.

Cyclin-dependent kinases (CDKs)

In addition, cyclin-dependent kinases (CDKs) also play a key role in the cellular response to stress or under normal conditions, such as the cell cycle. CDKs respond to extracellular and intracellular signals to regulate cell division by forming a heterodimer complex with cyclins, which function as regulatory subunits. There are 20 CDKs and 29 cyclins in human cells. CDK1, CDK2, CDK3, CDK4, CDK6, and CDK7 are directly involved in cell-cycle transitions and division, whereas CDK7–11 are involved in gene transcription. Throughout the cell cycle, the expression of CDKs varies cyclically.

The eukaryotic cell cycle is a sequence of events that occur in a cell as it grows and divides, including DNA replication, chromosome segregation, and cell division. Typically, cells spend most of their time in interphase, in which it grows, replicates their chromosomes, and prepares for cell division. Then cells exit interphase, go through mitosis and completely divide into two daughter cells. The resulting daughter cells enter their own interphase and begin a new round of the cell cycle. The cell cycle usually is divided into four phases: the gap phase 1 (G1), controlling cell growth; the S-phase, hosting DNA replication; the gap phase 2 (G2), organizing the genetic material and preparing for the cell division; and the mitosis (M) phase, where the cell divides the two copies of the genetic material into two daughter cells. Cells can also exit the cell cycle and enter a dormant state known as the G0 phase (**Figure 19**). Among the five stages in a 24-hour cell cycle in an animal cell, interphase lasts the longest, while the other stage, M-phase lasts only 30 minutes to 1 hour (**Figure 19**). However, embryonic cells complete the cell cycle in 15–30 minutes or less; in this case, the cell enters S-phase immediately after M-phase. It also implies that all of the inputs required for the next stage are already present because they are constantly synthesized in embryonic cells.

The activities of complexes composed of cyclins (A, B, D, E) bound to cyclin-dependent protein kinases control cell cycle progression (CDKs). The D-type cyclins activate the CDK4 and CDK6, which are required to enter and progress cells into the G1-phase. Cyclin E associates with CDK2 to progress from the G1 to the S phase. Cyclin A, which is linked to CDK2, allows progression through the S phase. Cyclin A, which is associated with CDK1, initiates the M phase during the G2 phase. Cyclin B then activates CDK1 and promotes the M phase of the cell cycle (**Figure 19**). The synthesis and degradation of cyclins during cell cycle progression, the phosphorylation status of CDKs, or the binding of CDK

inhibitory proteins to cyclin-CDK complexes all influence the formation and activity of cyclin-CDK complexes.

When cells contain DNA damage that needs repair, they activate a DNA damage checkpoint, which stops the cell cycle until DNA damage is removed. Checkpoint arrests occur at various stages of the cell cycle, including the G1/S transition (the G1 checkpoint), S phase progression (the intra-S phase checkpoint), the G2/M boundary (the G2/M checkpoint) (**Figure 19**). Different types of damage activate ATR-CHK1 and ATM-CHK2 pathways at checkpoints. The degradation or inhibition of CDC25 phosphorylation and p53-dependent induction of the p21 CDK inhibitor both inhibit Cdk activity. CDK-dependent events inhibit the origin firing of replication fork by inhibiting CDK and CDC7 kinase activity. Checkpoint activation causes either cell death or increased cell survival, and disruption of these critical signaling pathways may result in the disruption of essential cellular functions. More checkpoints exist, such as the Spindle checkpoint and the Morphogenesis checkpoint. The spindle checkpoint causes the cell cycle to stop at M-phase until all chromosomes are aligned on the spindle. This checkpoint is critical for chromosome distribution equality. Morphogenesis checkpoint detects cytoskeleton abnormalities and arrests cell cycle at G2/M transition.

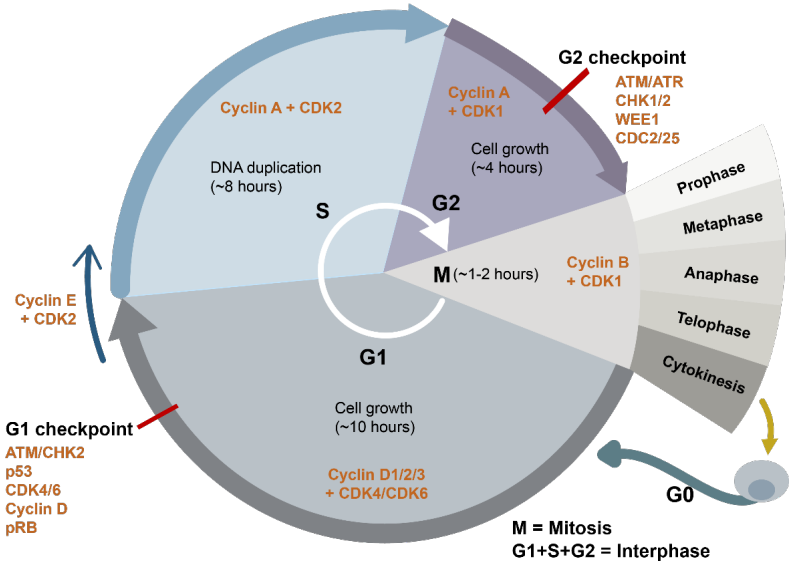

Figure 19: The regulation of cell cycle by kinases. The eukaryotic cell cycle consists of the G1 and G2 phases, the S-phase, and the M (mitosis) phase. Cells can also enter a dormant state known as the G0 phase. M phase includes Prophase, Metaphase, Anaphase, Telophase, and Cytokinesis. Colored arrows indicate different stages of the cell cycle. The cell cycle is controlled by complexes of cyclins binding to cyclin-dependent protein kinases (CDKs). Cyclin-CDK complexes are shown aligned with the arrow denoting the corresponding cell cycle phase. Cyclin-CDK complexes are regulated by checkpoint pathways, which prevent cells from progressing to the next stage when it is not permitted [484].

Mitogen-activated protein kinase (MAPK) pathway

Besides ATM, ATR, DNA-PK and CDKs, another kinase family has been characterized to be important in cellular processes, the mitogen-activated protein kinase (MAPK) family. Environmental stresses, such as ultraviolet irradiation, heat shock, genotoxic agents, growth factors, and inflammatory cytokine stimulation, can activate MAPK pathways [485], [486]. These MAPK signaling cascades transduce signals by sequentially activating three to five layers of protein kinases known as MAPK

kinase kinase kinase (MAP4K), MAPK kinase kinase (MAP3K), MAPK kinase (MAP2K), MAPK and MAPK-activated protein kinases (MAPKAPK). MAPKs are evolutionarily conserved enzymes that require dual phosphorylation on the threonine-X-tyrosine catalyzed by MAP2K kinases. MAPKs phosphorylate specific serine (S) and threonine (T) residues of target substrates, including other protein kinases and many transcription factors [487]. The first three central layers are considered a fundamental core unit, whereas the last two layers appear in some cascades and vary depending on cells and stimuli. Based on the components of the MAPK layer, four MAPK cascades have been defined: ERK1/2, c-Jun N-terminal kinase (JNK), p38 MAPK, and ERK5 [488], [489], [490], [491]. The JNK and p38 MAPK pathways are primarily associated with cell stress and apoptosis, whereas the ERK/MAPK signaling pathway is associated with cell proliferation and differentiation and plays a critical role in the cell signal transduction network [492], [493], [494], [495].

Phosphatases tightly regulate the MAPK pathways. MAPK phosphatases (MKPs) or dual-specificity phosphatases (DUSPs) control the inactivation of MAPKs [496]. The majority of MKPs have phosphatase activity toward p38 MAPKs and ERK. MKP1/DUSP1, MKP5/DUSP10, MKP7/DUSP16, and DUSP8 have been reported to dephosphorylate p38α and p38β MAPKs [496]. In addition, the protein phosphatase (PP) 2A and PP2C families can also regulate the phosphorylation of p38 MAPK [496]. In response to UV-light, WIP1/PPM1D, a member of the PP2C family, deactivates p38 and thus inhibits the p53 pathway [497], [498]. The UV-light response activates MKP1 through p38, which dephosphorylates JNK and thus suppresses apoptosis [499].

p38 MAPK pathway

The p38 MAPK is composed of four isoforms: p38α (MAPK14), p38β (MAPK11), p38γ (MAPK12), and p38δ (MAPK13), with ~60% sequence similarity among the four isoforms [500], [501]. p38α is ubiquitously expressed at a high level in most tissues, whereas p38β appears to be expressed at a very low level. p38γ and p38δ tend to have restricted expression patterns and may have specialized functions. The activity of p38 kinases is tightly regulated and involves activation by a dedicated kinase. The p38 MAPK can be activated by MKK3 and MKK6 kinases and also be phosphorylated by MKK4, which is well known as a JNK activator, via dual phosphorylation on threonine-180 and tyrosine-182 [502]. In addition, autophosphorylation may also contribute to p38 MAPK activation [503]. Upon activation, p38 proteins can translocate from the cytosol to the nucleus, where they orchestrate cellular responses by mediating the phosphorylation of their downstream transcription factors (e.g., p53 and ATF2), RNA binding protein (e.g., NELFE, RBM7), and other kinases (e.g., MNK1/MNK2, and MSK1/MSK2) which in turn phosphorylate other important proteins (e.g., HSP27, and eIF-4E) [504], [505], [506]. The p38 pathway is important in the regulation of cell survival in response to stress by regulating many protein phosphorylation changes, allowing rapid control of processes such as cell cycle progression, DNA damage repair, or mRNA processing. P38 regulates the phosphorylation of kinases, transcription factors, and mRNA stability regulators in the immune response, which collectively regulates the expression of cytokines and other factors involved in inflammatory processes. P38 MAPK

causes cell cycle arrest by upregulating cyclin-dependent kinase (CDK) inhibitors, such as p53 or GADD45, or by downregulating cyclin D or CDC25 via a variety of mechanisms [507], [508], [509], [510]. P38 has a number of targets including MAPK-activated protein kinase 2 (MK2). P38 MAPK activates MK2 via phosphorylation at Thr-222, Ser-272, and Thr-334 [511]. MK2 has several phosphorylation targets, but its primary function is as a master regulator of RNA-binding proteins, indirectly controlling gene expression at the translational level [512].

In response to UV irradiation, the p38-MK2 cascade regulates the phosphorylation of NELFE or RBM7, allowing RNA Pol II to release and increase the transcription of specific genes, including genes required for telomere maintenance or DNA repair [296]. Additionally, following UV irradiation, p38-MK2 will arrest the cell cycle by inducing the G2/M checkpoint by deactivating CDC25B and CDC25C phosphorylation. Furthermore, p38 MAPK can directly phosphorylate DNA repair regulators like CtIP, which helps to coordinate the DNA damage response while also limiting replication stress and chromosome instability [513], [514]. In addition, it has been shown that p38 MAPK signaling is involved in the phosphorylation of XRCC1 at T358 and T367, thus regulating the recruitment of XRCC1 to oxidative stress sites [515]. It has been proposed that p38 can act as a tumor suppressor via activating the p38 MAPK, increasing cell apoptosis in response to chemotherapeutic agents [516], [517].

JNK MAPK pathway

JNK is another subtype of the MAPK signaling pathway and is known as stress-activated kinase (SAK) in mammalian cells. JNK proteins are encoded by three genes: MAPK8 (encodes JNK1), MAPK9 (encodes JNK2), and MAPK10 (encodes JNK3), sharing 85% identity, which is alternatively spliced to produce at least ten isoforms. JNK1 and JNK2 are found in almost every cell, whereas JNK3 is found primarily in the testis, heart, and brain [518]. JNK is activated by a cascade reaction. Stress signals are delivered by small GTPases (Rac, Rho, and cdc42) to a series of kinase cascades, and eventually, JNK is activated by the upstream kinases MKK4 and MKK7 (MAPK4 and MAPK7) [519]. The activated JNKs are translocated from cytoplasm to nucleus, where it regulates the activity of multiple transcription factors (e.g., c-Jun, ATF2, ELK1, p53, and NFAT4). In addition, activated JNK also phosphorylates many cytoplasmic substrates such as cytoskeletal proteins and mitochondrial proteins like Bcl-2 and Bcl-xl [520]. Consequently, JNK can regulate gene expression, subsequently regulating cell proliferation, differentiation, survival, and apoptosis, as well as autophagy and cell motility (

Figure 20). The cellular response can range from induction of apoptosis to increased survival and altered proliferation depending on the stimuli and the strength and duration of JNK activation. Under normal conditions, JNK2 primarily targets JUN for degradation, whereas JNK1 phosphorylates and stabilizes JUN following stimulation, leading to transcriptional activation [521]–[523].

The role of JNK has been very well unraveled in the regulation of cell apoptosis, and their inhibition has typically been associated with a resistant phenotype to various genotoxic stimuli, such as

chemotherapeutic drugs [524]. For instance, the transcription factor AP1, which is made up of Fos and Jun family members, is a major JNK target. JNK promotes c-Jun and ATF-2 phosphorylation, which results in the activation of AP-1 (Activator Protein 1) and the expression of Fas/FasL signaling pathway-related proteins. The binding of FasL onto Fas can mediate the activation of caspase 8, and thus prototypically activate downstream factors of caspase 3 to cause apoptosis [525], [526]. On the other hand, JNK mediates the phosphorylation of the anti-apoptotic proteins Bcl-2/Bcl-xL at multisite to change the mitochondrial membrane potential, resulting in the release of cytochrome C, the activation of caspase 9 and further caspase 3 to induce apoptosis [527], [528]. JNK can sensitize cancer cells to genotoxic stress-induced cell death. For example, quinuclidine derivative 6 induces the apoptosis of human breast cancer by reducing the expression level of Bcl-2, Bcl-xL, and increasing mitochondrial apoptotic pathways by activating the release of cytochrome C [529]. JNKs' oncogenic functions are primarily based on their ability to phosphorylate JUN and activate AP1, whereas their tumor-suppressive functions are most likely related to their pro-apoptotic activity.

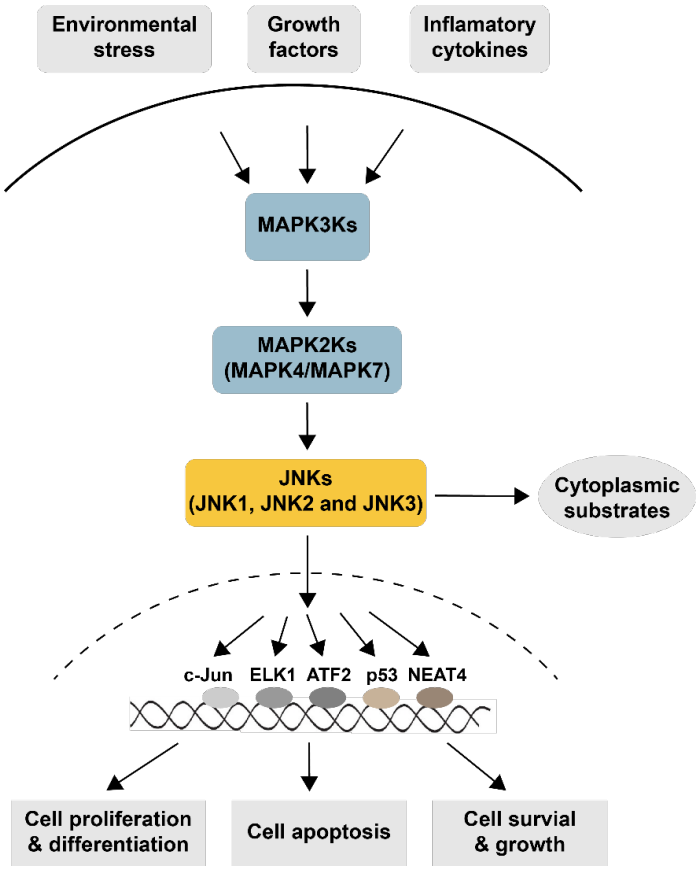

Figure 20: The scheme of JNK MAPK signaling cascade. Environmental stresses (such as UV), inflammatory cytokines, and growth factors can all activate the JNK MAPK pathway. The Rho family of small molecule GTPases transmits stress response signals to this cascade. Activated JNK regulates the activity of various transcription factors after being transported into the nucleus, thereby mediating the transcriptional activation of c-JUN, ATF2, ELK1, p53, and other transcription factors that play important roles in physiological and pathological processes such as cell cycle, proliferation, apoptosis, and cellular stress.

In addition, factors such as the signal intensity and crosstalk between different JNK isoforms are important for pro- and anti-oncogenic functions in different cell types and stages of tumor development, such as human hepatocellular carcinoma (HCC) and head and neck squamous cell

carcinoma (HNSCC). Several studies have shown that JUN is required for the development of HCC by inhibiting the pro-apoptotic function of p53 [530]. Similarly, JNK1 deficiency (but not JNK2 deficiency) has been shown to reduce susceptibility to diethylnitrosamine (DEN)-induced HCC formation [531]. Impairment in liver cell proliferation and tumor formation is caused by JNK1 downregulation caused by decreased expression of MYC and increased expression of the CDK inhibitor p21 [532]. Pharmacological JNKs inhibition with an inhibitory peptide also reduced HCC development in mice by a DEN–phenobarbital protocol and in xenografted HCC cells, implying that JNK1 targeting should be considered as a new therapeutic approach for HCC treatment [533]. The DMBA–TPA analysis has revealed that JNK1 and JNK2 to tumor development of mouse skin carcinogenesis, which is a widely used and well-established two-stage skin carcinogenesis protocol that depends on tumor initiation with DMBA (7,12-dimethylbenz(a)anthracene) and promotion with TPA (12-O-tetradecanoylphorbol-13-acetate). Unlike in liver cells, *JNK1*-knockout mice appeared to be more susceptible to skin tumors, whereas papilloma formation was found to be significantly reduced in *JNK2*-knockout mice [534], [535]. These findings suggest that JNK2 has an oncogenic function, whereas JNK1 appears to be a suppressor of skin tumor development. At the molecular level, it was proposed that the specific functions of JNK1 and JNK2 in skin tumors could be explained by differential regulation of ERK and AKT signaling, as well as altered AP1 DNA binding activities. Future research should look into whether these differences are the molecular basis for the cell context-dependent JNK phenotypes observed in skin versus liver.

JNK activation also contributes to autophagic induction. High-mobility group box 1 protein (HMGB1), makes myeloid leukemia cells resistant to conventional anticancer treatments by increasing the transcriptional activity of JNK and inducing the JNK-dependent autophagy. In addition, JNK signaling is required for the upregulation of LC3 in human nasopharyngeal carcinoma cells during ceramide-induced autophagy.

Recent research has elegantly demonstrated JNK plays a role in response to stress. Oxidative stress can trigger a variety of responses, including autophagy and cell death via the JNK pathway. In response to ROS, the JNK-AP-1 signaling pathway is activated and plays an essential role in the regulation of JNK-dependent autophagy [536]–[538]. Most recently, JNK was shown to promote the repair of UV-induced photolesions by JNK activating the phosphorylation of the microRNA biogenesis protein DGCR8 upon RNA Pol II stalling [539]. However, the clear mechanism is unknown. Therefore, it is meaningful to uncover the role of the JNK MAPK pathway in response to TBLs.

1.5.4 Acetylation in NER

Acetylation is reported to be involved in regulating transcriptional activity and protein stability [540], [541]. XPA plays an essential role in correctly positioning the DNA repair machinery at DNA damage sites by interacting with other NER factors. Recent studies showed that UV irradiation induced the acetylation of XPA at lysines 63 and 67. The acetylated XPA can be deacetylated by Sirtuin 1

(SIRT1) [542]. In addition, in response to UV irradiation, the acetylation of histone H3 lysine 9 (H3K9) increases the accessibility of the NER machinery to DNA lesions [543], [544].

1.5.5 Methylation in NER

Methylation primarily modifies proteins on arginine and lysine residues, which is catalyzed by S-adenosylmethionine- (AdoMet-) dependent enzymes that donate a methyl group to these residues' side-chain nitrogen atoms. Arginine methylation affects a variety of cellular functions, including RNA processing, signal transduction, transcriptional regulation, and DNA repair [545]–[547]. In the last step of NER, Flap endonuclease-1 (PEN1) plays a vital role. The methylation of FEN1 promotes the interaction between FEN1 and PCNA, thus increasing the NER efficiency [548].

1.5.6 SUMOylation in NER

SUMOylation is a process similar to ubiquitinylation with a three-step enzymatic process. This process attaches SUMO to a substrate via an isopeptide bond between SUMO's C-terminal carboxyl group and a lysine residue in the substrate, similar to ubiquitylation. Although the enzymes of the SUMO pathway are similar to those of the ubiquitin pathway, they are specific for SUMO. SUMOylation begins with the activation of the SUMO C-terminus by a SUMO-activating enzyme (E1) in an ATP-dependent manner. A SUMO that has been activated is transferred to the SUMO-conjugating enzyme Ubc9 (E2). Then SUMO is transferred from Ubc9 to the substrate with the help of one of several SUMO-protein ligases (E3s). UBC9 and the E3s ensure substrate specificity [549], [550]. Many proteins have been reported to be SUMOylated in the NER pathway, such as DDB2, CSB, and XPC [549]. In response to UV irradiation, the SUMOylation of XPC by SUMO-1 increases its stability by inhibiting degradation via a ubiquitin-proteasome system to initiate DNA damage recognition in the NER pathway [551]. Inhibiting SUMOylation reduced the recruitment of CSB at UV-irradiated DNA lesions and the recovery of transcription.

1.5.7 PARylation in NER

PARylation is a reversible PTM catalyzed by poly(ADP-ribose) polymerases (PARPs), which form an ester bond between ADP-ribose and the carboxyl-group of acidic amino acids [552], [553]. The poly(ADP-ribosyl)ation of histone H1 relaxes chromatin fiber at DNA lesion and enables DNA repair machinery to access damaged DNA [554]. In response to UVR-induced DNA damage, XPA has been shown to associate with PARP-1 and PAR to be recruited to DNA lesions [555].

1.6 Mass spectrometry-based proteomics

DNA damage is unavoidable, regardless of whether the physiological or abnormal conditions. In response to DNA damage, DNA repair pathways are activated to remove the damaged DNA or induce cell cycle arrest or apoptosis. During the process, PTMs modulate enzymatic activities and regulate protein stability, localization, and protein-protein interactions in cells. The development of specific detection and purification methods are the main technical challenges in studying all these biological processes. PTMs will also be better studied with the advancement of life science technology and chemical synthesis and the application of computer image modeling, which will contribute to a better understanding of the role of PTMs in DNA repair. Elucidating the molecular mechanisms of DNA repair pathways involving PTM signaling could reveal new and selective therapeutic approaches to target cancers.

Fortunately, a variety of new and refined proteomics technologies are overcoming these technical challenges, such as quantitative proteomics analysis based on mass spectrometry (MS). MS was first used to trace heavy isotopes through biological systems, nearly 100 years after it was developed to measure elemental atomic weights and the natural abundance of specific isotopes. MS was later used to sequence oligonucleotides and peptides and also nucleotide structures [556].

Mass spectrometry is a powerful analytical technique that can be used to determine the mass-to-charge ratio (m/z) of one or more molecules in a sample. These measurements are frequently used to calculate the precise molecular weight of the sample components. Therefore, mass spectrometers are typically used to identify known material, quantify unknown or known compounds, and elucidate the structure and chemical properties of various molecules through molecular weight determination [557]–[559]. The entire process entails converting the sample into gas-phase analyte ions, with or without fragmentation, which is then classified based on their specific m/z ratios and relative abundances. To achieve this, all mass spectrometers have three vital parts: the ion source, the mass analyzer, and the ion detector. The nature of these components differs depending on the mass spectrometer's purpose, the type of data required, and the physical properties of the sample [559].

MS-based proteomics usually is performed with two strategies: Top-down or bottom-up. In "top-down" proteomics, intact proteins are analyzed, which theoretically allows for simultaneous detection of all existing modifications as well as correlations between these modifications. In contrast, in "bottom-up" proteomics, analytes are peptides, not entire proteins [560].

1.6.1 Sample preparation and electrospray ionization

In "bottom-up" proteomics, sample preparation differs depending on the purpose and complexity. But there are still some common steps: Proteins are isolated from biological material, such as cells and tissues, then disulfide bridges are reduced, and the resulting free cysteines are alkylated. Then proteins are digested with a sequence-specific protease, typically trypsin, which is highly active and tolerant of many additives like 2 M Urea. Trypsin cuts C-terminally of lysine and arginine, leaving

positive charges on the peptides and making them detectable by MS. Protein digestion produces a complex mixture of peptides [561]. In order to reduce the sample complexity, fractionation is usually applied to separate peptides based on their physicochemical properties, including polarity, hydrophobicity, charge, and size. There is offline and online fractionation [562]. For example, the Fe-IMAC or TiO_2 based enrichment or immunoprecipitation with diGly remnant specific antibodies steps followed with strong cation exchange (SCX) are commonly performed to enrich phosphorylated peptides or ubiquitinated peptides respectively and improve the detection of PTM-modified peptides, which is more complex and challenging [562], [563]. To further decrease the sample complexity before introducing to MS, peptide mixtures are commonly separated by ion-pair reversed-phase high-performance chromatography (RP-HPLC) at low pH. High peak capacities and peptide resolution make this technique highly suitable for bottom-up approaches. It is achieved through differential solvophobic interactions of unipolar side chains with the non-polar stationary (e.g., octadodecyl alkane chains, C_{18}) and mobile phase. Peptide retention and thereby resolution can be further increased by the addition of ion-paring reagents (amphiphilic molecules) that mediate the interaction of polar peptide chains with the stationary phase. In HPLC, the concentration of organic solvents in the mobile phase is gradually increased to achieve the consecutive elution of peptides. As a result, peptide retention increases as peptide hydrophobicity increases (**Figure 21**) [563].

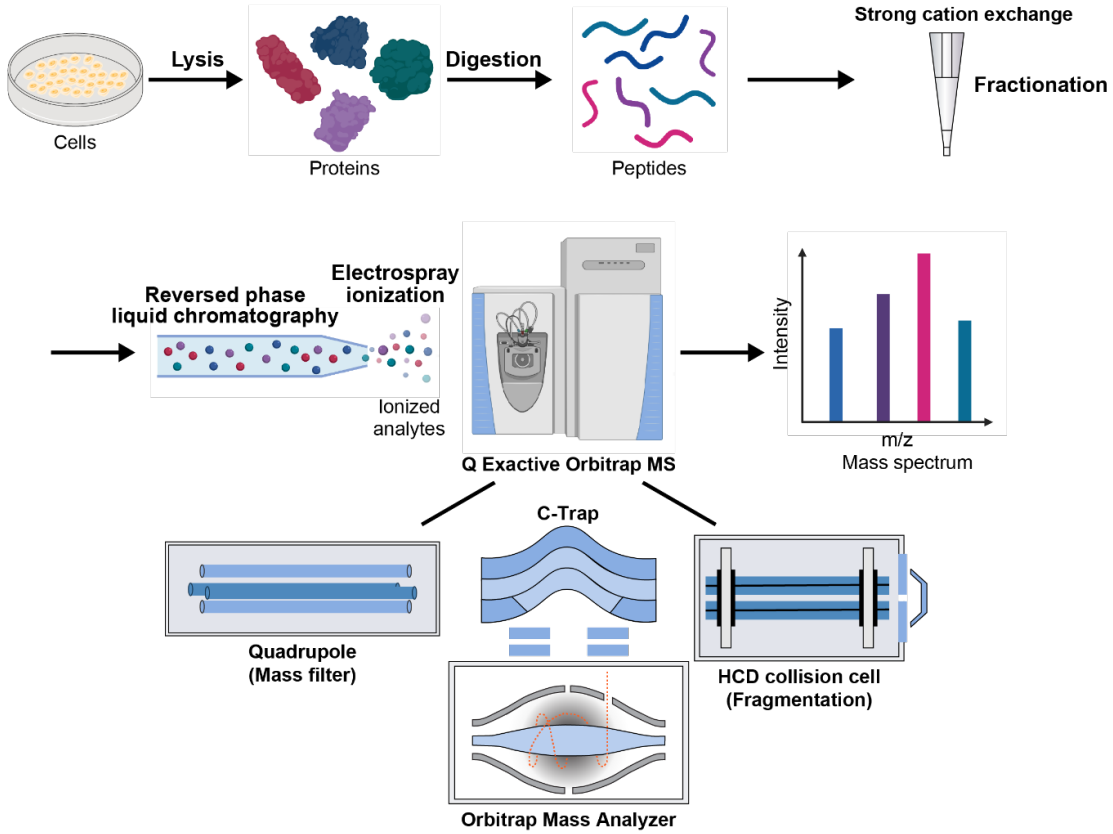

Figure 21: Schematic of the workflow of bottom-up proteomics and Orbitrap Q Exactive mass spectrometer. Proteins are extracted from cells and digested by trypsin into a peptide mixture. Peptides are further fractionated through enrichment followed by strong cation exchange. Peptides are ionized and transferred into the gas phase by ESI. Ionized peptides will be introduced into the mass spectrometer and accelerated forward towards the quadrupole mass analyzer, functioning as a mass filter. Then peptides are stored and focused in the C-trap before being channeled to the HCD collision cell for fragmentation or the orbitrap for MS1 and MS2 spectra acquisition. Some parts of the image were created with BioRender.com.

Samples are loaded into the mass spectrometer as liquids, gases, or dry particles, which are then vaporized and ionized by the ion source such as ElectroSpray Ionization (ESI), Atmospheric Pressure Chemical Ionization (APCI), Matrix-Assisted Laser Desorption Ionization (MALDI) [564]. Our lab uses ESI to get gas-phase ions and introduce these analytes into MS. The peptide solution exits the reverse phase column to a thin capillary (emitter) via an electric field. As a result, peptides are in the formation of the tailor cone. The liquid is dispersed into small, multiply charged droplets as the Coulomb forces become stronger. The liquid evaporates as it moves through the electrostatic field towards the mass spectrometer's vacuum system, increasing the charge density on the droplets. When the Coulomb forces exceed the surface tension (Rayleigh limit), the charges repel each other, causing the droplets to explode into even smaller ones (Coulomb explosion) and subsequently introduce peptides into MS as single ions [565].

1.6.2 Mass analyzer

Once ionized, ions are sorted and separated based on m/z ratios via a mass analyzer, which is the heart part of the mass spectrometer [566]. To date, there are various types of mass analyzers, including quadrupole, ion trap, orbitrap, Fourier transformation, time of flight (TOF), and magnetic sector analyzers. Each of them has its own set of trade-offs in terms of operation speed, separation resolution, and other operational requirements [564]. Common to all mass analyzers is the manipulation of analyte trajectories in an electromagnetic field through the application of direct (DC) and alternating (AC) currents. They differ in terms of resolution, mass accuracy, sensitivity, and dynamic range. The mass range determines the m/z limit for ion measurement by a mass analyzer. The rate measured by the analyzer in a specific mass range is referred to as the scan speed. The mass accuracy is defined as the difference between the measured m/z and the exact m/z of an ion. And resolution is the ability to distinguish between two peaks with slightly different m/z values. Modern mass spectrometers frequently combine at least two different mass analyzers, allowing for flexible data acquisition [567]. Two such hybrid mass spectrometers (Q Exactive plus and Exploris 480), incorporating both a quadrupole and an orbitrap mass analyzer, were used in our lab (**Figure 21**). The mass analyzer is frequently used in conjunction with the ion detection system, determining the reliability and quality of the analysis [567].

1.6.3 Peptide detection

After passing through the mass analyzer, ionized peptides are detected and converted into an analytical signal that shows *m/z* ratios along with their relative abundance, resulting in a mass spectrum by comparing them to known m/z standards. Detectors are capable of producing an electric current proportional to the abundance of incident ions. As the number of ions leaving the mass analyzer at any given time is usually quite small, amplification is commonly used to obtain a usable signal. However, some detectors are designed to count ions of a single mass at a time. Therefore, they detect the arrival of all ions sequentially at one point. Other detectors (such as photographic plates or image current detectors) can count multiple masses and detect the arrival of all ions along a plane at the same time [568].

1.6.4 Peptide and protein identification using tandem mass spectrometry

Modern mass spectrometers automatically select the most intensely ionized peptides from the top N parent or precursor ions (MS1), followed by fragmenting them in a high collision-induced fragmentation cell, allowing the specific peptides to be recorded (MS2) [568]. Fragmentation itself is commonly achieved by collision-induced dissociation (CID) or higher-energy c-trap/collision dissociation (HCD), in which peptide ions collide multiple times with inert neutral gas molecules (Ar, He, N_2). The kinetic energy is converted into internal vibrational energy, resulting in the bond breakage. In both methods, fragmentation occurs primarily at the peptide bond, which is advantageous for sequence determination. However, CID has a higher preference for low energy bonds, resulting in more neutral loss ions, and it also suffers from peptide loss in the low mass region. In the Exploris, ions with a specific m/z value are first accumulated, focused in a C-trap, and then accelerated towards the Ion-Routing Multipole for HCD. HCD-induced fragmentation typically produces b and y ions. In an ideal scenario, fragments differ by only one amino acid, allowing the peptide sequence to be determined. This procedure is referred to as tandem MS (MS/MS) [569].

The acquired mass spectra lay the groundwork for peptide and thus protein identification. The MS1 spectrum provides information about the peptide mass, whereas the MS2 spectrum allows the peptide sequence to be determined. In theory, this allows for the derivation of the entire peptide sequence (de-novo-sequencing). MS/MS-based peptide identification is typically performed via database searching in data analysis programs like MaxQuant. MS2 spectrums are matched to theoretical ones, generated by *in silico* digestion of proteins from a chosen proteome database or a spectral library. Probabilities for the best matches are calculated, and peptide identities are assigned based on software and user-specific parameters. A target-decoy approach is used to assess the accuracy of the assignments. Spectra matching to the decoy database is wrong by definition, allowing for the calculation of a false discovery rate (FDR) [570]. High-quality peptide identifications can be obtained by combining both pieces of information and using stringent cut-offs. An identified peptide should ideally match a specific protein. In reality, proteins share parts of their primary sequence naturally. Therefore, a protein can only be identified if a unique peptide is measured. This is known as the protein inference problem [571].

1.6.5 Peptide and protein quantification

To accomplish the quantification of proteins and peptides, different techniques can be applied depending on the demand, such as label-free quantification (LFQ), stable isotope labeling with amino acids in cell culture (SILAC), and isobaric isotope labeling with tandem mass tag (TMT) [572]. One disadvantage of LFQ is that samples are processed and measured separately, resulting in low reproducibility. In contrast, SILAC allows for early mixing and measurement of multiple samples at the same time. However, SILAC is typically limited to three conditions and produces more complex MS1 spectra compared to LFQ [573]. TMT is added to samples at the peptide level, allowing for peptide multiplexing up to 18 conditions and a direct comparison between different them in replicates [574]. Additionally, multiplexing allows deep peptide coverage in a reasonable time and results in a low

number of missing values compared to LFQ. TMT tags consist of a reactive group, a mass normalization spacer, and a reporter group. The reactive N-hydroxysuccinimide (NHS) ester reacts with primary amino groups such as peptide N-termini or lysine-amino groups. Following that, the labeled peptides are pooled and processed together to reduce technical variation in the downstream process. Therefore, in different samples, peptides with identical sequences share the same chemical structure and mass, while they are made up of distinct combinations of heavy carbon and nitrogen isotopes due to the TMT tags. These peptides will contribute to the same MS1 signal in MS. After fragmentation, the sample-specific reporter groups with unique m/z ratios are released and used for relative quantification in the same MS2 scan. Therefore, TMT quantification will suffer co-isolation interference when more than one peptide is fragmented from MS1 and subsequently induce reporter quantification interference. Several methods have been used to address these issues, including software corrections, additional gas-phase manipulation, measuring reporter ions in the third dimensions (MS3), and quantifying peptide fragments complementary to the reporter ions (TMTc) [574].

1.7 Aims of the study

UV irradiation has been previously shown to produce reversible transcription-blocking lesions, including pyrimidine dimers (CPDs) and the 6,4-pyrimidine dimers (6,4-PPs). Protein phosphorylation is central for spatial and temporal response to these transcription-coupled DNA damages to achieve the regulation of different processes. However, a systematic investigation of these changes on the cellular proteomics scale remains poorly studied. In this doctoral thesis, we used skin cell, keratinocytes as a model system. We employed tandem mass-tag-based quantitative mass spectrometry (TMT-MS) to quantify the changes of global proteome and phosphoproteome with a great spatial and temporal resolution, at different time points spanning from 30 min to 18 hours after UV irradiation. This permitted to characterize phosphorylation signaling that regulates different cellular responses to UV: (i) transcriptional shutdown, (ii) the recognition and repair of DNA lesions, (iii) the early and the late transcriptional restart, as well as (iv) when transcription is fully recovered. We especially hope to uncover specific phosphorylation signaling cascades associated with these 4 distinct stages of the cellular response to UV irradiation. We aim to provide a resource for identifying players required for a variety of cellular mechanisms in transcriptional regulation and DNA repair.

By using a specific JNK1 kinase inhibitor, we are able to uncover the dependencies of phosphorylation events during response to UV-induced DNA damage on the JNK1 MAPK pathway. Further, with various biochemistry methods, we try to mechanistically reveal the role of the transcription-related kinase JNK in response to UV stress.

2 Results

2.1 Analysis of the phosphoproteome and proteome landscapes in response to UV irradiation-induced DNA lesions by TMT-based mass spectrometry

To explore the dynamics of the cellular response following UV irradiation, we measured changes in global protein levels, protein phosphorylation, and chromatin-associated proteomes across different time points post sub-lethal UV irradiation of HaCaT cells. For this multi-sample analysis, we took advantage of TMT labeling coupled with mass spectrometry (MS).

2.1.1 UV induces conformational changes in DNA structure changes and transcription inhibition

UV irradiation of HaCat cells resulted in the formation of both CPD and 6-4PP photolesions in a dose-dependent manner (**Figure 22**A, B). Along with the formation of lesions, UV irradiation reduced cell viability with LC50 of 15 J/m^2 at 48 hours post-irradiation (**Figure 22**C). The two photolesions were resolved with different kinetics, with the 6,4-PP levels returning to undetectable baseline levels by 12 hours post-UV irradiation. By contrast, the CPD photolesions persisted up to 24 hours post-irradiation (**Figure 22**E, F). Furthermore, UV irradiation rapidly inhibited RNA Pol II-dependent transcription as measured by 5-EU incorporation. Transcription started to recover 1 hour post-UV and returned to normal levels 18-24 hours post-UV (**Figure 22**D).

For a comprehensive screen, we selected 4 different time points post-UV to analyze the (phospho)proteome: at an early time point (0.5h) when transcription is inhibited, at later time points (2 and 6h) when cells are focusing on the DNA repair and transcription restart, and at a later time point (18h) post-UV irradiation when transcription is recovered completely (**Figure 23**). Briefly, human keratinocyte HaCaT cells were treated with a median lethal dose of UV irradiation (15 J/m^2), and left for recovery for 30 minutes, 2 hours, 6 hours or 18 hours. Then cells were processed for chromatin fractionation or lysed in modified RIPA buffer for total proteome and phosphorylated peptides enrichment (**Figure 23**). Proteins were precipitated in acetone followed by in-solution digestion into peptides using both LysC and trypsin. The peptides originating from differentially treated cells were labeled with the TMT11-plex kit and mixed accordingly. Phosphopeptides were enriched using TiO2 beads after sample labeling. To allow for downstream data normalization, a reference sample, mixed from all other samples, was included in each TMT batch. The samples were fractionated by micro SCX (strong cation exchange) chromatography and applied to an easy-nLC RP chromatography coupled to the Orbitrap Exploris mass spectrometer (**Figure 23**). Following measurement, raw intensities were normalized and used for statistical analysis. Of these steps, protein digestion, peptide labeling, peptide fractionation, and data normalization are described in more detail in the data analysis part.

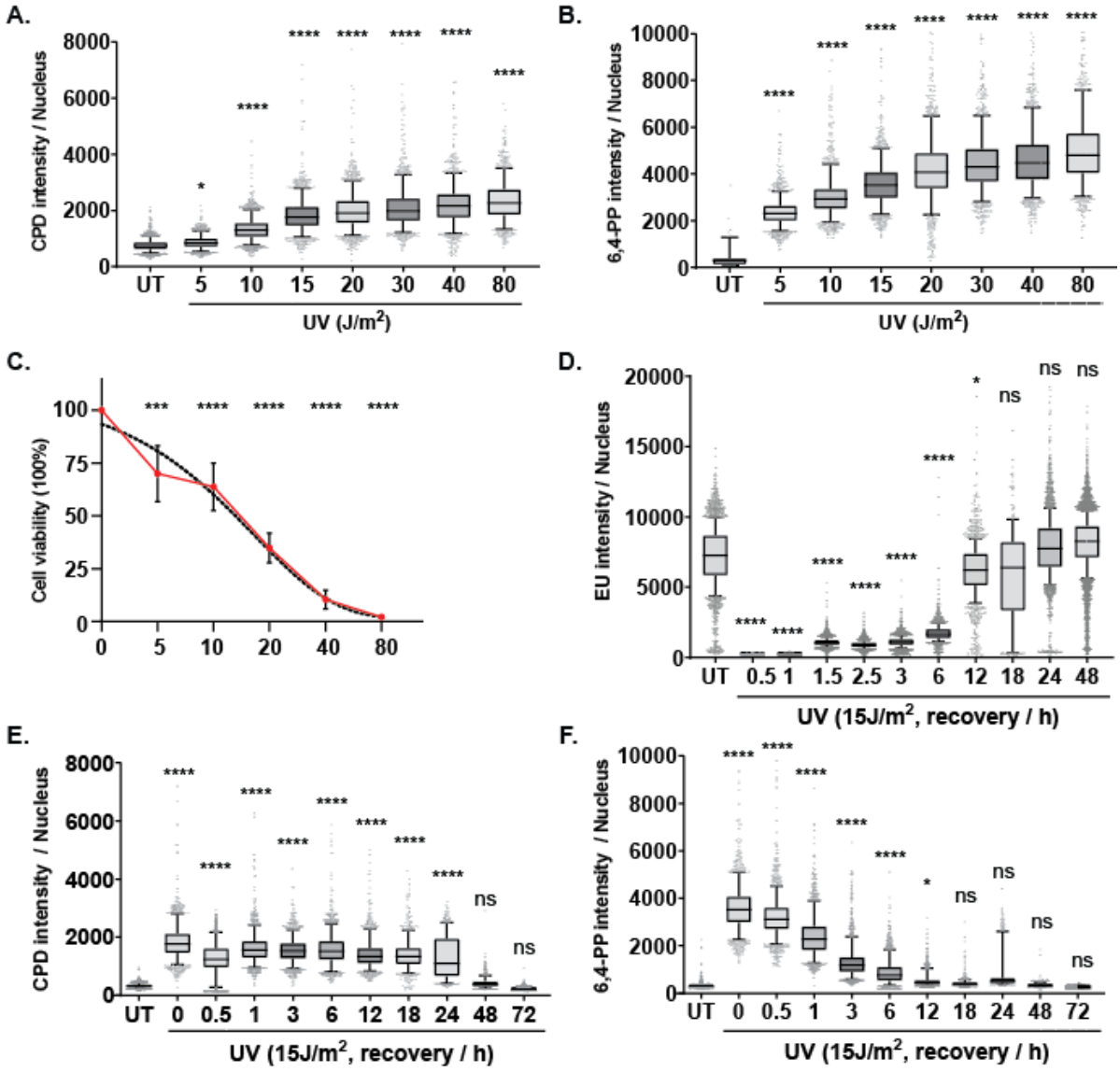

Figure 22: UV irradiation induces DNA dimers and transcription inhibition. HaCaT cells were treated with the indicated dosages of UV light and cells were immediately fixed for immunofluorescence microscopy (**A and B**), or cells were left to recover for indicated hours and then fixed for immunofluorescence microscopy (**D, E, F**). Cell viability was evaluated using the CellTiter-Blue viability assay. Titration curves are shown for three replicates. Reported LC_{50} values were calculated from all replicates (n = 3). One-way ANOVA was used to calculate the significance (****p-value < 0.0001) (**C**). This experiment was done by Claudia Scalera. Boxplot displaying the quantification of the mean 6,4-PP, CPD or EU intensity per nucleus for cells. Center of boxplots indicates the median and whiskers the 10th-90th percentile. ****P-value < 0.0001, One-way ANOVA.

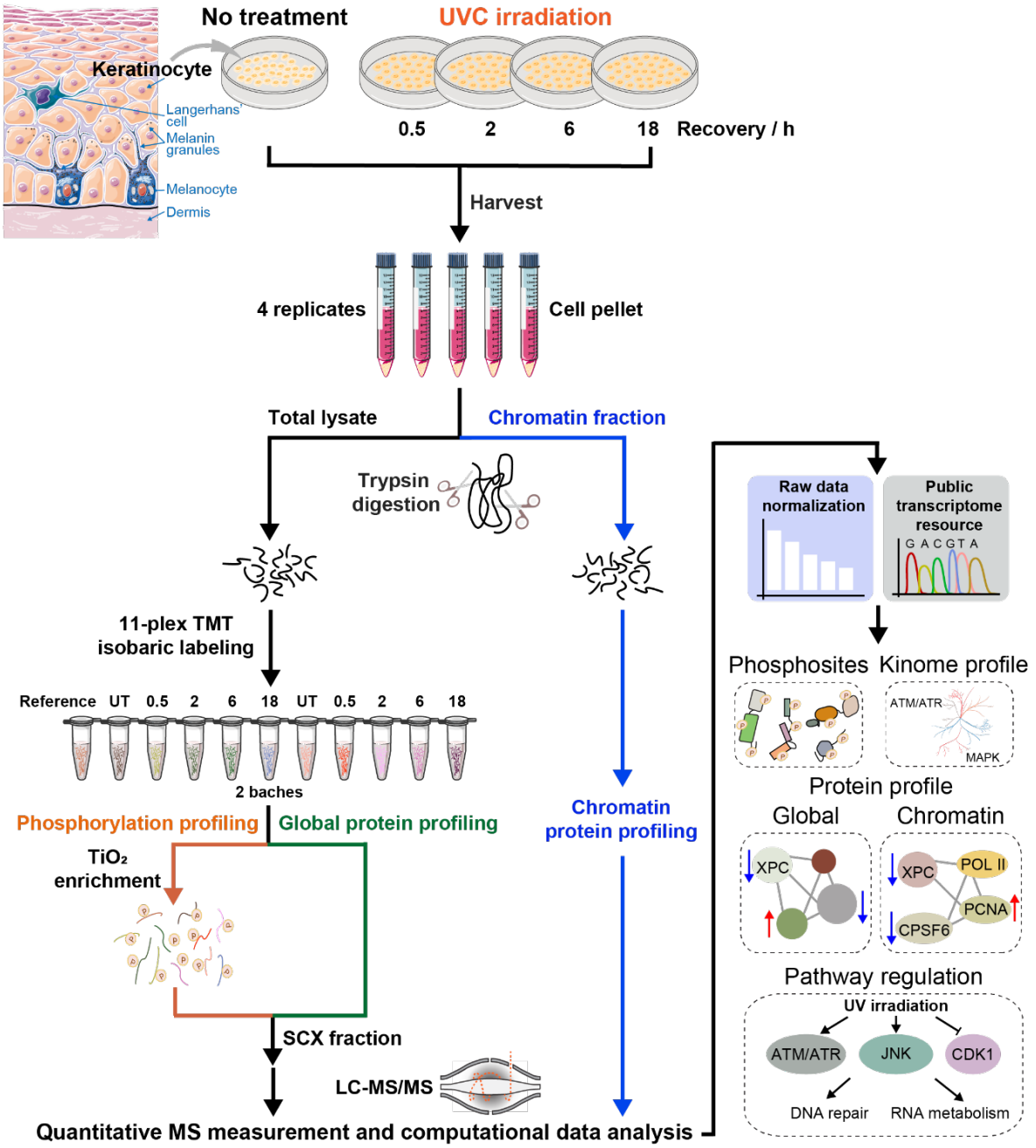

Figure 23: Proteomics analysis of the cellular response to UV irradiation-induced DNA damage. HaCaT cells were treated with UV(15J/m^2). Cells were lysed with fractionation procedure for chromatin-bound proteins (blue) or lysed in RIPA buffer, precipitated in acetone, and digested with both LysC and trypsin. Peptides were labeled with TMT, fractionated by SCX, and analyzed by LC-MS/MS (orange). Alternatively, phosphopeptides were enriched using TiO_2 beads after TMT labeling (black). The phosphorylation dataset analysis was done by Matthias Ostermaier.

2.1.2 UV stress-induced changes in protein levels

In the global proteome analysis, we identified 10,110 protein groups. For high confidence protein identification, we further excluded proteins identified with no proteotypic and less than two peptide identifications. Further, only proteins identified in all four replicates were included. After filtering with this criteria, 6,083 protein groups were left to be analyzed statistically (**Figure 24**B). Uniform Manifold Approximation and Projection (UMAP) showed high reproducibility and the replicates clustered according to treatment (**Figure 24**A). As expected, protein levels remained largely

unchanged at the early recovery time point (0.5h) after UV irradiation. Only 51 proteins showed significant (1.5 fold change, FDR ≤ 0.05) change in levels, which explains the relatively poor clustering of the treatment replicates (**Figure 24**B). We observed an increase of proteins that show significant changes in levels after longer recovery times, especially after 18 hours of recovery (**Figure 24**A).

In contrast, after a longer recovery post-UV irradiation, more proteins are found to be regulated that are involved in protein modification and mRNA metabolic processes, especially after 18h (**Figure 25**). Network analysis (Confidence = 0.7) of the 142 proteins using the STRING database, Ensemble of Gene Set Enrichment Analyses (EGSEA) with REACTOME and GO annotation revealed that these proteins fall into several categories, such as "DNA repair and cell cycle," "mRNA metabolic process," "cytokine response in the immune system" and "cellular organization" (**Figure 25** and **Figure 26**). Surprisingly, proteins involved in the immune response are upregulated when transcription is fully recovered and DNA lesions are removed completely, for example, Interferon-stimulated genes coded proteins including ubiquitin-like protein ISG15, and Interferon-induced protein with tetratricopeptide repeats (IFIT) like IFITM3, IFIT1 and IFIT3. Other studies also revealed that these genes are more transcripted upon UV stress, which are indicated with purple rings in network (**Figure 26**). This strongly suggested that UV stress induces an inflammatory response in primary cells.

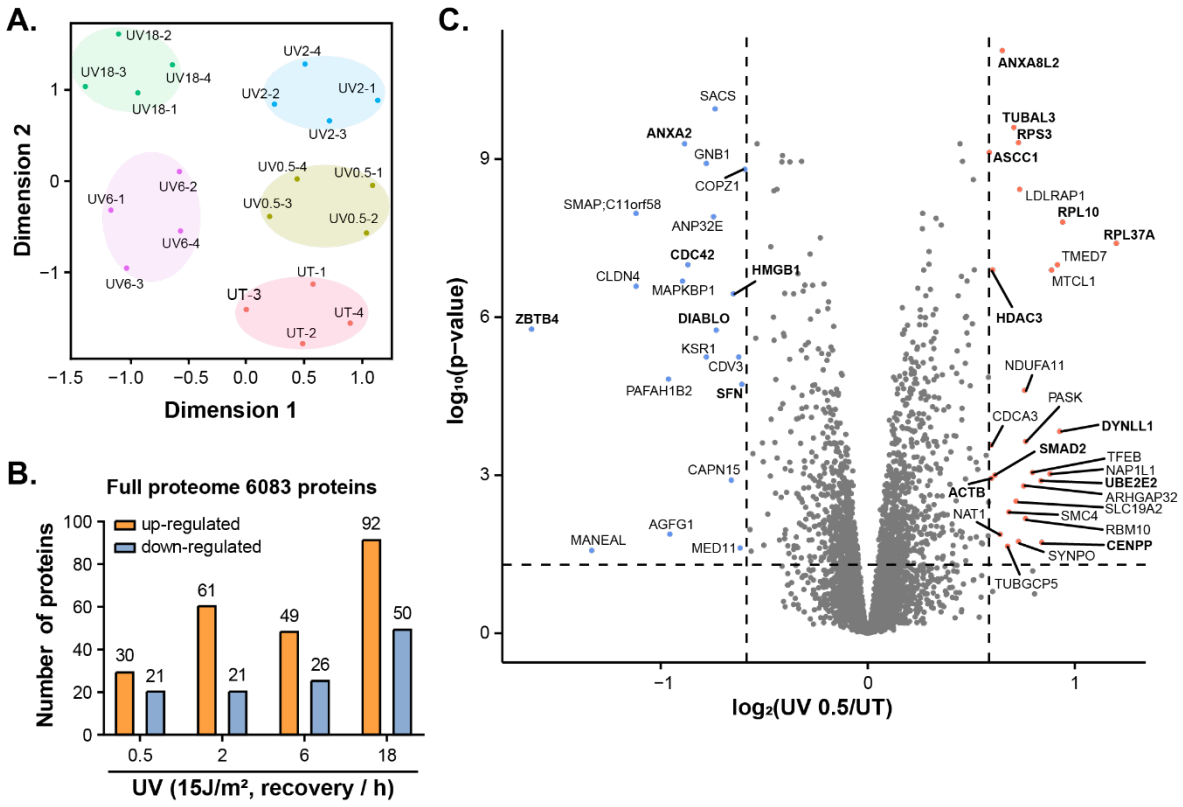

Figure 24: UV stress induces protein level changes. A. UMAP for the correlation of global proteome analysis between different replicates. **B.** The number of proteins with significantly regulated levels for each time point is depicted in the bar plots after different time point recoveries from UV light-induced DNA damage in HaCaT cells. The overlap of regulated proteins for 2h, 6h and 18h (n= 4; FDR ≤ 0.05; fold-change ≥ 1.5). **C.** Volcano plot of regulated proteins at 0.5h post UV treatment. (n= 4; FDR ≤ 0.05; fold-change ≥ 1.5).

Figure 25: Heatmap of GSEA analysis for proteins with regulated protein levels after different recovery time points post-UV irradiation. The color panel indicates the significant score of enrichment. (n= 4; fold-change ≥ 1.5, or ≤ 1.5; p-value ≤ 0.01 is indicated with an asterisk).

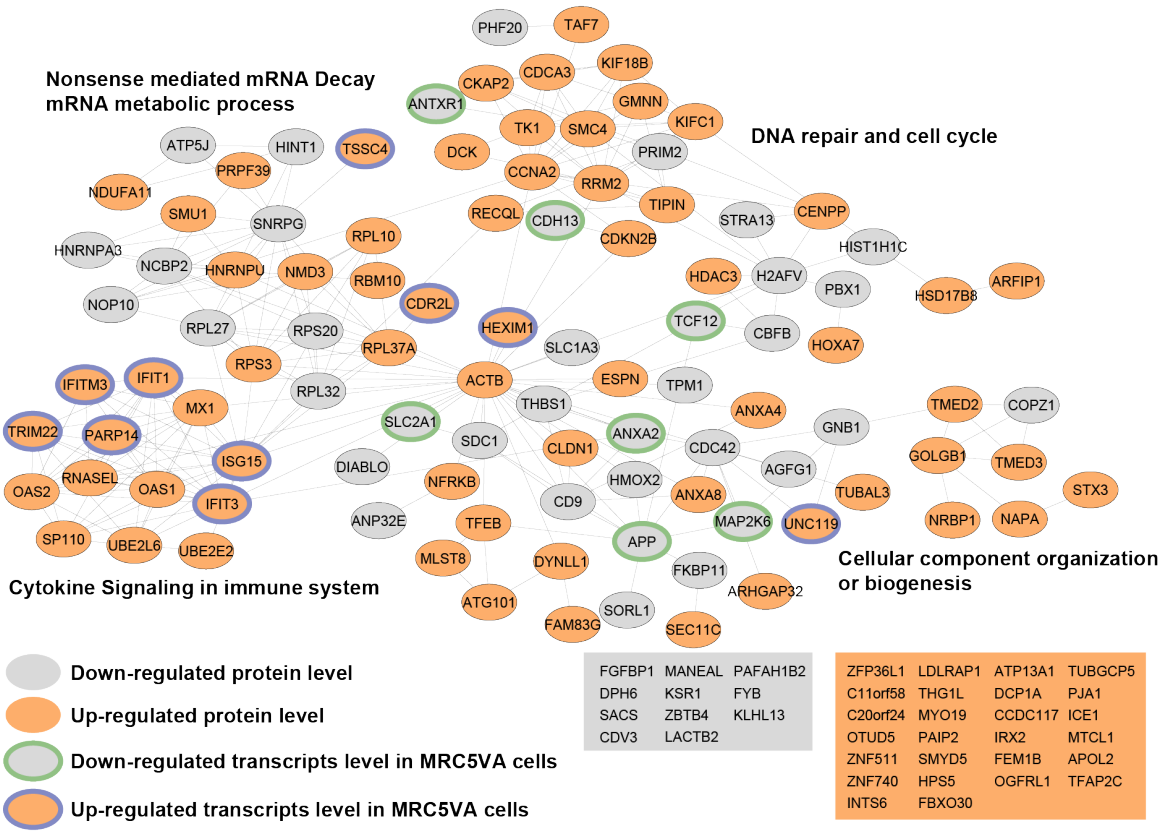

Figure 26: Network of proteins with regulated levels at 18h after UV treatment. Proteins with upregulated levels at 18h are indicated with orange, and with downregulated levels are labeled with grey (FDR ≤ 0.05, fold change≥1.5 or ≤1.5) (Confidence = 0.7). Green and purple circles indicate downregulated or upregulated transcription levels in MRC5VA cells at 20h after UV exposure (15J/m^2) [575].

2.1.3 UV-stress induced protein localization on chromatin

In total, we identified 4,790 protein groups from the isolated chromatin fractions. Missing values were replaced by random numbers drawn from a normal distribution with a width of 0.3 and down shift of 1.8. Similar to the global proteome, for high confidence protein identification, we further excluded proteins identified with no proteotypic and less than two peptide identifications. Only proteins identified in three out of four replicates in at least one treatment group were included. After filtering with this criteria, 3,542 protein groups entered the statistical analysis (**Figure 27**B). Statistical analysis shows regulated proteins (1.5 fold change, FDR ≤ 0.05) in at least two replicates out of four replicates for each condition.

Similar to the proteome analysis, relatively good clustering of the treatment and four replicates is indicated in UMAP (**Figure 27**A). During the short recovery phase from UV irradiation (0.5h), only 8 proteins are recruited to chromatin, which includes single-stranded DNA protein RPA3, transmembrane BAX inhibitor motif containing 1 (TMBIM1), Peroxiredoxin-4 (PRDX4), eukaryotic translation initiation factor 4E (EIF4E2) and GRB10 Interacting GYF Protein 2 (GIGYF2) (**Figure 27**B and C). Of note, the general transcription initiation factor TFIIH and its associated cyclin-dependent kinases 7 (CDK7) are rapidly removed from chromatin upon UV stress (**Fig. S2B**). 7SK snRNP (small nuclear ribonucleoprotein) component, HEXIM1 is also removed from chromatin. The re-localized HEXIM1 may bind to the proximal and distal parts of 7SK, concluding the formation of the canonical

7SK snRNP to inhibit the activity of P-TEFb. This will further inhibits the release of paused RNA Pol II to the gene body.

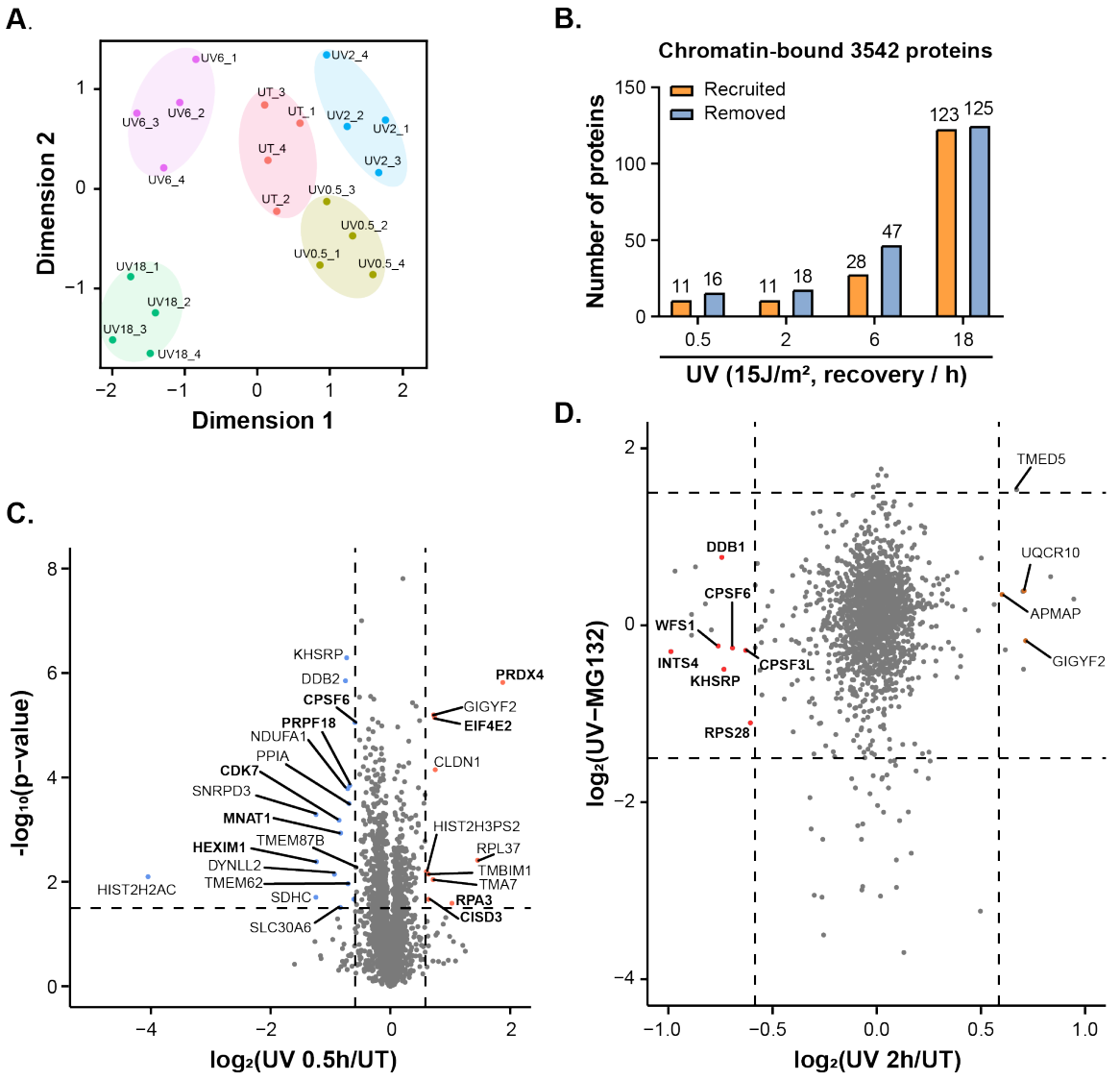

Figure 27: Analysis of proteins that are removed from or recruited to chromatin during the response to UV stress. A. UMAP analysis of biological replicates for quantitative mass spectrometric analysis for chromatin-bound proteins after different recovery time points post UV light-induced DNA damage in HaCaT cells. **B.** The number of proteins that are significantly removed from or recruited to chromatin at each time point is depicted in the bar plots (n= 4; FDR ≤ 0.05; fold-change ≥ 1.5, or ≤ 1.5). **C.** Scatter plot for chromatin-bound proteins' profile at 2h against the chromatin protein landscape with proteosome inhibitor MG132 treatment from the Svejstrup lab [576]. Proteins removed from chromatin (n= 4; FDR ≤ 0.05; fold-change ≥ 1.5, or ≤ 1.5) and not regulated upon MG132 treatment are labeled (n=2; FDR ≤ 0.05, fold-change ≥ 1.5, or ≤ 1.5). **D.** Volcano plot of the logarithmized ratios UV/UT for proteins recruited to or removed from chromatin after 0.5h (left) and 6h (right) recovery post UV irradiation (n= 4; FDR ≤ 0.05; fold-change ≥ 1.5, or ≤ 1.5).

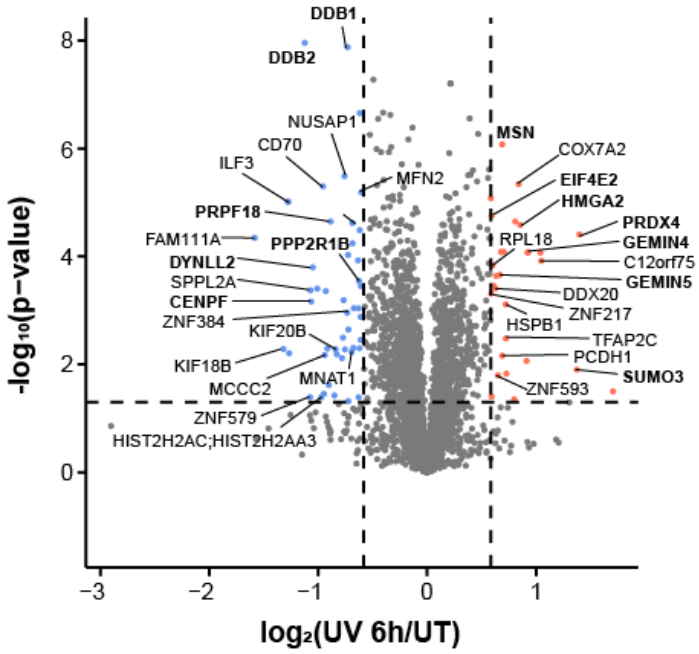

Figure 28: Volcano plot of the logarithmized ratios UV/UT for proteins recruited to or removed from chromatin after 6h recovery post UV irradiation (n= 4; FDR ≤ 0.05; fold-change ≥ 1.5, or ≤ 1.5).

The restart of transcription, at 6h after UV stress, leads to different patterns of protein association to chromatin compared to early time points that reflect transcription inhibition. The survival of motor neurons (SMN) complex components (DDX20, GEMIN4, and GEMIN5) are recruited to chromatin, which are essential for the biogenesis of small nuclear ribonucleoproteins (snRNPs) in cells (**Figure 28**) [577]–[580]. Notably, HEXIM1 comes back to chromatin. All of these changes indicate the disassembling of 7SK snRNP, aiding the release of paused RNA Pol II and the restart of transcription (**Figure 28**). Moreover, the UV-lesions detection DDB1-DDB2 complex dissociates from chromatin (**Figure 28**). It suggested that UV-lesions initial detection has been finished. Notably, NADH dehydrogenases, NDUFB2 and NDUFA1 are removed from chromatin. It could be the result of ROS induction by UV stress [581], [582]. After 18 hours of recovery post UV stress, we found that most of the recruited proteins are involved in the regulation of the cell cycle during G1/S transition. Subsequently, these changes result in the cell cycle arrest and DNA synthesis block, as we also observed (**Figure 29** and **Figure 30**).

To identify overrepresented signatures among the different proteins that are differentially recruited to or removed from chromatin at different recovery phases, we carried out EGSEA, similar to the global proteome dataset, which integrates 12 prominent gene set testing algorithms to obtain biologically relevant results [583]. The significantly enriched GO, KEGG, and REACTOME signatures were further grouped based on hierarchical clustering and named according to the most prominent feature (**Figure 30**). Immediate response pathways to UV light are involved in "transcription regulation," "NER" and "TC-NER." In later responses, proteins recruited to chromatin fall into

categories related to "cell cycle," "DNA damage checkpoint," and "proteasome," and proteins removed from chromatin fall into the category of "translation initiation" (**Figure 30**).

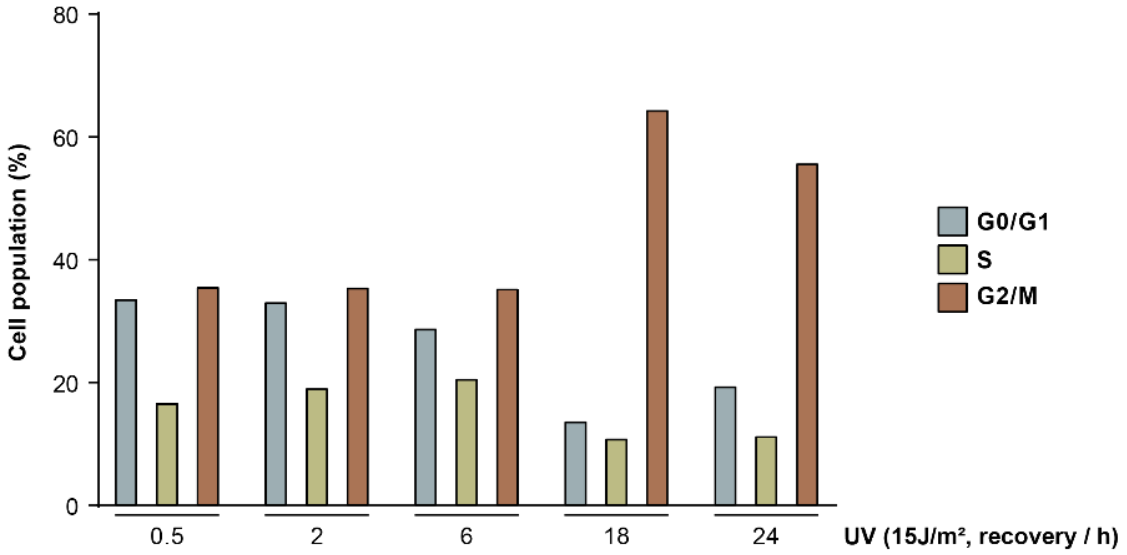

Figure 29: Cell cycle is arrested at G2/M phase after 18 h recovery post UV irradiation. FUCCI HaCaT Cells were treated with UV (15J/m^2), then left to recover for different time points. Nuclei were stained with DAPI, and flow cytometry analysis was performed. Around 21,000 cells were mesured.

Together, all these significant changes suggest that UV irradiation induces the removal of transcription regulators and translation initiation complex; and the recruitment of some splicing factors to chromatin after a short recovery time while cells are dealing with transcription inhibition. Proteins involved in other subsequent effects, such as ROS response and cell cycle arrest, are recruited to chromatin after the longer-term recovery from UV irradiation.

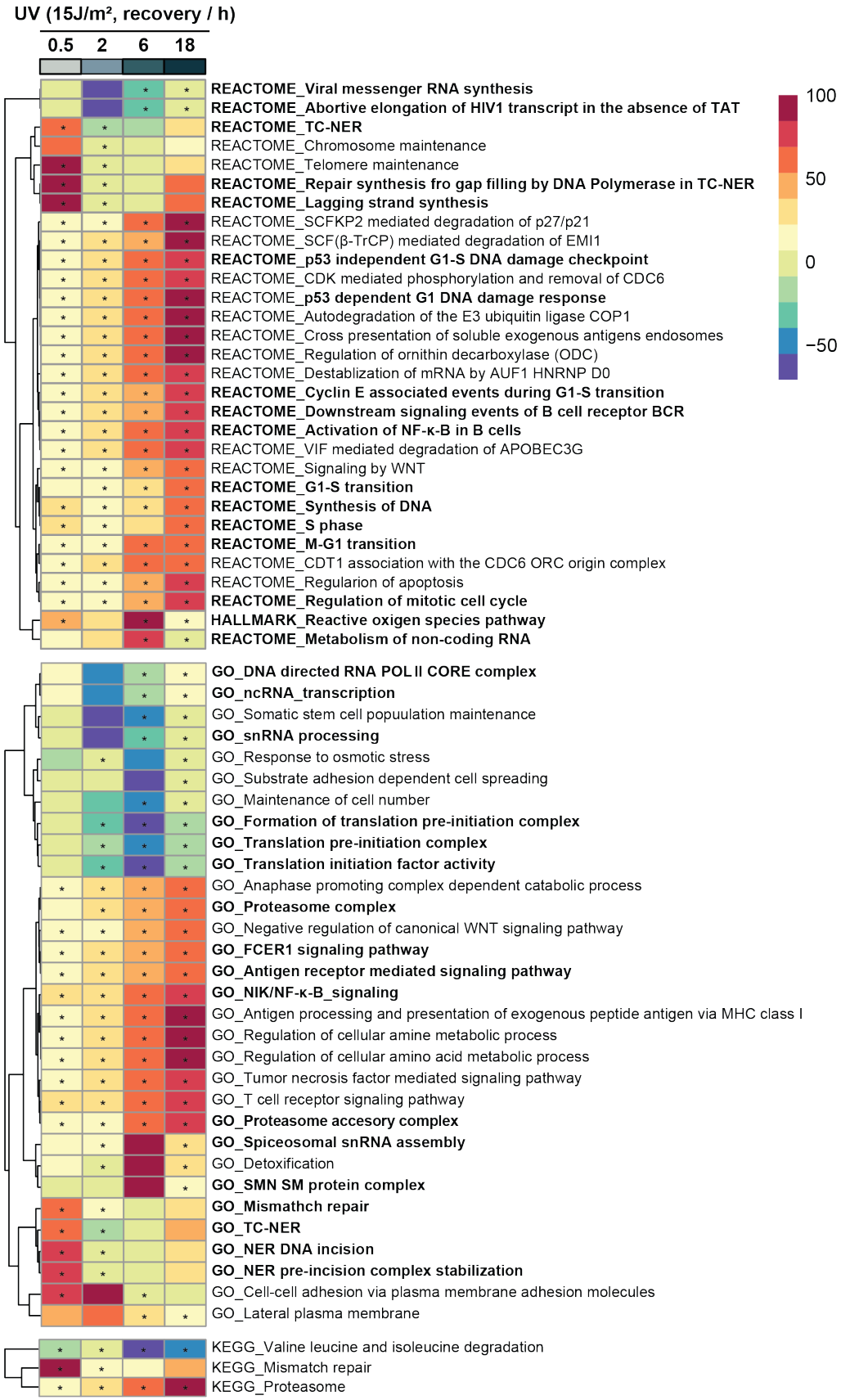

Figure 30: Heatmap of GSEA analysis (GO-terms) for proteins that are removed from or recruited to chromatin after different recovery time points post-UV induction. The color panel indicates the significant score of enrichment. (n= 4; fold-change ≥ 1.5, or ≤ 1.5; p-value ≤ 0.01 is indicated with an asterisk).

2.1.4 Different kinase modules are regulated in response to UV stress

The phosphoproteome clustered nicely according to the different recovery time points as shown by UMAP (**Figure 31**A). In total, we identified 37,450 phosphorylation sites. Of which, 26,500 sites were of high confidence (localization probability > 0.75). 17,848 phosphorylation sites on 4,775 proteins were quantified in four replicates. Among these quantified phosphorylation sites, 1,886 (10.6%) sites have not previously been identified within the PhosphoSitePlus database (https://www.phosphosite.org/) (04/2021). In comparison with the proteins level, protein phosphorylation changes occur rapidly in response to UV stress. The number of regulated phosphorylation sites increased further with longer recovery times, peaking after 6 hours. Interestingly, more downregulated phosphorylation sites were also observed at this time point (**Figure 31**B). After 18 hours of recovery post UV irradiation, there is a decrease in differential phosphorylation of downregulated sites (**Figure 31**B).

To explore which kinases respond to UV stress, we compared kinase activities using two different approaches. First, linear motif enrichment analysis was carried out using a stretch of 13-amino acids surrounding the differentially regulated phosphorylation sites. Overrepresentation of amino acids was displayed using IceLogo and enriched motifs were identified by sequence annotation followed by Fisher's exact tests. At all recovery phaes, upregulated phosphorylation (peptide fold-change ≥ 2, FDR ≤ 0.05) shows a similar overrepresentation of amino acids. A significant overrepresentation of proline (P) and glutamine (Q) on position +1 were displayed (**Figure 32**). However, down-regulated phosphorylation shows a different pattern at later time points compared to 0.5h recovery post UV stress, with a significant overrepresentation of proline (P) on position +1 within a S/T-P motif (**Figure 32**).

In order to better understand the dynamic of kinases activity, we performed the Kinase-Substrate Enrichment analysis (KSEA), which calculates Z-scores based on collective phosphorylation changes [375]. Kinase prediction by KSEA revealed that kinases ATM, ATR, and PRKDC (DNA-PK) were strongly activated at all time points, leading to an overrepresentation of the indicative S/T-Q motif and the activity of the downstream checkpoint kinases CHK1 and CHK2 (**Figure 31**C) [377]. Complementary to this, the most substantial reduction in kinase activity was reported for the cyclin-dependent kinases CDK1 and CDK2 in all conditions. A similarly substantial reduction in the activity of the kinases CDK5, NEK2 and AURKA/B, and the partial recovery after 18 hours as we observed are expected (**Figure 31**C), as it is also indicated in the linear-motif analysis for down-regulated phosphorylation sites (**Figure 31**D). Interestingly, MAPK is rapidly activated, with its activity decreasing over time. Especially the whole MAPK cascade is predicted to be activated, including MAPK8, MAPK9 and MAPK11. Their activity decreases to the average level at later time points, at 6h and 18h post UV irradiation, as we also observed in western blotting (**Figure 31**C, E). This pattern was also shared with predicted kinase activation, including RPS6KA1, RPS6KB1/2, CAMK2D, PKD1, PRKD1, PRKG1 and PRKCE.

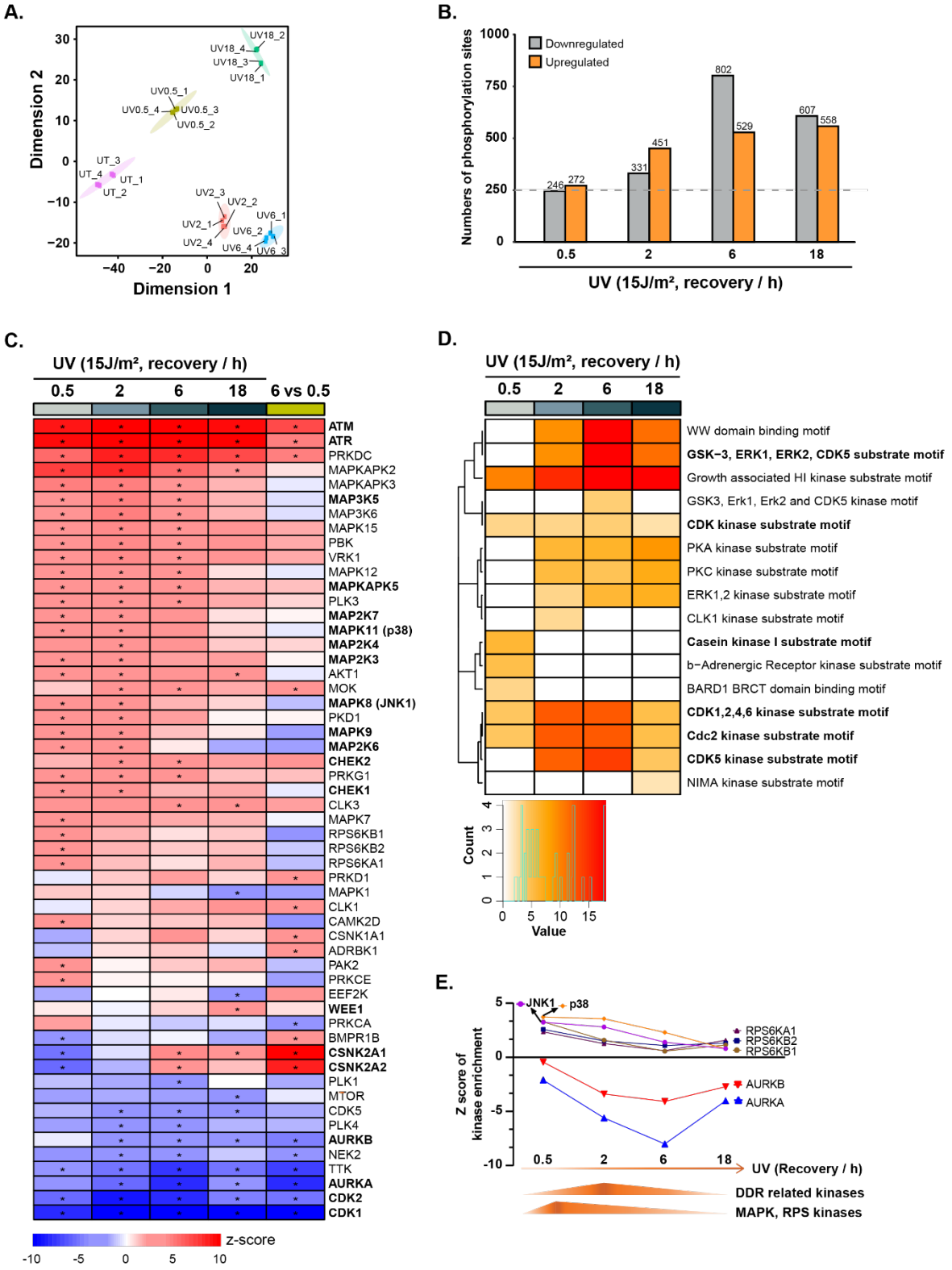

Figure 31: Different kinases are responding to UV stress at different time points. Phosphorylation modified peptides were filtered for missing values. Additionally, site numbers are reported (phosphorylation sites ≥ 0.75 localization probability). **A.** UMAP analysis of biological replicates for quantitative mass spectrometric analysis of phosphorylation sites after different recovery time points from UV light-induced DNA damage in HaCaT cells. **B.** The number of significantly regulated phosphorylation sites for each time point is depicted in the bar plots (n= 4; FDR ≤ 0.05; peptide fold-change ≥ 2). **C.** Kinase prediction by KSEA. The comparison between 6h and 0.5h is named as "6 vs 0.5". (Z-score ≥ 1.5, p-value ≤ 0.01 are indicated with an asterisk). **D.** Linear motif analysis for down-regulated phosphorylation sites. **E.** Line plot shows the dynamic changes in selected kinases (n= 4). Experiment was done by me and data analysis was done by Matthias Ostersmaier.

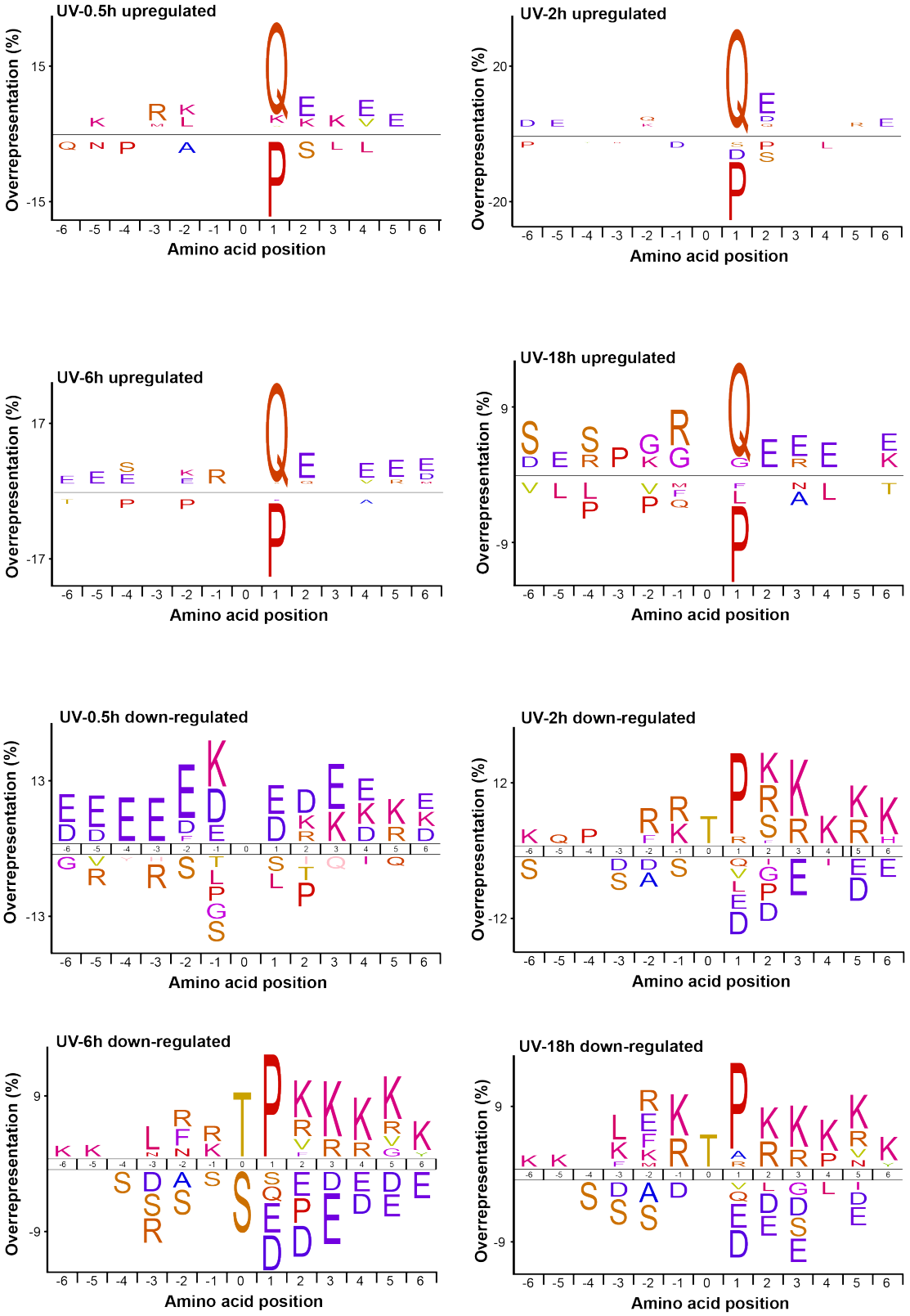

Figure 32: Sequence motif analysis of treatment-specific phosphorylation sites by IceLogo. Sequence windows of significantly down- or upregulated phosphorylation sites were compared to all quantified phosphorylation sites. Overrepresentation of amino acids at specific conditions is displayed in percent (n= 4 FDR ≤ 0.05; peptide fold-change ≥ 2). The experiment was done by me, and data analysis was done by Matthias Ostersmaier.

Notably, the complementary pattern of kinase activation for protein kinase CK2 (isoforms CSNK2A1 or CSNK2A2) is predicted by KSEA (**Figure 31**C). The activation of CK2 was only

identified for later time points but not for the early time point where MAPK shows high levels of activity (**Figure 31**C). This potentially correlates with cell cycle progression in response to UV irradiation. Moreover, the decreased activity of mTOR is observed at 18h (**Figure 31**C).

KSEA analysis allows us to investigate the known and potentially novel substrates of specific kinases. In the well-studied ATR/ATM cascade, we found some well-known phosphorylation sites to be regulated after UV irradiation, such as BRCA1 (pS1466) and exonuclease 1 (EXO1) (pS714) [584], [585], [280], [586]. In addition, many new phosphorylation sites are regulated; for example, sites on NER initiation factor XPC, structural maintenance of chromosomes 3 (SMC3), and RNA binding proteins MATR3 were quantified (**Figure 33**B). These events potentially expand the role of ATR/ATM in response to DNA damage. For the MAPK family, JNK1 kinase substrates are mainly transcription factors, such as ATF2 (pT69/T51) and JUN (pS63). AURKB regulates the phosphorylation of TPX2 (pS121 and pS125) in response to DNA damage [587], [588], [589]. Notably, UV light induced the reduction of mTOR activity which phosphorylates translation initiation factor EIF4EBP1 (pT46, pT70 and pS65) (**Figure 33**A).

By applying GO, KEGG and REACTOME GSEA analysis, we found that the terms "transcription regulation" and "mRNA processing" were enriched among proteins with upregulated phosphorylation sites at early recovery time post UV irradiation (**Figure 34**). However, "cell cycle", "chromosome organization and segregation," and "nuclear division" terms were enriched among proteins with down-regulated phosphorylation sites at 18 hours (**Figure 35**). This emphasizes the cell cycle arrest at 18h recovery after UV irradiation. Many NER and translation initiation factors are also dephosphorylated (**Figure 33**C). Downregulated phosphorylation sites also emphasize the dephosphorylation events that play a crucial role in response to UV-induced DNA damage (**Figure 35**). When comparing 6h to 0.5h for the up-regulated phosphorylation sites on proteins, we observed a significant increase of enrichment of "alternative splicing" and "5'-RNA processing" terms, highlighting the regulation of RNA metabolism in response to UV-induced DNA damage, especially during transcription restart at 6h (**Figure 37**A). As we also observed, the phosphorylation of alternative polyadenylation factors increased at 6h comparing 0.5h (**Figure 37**B). We assessed the alternative polyadenylation (APA) in response to UV stress by using a dual fluorescence reporter [318], [590]. In this construct, both GFP and dsRed genes are expressed in a bicistronic mRNA and both can be translated (GFP translation is driven by an internal ribosomal entry site (IRES)). One PAS is inserted between the two GFP and dsRed, and another PAS is behind GFP (**Figure 37**C). By monitoring the intensity of dsRed, we could evaluate the polyadenylation efficiency and/or relative usage of two APA sites. Therefore, the dsRed/GFP ratio provides a quantitative measurement of APA activity (**Figure 37**C). We observed significantly increased APA after UV irradiation that peaked at 6h. As a positive control, the deficiency of CPSF6 induced a higher APA activity (**Figure 37**C) [591].

In summary, these results show that UV irradiation immediately induced transcription inhibition and the phosphorylation of transcription factors. After a short recovery, while transcription was

recovering and restarting, mRNA processing factors were found to be more phosphorylated by MAPKs to regulate the RNA metabolism, such as alternative splicing and RNA 5'-end processing. After a longer recovery, when transcription was completely recovered, proteins involved in the cell cycle and chromosome organization were observed to be less phosphorylated due to the decreased activity of the CDKs and AURKA/B and the cell cycle arrest induced by UV irradiation.

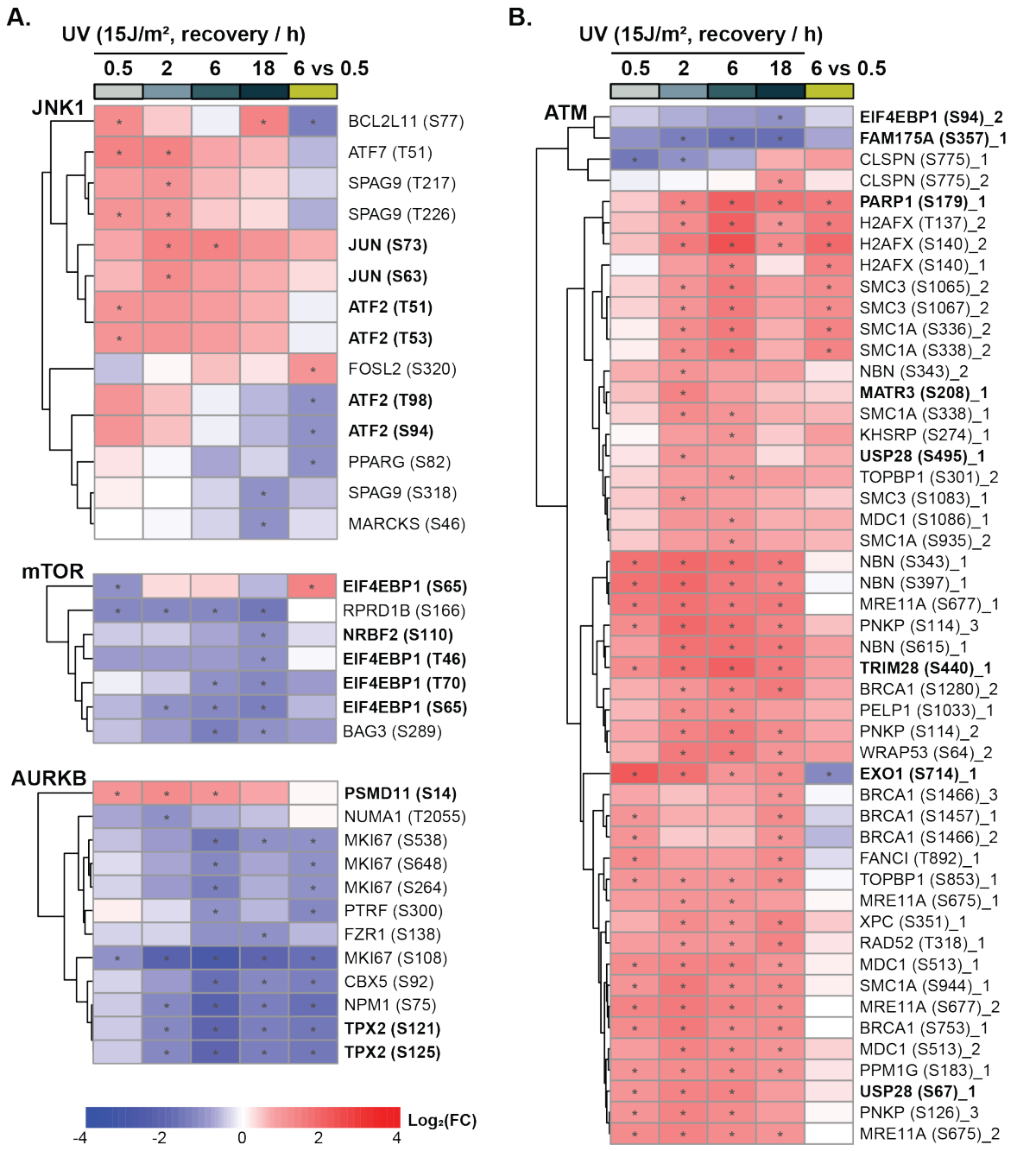

Figure 33: Heatmaps of proteins with regulated phosphorylation sites which are predicted to be phosphorylated by different kinases in response to UV light. A. Phosphorylation sites are predicted in kinases JNK1, mTOR and AURKB cascades. **B.** Phosphorylation sites are predicted in the kinase ATM cascade. The comparison between 6h and 0.5h is named as "6 vs 0.5". (FDR ≤ 0.05 is indicated with an asterisk, peptide fold-change ≥ 2).

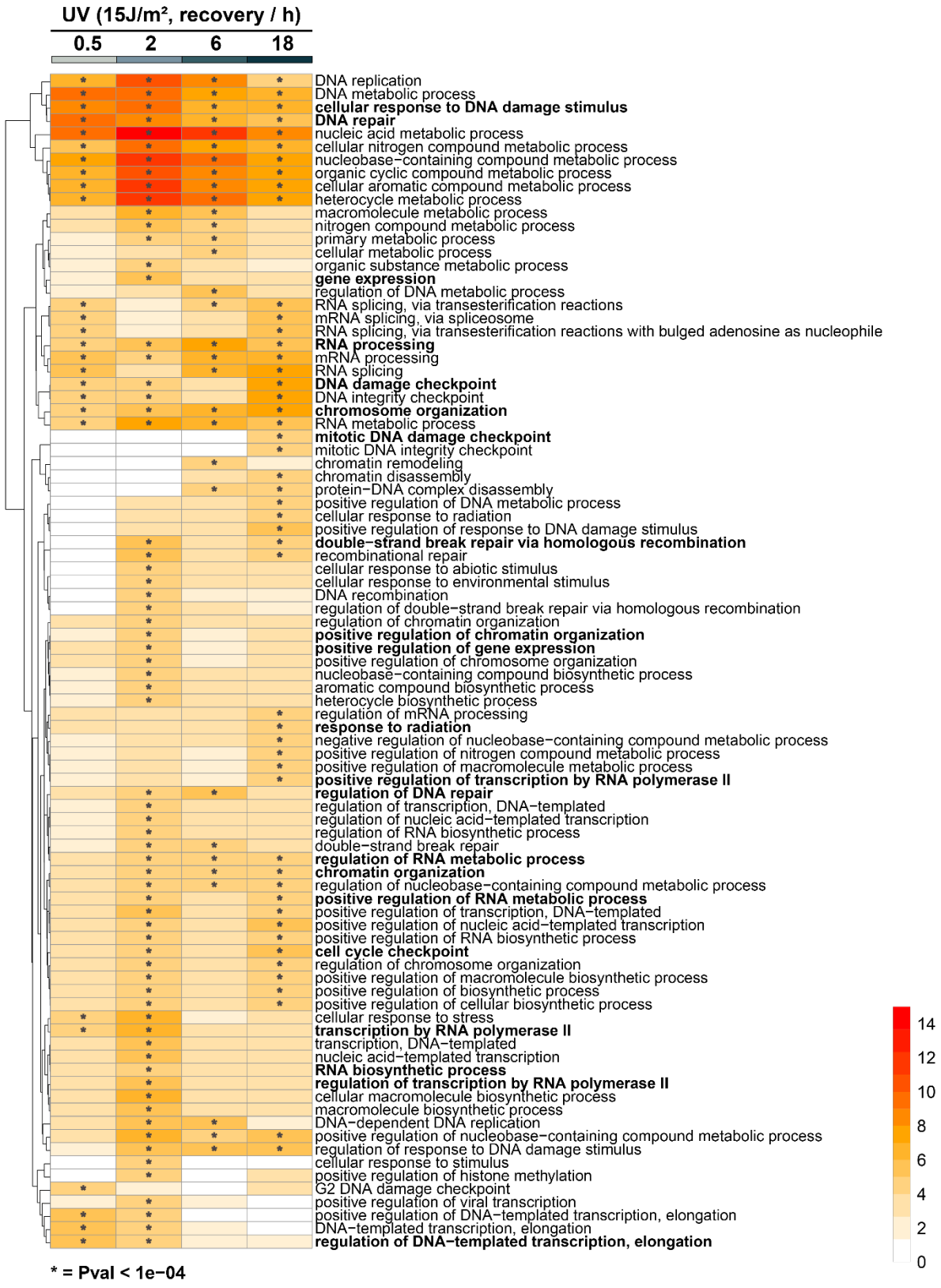

Figure 34: GO terms enrichment analysis was carried out for proteins with upregulated phosphorylation sites after UV irradiation and the highest scoring GO terms for the biological process are displayed. The color panel indicates the significant score of enrichment. (n=4; peptide fold-change ≥ 2; p-value ≤ $1x10^{-4}$ are indicated with an asterisk).

Figure 35: GO terms enrichment analysis was carried out for proteins with downregulated phosphorylation sites and the highest scoring GO terms for the biological process are displayed. The color panel indicates the significant score of enrichment. (n=4; peptide fold-change ≤ 2; p-value ≤ $1x10^{-4}$ are indicated with an asterisk).

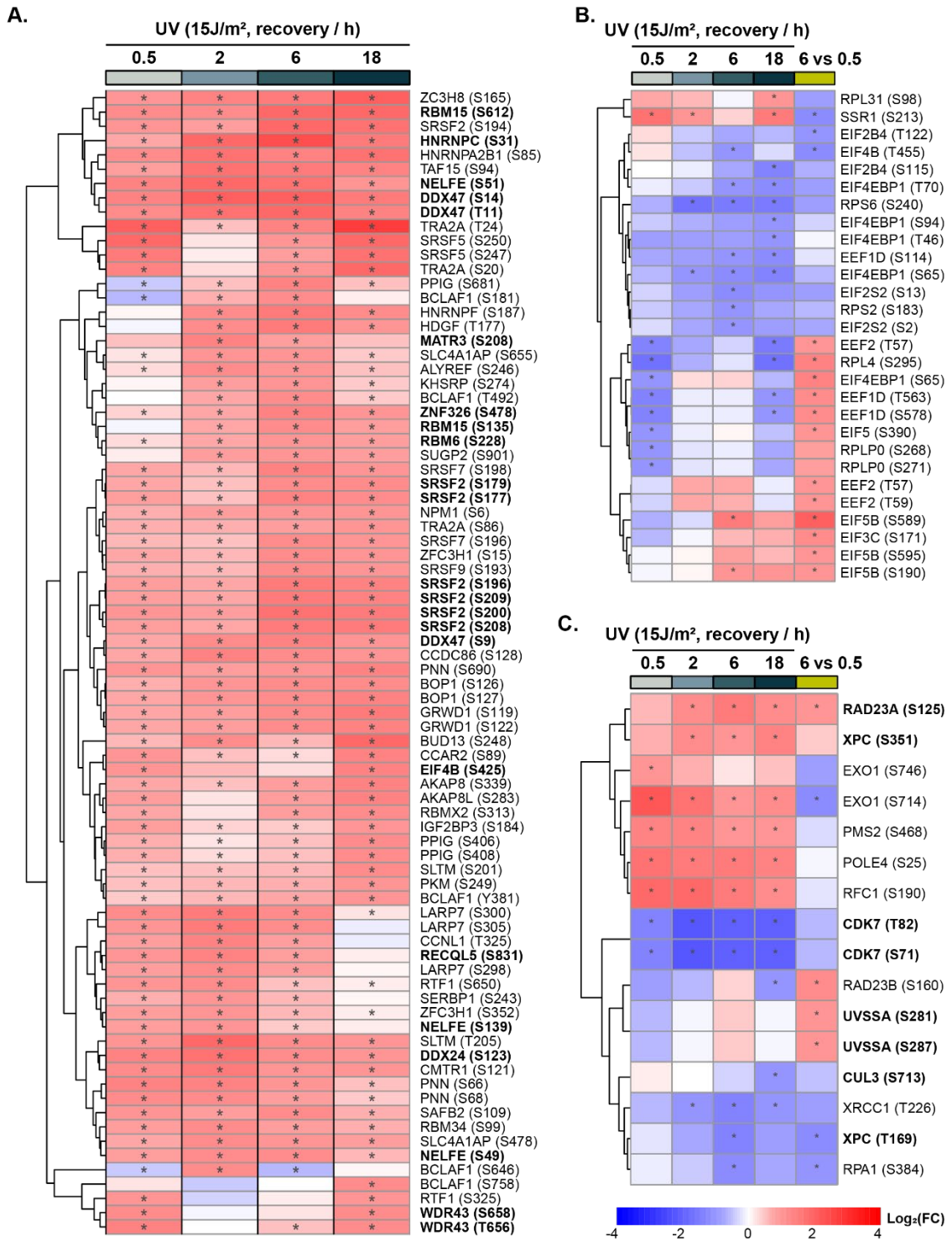

Figure 36: Heatmaps of protein groups in different pathways with regulated phosphorylation sites. A. RNA binding proteins **B**. Translation factors **C.** NER factors (n=4; FDR ≤ 0.05 is indicated with an asterisk, peptide fold-change ≥ 2)

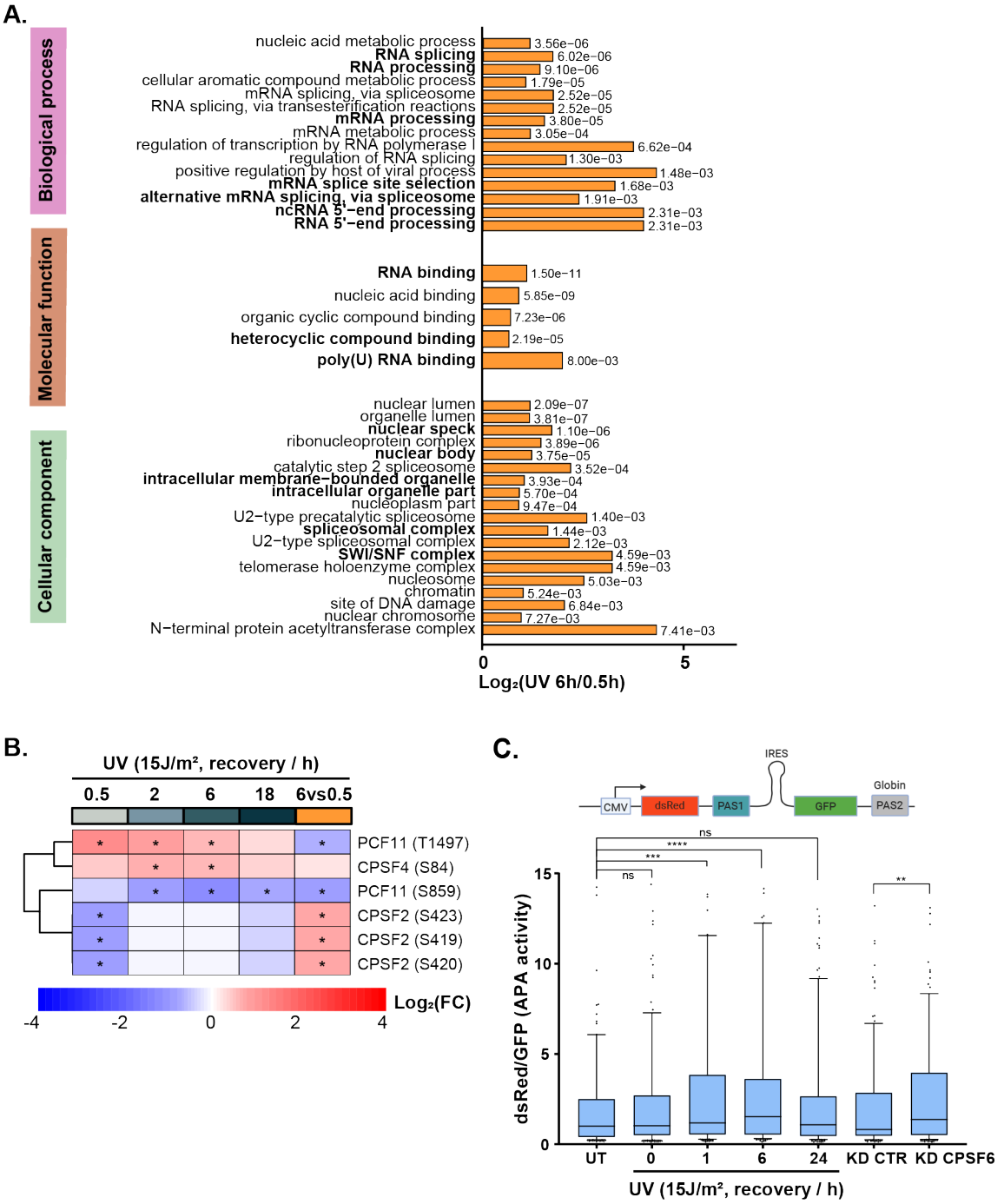

Figure 37: UV stress induced increased alternative polyadenylation activity. A. GO terms enrichment analysis was carried out for proteins with upregulated phosphorylation sites after the comparison of 6h to 0.5h, and the highest-scoring GO terms for the biological process are displayed. The terms with the highest scores for cellular compartment, biological process and molecular function are shown here. This analysis was done by Matthias Ostersmaier. **B. Heatmap of alternative polyadenylation factors with regulated phosphorylation sites** (n=4; FDR ≤ 0.05 is indicated with an asterisk, peptide fold-change ≥ 2). **C. The dual fluorescence reporter construct was used to monitor the usage of the two different PAS (upper). Quantification analysis of APA activity per nucleus during the response to UV irradiation (lower).** U2OS cells were transfected with CPSF6 siRNA or CTR siRNA. Then cells were transfected with a vector containing the dual fluorescence reporter, dsRed and GFP. After 48 hours post-transfection, cells were irradiated with UV light (15J/m^2) and fixed, followed by imaging analysis. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). A 20X water immersion objective (numerical aperture (NA) 1.1) was used. The nuclear intensity was analyzed by automated "Harmony High Content Imaging and Analysis Software" (PerkinElmer). Cells in the periphery of the image were excluded from further analysis. One-way ANOVA was used to calculate the significance (n=3; 300 cells were measured per replicate, ****p-value < 0.0001).

2.1.5 Integration and summary of the proteomics analysis

UV irradiation causes conformational changes in the DNA structure, which subsequently induces transcription blocking in cells. Thus, DDR is triggered in cells, which results in changes in transcription level, RNA modifications, protein level, protein localization, posttranslational modifications, protein-protein interactions, and proteins' degradation. While studying these changes separately can already provide significant insight into the cellular response to DNA damage, it is challenging to highlight the most meaningful factors with this long list of proteins changes.

To illustrate this point, we combined information from our phosphorylation, the global proteome and the chromatin datasets with the transcriptome upon UV treatment in Human MRC5VA cells, the phosphoproteome, the chromatin-bound proteins profile with MG132 treatment, the ubiquitination and RNA Pol II-IP dataset (http://www.biologic-db.org) from the Svejstrup lab upon UV treatment in HEK293T cells [592], [593].

First, we compared the datasets that we generated for different recovery time points (e.g., phosphoproteome, global proteome and chromatin-bound proteome) (**Figure 38** and **Figure 39**). We did not observe much overlap for the datasets after 0.5h recovery from UV irradiation (**Figure 38**A). Surprisingly, we found more overlaps between the phosphorylation dataset at 6h and 18h of proteome or chromatin-bound proteome datasets (**Figure 38**B). Some proteins were phosphorylated at early time points and then recruited to chromatin after 6h or 18h post UV irradiation. They are mainly involved in chromosome organization and cell cycle, for example, kinesin family member KIF18B, cyclin-A (CCNA2) (**Figure 39**A, **Figure 38**A and B). Notably, the protein level of CCNA2 and KIF18B also increased upon UV induction. This suggests that the increased level of the proteins likely results in a more phosphorylated form and an increase in chromatin recruitment.

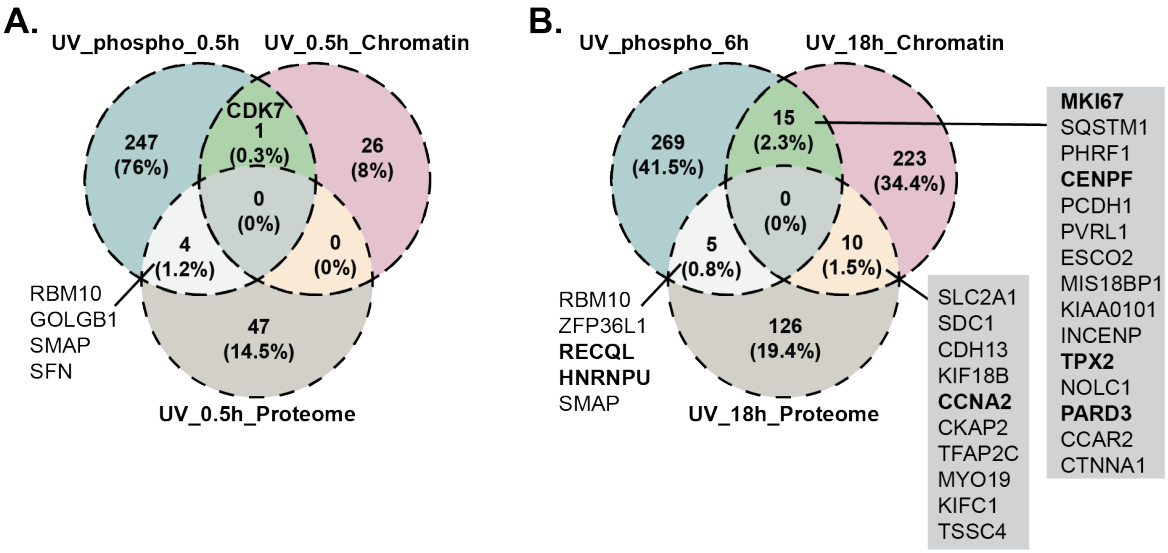

Figure 38: Proteins with regulated phosphorylation sites overlap with regulated proteome and chromatin-bound proteome.

Interestingly, DNA repair factor TPX2 was dephosphorylated due to the decreased AURKB activity and recruited to chromatin at after 18 hours (**Figure 39**A and **Figure 38**B). Proliferation marker, MKI67 coating chromosomes during mitosis and nucleolar and coiled-body phosphoprotein 1 (NOLC1), are recruited to chromatin, but less phosphorylated at 18h recovery post UV irradiation. NOLC1 has been reported to connect to RNA polymerase I to regulate rRNA processing and cell organization (**Figure 39**A and **Figure 38**B) [594].

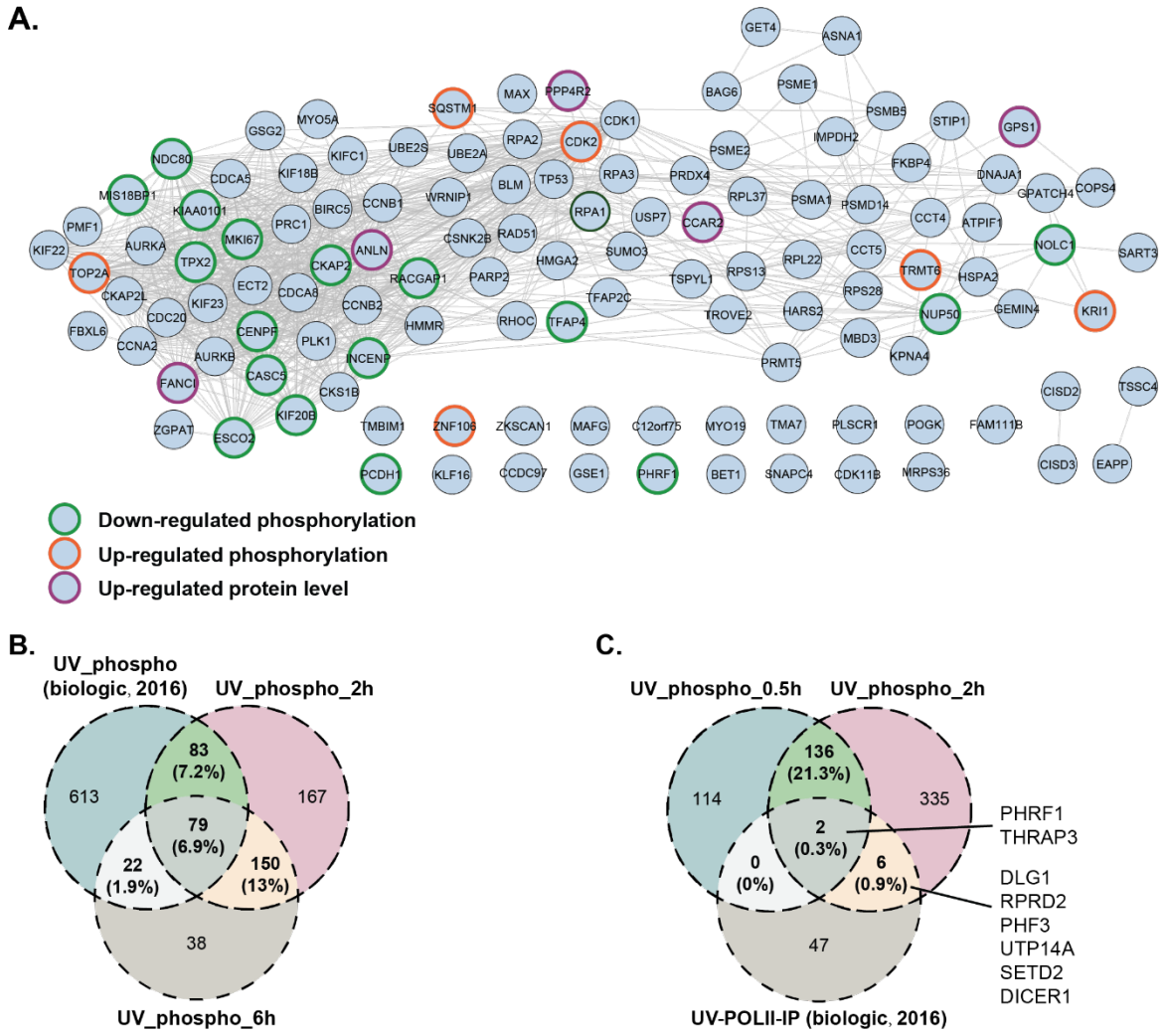

Figure 39: Overlaps of our proteome, chromatin-proteome and phosphorylation datasets (15 J/m2, 0.5, 2, 6, 18h recovery, HaCaT cell) with the dataset (the phosphoproteome and RNA Pol II-IP) from the Svejstrup lab upon UV treatment in HEK293T cells (15 J/m2, 3h recovery) (http://www.biologic-db.org) [592]. **A. Network (Confidence = 0.7) for proteins recruited to chromatin at 18h post UV irradiation.** The red circle labeled the proteins with upregulated phosphorylation sites. The green circle labeled the proteins with downregulated phosphorylation sites. Asterisk (*) indicates upregulated protein levels. **B.** The overlap of our phosphoproteome dataset at 2h and 6h with the phosphorylation dataset, named as "UV_phospho (biologic, 2016)", and with the RNA Pol II interactome (UV-POLII-IP) (**C**) from the Svejstrup lab**.**

In contrast, scaffolding protein PARD3, catenin CTNNA1 and centromere protein F (CENPF) were dephosphorylated and disassociated from chromatin at 18 hours (**Figure 39**B). However, their protein levels did not change (**Figure 69**). Therefore, it seems that the phosphorylation states of these proteins regulate their re-localization. Remarkably, mRNA splicing factor RBM10 was highly phosphorylated at 18h recovery after UV light treatment and the protein level was also increased.

However, its transcription level and ubiquitylation did not change upon UV stress (**Figure 24**D and **Figure 69**). The phosphorylation level of RECQL (ATP-dependent DNA helicase Q1) was decreased at 6h and 18h. However, the protein level increased at 18h. The Svejstrup lab has reported the transcription level of RECQL increasing after 24 hour recovery post UV irradiation (15 J/m^2), but not after 8 hour recovery (**Figure 69**). Thus, clearly the increased protein level of RECQL (at 18h in our dataset) results from this reported increase in transcription observed after 24h recovery. Ubiquitylation is a trigger to degrade proteins [595]. We compared our global proteome data sets (proteome data) with the ubiquitylation data set (**Figure 69**) from the Svejstrup lab. We observed that PCNA, DDB2 and RPS20 protein levels were decreased with increased ubiquitylation levels. This corresponds to previous research that shows UV stress induces ubiquitylation of PCNA and DDB2 [596]–[598].

In addition, comparing our phosphoproteome dataset (UV_phospho_2h and UV_phospho_6h) with the phosphoproteome dataset of the Svejstrup lab (UV_phospho (biologic, 2016)), proteins with regulated phosphorylation sites at 2h have around 10% overlap, and a higher overlap is observed for the 6h dataset (**Figure 39**C). Notably, some proteins are phosphorylated and have UV-induced interaction with RNA pol II after 3 hours post UV treatment (**Figure 39**D), such as PHD finger protein 3, PHF3. PHF3 is reported to accumulate at DNA lesions and promote G2 checkpoint recovery [599]. Recently, it has been shown that the PHF3 SPOC domain binds to the RNA Pol II Ser2 to regulate transcription and mRNA stability, thus playing a key role in proper neuronal differentiation [600].

Taken together, the results presented in this chapter show that UV irradiation induces not only the activation of some well-known kinases (such as ATM and ATR), but also the activation of transcription-related kinases, MAPK and RPSKs. In addition, UV irradiation also induces the deactivation of cell cycle-related kinases, for example, CDKs and AURKs (**Figure 31** and **Figure 40**). The dynamic regulation of phosphorylation states results in changing protein levels. Subsequently, it also influences the proteins' relocalization, particularly regarding the chromatin-bound proteins. All these cooperative spatial and temporal dynamic changes regulate many cellular processes to assist cells in recovering from the transcription block induced by UV irradiation. At an early recovery time (0.5h), when transcription is shut down, DNA repair and transcription processes are highly regulated. While transcription recovers and restarts (2h and 6h); gene expression, RNA processing including mRNA and ncRNA process, and translation initiation are regulated by the changes in the associated proteins' levels and phosphorylation. All these changes result in the cell cycle arrest at a later time (18h), when transcription is completely recovered. Of note, there is a more inflammatory response at 18h when transcription blocking lesions are obliterated completely (**Figure 40**).

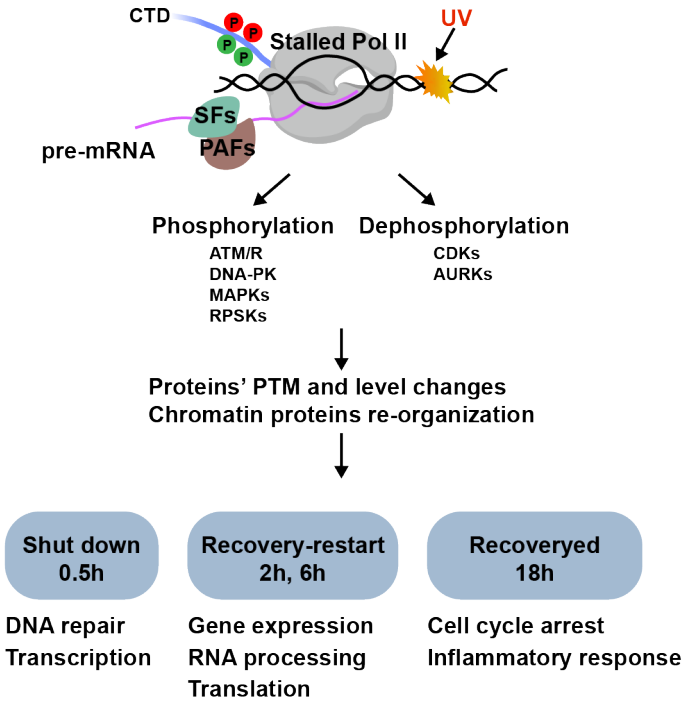

Figure 40: A simplified results scheme summarized based on our quantitative mass spectrometry screening data.

For our interests, looking into the transcription-related kinase MAPK pathway, we focused on identifying the substrates of one of its members, the JNK1 pathway, as described in the next chapter.

2.2 Identification of the JNK pathway cascade by combining chemical JNK inhibitor and quantitative phosphoproteome analysis

2.2.1 JNK1 is activated in response to UV stress

To identify substrates and functions of the JNK signaling cascade, we performed phosphoproteome analysis with the chemical inhibition of JNK1 (JNK1-IN-8), which is the first irreversible JNK inhibitor for JNK1, JNK2 and JNK3 [601], [602].

As we monitored above, JNK1 activation is mainly at early time recovery from UV irradiation. Therefore, we decided to investigate the cascade after short recovery post UV irradiation to identify most substrates of JNK signaling, where we observed the highest activity of JNK. We observed its activity after different recovery times ranging from 30 minutes to 24 hours. We could observe strong activation of JNK after 30 minutes and the activation showed a peak between 30 and 60 minutes post UV irradiation. After 60 minutes, the phosphorylation decreased (**Figure 41**A and **Figure 73**). So we decided to study the JNK activity at 1h recovery post UV irradiation.

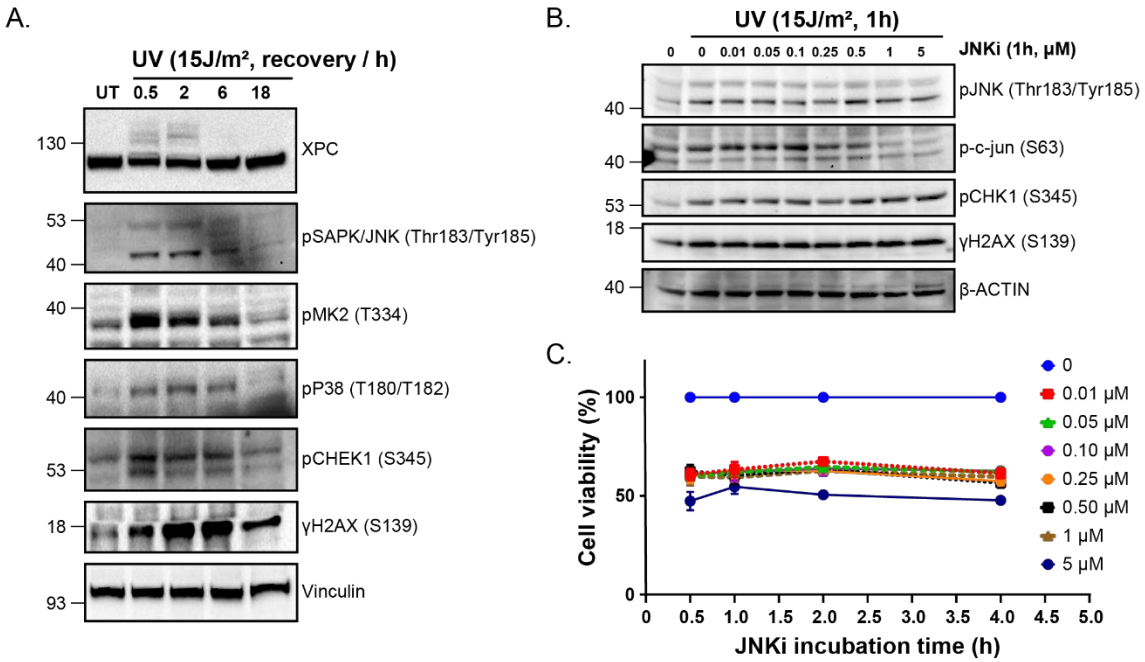

Figure 41: The JNK signaling is activated in response to UV light and determination of JNK1 inhibitor concentration based on target phosphorylation and cell viability. A. HaCaT cells were treated with UV light (15 J/m^2) and left to recover for different times. The whole-cell lysate is analyzed with western blotting. **B** and **C**. HaCaT cells were pre-treated with indicated JNK1 inhibitor (JNKi, JNK1-IN-8) concentration or mock treatment for indicated incubation time points and then irradiated by UV light (15 J/m^2), followed by 1h recovery post UV irradiation. Then cell viability was evaluated using the CellTiter-Blue viability assay (n=3) (**C**). Whole-cell lysates were analyzed with western blotting (**B**).

Quantitative methods offer an effective way for determining the concentration of JNK1 inhibitor. We pre-treated HaCaT cells with different concentrations of JNK1 inhibitor before UV irradiation, and then cell viability was measured by Cell Titer Blue assay to determine the LC50. We found 0.01µM of JNK1 inhibitor already induced 50% cell death (**Figure 41**C), however it did not successfully inhibit the phosphorylation of its well-known substrate c-JUN (**Figure 41**B). Since a higher concentration of JNK1 inhibitor did not induce more cell death than 50% observed with 0.01µM (**Figure**

41C), and based on our western blot analysis, we decided to use 1µM and 5µM of JNK1 inhibitor. The phosphorylation of c-JUN is inhibited at these concentrations.

2.2.2 JNK activity is essential during recovery from UV stress

In order to assess the importance of JNK1 activity in the cellular response to UV light-induced DNA damage, we first checked the removal of UV-induced CPD dimers with or without JNK1 activity after recovery from UV irradiation in HaCaT cells. We observed that CPD dimers were completely removed around 48 to 72 hours post UV irradiation. In contrast, CPD foci still persisted after 72 hours

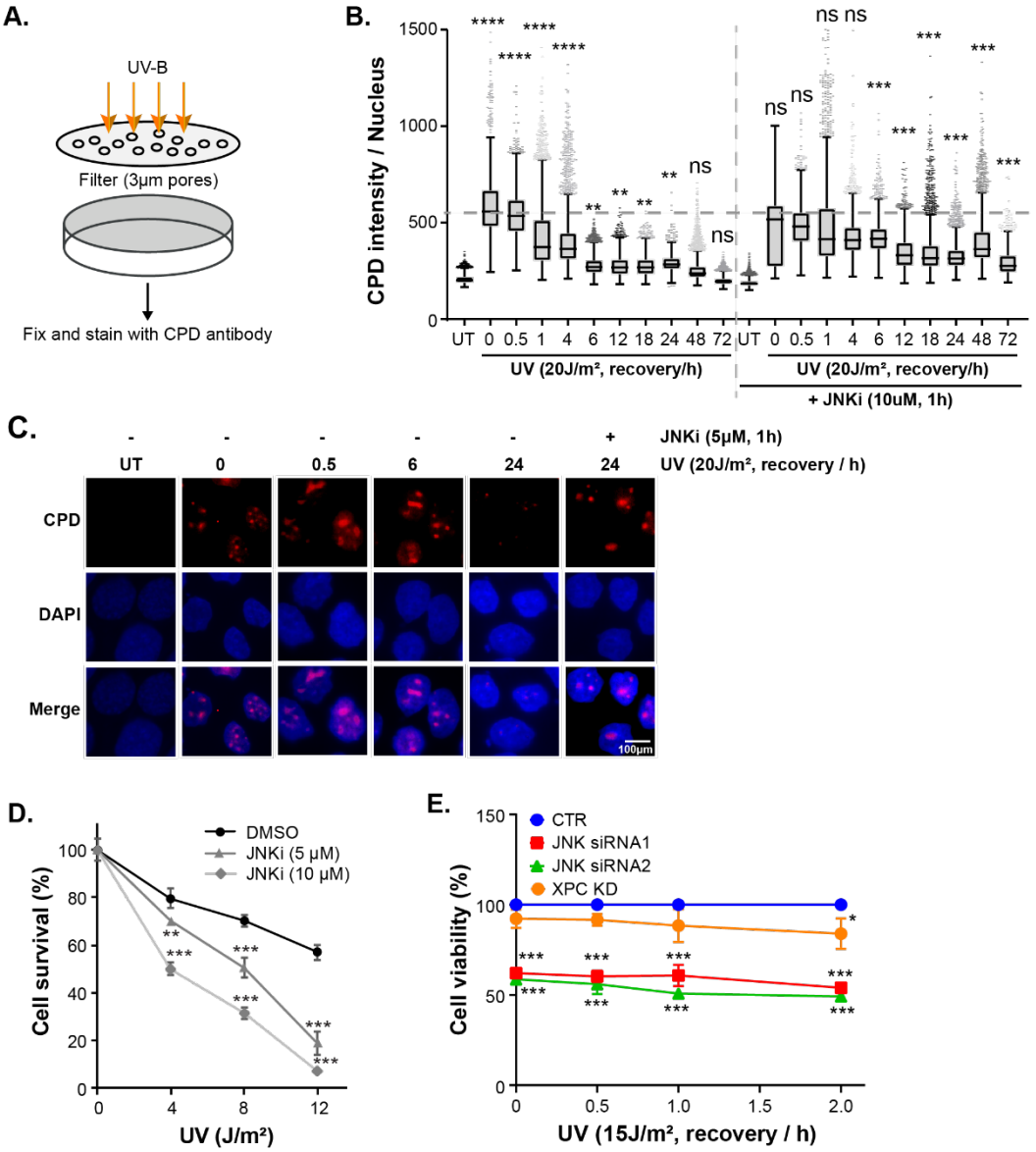

Figure 42: JNK1 is essential for the cellular response to UV irradiation. A. Experimental flow for the micropore UV irradiation using a combination of a microfilter mask and UV light. HaCaT cells were pre-treated with JNK1 inhibitor (5µM) or mock-treated for 1h and then irradiated by UV light (15 J/m^2). Cells were left to recover for the indicated time points post UV irradiation. Cells were fixed and stained with the specific antibody for CPD dimers. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). Quantification of the CPD intensity in the nucleus during recovery from UV treatment is shown in a boxplot (**B**), and representative images are shown in (**C**). One way ANOVA was used to calculate the significance (****p-value < 0.0001). **D.** HaCaT cells were pre-treated with JNK1 inhibitor (5µM and 10µM) or mock-treated for 1h and then irradiated by UV light (4, 8 and 12 J/m^2), then left to grow for 10 days. Cells are stained with Crystal violet solution. The error bars show the mean and the SD of results obtained in three technical replicates. A two-sided Student's t-test was used to calculate the significance (n=3; ***p-value < 0.0001). **E.** HaCaT cells were transfected with two different JNK1 siRNAs or a control (CTR) siRNA for 72h and then irradiated by UV light (15 J/m^2), then left to recover for the indicated time points post UV irradiation. Then cell viability was evaluated using the CellTiter-Blue viability assay. One way ANOVA was used to calculate the significance (n=3; ****p-value < 0.0001).

when JNK1 activity was inhibited (**Figure 42**B and C). This is also reflected in cell survival and cell viability. JNK1 inhibition decreased the ability of cells to form colonies. The higher concentration of JNK1 inhibitor resulted in a further decreased ability to form colonies (**Figure 42**D). In line with this, cell titer blue assay also showed reduced cell viability with around 50% decrease after JNK1 depletion with different siRNAs. However, XPC knockdown did not reduce cell viability (**Figure 42**E). These results show that JNK1 activity plays a crucial role in both the short-term response and the long-term response to UV irradiation.

2.2.3 Quantitative proteomics identifies UV stress-induced, JNK-dependent signaling

We carried out multiple phosphoproteome analyses using the specific JNK inhibitor combined with either TMT-labeling-based or SILAC-labeling-based quantification.

In the TMT-labeling phosphoproteomics analysis, HaCaT cells were pre-treated with 1μM JNK inhibitor or with DMSO (vehicle) for 1h before irradiating cells with UV (15 J/m^2), after which cells were left to recover for 30min. Then we harvested cells and processed the lysates following the TMT-phosphoproteomics workflow described above (**Figure 43**A). Pearson analysis showed a nice correlation between replicates and clear clustering of different treatments (**Figure 43**B). We quantified 31,900 phosphorylation sites. After applying all quality filters as explained in the first chapter, 21,476 sites (67.3%) were used for the following statistical analyses (**Figure 43**C). Quantitative and statistical analysis revealed that 467 phospho-sites (2.2%) on 284 proteins were significantly upregulated and 202 phosphorylation sites (0.9%) on 134 proteins were downregulated upon UV stress. When JNK1 activity is inhibited, 87 phosphorylation sites (18.6%) on 67 proteins are downregulated after UV stress (**Figure 43**C). JNK1 is a proline-directed serine/threonine-protein kinase. In line with this, sequence motif analysis of peptides phosphorylated in a JNK-dependent manner revealed a significant overrepresentation of proline (P) on position +1 within an S/T-P motif (**Figure 43**D).

To better understand the biological processes associated with the proteins with JNK-dependent phosphorylation sites upon UV induction, GO terms analysis was carried out. Most proteins are RNA-binding proteins and fall into the “gene expression” and “mRNA processing” categories in the String database (**Figure 44**A and B).

Among these proteins, we found well-known JNK substrates, such as transcription factors ATF7 and c-Jun (**Figure 44** C) [603], [604]. Putative JNK1 substrates after UV stress are primarily RNA-binding proteins involved in gene expression and RNA processing (e.g., transcription regulator BRD4, translation factor EIF4H, splicing factor HNRNPA1 and SRSF2) and form a nice interaction network (**Figure 44**C). Some DNA repair factors were observed to get less phosphorylated after JNK inhibition, such as Flap endonuclease 1 (FEN1) and CDKN2A interacting protein (CDKN2AIP). Interestingly, Protein Phosphatase Methylesterase 1 (PPME1), DNA Methyltransferase 1 (DNMT1), Lysine Demethylase 2B/3A (KDM2B, KDM3A) and Ribosomal Protein S6 Kinase A4 (RPS6KA4) were less phosphorylated upon JNK inhibition (**Figure 44**C).

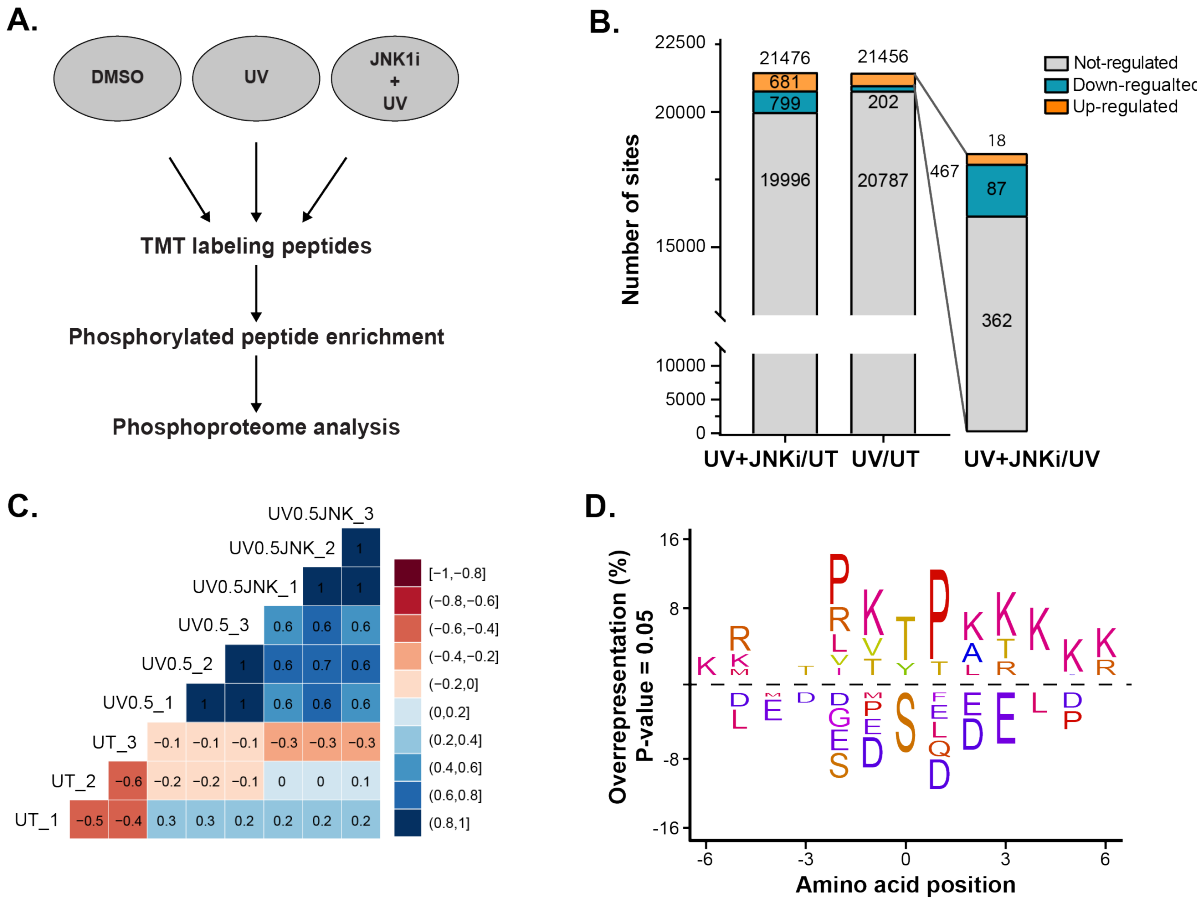

Figure 43: Phosphoproteomics analysis of JNK substrates upon UV stress in HaCaT cells based on the TMT-labeling MS. A. Experimental workflow for quantitative mass spectrometric analysis of phosphorylation sites after inhibition of c-Jun amino-terminal kinase 1 (JNK1i) (1µM) upon UV (15J/m^2) light-induced DNA damage in HaCaT cells with TMT-labeling. **B.** The bar graph shows the number of significantly up-, non-, and downregulated phosphorylation sites after JNK inhibition in TMT labeled phosphoproteomics analysis in HaCaT cells (n=4; fold change ≥ 1.5, FDR ≤ 0.05). **C.** Correlation of phosphoproteome with JNK1 inhibitor between replicates. For each experiment intensity, a ratio was calculated against the untreated intensity mean. Pearson correlations were calculated from the $\log_2$ transformed ratios. **D.** JNK phosphorylates proteins on a S/T-P motif after UV light. Sequence motif analysis of the JNK-dependent and UV-induced phosphorylation sites was carried out with IceLogo (2022). This analysis was done by Matthias Ostersmaier.

In the SILAC-based phosphoproteomics analysis, light-labeled U2OS cells were mock-treated, medium-labeled U2OS cells were treated with UV light (40 J/m^2, 1h recovery), and heavy-labeled U2OS cells were pre-treated with the specific 5µM JNK1 inhibitor followed by UV light irradiation (**Figure 45**A). This analysis revealed that 741 (4.22%) and 153 (1.72%) out of 17,546 phosphorylation sites were significantly upregulated and downregulated after UV light, respectively (**Figure 45**B). Notably, 264 phosphorylation sites (35.6%) significantly decreased after JNK inhibition (**Figure 45**B). Similar to phosphoproteome analysis, we also carried out sequence motif analysis for phosphorylated peptides in a JNK-dependent manner upon UV stress. It also revealed that JNK1 phosphorylates proteins on an S/T-P motif (**Figure 45**C).

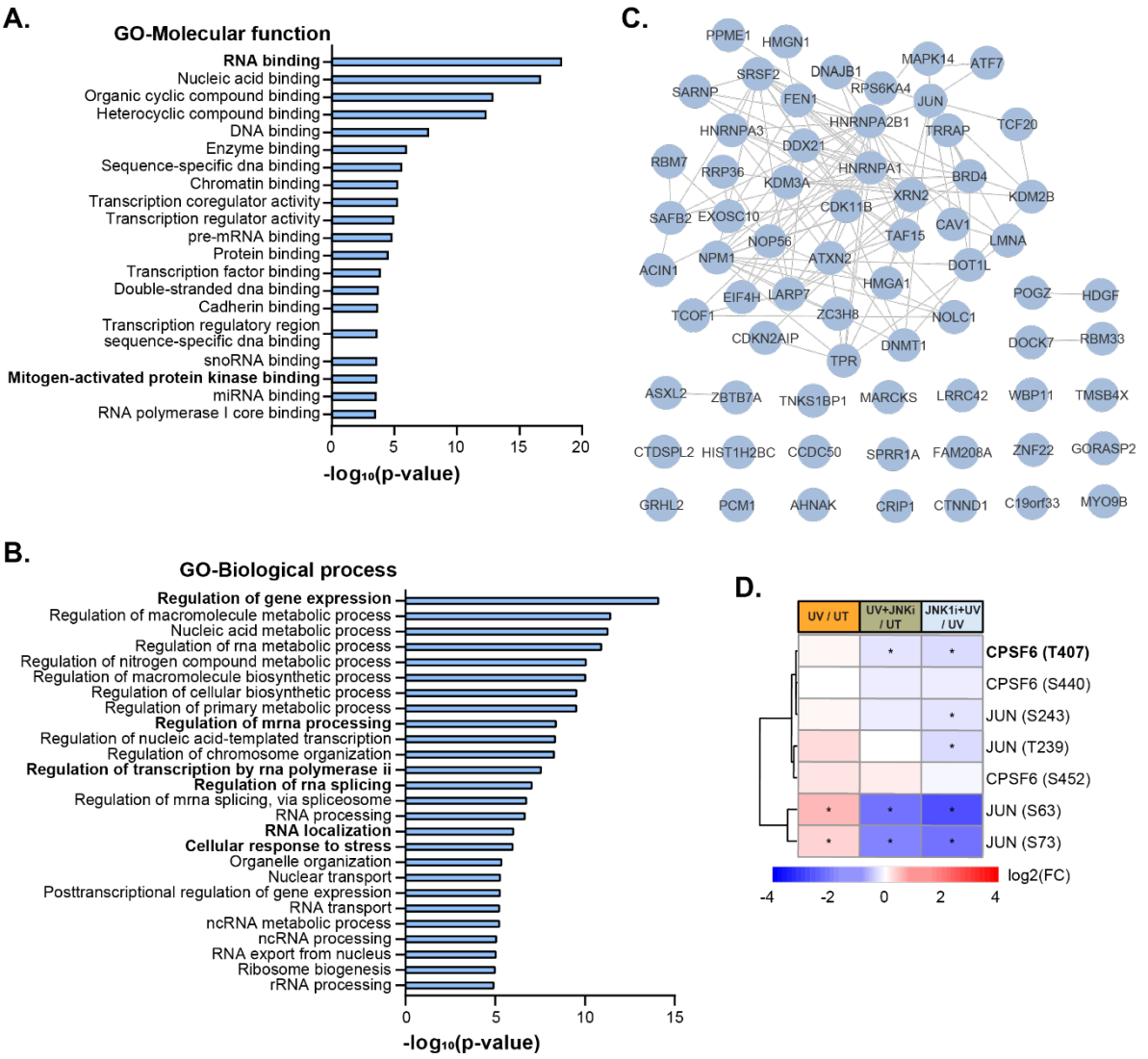

Figure 44: A and **B. GO terms enrichment analysis was carried out for proteins with with JNK1 activity-dependent phosphorylation sites in response to UV stress.** The highest-scoring GO terms for the molecular function and biological process are displayed in (**A**) or in (**B**). **C.** Network of proteins with JNK1-dependent phosphorylation sites in response to UV stress (Confidence = 0.7). **D.** Heatmap of JNK substrates in all conditions. (n=3; peptide fold-change ≥ 1.5; FDR ≤ 0.05 are indicated with an asterisk). This analysis was done by Matthias Ostersmaier.

To further determine the substrate of JNK1 for our subsequent studies, we combined our SILAC-JNKi phosphoproteomics data with chromatin landscape changes (at 2h post UV irradiation), which may be induced by the differential phosphorylation state of proteins. Within these 2 datasets, only one protein overlapped: the Cleavage factor Im (CFIm) complex component, CPSF6, with two phosphorylation sites on T404 and T407 is a putative JNK substrate that shows dissociation from chromatin at 2 hours post UV stress (**Figure 44**D, **Figure 45**D and **Figure 46**).

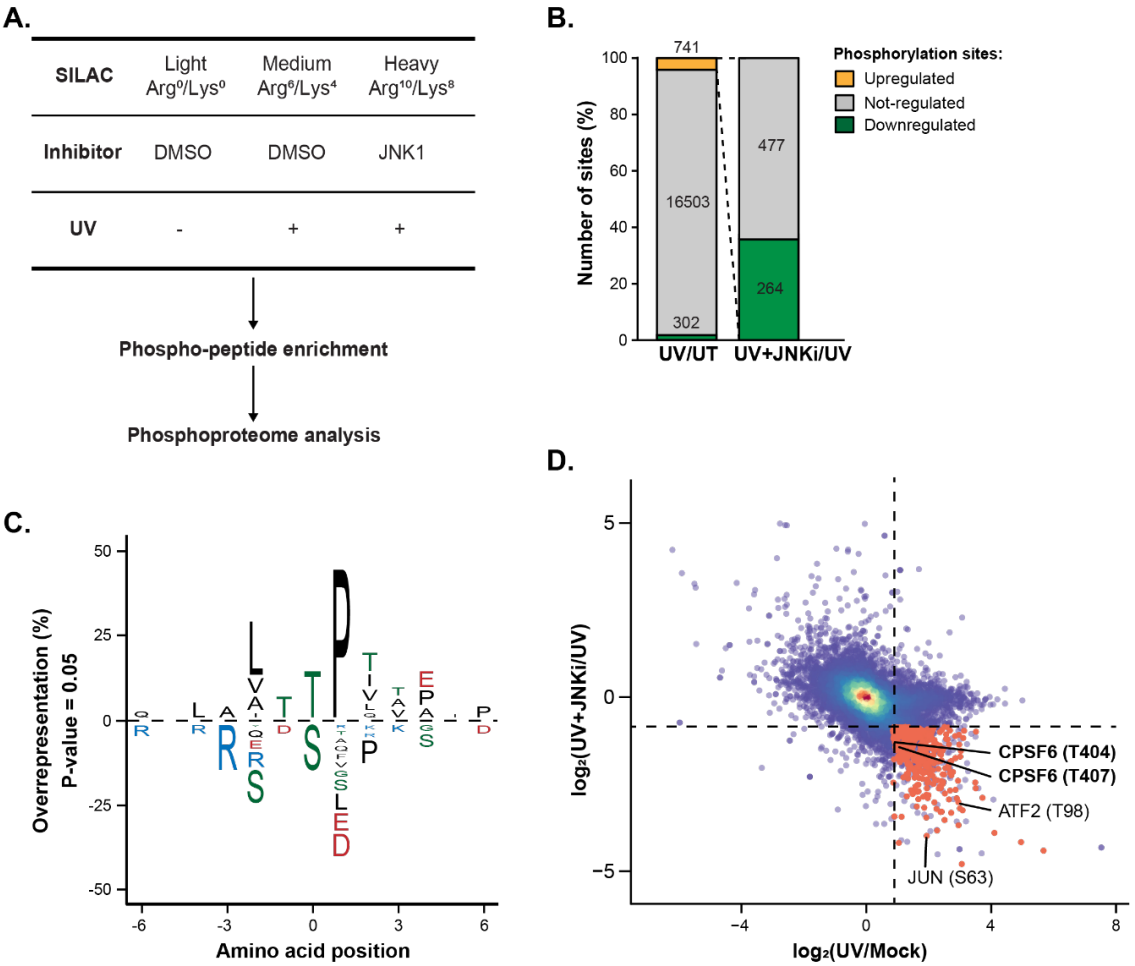

Figure 45: Phosphoproteomics analysis of JNK substrates upon UV stress in SILAC-U2OS cells. A. Experimental workflow for quantitative mass spectrometric analysis of phosphorylation sites after inhibition of c-Jun amino-terminal kinase 1 (JNK1i) (5µM) pre-treatment and UV (40J/m^2) irradiation in U2OS cells with SILAC labeling. **B.** The bar graph shows the number of significantly up-, non-, and downregulated phosphorylation sites after JNK inhibition in SILAC-labeled U2OS cells (n=5, fold change ≥ 2, FDR ≤ 0.05). **C.** JNK phosphorylates proteins on a S/T-P motif after UV light. Sequence motif analysis of the JNK-dependent and UV-induced phosphorylation sites was carried out with IceLogo (2018). **D.** Scatter plot shows the logarithmized SILAC ratios UV/Mock and UV +JNKi /UV+DMSO of quantified phosphorylation sites. UV light-induced, JNK-dependent phosphorylation sites are shown in red (n=5, fold change ≥ 2, FDR ≤ 0.05). The experiment was done by me, and the data analysis was done by Matthias Ostersmaier.

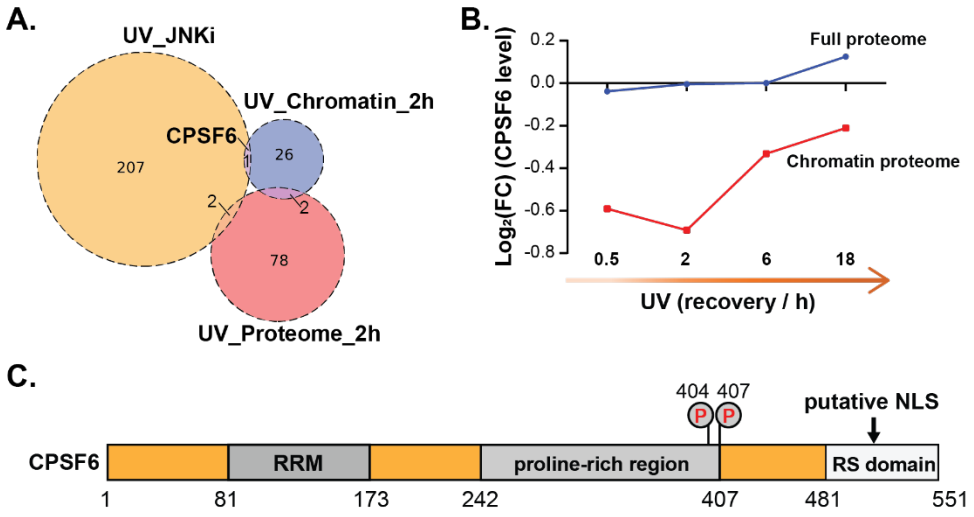

Figure 46: The overlap of JNKi phosphoproteomic data with the full proteome and the chromatin-proteome at 2h post UV stress (A) and CPSF6 protein domains scheme with indicated phosphorylation sites at T404 and T407 (C). RRM stands for RNA recognition motif. NLS stands for nuclear localization sequence. RS domain stands for arginine/serine repeat domain. **B. Line plot shows CPSF6 level dynamic in the full proteome and the chromatin-proteome from the first chapter.**

2.3 CPSF6 - a new putative substrate of the JNK pathway in response to UV irradiation

2.3.1 CPSF6 is proximal to JNK1 after UV irradiation

Since phosphorylation events require kinases to have direct contact with substrates, we carried out a proximity ligation assay (PLA), which can generate strong signals *in situ*, when two molecules exist at most within 40-nm of each other in cells [605]. This assay allowed the investigation of CPSF6 and JNK proximity after UV stress. It revealed that UV light significantly increased the number of PLA signals between CPSF6 and JNK1 or p-JNK in HaCaT cells. With inhibited JNK1 activity, the signal was reduced (**Figure 47**).

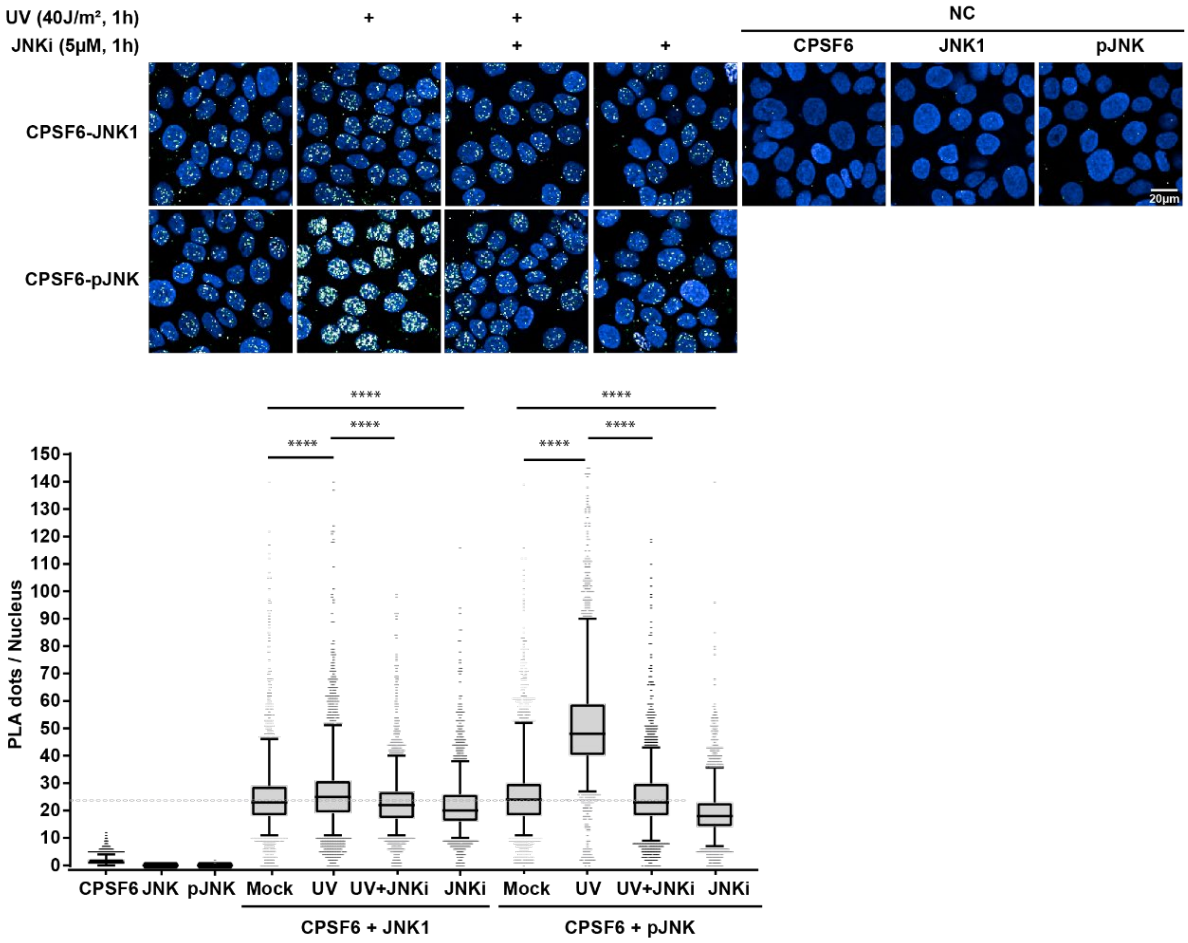

Figure 47: CPSF6 is proximal to the activated JNK1. Proximity ligation assay (PLA) for CPSF6 and JNK1 or pJNK1. HaCaT cells were pre-treated with 5μM JNK inhibitor or with DMSO, and then subjected to UV (40J/m^2) irradiation. Cells were left for 1h recovery and fixed with 4% PFA, then processed according to the manual of PLA kit using CPSF6, JNK1 and phosphorylated JNK1 antibodies. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). One-way ANOVA was used to calculate the significance (n=3; ****p-value < 0.0001).

As a complementary method, we employed proteome analysis based on proximity labeling. JNK1 was fused to an engineered biotin ligase, TurboID, which uses ATP to convert biotin into biotin–AMP, a reactive intermediate that covalently labels proximal proteins with biotin [606]–[608]. These biotinylated proteins can be enriched using Neutravidin beads and analyzed by MS analysis, (the "TurboID-MS" approach) (**Figure 53**A). We observed that CPSF6 had a significantly increased

enrichment by JNK1 proximal labeling upon UV treatment. This result confirms that CPSF6 is proximal to activated JNK1 after UV irradiation (**Figure 47** and **Figure 48**).

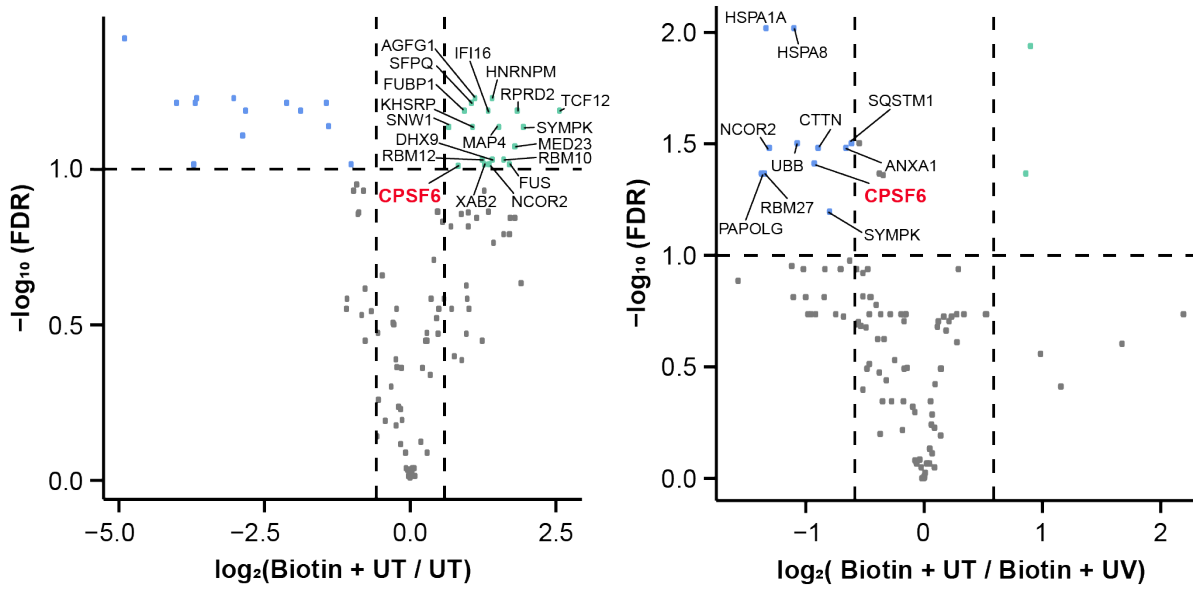

Figure 48: TurboID-JNK1 proteomic analysis in U2OS cells. A. The scatter plot shows the logarithmized SILAC ratios of JNK1 proximal interactors in U2OS cells, with biotin condition compared to without biotin (**A**) and UV light irradiation compared to mock-treated cells (**B**). Significantly enriched proteins are labeled in green and proteins with less biotinylation are marked in blue (n=3; fold change≥1.5, or ≤1.5, FDR ≤ 0.05). Cells were transfected with TurboID-JNK1-NLS. After 48h of transfection, cells were treated with UV (40J/m^2) irradiation and labeled with 50 µM biotin after 30 mins recovery post UV treatment, followed with the same analysis as **Figure 53**.

2.3.2 CPSF6 re-localizes into paraspeckles in a JNK-dependent manner

As part of the CFIm complex, CPSF6 acts as a key regulator determining mRNA length by affecting APA [318], [339]. It has been reported that low levels of CPSF6 lead to preferential cleavage of the pre-mRNA at the proximal APA sites, thus producing more short mRNA [318]. Therefore, we first checked whether the CPSF6 protein level is sensitive to UV light by western blotting. We could not see any changes in whole-cell lysates consistent with our proteome data upon UV treatment (**Figure 49**B). However, we observed the removal of CPSF6 from chromatin 0.5h post UV irradiation in HaCaT cells, which reached the peak after 6h, as we did observe from the chromatin-bound proteome analysis upon UV irradiation. In contrast, PCNA was recruited to chromatin after 0.5h recovery post UV irradiation. The levels of XPC decreased with more ubiquitylation at 0.5h and then slowly started recovering after a longer recovery time (**Figure 49**A).

As we observed re-localization of CPSF6 from chromatin after UV light irradiation, we investigated whether CPSF6 interactome also changes in these conditions. We employed SILAC-based quantitative MS to compare the interaction profile of CPSF6 upon UV stress and in mock-treated cells (**Figure 49**C). This analysis revealed 320 proteins putatively interacting with CPSF6 (**Figure 49**D, **Figure 50**). CPSF6 interacts with many histone proteins, such as HIST1H. Besides the CFIm complex components and known interactors CPSF5, NUDT21 and FIP1L1, many RBPs were detected, including transcription-related factors (RNA Pol II, CDK9, RNABP2) and RNA splicing factors (HNRNPC, HNRNPA3, SRSF1/7/9/10) (**Figure 50**) [21]. The CFIm complex might be stabilized as CPSF5 showed increased interaction with CPSF6 after UV irradiation (**Figure 49**D). This, in turn, might regulate pre-

mRNA processing, accelerating the release of nascent transcripts from chromatin, and subsequently protecting cells from UV-induced DNA damage.

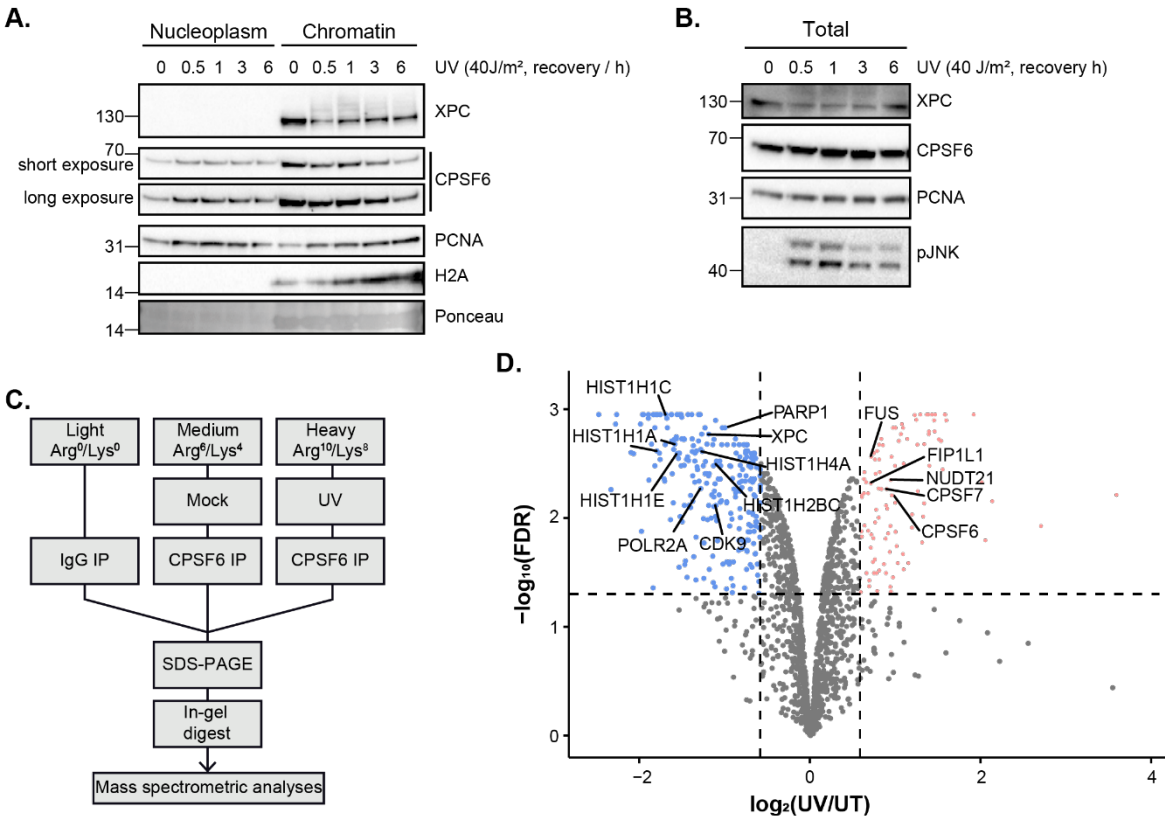

Figure 49: CPSF6 dissociates from chromatin in response to UV light and interaction partners of CPSF6 are regulated during the response to UV light. A. CPSF6 is removed from chromatin after UV light irradiation. Protein fractions are from UV light (40J/m^2, different recovery times) treated HaCaT cells. **B.** Total proteins are analyzed by western blotting after UV light (40 J/m^2, different recovery times) treatment in HaCaT cells.**C.** Experimental workflow for quantitative mass spectrometric analysis of CPSF6 interaction partners after UV light (40J/m^2, 1h recovery). SILAC-labeled cells were lysed and followed by incubation with IgG or CPSF6 antibody-coupled protein G Sepharose beads. Immunoprecipitated proteins were resolved on SDS-PAGE and digested in-gel into peptides, followed by LC-MS/MS. **D.** The scatter plot shows the logarithmized SILAC ratios of CPSF6 interaction partners in cells with UV light irradiation compared to mock-treated cells. Significantly enriched interactors of CPSF6 are labeled in red and proteins with less abundance are marked in blue (n=3; fold change≥1.5, or ≤1.5; FDR ≤ 0.05).

UV stress also induced many interactions with higher affinity with CPSF6. These proteins interact tightly in a network, mainly falling into the cellular organization “tight junction” (**Figure 51**). Notably, UV stress induced FUS to interact with CPSF6 (**Figure 49**D), suggesting that FUS might be involved in CPSF6 re-localization from chromatin in response to UV stress.

In order to track the subcellular localization of CPSF6 during the response to UV stress, we employed proximity labeling proteomics analysis. CPSF6 was fused to an engineered biotin ligase, TurboID, which uses ATP to convert biotin into biotin–AMP, a reactive intermediate that covalently labels proximal proteins with biotin [606]–[608]. These biotinylated proteins can be enriched using streptavidin and analyzed by MS (**Figure 53**A and B).

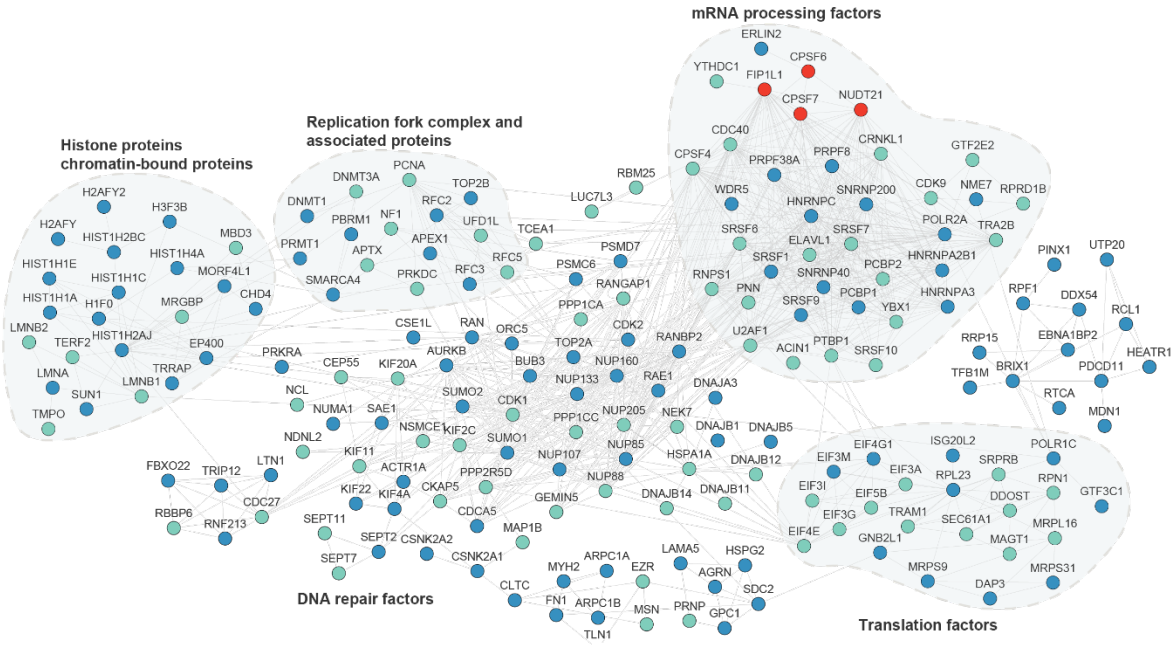

Figure 50: Network of CPSF6 interactors in untreated U2OS cells (Confidence = 0.7). Significantly enriched interactors of CPSF6 after UV treatment are labeled in red, and proteins with less abundance are marked in blue, singletons are left out (n=3; fold change ≤1.5, FDR ≤0.05).

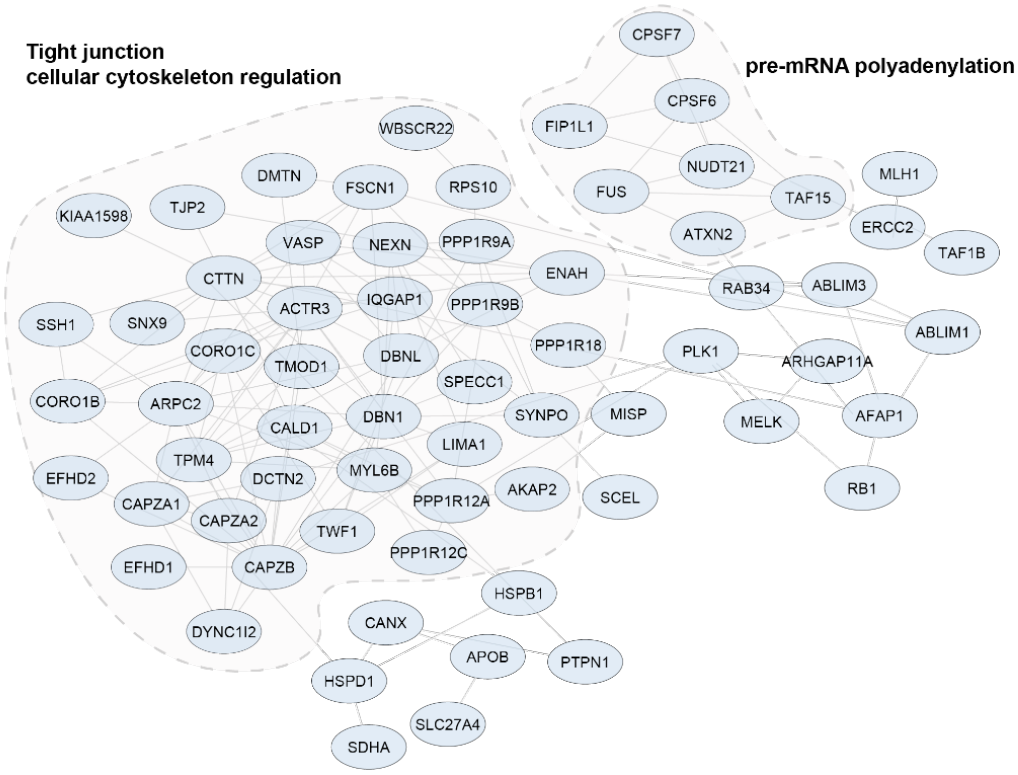

Figure 51: Network for proteins with which an increased interaction with CPSF6 is observed after UV irradiation (Confidence = 0.7) (n=3; fold change≥1.5; FDR ≤ 0.05).

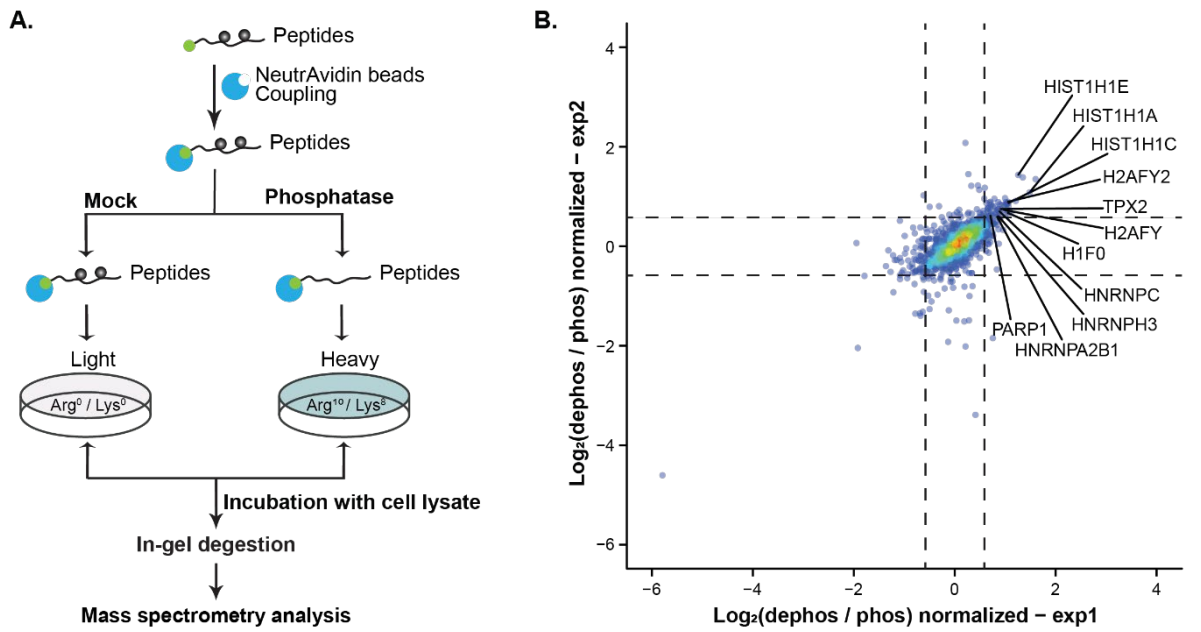

Figure 52: CPSF6 phosphopeptide analysis revealed reduced interaction with histones. A. Experimental workflow for quantitative mass spectrometric analysis of CPSF6 peptide interactors after UV light (40J/m^2,1h recovery). Biotinylated phosphorylated CPSF6 peptides corresponding to T404 and T407 were bound to NeutrAvidin agarose. Phosphorylated and dephosphorylated peptides were incubated with purified cell lysate. **B.** The scatter plot shows the logarithmized SILAC ratios of CPSF6 peptides between replicates for dephosphorylated peptides compared to phosphorylated peptides. Proteins that interact more with dephosphorylated peptides are highlighted (n=2; fold change≥1.5, FDR ≤0.05).

TurboID-tagged CPSF6 successfully generated biotinylation in cells, monitored with Alexa-streptavidin antibody (**Figure 53**C). We established that TurboID-CPSF6 biotinylated proteins (1.5 fold change, FDR ≤ 0.05), including CFIm complex, many spliceosome components and mRNA transport factors (**Figure 53**D). Upon UV stress, core components of paraspeckles were enriched (1.5 fold change, FDR ≤ 0.05), such as SF3B3, RBM25 and HNRNPA1 (**Figure 53**E). The interaction of CPSF6 to other CFIm complex components increased in response to UV stress as the complex components were more biotinylated by Turbo-ID-CPSF6. Very similar results were observed in the "APEX-MS" approach with overexpressing APEX-CPSF6 in U2OS cells (**Figure 75**). These results suggest that CPSF6, with other CFIm complex components, is translocated from chromatin into paraspeckles after UV stress (**Figure 53**C).

To confirm the relocalization of CPSF6 during the response to UV stress, we applied immunofluorescence analysis in HaCaT cells. After UV stress, we observed that CPSF6 forms "rosettes" in the nucleus, usually recognized as paraspeckles in the nucleus. At the same time, we observed the co-localization of CPSF6 and PSPC1, which is a core component of paraspeckles (**Figure 53**F) [19]. Notably, following JNK inhibition, CPSF6 formed fewer rosettes in the nucleoplasm after UV light irradiation (**Figure 53**F and **Figure 54**).

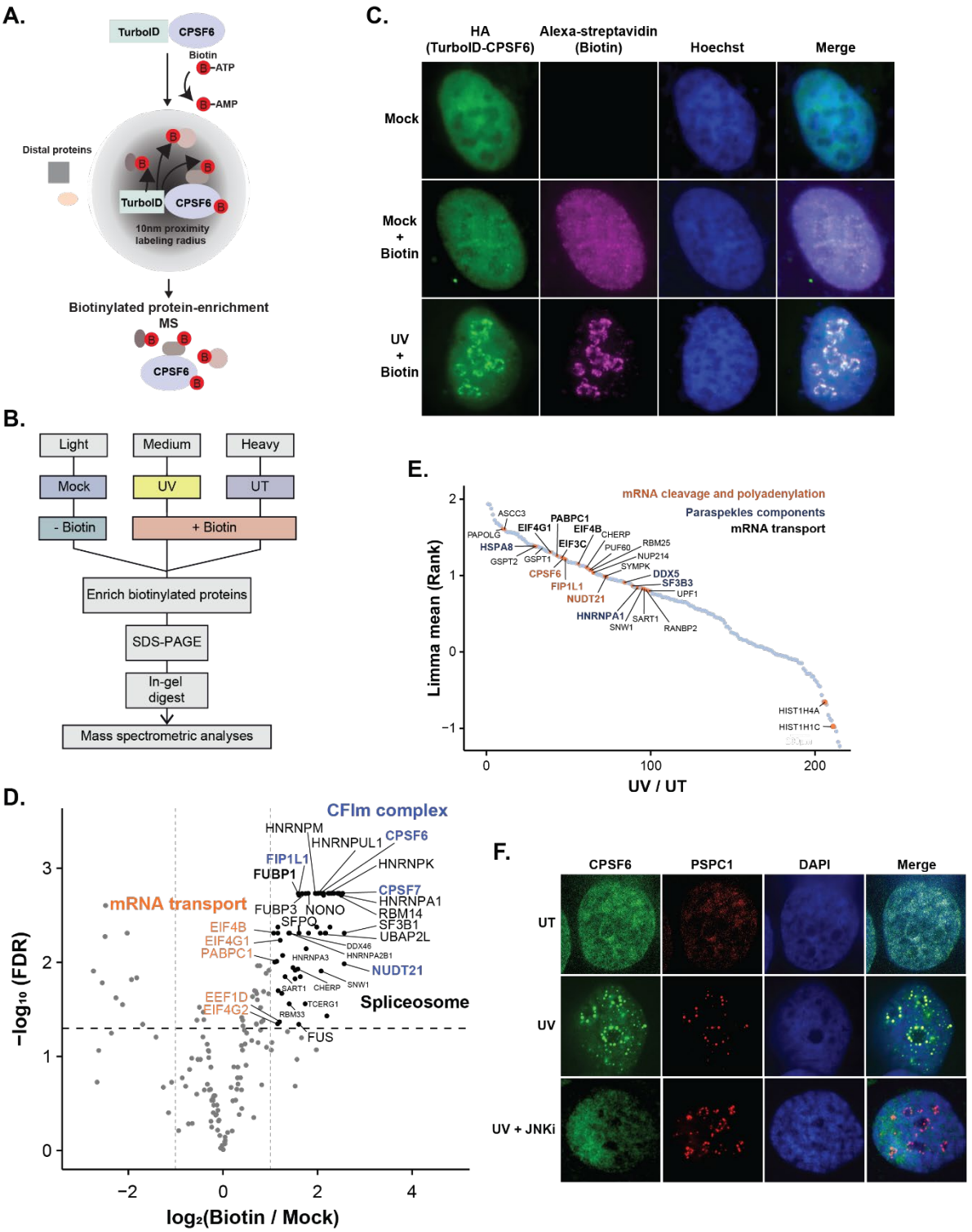

Figure 53: CPSF6 re-localizes from chromatin into paraspeckles in response to UV light. A. A scheme for the principle workflow for identifying components of the TurboID-tagged CPSF6 complex. **B.** Experimental workflow for quantitative mass spectrometric analysis of TurboID-tagged CPSF6 complex after UV light (15J/m^2, 6h recovery). SILAC-labeled cells were transfected with TurboID-tagged CPSF6. After 48 hours of transfection, cells were treated with UV light (15J/m^2, 6h recovery). After 5.5h recovery post UV irradiation, we added biotin (50µM) to medium and incubated the cells for 30 minutes. Cells were lysed and then incubated with NeutrAvidin beads to enrich biotinylated proteins. Immunoprecipitated proteins were resolved on SDS-PAGE and digested in-gel into peptides, followed by LC-MS/MS. **C.** Immunofluorescence analysis for TurboID-CPSF6 overexpression and biotinylation in SILAC U2OS cells**. D.** The scatter plot shows the logarithmized SILAC ratios of CPSF6 proximal interactors in cells by comparing the biotin treatment with the mock treatment. **E.** Rank plot of enriched proteins in UV-treated cells compared to untreated cells with biotin labeling. (n=3; fold change≥1.5, FDR ≤ 0.05). **F.** Immunofluorescence analysis in HaCaT cells shows CPSF6 forms "rosettes" after UV light irradiation. Cells were irradiated with UV light (15J/m^2, 6h recovery), or pre-treated with JNK1 inhibitor (5µM, 1h) followed by UV light irradiation (15 J/m^2, 6h recovery). Cells were fixed with 4% PFA and nuclei were stained with DAPI. Representative images are shown after Deconvolution with Imaris.

These results indicate that JNK activity regulates the re-localization of CPSF6 into paraspeckles in response to UV stress. Recently, it has been shown that CPSF6 is one of the core components of paraspeckles and re-localizes into paraspeckles after transcription inhibition [58], [610], [611]. Paraspeckles are reported as a site in which long noncoding RNAs (lncRNAs) are sequestered and edited [59], [612]. It is plausible that after CPSF6 is translocated from chromatin into paraspeckles after UV light, it takes part in regulating pre-mRNA cleavage and polyadenylation within paraspeckles.

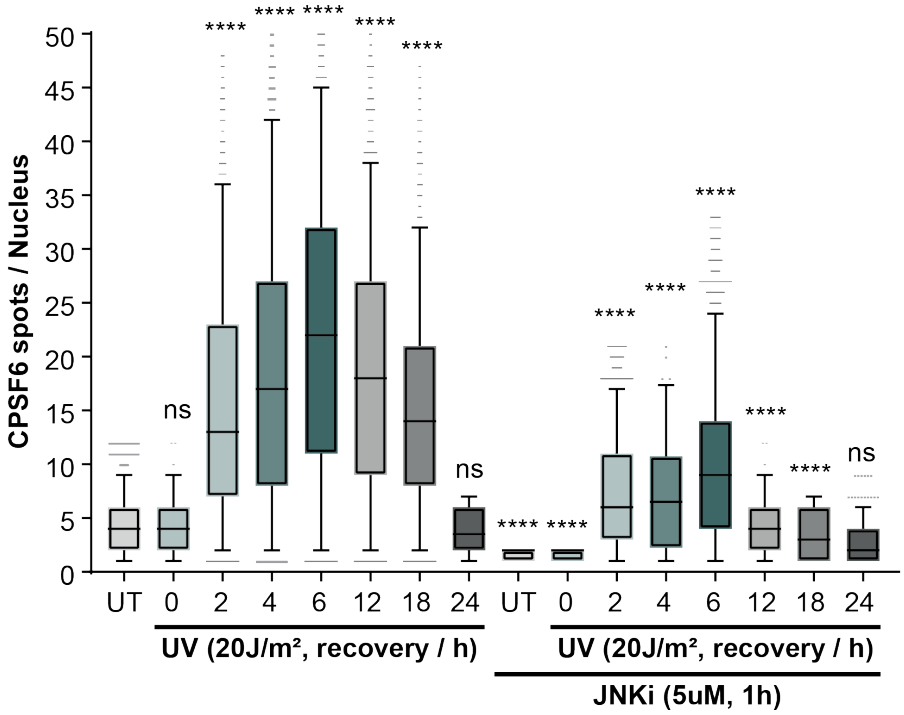

Figure 54: The quantification analysis of CPSF6 nuclear foci in HaCaT cells. 100 cells were quantified in each condition. One-way ANOVA was used to calculate the significance (n=3; ****p-value < 0.0001).

To further investigate whether the phosphorylation of CPSF6 by JNK interferes with its localization, we checked the localization of CPSF6 phosphorylation mutants. We studied the localization of phosphorylation-defective mutants by microscopy analysis. We established doxycycline-inducible cell lines that overexpress GFP-tagged CPSF6 wild-type and JNK targeted-phosphorylation site mutants T404A-, T407A-, T404/407A-CPSF6 (**Figure 55**A). GFP-tagged CPSF6 WT was still removed from chromatin as endogenous CPSF6 after UV irradiation. IF imaging also revealed GFP-tagged CPSF6 WT was mainly located in the nucleus and formed “rosettes” after UV light treatment (**Figure 55**C, D). In contrast, CPSF6 T407A did localize in the nucleus and the cytoplasm did not form “rosettes” in response to UV stress. However, the CPSF6 T404A mutant still behaved very similarly to the wild type of CPSF6 (**Figure 55**D). From the predicted structure of CPSF6 with AlphFold2 (December 2021), the phosphorylation site T407 is close to the N-terminal, which has been reported to be the potential nuclear sequence (**Figure 55**E). It hints that the phosphorylation site on 407 plays a key role in the localization of CPSF6 in the nucleus and its re-localization into paraspeckles in response to UV stress.

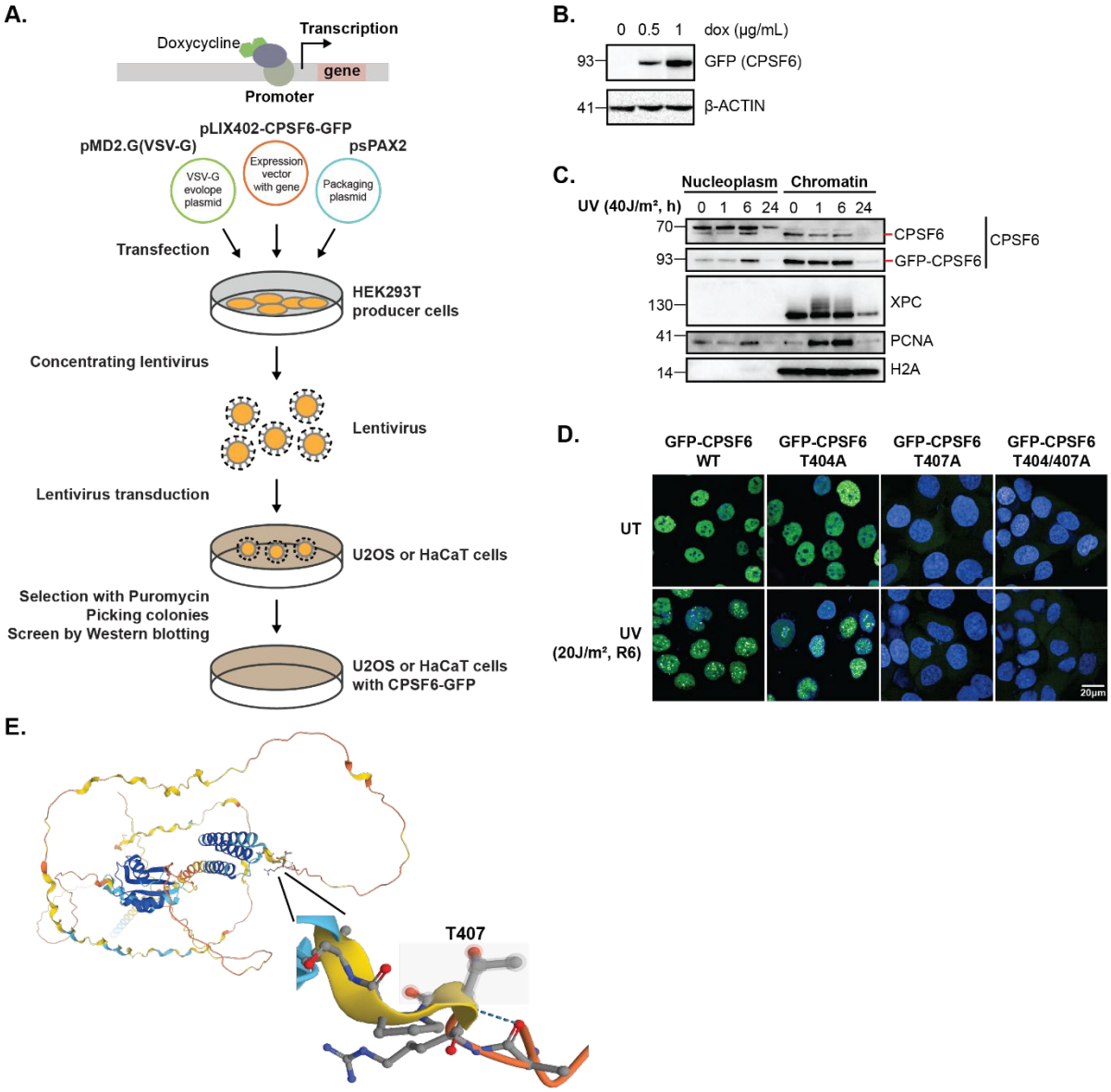

Figure 55: GFP-CPSF6 dissociates from chromatin in response to UV light. A. Tet-on system and the generation of doxycycline (dox) inducible CPSF6-GFP cell lines. **B.** Western blotting shows the overexpression of GFP-CPSF6 in doxycycline (dox) inducible GFP-CPSF6 HaCaT cells after 3 days with dox induction. **C and D.** GFP-tagged CPSF6 is removed from chromatin after UV light irradiation. Protein fractions are from UV light (40 J/m^2, different recovery times) treated HaCaT cells. **D.** Immunofluorescence analysis in dox inducible GFP-CPSF6 HaCaT cells shows CPSF6 forms nuclear foci after UV light irradiation. Cells are irradiated with UV light (15J/m^2, 6h recovery). Cells were fixed with 4% PFA and nuclei were stained with Hoechst. **E.** CPSF6 structure predicted with AlphaFold (Feburary, 2022). T407 phosphorylation site is labeled.

2.3.3 CPSF6 "rosettes" show LLPS characteristics

Paraspeckles are non-membrane organelles in the nucleus generated by protein liquid-liquid phase separation (LLPS) using RNAs as seed [58], [613]. Phase-separated compartments appear as liquid droplets through multivalent and weak interactions between biological polymers. Multivalent interactions can be provided by low-complexity intrinsically disordered domains (IDRs) or structured domains [614], [615].

To find out whether CPSF6 contains IDRs, we applied computational analysis of amino acid sequences of CPSF6 using Predictor of Natural Disordered Region (PONDR, http://www.pondr.com/, 2019). We found 64.61% disordered sequence overall, which means CPSF6 contains low-complexity

regions and long intrinsically disordered regions with proline-rich domains (**Figure 56**A). Previous studies show that a high content of proline-rich domains facilitates the condensate formation of proteins [616], [617]. This further indicates that CPSF6 is likely to undergo phase separation. To confirm whether CPSF6 "rosettes" have LLPS characteristics, we treated cells with 1,6-hexanediol, which has been used to disturb liquid-like droplets[618]. We observed CPSF6 "rosettes" completely dissolved within 30 minutes after adding 5% of 1, 6-hexanediol to the medium (**Figure 56**B). Ammonium acetate (NH_4OAc) has been reported to disturb RNA foci [619], [620]. After treatment with NH_4OAc, CPSF6 "rosettes" completely dissolved within 5 minutes, similar to what we observed after 1, 6-hexanediol treatment (**Figure 56**B).

To determine the biophysical properties of CPSF6 rosettes, we performed a live-cell imaging analysis of the dynamics and morphology of individual rosettes post UV. The number and average focal radius of CPSF6 rosettes increased over time during the first 3 hours, followed by a decreased rosette number but increased size between 6 to 8 hours, resulting from the coarsening of CPSF6rossetes (**Figure 56**C). The progression of growth and coarsening is also a characteristic of LLPS.

Another important characteristic of liquid-liquid-like condensates is their rapid dynamic internal reorganization and rapid exchange between phases, which can be revealed by fluorescence recovery after photobleaching (FRAP) [621], [33]. Therefore, we used the cell line expressing GFP-tagged CPSF6 in HaCaT cells and photobleached individual GFP-CPSF6 foci at different time points after UV to check whether CPSF6 shows LLPS characteristics. We observed a fast and homogeneous recovery within 10-12s after 2h recovery from UV stress. An increase in FRAP recovery time was observed after bleaching in 6h recovery cells where CPSF6 rosette number peaked and increased in size compared to 2h. This behavior confirms a progressive increase in internal viscosity as previously reported for 53BP1 foci and other nuclear bodies (**Figure 56**D) [620].

Together, these results demonstrated that CPSF6 undergoes LLPS in response to UV stress.

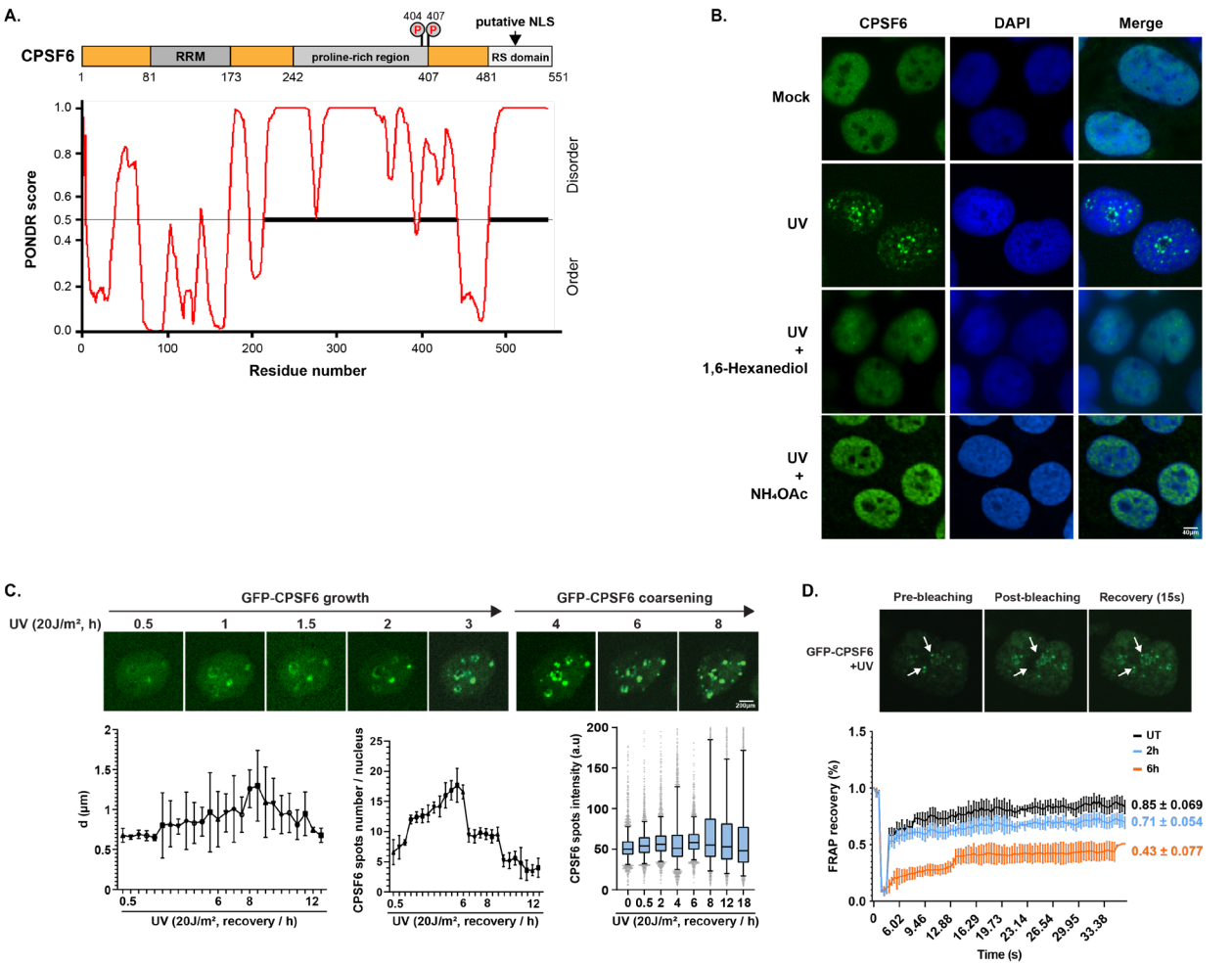

Figure 56: CPSF6 "rosettes" shows LLPS characteristics. **A.** Computational analysis of amino acid sequence of CPSF6, in Predictors of Natural Disordered Regions (PONDRs). Score of >0.5 is predicted to be a disordered domain. **B.** Immunofluorescence analysis in HaCaT cells shows CPSF6 rosettes that have LLPS characteristics after UV light irradiation. Cells were irradiated with UV light (40J/m^2, 6h recovery) followed by treatment with 5% final concentration of 1, 6-hexanediol for 30 minutes or 50 mM NH4OAC for 2 minutes. Cells were fixed with 4% PFA and nuclei were stained with DAPI. **C.** Live imaging analysis for GFP-CPSF6 foci in HaCaT cells shows CPSF6 rosettes have LLPS characteristics after UV light irradiation. Cells were irradiated with UV light (40J/m^2) and then monitored under VisiScope Spinning Disk Confocal Microscope. Quantification analysis was done in Imaris and plotted in GraphPad (n=3). **D.** FRAP analysis of GFP-CPSF6. Fluorescence recovery kinetics of GFP-CPSF6 in non-irradiated HaCaT cells and UV treated cells. Images before bleaching, immediately after the photobleach event (0s), and after recovery (15s). The photobleached region is indicated by an arrowhead. Relative intensities are plotted versus time in seconds. The FRAP curve shows the mean and standard error across three independent replicates (n=50 in each replicate).

2.3.4 The phosphorylation of CPSF6 is essential for alternative polyadenylation in response to UV irradiation

UV light not only induced CPSF6 re-localization into paraspeckles, but also the re-localization of other RNA binding proteins into nuclear granules, such as pre-mRNA splicing factor U2AF65, phosphorylated splicing factor SF3B1 (pSF3b155), RNA Pol II and RNA Pol II pSer2 (**Figure 57**). Previous studies have proposed that the RS domain of U2AF65 drives a liquid-liquid phase separation and further contributes to the complex mechanisms leading to specific splice site choice on RNA [622]. In cells, the carboxy-terminal domain (CTD) of RNA Pol II is an intrinsically disordered low-complexity region that drives RNA Pol II hub formation that regulates chromatin organization, pre-mRNA transcription and processing [623]–[625]. Therefore, we propose that UV-induced transcription inhibition may induce re-localization of nuclear RBPs to membraneless compartments, regulating splicing and polyadenylation.

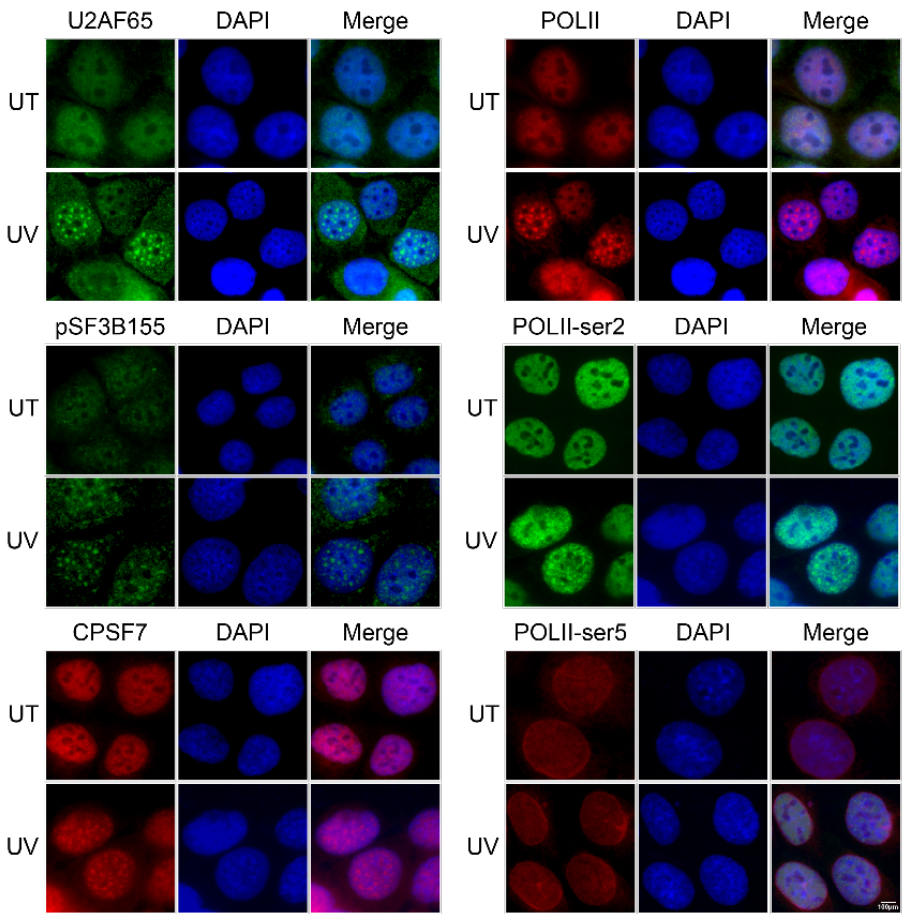

Figure 57: UV induced other RBPs' re-localization. HaCaT cells were treated with UV light (20J/m^2) and left for 6h recovery. Cells were fixed with 4% PFA and respective antibodies were applied for staining. Nuclei were stained with DAPI. Representative images are shown after Deconvolution with Imaris (n=3).

To investigate the regulation of alternative polyadenylation post UV irradiation, we checked the polyadenylated RNA levels in cells by measuring the level of oligo(dT) incorporation. Globally, we observed a decreased total level of polyadenylation (**Figure 58**A). This could be the result of transcription shut down induced by UV light. Interestingly, we observed the re-localization of polyA RNA. PolyA RNA was equally distributed in cells in the normal state. After UV irradiation, polyA RNA formed condensates in the nucleus and co-localized with SRSF2 and partially co-localized with PSPC1. Thus, UV light also induced the re-localization of RNAs together with RBPs (**Figure 58**B). Based on this observation, we hypothesized that CPSF6 re-localizes into paraspeckles together with RNA as a polyadenylation factor.

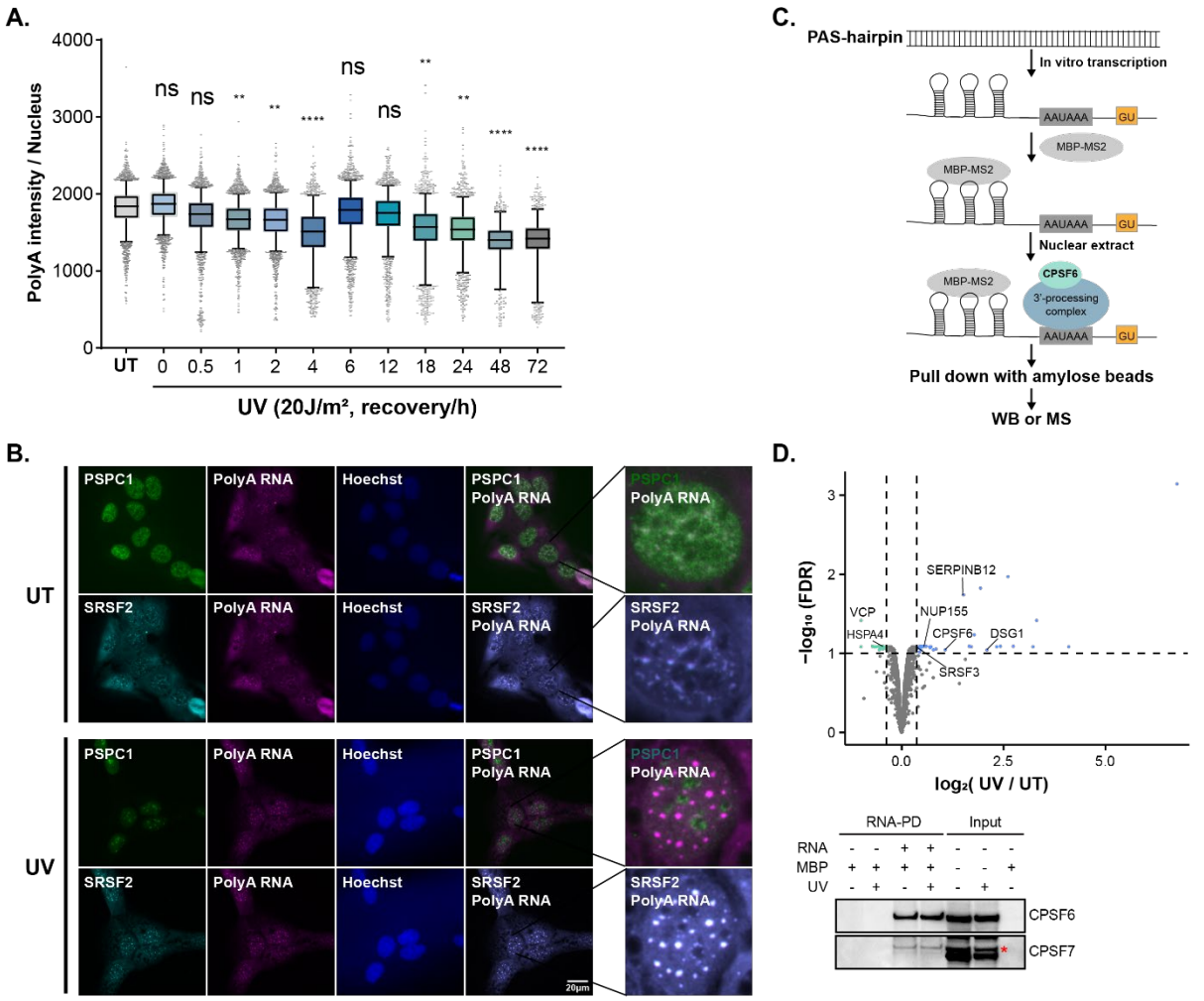

Figure 58: Polyadenylation is regulated during the response to UV light irradiation. A. Quantification analysis of polyA RNA in every nucleus during the response to UV irradiation**.** HaCaT cells were irradiated with of UV light (20J/m^2) and cells were fixed, followed by RNA fluorescence in situ hybridization (FISH) with oligo(dT) to mark polyA RNA and with the indicated antibodies. Nuclei were stained with DAPI. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). The nuclear intensity was analyzed by automated "Harmony High Content Imaging and Analysis Software" (PerkinElmer). Cells in the periphery of the image were excluded from further analysis. One-way ANOVA was used to calculate the significance (n=3; ****p-value < 0.0001). Representative images are shown after Deconvolution with Imaris. The close-up view is shown on the right side (**B**). **C.** Experimental design of RNA-protein complex purification using MS2-MBP. Synthesized DNA contains sequences generating pre-mRNA with PAS sites and three MS2 binding sites (hairpins) at the 3'-end. After *in vitro* transcription, pre-mRNA were bound to MS2-MBP protein into a form (MS2-MBP-pre-mRNA). We applied HaCaT nuclear extracts to MS-MBP-pre-mRNA to capture proteins on pre-mRNA and coupled the complex to amylose beads via MS2-MBP. Then we eluted the captured proteins and performed western blotting or LC-MS analysis. **D.** The scatter plot shows the logarithmized SILAC ratios of captured proteins by pre-mRNA in cells with UV light irradiation compared to mock-treated cells. Significantly enriched proteins are labeled in blue and proteins with less abundance are marked in green (n=3; fold change≥1.5, or ≤1.5, FDR ≤ 0.05).

To test our hypothesis, we checked the binding affinity of CPSF6 to RNA that contains polyadenylation sites. We synthesized an RNA sequence that contains polyadenylation sites and a hairpin structure, allowing maltose-binding protein (MBP) affinity-purification from mock-treated or UV-irradiated cells [626]. Then we quantified the co-purified proteins either with MS or western blotting. There was a significantly increased binding of CPSF6 to RNA after UV irradiation (**Figure 58**C and D). It has been shown that CPSF6 regulates alternative polyadenylation (APA) activity by interacting with FIP1L1 [318]. Interestingly, we observed a stronger interaction of CPSF6 with FIP1L1 and other CFIm complex components after UV irradiation in the above result (**Figure 49**D). An implication of this is the possibility that the CFIm complex re-localizes into paraspeckles together with

RNA upon UV irradiation, thus regulating the alternative polyadenylation in response to UV light-induced stress.

In mammalian cells, UV light induces more short mRNA production by the regulation of distal- or proximal-polyadenylation site (PAS) usage [627], [592]. To study if the re-localization of CPSF6 plays a role in regulating the alternative polyadenylation (APA) in response to UV stress, we used a dual fluorescence reporter, as previously explained [318], [590]. We observed significantly increased APA after UV irradiation that peaked at 6h, which is also the time point with the highest abundance of CPSF6 condensates. Consistent with a previous study, the deficiency of CPSF6 induced a higher APA activity (**Figure 59**A) [591]. To connect the role of CPSF6 with UV-induced higher APA activity, we checked if UV-induced phosphorylation is needed for CPSF6 functioning in APA regulation during the response to UV irradiation. We checked the APA activity using the reporter in CPSF6 WT and phosphomutant cells after UV treatment. Surprisingly, we found that the overexpression of wild-type CPSF6 and T404A

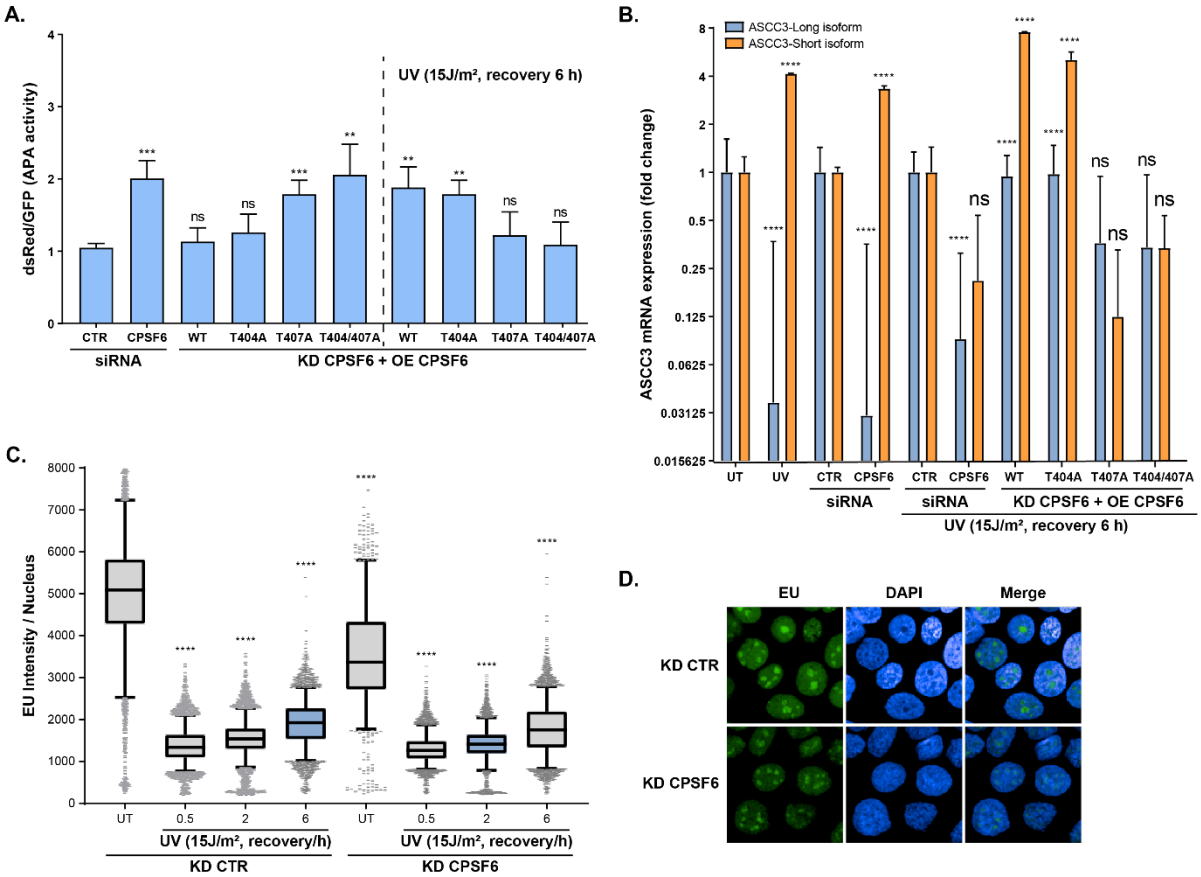

Figure 59: Knockdown of CPSF6 affects alternative polyadenylation in response to UV stress. C. Quantification analysis of the APA activity mean during the response to UV irradiation in CPSF6-deficient cells and CPSF6-overexpressing cells**.** Unpaired t test was used to calculate the significance (n=3; ***p-value < 0.05). **D.** ASCC3 long and short mRNA levels after qRT-PCR analysis in CPSF6-deficient cells and CPSF6-overexpressing cells. Cells were transfected with CPSF6 siRNA or CTR siRNA. Then cells were transfected with CPSF6 overexpression vectors. After 48 hours of transfection, mRNA was extracted according to the RNA extraction kit, followed by qRT-PCR according to the instructions of the quantitative PCR kit. ANOVA was used to calculate the significance (n=4; ****p-value < 0.0001). C. **Immunofluorescence microscopy analysis of EU incorporation in CTR and CPSF6 depleted cells, see Figure 59.** Representative images for the untreated cells are shown here **(D)**. One-way ANOVA was used to calculate the significance (n=3; ****p-value < 0.0001).

mutant increased APA activity similar to the deficiency of CPSF6. In contrast, CPSF6 with T407A mutation showed a different effect, in which CPSF6 was not capable of re-localizing into paraspeckles and instead leaked into the cytoplasm (**Figure 59**A and **Figure 55**D).

The previous study has shown that UV induces more ASCC3 short mRNA production to help cells recover from UV irradiation [592]. We used ASCC3 mRNA in its short and long forms as a read-out to confirm that CPSF6 regulates the APA activity in response to UV irradiation. Quantitative real-time PCR revealed that upon CPSF6 knockdown in untreated cells, an increased level of ASCC3 short mRNA and a decreased level of the long mRNA were observed. However, when CPSF6 was depleted and the cells were treated with UV irradiation, short ASCC3 was not increased. In WT CPSF6 overexpressing cells, an increased level of ASCC3 short mRNA was observed, but not in CPSF6 T407A or T404/407A mutants' cells (**Figure 59**B). Moreover, we observed a significant transcription inhibition in CPSF6-deficient cells (**Figure 59**C and D).

Taken together, all this evidence suggests that the re-localization of CPSF6 plays a key role in the regulation of APA activity during the response to UV-induced stress.

2.3.5 The phosphorylation of CPSF6 is essential for cell survival in response to UV stress

To establish the importance of CPSF6 in response to UV light, we checked the DNA damage foci removal efficiency in CPSF6-deficient cells. We found knocking down of CPSF6 significantly increased γH2AX level in cells; knockdown of known DNA repair factors, XPC and AQR were used as positive controls (**Figure 60**D). Moreover, CPSF6-deficient cells still contain a high level of DNA damage foci (6,4-PPs) and DNA damage response foci (γH2AX) after 24 hours of recovery post UV irradiation compared with control cells where 6,4-PPs and γH2AX foci are removed after 12h recovery. This finding indicates that knocking down of CPSF6 slows down the removal of DNA damage foci during the response to UV light (**Figure 60**A, B and C). Notably, overexpressing wild-type CPSF6 or T404 mutant in CPSF6-deficient cells rescued this effect, but not the overexpression of the T407A or the T404/407A phosphomutants (**Figure 60**C).

We performed cell survival assays to investigate if the effect of CPSF6 we observed during the DDR was reflected in cell viability. CPSF6-deficient cells were more sensitive to UV light and had less ability to form colonies. Similar to the observation of γH2AX levels, overexpressing the phosphorylation mutant on the 407 site of CPSF6 did not rescue this decreased viability, whereas overexpressing wild-type CPSF6 restored colony formation levels (**Figure 61**). Taken together, the phosphorylation of CPSF6 is essential for cell survival during the response to UV light-induced DNA damage.

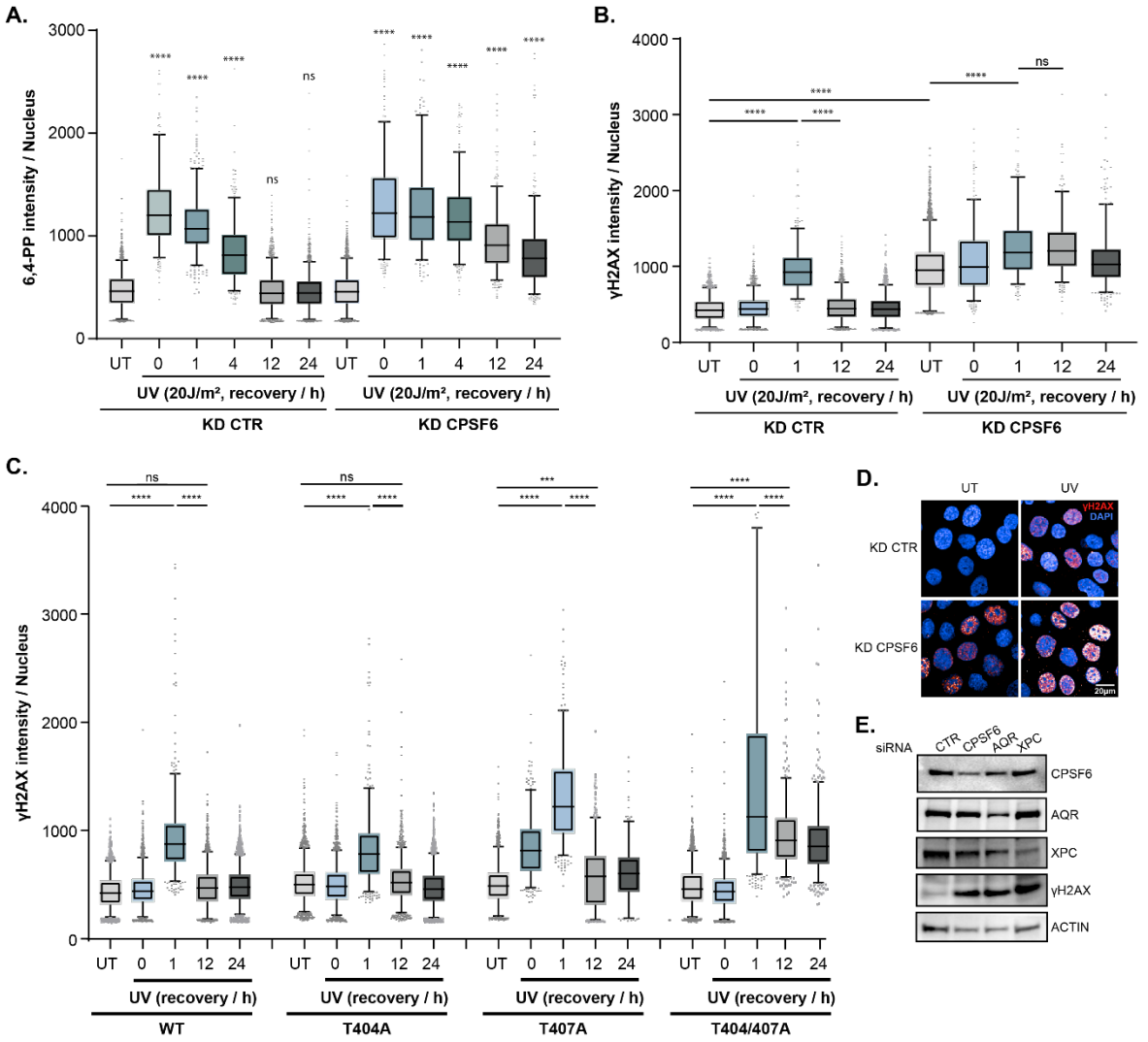

Figure 60: The phosphorylation of CPSF6 is essential for γH2AX removal in response to UV light irradiation. A, B and C. Quantification analysis of 6,4-PP and γH2AX levels per nucleus during the response to UV irradiation in HaCaT cells**.** Doxycycline (dox) inducible GFP-tagged CPSF6 HaCaT cells were transfected with CPSF6 siRNA or CTR siRNA. Then cells were added dox (0.5µg/mL) for 48 hours. Cells were irradiated with UV light (20J/m^2) and left for the indicated time to recover. Cells were fixed and then immunofluorescence staining and imaging analysis were performed. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). The nuclear intensity was analyzed by the automated "Harmony High Content Imaging and Analysis Software" (PerkinElmer). Cells in the periphery of the images were excluded from further analysis. One-way ANOVA was used to calculate the significance (****p-value < 0.0001). **D.** Representative images for untreated or 1h post UV-treated cells from **A** are shown. **E.** HaCaT cells were transfected with CPSF6, XPC and AQR siRNA. 72 hours after transfection, cells were lysed and western blotting analysis was performed with the indicated antibodies.

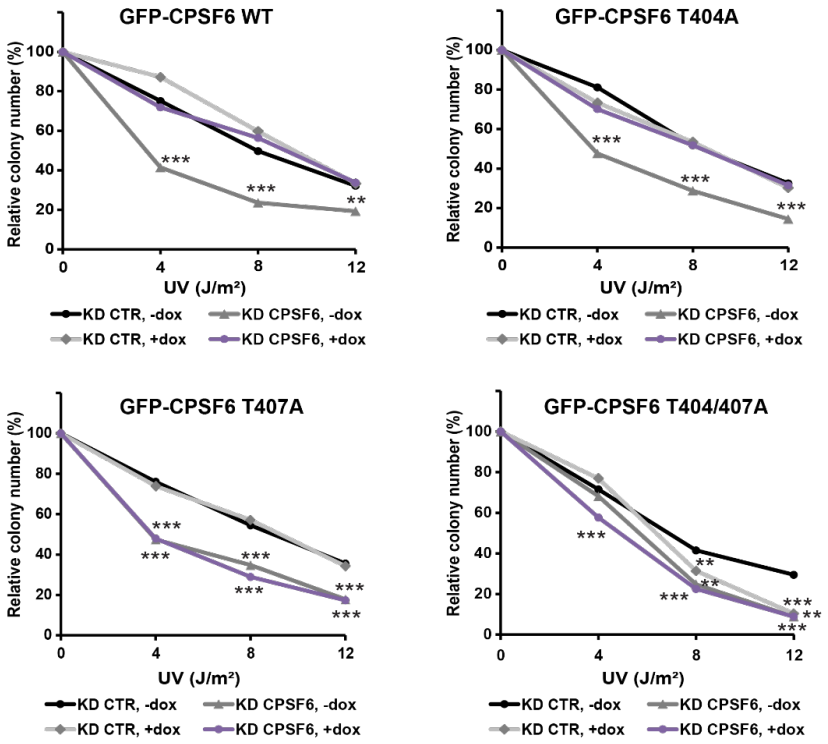

Figure 61: Knockdown of CPSF6 reduced the ability of HaCaT cells to form colonies after UV light irradiation. Cells were from the same pool in B and C of **Figure 60**. Cells were irradiated by UV light (4, 8 and 12 J/m^2), and left to grow for 10 days. Cells were stained with Crystal violet solution. The error bars show the mean and SD of results obtained in three technical replicates. A two-sided Student's t-test was used to calculate the significance (n=3; ***p-value < 0.05).

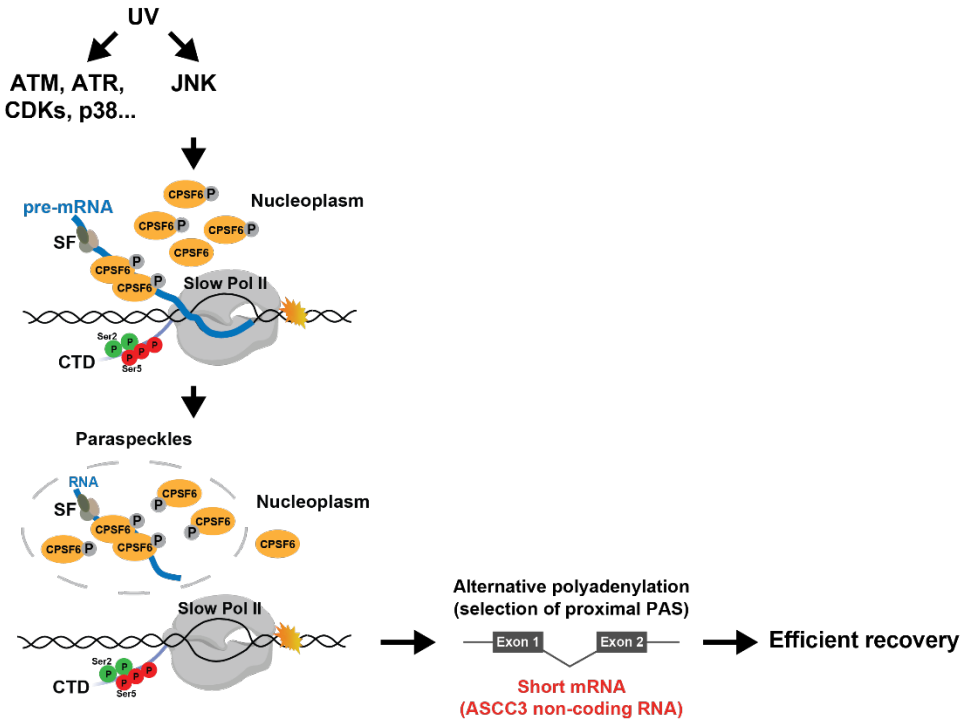

Figure 62: Model of the JNK1-CPSF6 axis regulating in response to UV stress. UV stress induced JNK activation immediately. The activated JNK phosphorylates CPSF6 on 407 site. The phosphorylated CPSF6 disassociated from transcription machinery on chromatin and re-localized into paraspeckles with mRNA via liquid-liquid phase separation. This re-localization regulates alternative polyadenylation and transcription recovery in response to UV stress-induced transcription blocking lesions.

2.4 R-loop regulation in response to UV stress

2.4.1 UV stress induces the phosphorylation of R-loop factors

Recent studies have shown that UV-induced TBLs induce displacement of co-transcriptional spliceosomes from stalled RNA Pol II, resulting in R-loop formation at DNA damage sites [371].

To learn more about the mechanism of R-loop regulation during the response to UV stress, we consulted our phosphorylation dataset. Indeed, we found that phosphorylation of many R-loop regulators are regulated by UV light induction, such as ATAD5, senataxin (SETX), DExH-Box Helicase 9 (DHX9), and a large increase in phosphorylation of splicing factor SRSF2 was observed [628], [629],[630], [631] [368], [632], [633], [634], [635] (**Figure 63**A). At the same time, we also observed the re-localization of SRSF2 (**Figure 57**). It has been shown that RNA processing proteins inhibit R-loop formation by occupying RNA transcripts, reducing the chance of RNA invading DNA [636], [637]. Among all R-loop regulators with regulated phosphorylation sites, after network and GO-enrichment analysis, mainly RNA binding proteins are found in our dataset that fall into categories related to "cell cycle," "chromatin organization," and "RNA processing" (**Figure 63**B). ATPase Family AAA Domain Containing 5 (ATAD5) has been shown to prevent R-loop formation through unloading PCNA and RNA helicases [628]. Nuclear DNA Helicase II, DExH-Box Helicase 9 (DHX9) promotes the formation of both pathological and non-pathological R-loops with impaired RNA splicing. When splicing factors are absent in cells, the occurrence of R-loops coincides with the extended interaction of DHX9 with RNA Pol II. This results in forming a DNA–RNA hybrid, which traps RNA Pol II on chromatin and may then prevent DNA replication [632].

While it is intriguing that we find these factors to be differentially phosphorylated upon UV treatment, how their phosphorylation affects their function in the regulation of R-loop formation during the response to UV-induced DNA damage remains uncharacterized. It might influence their localization on chromatin and subsequently induce R-loop formation or protect cells via resolving R-loops on DNA upon UV irradiation.

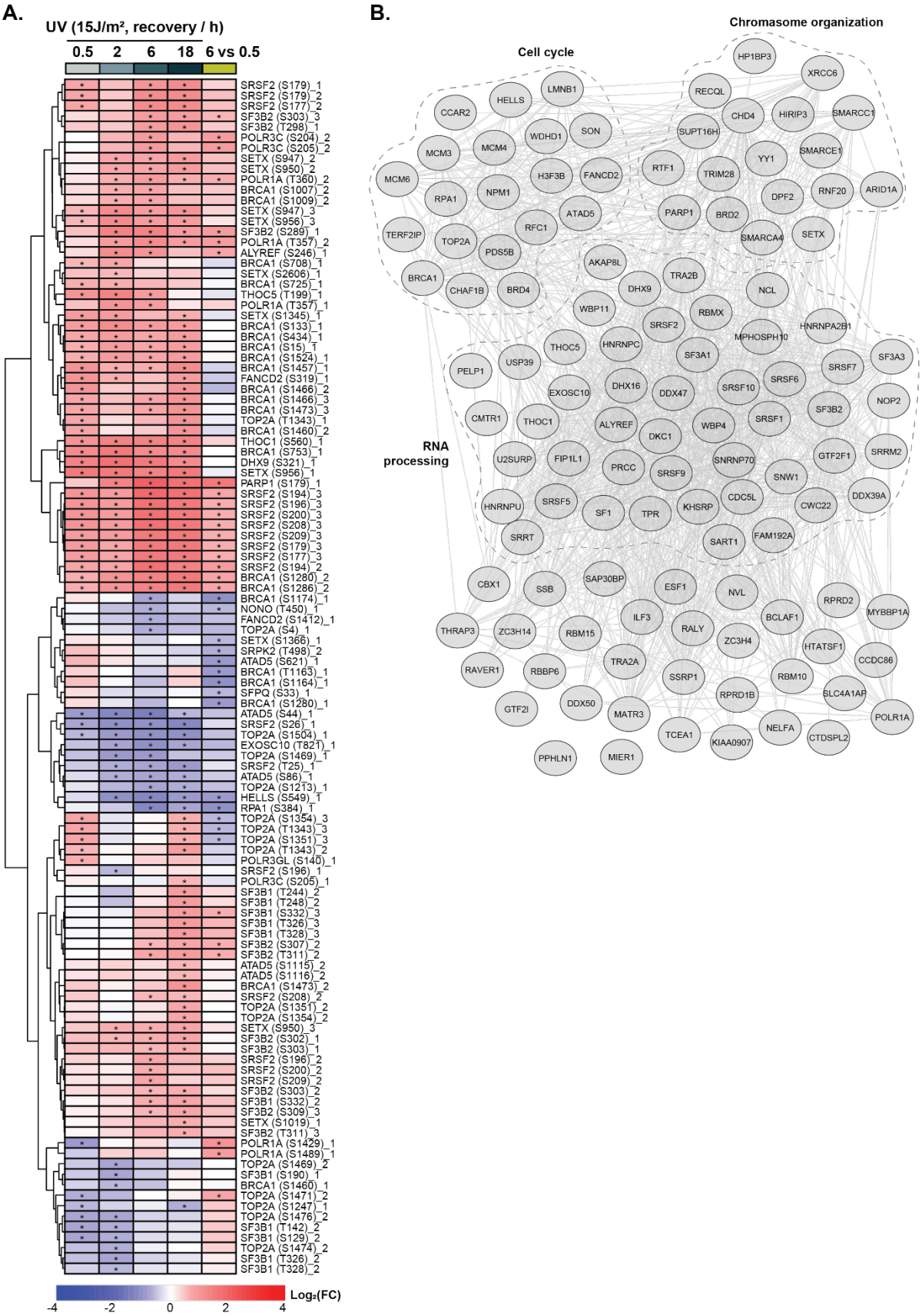

Figure 63: R-loop factors' phosphorylation is regulated upon UV light induction. A. Heatmap of published R-loop factors with regulated phosphorylation sites in response to UV light. (peptide fold-change ≥ 2, FDR < 0.05 are indicated with an asterisk). **B.** Interaction network of all potential R-loop regulators identified in the APEX-HBD screen from our lab with regulated phosphorylation sites (confidence = 0.7) [638].

2.4.2 UV stress induces R-loop formation

To test R-loop formation in HaCaT cells, we performed a typical method, dot-blot, to quantify the R-loop levels with the widely-used monoclonal antibody S9.6 [639], [640]. R-loop levels were increased after UV irradiation, and we used the knockdown of AQR as a positive control. These dots were resolved by RNaseH1 digestion (**Figure 64**D) [641]. Furthermore, we carried out immunofluorescence imaging analysis using the DNA:RNA hybrid binding domain (HBD) antibody and its catalytically-dead mutant (WKK) antibody, which was developed in our lab [362]. This analysis demonstrated that HBD intensity decreased after 30min recovery post UV treatment and increased later, at 6h and 8h when transcription restarted after UV irradiation. At these later time points, the HBD intensity is higher than the normal state in HaCaT cells. It is also sensitive to RNaseH1 treatment and can not be recognized by the WKK antibody (**Figure 64**A). As a positive control, knockdown of AQR

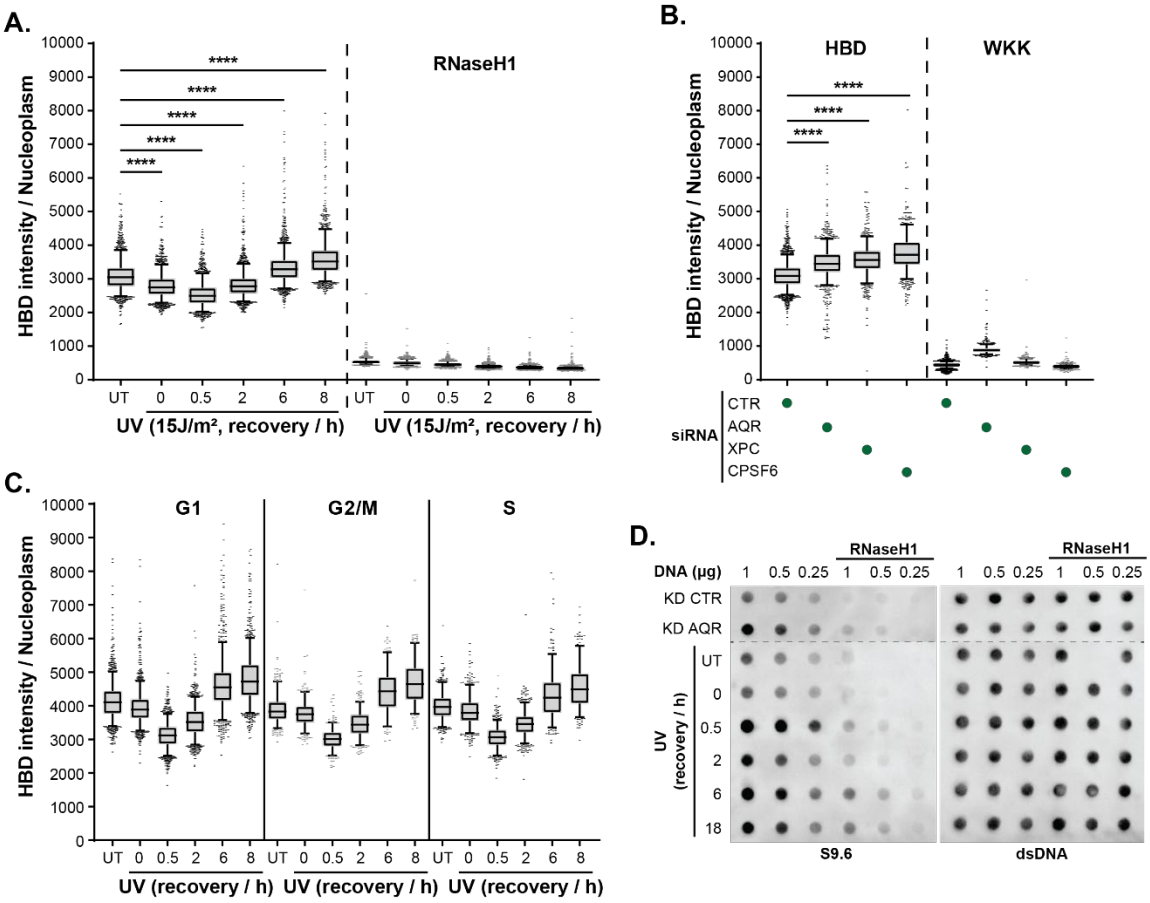

Figure 64: UV irradiation induces R-loop formation. A. Quantification analysis of HBD levels per nucleus during the response to UV irradiation in HaCaT cells**.** Cells were irradiated with UV light (15J/m^2) and left for the indicated time to recover. Cells were fixed with ice-cold methanol, followed by RNaseH1 or mock treatment at 37°C for 1h. Then cells were processed with immunofluorescence staining and imaging analysis. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). Cells in the periphery of the image were excluded from further analysis. One-way ANOVA was used to calculate the significance (****p-value < 0.0001). **B.** Quantification analysis of HBD levels per nucleus in AQR-deficient cells. Cells were transfected with CPSF6, XPC and AQR siRNA or CTR siRNA and the same analysis was performed as described in **A**. **C.** Cell cycle analysis of HBD intensity per nucleus during the response to UV light. Cell cycle analysis was based on DAPI sum intensity of each nucleus. **D.** Dot-blot analysis in HaCaT cells upon knockdown of AQR and UV irradiation. Genomic DNA was collected after different recovery times post UV treatment or 48 h post-knockdown of AQR. Half of the sample was treated with RNaseH1 to control signal specificity. The gDNA was spotted in different concentrations and the membrane was probed with the S9.6 antibody before being stripped and consequently probed with dsDNA antibody as a loading control. Representative images of membranes probed with S9.6 and dsDNA antibodies.

also increased HBD intensity in the nucleus, which is not recognized by the WKK antibody (**Figure 64**B).

Finally, since R-loops have been reported to be regulated throughout the cell cycle, we performed the immunofluorescence experiment looking at the HBD intensity during different phases of the cell cycle [642]. This experiment showed that the pattern of the HBD intensity observed in untreated cells and post UV irradiation did not show any cell cycle phase specificity (**Figure 64**C).

Moreover, we observed RNA Pol II co-localization with R-loop foci labeled by the HBD antibody post UV irradiation (**Figure 65**A). Interestingly, we found that the phosphorylated RNA Pol II at Ser2, which promotes transcription elongation, showed a similar behavior (**Figure 65**B) [168].

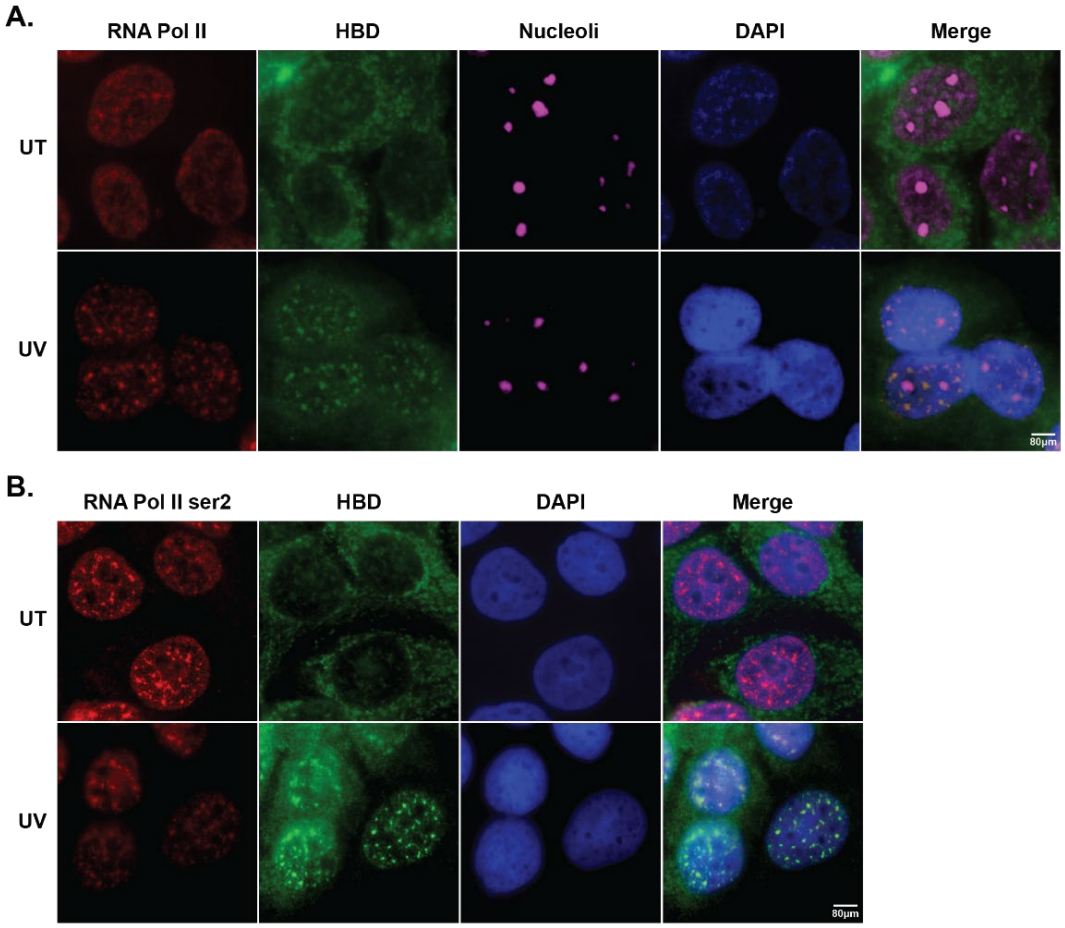

Figure 65: Immunofluorescence analysis of co-localization of HBD and RNA Pol II (A) or RNA Pol II phosphorylated at Ser2 (B). HaCaT cells were treated with UV (20J/m^2) irradiation. Cells were left for 6h recovery and fixed with ice-cold methanol, followed by staining with respective antibodies. Nuclei were stained with DAPI and images were captured under the confocal microscope. Representative images are shown after Deconvolution with Imaris.

In addition, R-loops may accumulate on DNA when the replication fork collides with the transcription machinery during the response to UV light-induced DNA damage [628], [643]. Thus, we monitored the interaction between the transcription and replication machinery using PLA. We used antibodies against RNA Pol II and PCNA to mark transcription and replication fork, respectively. Indeed, we observed an increase in the PLA signal between RNA Pol II and PCNA in HaCaT cells immediately after 30 min post UV irradiation (**Figure 66**A). To detect the PLA signal specifically on

chromatin, we removed soluble proteins in the nucleus by performing pre-extraction prior to fixation [644], [645]. With pre-extraction, the decrease in PLA signal that was observed after 30 min, which then increased again at 6h post UV irradiation, was consistent with the HBD antibody staining results (**Figure 66**B).

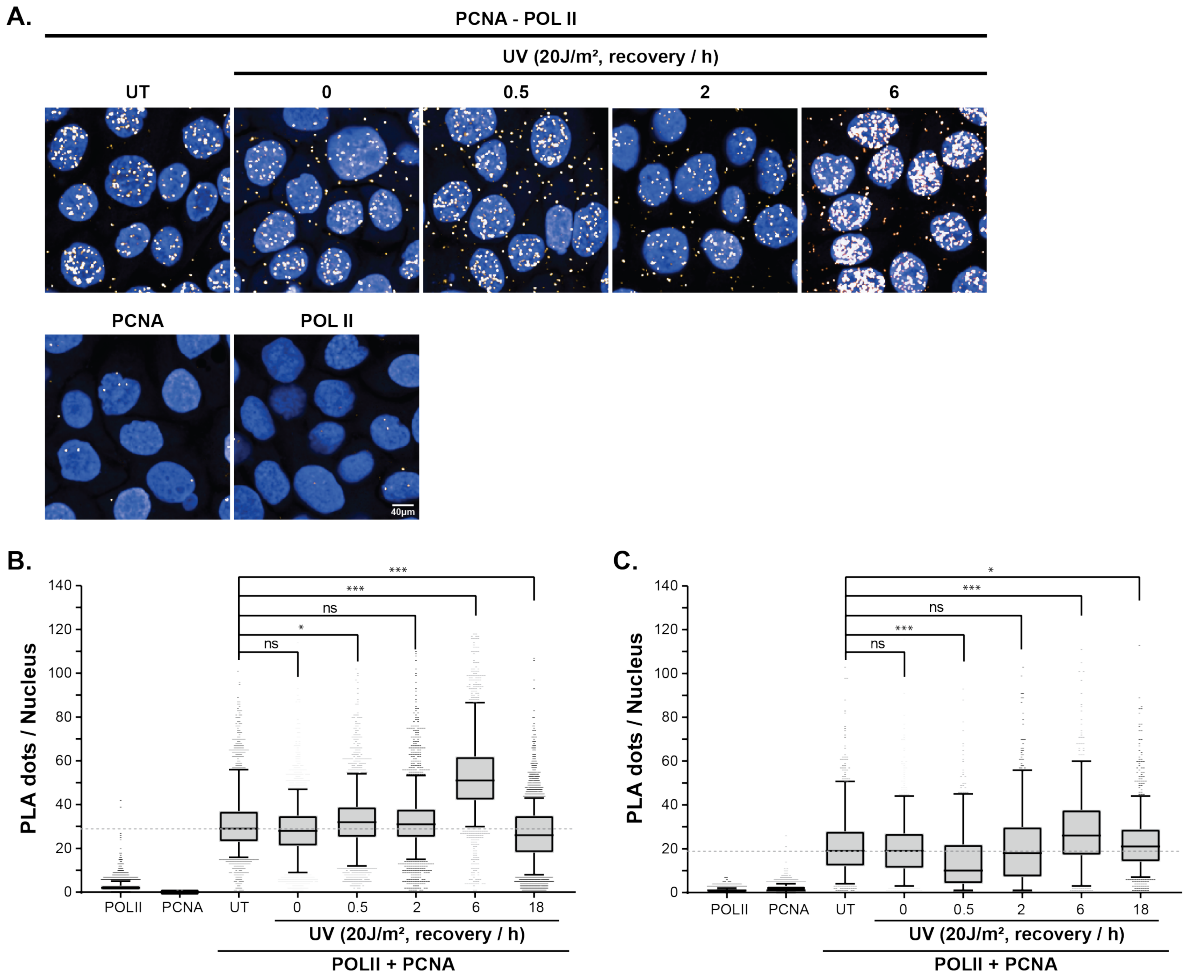

Figure 66: Proximity ligation assay (PLA) analysis for RNA Pol II and PCNA. HaCaT cells were treated with UV (20J/m^2) irradiation. Cells were left for different times of recovery and fixed directly with 4% FA **(B)** or first pre-extracted with 4% NP40 before fixation with 4% FA **(C)**, then processed according to the manual of PLA kit using RNA Pol II and PCNA antibodies. Nuclei were stained with DAPI. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). The nuclear intensity was analyzed by automated "Harmony High Content Imaging and Analysis Software" (PerkinElmer). Cells in the periphery of the image were excluded from further analysis**A.** Representative images are shown. **B.** Quantification analysis of PLA signals per nucleus for cells directly fixed with PFA. **C.** Quantification analysis of PLA signals per nucleus in cells that were pre-extracted first and then fixed with PFA. One-way ANOVA was used to calculate the significance (n=3; ***p-value < 0.0001).

In summary, all these results showed that R-loops form on transcription-replication collision sites during the response to UV light when transcription restarts. The increase we observed in R loop formation could also be due to the fact that RNA processing proteins are re-localized post UV irradiation. We have shown that UV-induced RNA processing factors re-localize into paraspeckles after 6h. This is consistent with the increase in R-loop levels observed after 6h recovery (**Figure 66**A). The exact cause for this increased R-loop formation remains to be elucidated.

2.4.3 Establishment of the proximity proteomics with split promiscuous biotin ligase (split-TurboID-MS)

As we observed UV stress induced dynamic level changing of R-loop formation. To investigate the exact mechanism in response to UV stress, we aim to capture proteins required for R-loop regulation at different recovery phases.

To get a more specific map of complexes that are involved in R-loop regulation, we took advantage of the more advanced biotin ligase, split-TurboID, a promiscuous biotinylation enzyme split into two inactive fragments. Co-expressed two fragments can be brought together by a drug, protein-protein interaction, or organelle contact to reconstitute TurboID enzymatic activity. It has been shown that RNaseH1 forms a dimer to resolve R-loop [646]. We fused HBD to two TurboID fragments (N or C), named split-TurboID-N/C, to map the protein composition of R-loop accumulation sites in SILAC U2OS cells (**Figure 67**A). They are successfully overexpressed, generating biotinylation signals in cells (**Figure 67**B and C). By comparing with other R-loop databases (APEX-HBD proteome and R-loopBase), we observed a significant overlap with known R-loop regulators (**Figure 67**D and E). In the future, we will apply the established "Split-TurboID-MS" method based on TMT labeling over HBD, or PCNA and RNA Pol II, to map the complex dynamics involved in R-loop regulation, at different time points spanning from 30 min to 18 hours after UV irradiation.

Collectively, the cooperative activities of replication fork progress, transcription bubble maintenance, and mRNA processing machinery contribute to preventing R-loop accumulation. How these individual components regulate R-loop homeostasis and how their respective regulation determines the local prevalence of these hybrid structures throughout the genome upon UV induction still need to be further investigated.

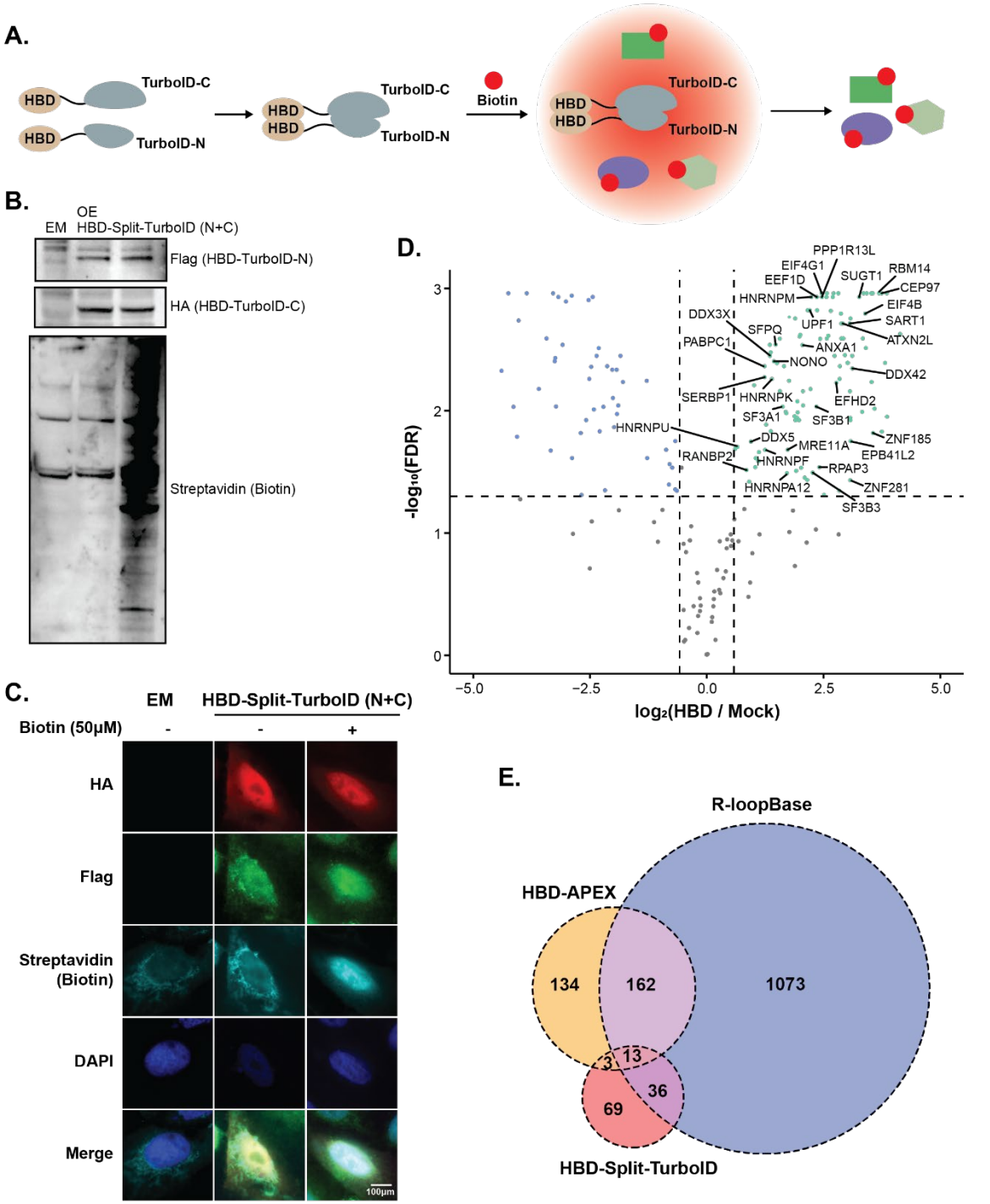

Figure 67: Proximal complex of HBD-split-TurboID-MS. A. A scheme for the principle workflow for identifying proximal components of split-TurboID reconstitution of HBD [647]. HBD fused with split-TurboID-N and C are co-transfected in heavy and medium conditions. After 24 hours of transfection, biotinylation labeling was induced by adding biotin (50µM) in medium and incubating for 12 hours. Cells were lysed and pooled together, followed by incubation with NeutrAvidin beads to enrich biotinylated proteins. Immunoprecipitated proteins were resolved on SDS-PAGE and digested in-gel into peptides, followed by LC-MS/MS. **D.** The scatter plot shows the logarithmized SILAC ratios of enriched proteins in heavy conditions compared to mock cells without any transfection. (n=3; fold change≥1.5, FDR ≤0.05). **B and C.** Western blotting and immunofluorescence analysis in U2OS cells for the overexpression of HBD fused with split-TurboID-N/C and biotinylation. Cells were fixed with 4% PFA and nuclei were stained with DAPI. Representative images are shown after Deconvolution with Imaris. **E.** Overlap of HBD-Split-TurboID and R-loop regulators in literature [363], [638].

3 Discussion

3.1 The landscape of temporal response to UV stress based on quantitative mass spectrometry analysis

Our study aimed to track the proteomic events in several phases of the process during the response to UV light-induced DNA damage: Immediate response-transcription shutdown, transcription recovery phase, and after transcription is recovered and DNA lesions are removed. We monitored the changes in protein levels, phosphorylation, and localization on chromatin by applying quantitative proteomics analysis. Phosphorylation can provide information for the understanding of molecular mechanisms underlying changes in protein level and localization upon UV stress. Different phases have different action tactics during the response to UV stress.

3.1.1 UV stress induced kinase activation and deactivation in a time-dependent manner

MAPKs are activated at an early timepoint post UV light treatment. The cascade can be activated by an array of stimuli like extracellular mitogens, growth factors, cytokines, as well as stress signals like interleukins, irradiation, or chemotherapeutics [648]–[650]. MAPKs are involved in several cellular processes such as transcription, translation, apoptosis, and DNA repair [651]–[655]. AKT1 shows a similar pattern with MAPKs. AKT1 can be activated through similar mechanisms and show extensive cross-talk with MAPKs [656]–[660]. Complementary to the MAPKs signaling, the increased CK2 (CSNK2A1/2) activity was observed at 6 and 18h post UV irradiation when MAPKs are no longer activated (**Figure 31**). This may signify that the strong MAPK activation (observed at early time points) suppresses CK2 activity. Some research has indicated that MAPKs and CK2 have extensive cross-talk, limiting or increasing each other's activity in rice, immune cells, and acral melanoma[661]–[666]. CK2 has been reported to regulate apoptosis and to promote DNA repair through the phosphorylation of MDC1 and RAD51 [667]–[670]. In yeast, CK2 is the terminal effector in a signal transduction cascade that downregulates TFIIB. Theoretically, conventional DNA damage sensors and transducers are likely to be upstream of CK2 [671], [672]. However, the mechanism of CK2 in response to UV irradiation-induced DNA damage has to be further studied in mammalian cells. We observed a strong deactivation of AURKA/B and CDK1/CDK2 after UV treatment (**Figure 31**C). Aurora kinases and CDKs play pivotal roles in regulating spindle assembly, chromosome segregation, and cytokinesis to ensure faithful segregation of chromosomes during mitotic cell division. AURKA favors the G2/M transition by promoting centrosome maturation and mitotic spindle assembly in a complex of AURKA-p53/p73. AURKB is part of the chromosomal passenger complex, which is crucial for microtubule binding to kinetochores and segregation of chromosomes [732]–[735]. The depletion of AURKA sensitizes esophageal squamous cell carcinoma to UV light [736]. The cross-talk between Aura kinases and MAPKs is also reported as the activated MAPK/ERK signaling pathway mediates robust AURKA promoter activation in melanoma [737].

3.1.2 UV stress induced cell cycle arrest

Notably, we observed the G2 checkpoint kinase WEE1 is activated at 18h post UV stress. WEE1 is reported to phosphorylate CDK1 on Tyr15, resulting in the inactivation of the CDK1/cyclin B1 complex and hence preventing entry into mitosis [738]. Indeed, we found CDK1 is phosphorylated with two sites, Thr14 and Tyr15, at 18h post UV stress but not at an early time point. In addition, CDK1 phosphorylation on these two sites also can be generated by another G2 checkpoint kinase, PKMYT1. Moreover, WEE1 directly interacts with APC/C components, composted of fizzy and cell division cycle 20 related 1 (CDH1), CDC20, cell division cycle 27 (CDC27), functioning as a negative regulator on the anaphase-promoting complex/cyclosome (APC/C) activity [739].

UV stress-induced all these kinases' activations or deactivation resulted in the cell cycle arrest at the G2/M phase at 18h.

3.1.3 Changes in global protein levels induced by UV stress

Very few significant changes were observed on the protein level at 0.5h, immediately after UV treatment, when transcription is shut down. With transcription shut down, there are much more changes at the transcriptome level, expectedly, as seen in the published transcriptome data [575], [673], [674]. At 0.5h, the upregulation of ASCC1 may indicate a potential effect from UV light-induced DNA alkylation as ASCC1 is important in the DNA alkylation repair by coordinating the recruitment of the activating signal cointegrator 1 complex (ASCC) to DNA damage sites. DNA binding proteins such as transcription repressors (ZBTB4 and HMGB1) are downregulated (**Figure 24**C). Zinc finger and BTB domain-containing protein 4, ZBTB4, is reported to bind to methylated CpG to repress the transcription of p21. The decrease of ZBTB4 subsequently induces the stimulation of the activator p53, which has been well-demonstrated to regulate the cellular response to stress and transcription [675], [676]. Human nonmelanoma skin cancers show a highly UV-specific mutation pattern at the p53 gene. These mutations also induce the UV-specific mutations at methylated CpG sites, which are part of a solar UV-specific mutation signature [677], [676]. Surprisingly, the kinase Ras-1 suppressor, KSR1, is downregulated (**Figure 24**C). This might lead to the ERK cascade activation, which is activated in response to cellular stress [678]. However, MAPK kinase binding protein MAPKBP1 is down-regulated. In contrast, upregulated proteins are involved in translation, such as ribosome large subunit RPL37A and small subunit RPS3 (**Figure 24C**). Some proteins are also involved in protein degradation, such as ubiquitylation ligase UBE2E. TGFβ regulator SMAD2 and NF-κ-B regulator, Activating Signal Cointegrator 1 Complex, Subunit 1, ASCC1, which are reported to be involved in DNA repair, are also upregulated [679], [680]. ASCC1 contains a putative RNA ligase-like domain near the C-terminus. ASCC1 interacts with the ASCC complex through the ASCC3 DNA helicase and coordinates the proper recruitment of the ASCC complex during alkylation damage response [681], [682]. These regulations of transcription factors and DNA repair factors indicate that in short-time recovery, DNA repair events are highly regulated and coupled with transcription events, which is expected.

More protein levels are significantly regulated with longer time recovery, particularly at 18h (**Figure 24**). From the heatmap for regulated proteins at different time points post UV irradiation, we can also see that the expression level of many proteins is continuously regulated during the DNA repair process (**Figure 69**). For instance, ANXA2 (Annexin A2) protein level is downregulated through all timepoints upon UV treatment. ANXA2 has been reported to be involved in many cellular processes. ANXA2 expression is also regulated in a cell cycle-dependent manner in the nucleus. UV induces ANXA2 nuclear aggregation to prevent genomic damage by reactive oxygen species signaling [682]. ANXA2 degradation was reported to be correlated with cellular apoptosis induced by p53-mediated pathways. It can be used to detect and diagnose cancer progression for some cancers [683], [684]. Upon the loss of ANXA2, tumor cells suffer apoptosis through proapoptotic p38MAPK JNK/Akt signaling due to hydrogen peroxide stimulation (e.g., induction of oxidative damage) [685]. Knockdown of ANXA2 also inhibits cell division and proliferation. We did observe cell cycle arrest after 18h recovery post UV irradiation. Cell cycle-related proteins are also downregulated, such as RECQL and CCNA2, which control both the G1/S and the G2/M transition phases of the cell cycle [686]. ANXA2 also helps reduce inflammation in mononuclear cells by acting as an endogenous inhibitor of phospholipase A2 (PLA2) [687]. Therefore, it has been hypnotized that inhibition of ANXA2 could be a potential treatment for inflammation-related disease progression [688], [689].

After a longer time recovery, many proteins with regulated protein levels are involved in DNA damage repair (**Figure 24**). As we also observed, many proteins involved in cytokine signaling in the immune system are upregulated, such as Interferon-induced protein with tetratricopeptide repeats, IFIT1 and IFIT3, Interferon-induced transmembrane protein 3, IFITM3 (**Figure 24**). IFITMs can restrict the early stages of replication for a wide variety of viruses, including Influenza virus, West Nile viruses, Ebola viruses, and Coronavirus [690], [691], [692]. And the high level of IFITM3 could decrease the inflammatory response through the TLR4 signaling pathway [693]. It has been shown UV irradiation results in alterations in immune surveillance at the local and systemic levels. Cytokines expressed are reported to play an essential role in this process [694].

Many RNA-binding proteins are upregulated at 18h, such as spliceosome proteins HNRNPU and RBM10. However, some splicing factors are down-regulated, such as HNRNPA3 and SNRPG (**Figure 69**). It might be because they belong to different sub-families during splicing events [695], [696]. It has been shown that in fibroblasts derived from patients, C9orf72 repeats and nuclear RNA foci were accumulated upon reducing the HNRNPA3 level, which regulates DNA damage response [697]. C9orf72 arginine-rich dipeptide repeats inhibit mRNA decay through recruiting proteins and mRNA into granules [698]–[701]. In contrast, HNRNPU is reported to preferentially bind to sense G-quadruplex RNA foci, which is generated from hexanucleotide repeat expansions in C9orf72 [702], [703], [704].

3.1.4 Re-localization of chromatin-associated proteins induced by UV stress

Among the regulated chromatin-bound proteins, at early time recovery, we found proteins involved in GG-NER and TC-NER are recruited to chromatin and transcription factors that positively regulate transcription are removed from chromatin. PRDX4 is recruited to chromatin upon UV stress. PRDX4 is known to protect cells from oxidative stress [705], [706]. UV stress may induce ROS in the nucleus, leading to PRDX4 recruitment. Interestingly, translation initiation factors EIF4E2 and GIGYF2 are recruited to chromatin which will inhibit translation initiation (**Figure 27**C) [707]–[709]. This is also in line with the rapid transcription inhibition after UV irradiation. Moreover, 7SK component, HEXIM1 are rapidly removed from chromatin upon UV stress (**Figure 27**C). CDK7 is one of the kinases which phosphorylates Ser5 and Ser7 of the heptapeptide (YSPTSPS) of the RNA Pol II CTD, promoting promoter clearance and the initiation of transcription [710]–[713]. It has been shown that HEXIM1 inhibits transcription by interacting with P-TEFb, and the HEXIM1 protein levels gradually decrease in response to UV stress [296].

During transcription recovery, translation initiation complex and RNA processing factors are released from chromatin (**Figure 30**). Their phosphorylation status might regulate the removal of translation factors from chromatin since we did observe these proteins to be highly dephosphorylated, and some of them are phosphorylated at a later time when transcription is recovered. Thus we can speculate that the regulation of transcription and translation are coupled along the journey of recovery from UV-induced DNA damage.

At 18h, when transcription is recovered, proteins involved in the cell cycle process are recruited to chromatin and proteins involved in cell junction are removed from chromatin (**Figure 30**). Interestingly, RPS20 and PCNA are removed from chromatin (**Figure 69**). The Svejstrup lab has published that the ubiquitylation of these 2 proteins is upregulated after UV treatment. In addition, their ubiquitylation can be inhibited by the proteasome inhibitor MG132 [592]. So, the removal of RPS20 and PCNA from chromatin may be due to their proteasomal degradation. However, we did not observe a change in their protein level in our total proteome data. In addition, the mono-ubiquitylation and poly-ubiquitylation of PCNA has been reported to regulate the assembly of the multi-protein complexes that promote damage bypass pathways [714]–[716]. As for RPS20, ZNF598 was identified to mediate the regulatory ubiquitylation of RPS20. The deficiency of RPS20 ubiquitylation results in defective resolution of stalled ribosomes and subsequent readthrough of poly(A)-containing stall sequences [417]. The ubiquitylation of ribosome subunits plays a critical role in mammalian ribosome-associated quality control (QC) pathways [717]. RPS20 is reported to bind to MDM2 and inhibit MDM2 E3 ligase activity, resulting in p53 stabilization and cell cycle arrest. This RPS20-Mdm2-p53 signaling pathway is critical for ribosome biogenesis surveillance [718]. Therefore, we can speculate that UV stress induced ribosomal biogenesis stress.

3.1.5 Chromatin-bound proteomics strategy challenges

Chromatin contains about twice as many proteins as DNA [719]. Chromatin-bound proteins have crucial roles in various biological processes in the eukaryotic cells. In our chromatin-bound proteome profile studies, the extraction protocol has a potential bias [720]. This approach is prone to contamination by cytoplasmic proteins [721]. Also, with this approach, mitotic chromatin could not be caught due to its much lower abundance in the cell population. Recently, the Chromatin Enrichment for Proteomics (ChEP) method was established that relies on formaldehyde crosslinking before the biochemical extraction of the chromatin proteome. It is a relatively simple procedure to reduce the contamination of cytoplasmic proteins [722]. It has been successfully used to study chromatin in several cell types of different organisms [723], [724], [725]. However, chromatin remains a biochemically troublesome organelle due to its highly charged nature. ChEP approach still results in cytosolic contamination, such as mitochondrial proteins [726]. A more recent application of crosslinked chromatin enrichment, Density-based enrichment for MS analysis of chromatin (DEMAC) has been developed, including crosslinking nuclei, lysis, and subsequent ultracentrifugation of the crosslinked lysate in a buoyant density gradient produced from cesium chloride [727]. Nuclear isolation before crosslinking and keeping chromatin in a soluble state may result in less contamination from cytoplasmic proteins crosslinking to chromatin. However, its application may be limited due to the demand for ultracentrifugation. A more feasible alternative is proximity labeling which has been widely used for the purification of subcellular component proteomes. More specifically, we can apply the most recently established "split-TurboID" over two chromatin markers. In this method, when these two proteins are proximal, the enzymatic activity is reconstituted, resulting in biotinylation of the proximal proteins *in vivo*.

Additional modifications during sample preparation may also be considered. Incorporating a single-pot solid-phase sample preparation process into chromatin proteomics applications could be one of the promising ways [728]. This method employs paramagnetic beads to collect proteins and peptides in a single tube efficiently, minimizing sample loss.

Finally, the MS data collection could be improved as well. Data-independent acquisition (DIA) is a potential route to investigate. Data-dependent acquisition has been the standard approach for MS for a long time. Peptides are sequenced depending on the peptide abundance in data-dependent acquisition. As a result, very low–abundant peptides are frequently hidden by other high-abundant peptides [729], [730]. Because of the high abundance of histone proteins in chromatin, this could be a difficulty in MS-based chromatin investigations. Peptides within a defined m/z range are fragmented comprehensively with DIA, regardless of their abundance within that m/z range [731]. As a result, DIA should theoretically provide a more extensive analysis of peptides, making it more suitable for chromatin proteome investigation in MS-based experiments.

Although we observed some interesting findings by combining several transcriptome datasets with our proteome datasets to get a complete view of the cellular response to UV light, there are still cautions to be taken. The transcriptome datasets were generated in different cell lines, unlike our proteome datasets generated using HaCaT cells. The best option would be to perform the transcriptome analysis in HaCaT cells during recovery from UV treatment which is not available in public resources.

In summary, our findings provide the following insights for future research: The regulation of mRNA processing, modification and translation, R-loop regulation and ribosomal biogenesis are significantly and tightly regulated in response to UV stress.

3.2 Identification of JNK cascade substrates in response to UV stress

A comparison of our phosphoproteomics analyses with TMT-labeling and SILAC-labeling revealed that much fewer JNK-dependent phosphorylation sites were quantified in the TMT-labeling dataset in response to UV light. One of the reasons could be tracked to the dosage of UV light and the concentration of JNK1 inhibitor. Indeed, we did observe stronger JNK activity with a higher dosage of UV light. SILAC-labeled phosphoproteome was done with stronger UV light ($40J/m^2$) and a higher concentration of JNKi (5μM). It would bring much more off-targets of JNK downstream signaling identification. On the other hand, these two experiments were performed in different cell lines, HaCaT cells and U2OS cells. As natural cellular biological processes vary in different tissues, this may also be a factor in the difference we observed between the two datasets. Indeed, we observed very little overlap between these two datasets (**Figure 73**). Compared to the TMT-labeling phosphoproteomics analysis with JNK1 inhibitor, more proteins (206 proteins) with UV light-induced and JNK-dependent phosphorylation sites were identified as potential JNK1 substrates in the SILAC-based quantification. Among these proteins, some well-known JNK substrates, such as transcription factors ATF2 and c-Jun, were found as well (**Figure 45**D) [602], [603].

The phosphorylation of CPSF6 is inhibited in both cell lines after JNK1 activity inhibition, especially on the 407 site. A standard way to validate CPSF6 as a substrate of JNK1 would be to perform an *in vitro* kinase assay with JNK1 protein to see if it shows kinase activity on CPSF6. However, this method did not work in our hands because we could not observe any JNK1 activity *in vitro,* not even for the positive control, c-JUN. However, we can still test, albeit in a less direct way, if the phosphorylation at the T407 site on CPSF6 is the target site of JNK1 by checking the PLA signal between CPSF6 WT and JNK1 versus CPSF6 T407A mutant and JNK1 in HaCaT cells.

3.3 The regulation of phase separation in response to UV stress

Nuclear proteins forming condensates is an adaptive stress response. Especially, poly(A)-binding proteins act as a physiological stress sensor [740]. We observed that UV induced several RBPs' translocations, for example, RNA Pol II and SR protein, SRSF2. The phosphorylation on Serine residues of SR proteins regulates their intranuclear movement and pre-mRNA splicing functionalities [741], [742], [743]. RNA Pol II undergoing phase separation is reported to take place via binding of its C-

terminal domain to FUS, which plays a key role in several steps of RNA metabolism due to regulating phase separation in cells [744]–[748]. In addition, we found many RBPs' phosphorylations are highly regulated upon UV treatment, such as SRSF2. It can be speculated that its translocation may have resulted from its phosphorylation. As we mentioned above, PTMs play an essential role in the formation of phase separation [37], [749]. Phase separation has been observed in response to double-strand breaks at DNA lesion sites [619], [750]–[755]. Nuclear proteins forming condensates is an adaptive stress response. Especially, poly(A)-binding proteins act as a physiological stress sensor [740]. We observed the global re-organization of RBPs in the nucleus after UV irradiation. We speculate that the formation of these nuclear granules is an adaptive system for cells to store proteins in condensed speckles during the DNA repair process. These speckles may protect proteins from degradation or mis-translocation into the cytoplasm, and the sequestering of the RBP proteins results in a clean and spatially more available environment for other enzymes to access DNA lesions. Moreover, the re-localization of splicing factors also indicates that the regulation of splicing events and splicing complexes re-organization in response to UV irradiation. UV stress induced paraspeckles formation. Paraspeckles are initiated by the specific long non-coding RNA (lncRNA) of NEAT1 (nuclear paraspeckle assembly transcript 1). Paraspeckles have been widely reported to be a storage of retained RNA in the nucleus [54], [61]–[63]. It suggested that UV induced-transcription blocking probably subsequently resulted in the retaining of RNA in the nucleus, such as non-coding RNA.

In addition, the formation of paraspeckles also highlighted the importance of lncRNA. However, to date, the role of lncRNA in response to DNA damage has been poorly studied. lncRNA has been reported to be induced by DNA damage through the recruited RNA polymerase II at double-strand break sites and plays an essential role in DNA damage repair foci formation, such as 53BP1 [756]. It might be valuable in future research on the function of lncRNA in response to transcription blocking lesions.

3.4 The regulation of APA by CPSF6 in response to UV stress

The cleavage factor I mammalian (CFIm) complex is composed of cleavage and polyadenylation specificity factor 5/7 (CPSF5/7) and serine/arginine-like protein CPSF6. It regulates alternative polyadenylation (APA) through the interaction of CPSF6 and FIP1L1 with mRNA downstream of cleavage sites. Loss of CFIm function results in higher usage of proximal polyadenylation sites and shortening of mRNA 3' untranslated regions (UTRs) [346], [757]–[761].

Our study found that another phosphorylation site on T407 near the proline-rich domain is crucial for CPSF6 localization in the nucleus in response to UV light. This translocation was the key effect on APA regulation under UV stress. Moreover, this translocation is mediated by liquid-liquid phase separation. We found FUS was induced to interact with CPSF6 upon UV treatment. It seems that UV irradiation enhanced the phase separation of CPSF6, which is potentially driven by FUS. This observation still needs to be studied further by checking the localization of CPSF6 after FUS depletion.

However, FUS depletion may be problematic to study because of all the complex biological processes FUS is involved in.

The C-terminal RS domain of CPSF6 is a binding domain for the β-karyopherin transportin 3 (TNPO3), contributing to its nuclear localization. CPSF6 T407A mutant displayed a mis-translocation into the cytoplasm. The mutation may have an effect on the C-terminal nuclear localization sequence (NLS). The arginine/serine (RS)-like domain (RSLD) of CPSF6 that mediates TNPO3 binding is critical for its nuclear import. TNPO3 might be engaged in the cytoplasm for nuclear transport when CPSF6 is phosphorylated at the 407 site. However, we did not observe TNPO3 interacting with CPSF6 in our interactome. This can be investigated by performing *in vivo* or *in vitro* interaction assays to test the interaction between CPSF6 or its phosphomutants and TNPO3 in the future.

We observed the wild-type CPSF6 cells have an increased APA activity post UV stress, whereas it is not observed in phosphorylation mutant on 407. However, without UV treatment, we observed contrasting results. CPSF6 mutants may act as a competitor with endogenous CPSF6 and interfere with its function in cells (**Figure 59**C). As we showed, the phosphorylation on 407 of CPSF6 mainly affects its re-localization into paraspeckles. Combining this information with the current data highlights the importance of CPSF6 re-localization in APA activity regulation during the response to UV light.

UV stress induced CPSF6 less interacting with RNA Pol II. CPSF6 knockdown affects transcription and transcription-related DNA damage foci 6,4-PP removal efficiency after UV stress. We could speculate that disassociation of phosphorylated CPSF6 from RNA Pol II, subsequently affects the stability of RNA Pol II at DNA damage sites. These reactions will result in the persistent RNA Pol II at DNA damage sites, thus impairing the transcription speed and the removal of UV stress-induced damage. We can check if the phosphorylation of CPSF6 plays a role in the interaction with RNA Pol II with PLA or immunoprecipitation in CPSF6 wild-type and phosphomutant cells.

CPSF6 knockdown also induced more short-mRNA production. It might be due to the lowered level of CPSF6 interfering with the interaction between FIP1L1 and mRNA [772]. It could also result from the altered phosphorylation status of CDK9 and RNA Pol II, which is regulated by protein phosphatase 2A (PP2A) stability as reported in CD4+ cells [773]. The phosphorylation of RNA Pol II also alters the production of ASCC3 short mRNA [774]. So we can check whether the depletion of CPSF6 also induces differential phosphorylation levels of RNA Pol II in HaCaT cells.

To get a global idea of polyadenylation events during the response to UV light, with an emphasis on the function of CPSF6, we plan to perform nascent RNA sequencing. An example for one experimental set-up would be performing 3T-seq in wild-type cells and CPSF6 phosphorylation mutant cells in response to UV light, which allows us to profile the transcriptome-wide APA sites.

It has been suggested that CPSF6 has a function in transcription during development, as mutations in the gene cause gene expression to be disrupted in embryos [610], [775], [776]. More recently, CPSF6 has been reported to play a role in HCC (hepatocellular carcinoma) progression by regulating the APA

of *NQO1,* which is a primary liver malignancy with few therapeutic options [346]. However, it is not revealed how CPSF6 regulates the APA in HCC. In addition, the JNK pathway has also been linked to HCC with double face roles: Tumor-suppressing and tumor-promoting [777], [778]. Recently, JNK inhibition has been reported to reduce LAMP2A (lysosomal-associated membrane protein 2A) stability in HCC cells specifically but not in immortalized liver cell lines, thus increasing the toxicity of lysosomotropic agents. This cancer-specific effect suggested that JNK1 inhibition could be for therapeutic treatment for liver cancer [779]. Our study could support another possibility that the potential function of JNK in HCC progression could have another module: JNK phosphorylates CPSF6 to regulate the APA in HCC, although we found this JNK-CPSF6 pathway in skin cells.

3.5 The R-loop formation regulation in response to UV stress

RBPs play a key role in preventing R-loop formation during transcription in eukaryotic cells, for instance, SR rich proteins SRSF1, SRSF2 and SRSF3 [780], [781], [782]. We also observed the phosphorylation alter of RBPs upon UV induction, for example, SRSF2 and SETX1. How their phosphorylation regulates R-loop formation or resolving during the response to UV-induced DNA damage still needs to be further characterized in the future.

A very recent study has found that elongating RNA Pol II accumulates together with RNA:DNA hybrids at sites of head-on transcription-replication collisions [783]. We observed R-loop formation is triggered upon UV treatment via PLA assay of PCNA and RNA Pol II. These R-loop foci co-localize with RNA Pol II and phosphorylated RNA Pol II at Ser2, promoting its elongation at 6h recovery after UV irradiation when transcription restarts, where transcription machinery might run into the replication fork. To further investigate which proteins are also spatially involved in the local R-loop formation or resolving during this process induced by UV irradiation, we can apply the established "Split-TurboID-MS" over PCNA and RNA Pol II, and combine it with TMT-labeling to capture the complexes at these transcription-replication conflict sites after different recovery times post UV treatment (**Figure 67**).

4 Materials and Methods

4.1 Lists of consumables, equipment, and software

Table 1: Buffers/Solutions/Consumables

Cell culture	**Composition / vendor**
Dialyzed FBS (10,000 molecular weight cut-off)	Thermo Fisher Scientific
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 keratinocytes cells (HaCaT)	ATCC
Human bone osteosarcoma epithelial cells (U2OS)	ATCC
Human embryonic kidney cells (HEK293/T)	ATCC
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
L-glutamine	Gibco
Penicillin/Streptomycin	Gibco
Polybrene	Sigma Aldrich
Puromycin	InvivoGen
Trypsin-EDTA (0.05%)	Gibco
Inhibitors and agents	
1,6-Hexanediol	Sigma Aldrich
4-nitroquinoline (4NQO)	Sigma Aldrich
5-Ethynyl-uridine (5-EU)	Jena Bioscience GmbH
Actinomycin D	Cell Signaling Technology
Ammonium acetate (NH_4OAC)	Sigma Aldrich
Biotin	Sigma Aldrich
Biotin phenol	Iris Biotech
5,6 dichloro 1 β D ribofuranosylbenzimidazole (DRB)	Sigma Aldrich
Hydrogen peroxide (H_2O_2)	Sigma Aldrich
JNK1-IN-8	Selleckchem
Transfection	
Linear polyethylenimine transfection (PEI, HCl Max, 40000)	Polysciences, Inc.
Lipofectamine RNAiMAX	Life Technologies
Opti-MEM with GlutaMAX	Gibco
Colony formation, cell proliferation, and comet assays	Gibco
0.4% Crystal violet	Sigma Aldrich
CellTiter-Blue Cell Viability Assay	Promega
CometAssay 2 Well ES Unit w/ Starter Kit	Trevigen
SYBR gold Nucleic Acid Gel stain	Thermo Fisher Scientific

Cell lysis	
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
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
Cell fractionation	
Fractionation buffer A	10 mM HEPFES pH 7.5, 10 mM KCl, 1.5 mM MgCl2, 0.34 M glucose, 10% glycerol, 1 mM DTT, 0.1% Triton-X100
Fractionation buffer B	3 mM EDTA, 0.2 mM EGTA, 1 mM DTT
Pull-downs	
GFP Trap agarose	Chromotek
Protein A/G beads	
Strep-Tactin Sepharose	IBA lifesciences
SDS-PAGE and western blotting	
0.45 µm nitrocellulose	Sigma Aldrich
Blocking buffer	10% skimmed milk solution 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) methanol, pH 8.3
Microscopy and flow cytometry	
4% paraformaldehyde (PFA) in PBS	Affymetrix
7-AAD Viability staining solution	eBioscience
Alexa Fluor 647 azide	Thermo Fisher Scientific
Blocking buffer	5% FBS in PBS-T
HCS CellMask Red stain	Thermo Fisher Scientific
Hoechst 33342	Thermo Fisher Scientific
Genomics	
RNase free water	Sigma Aldrich
Copper (II) sulfate pentahydrate in PBS	Sigma Aldrich
Fixation solution	4% FA in PBS
NeutraAvidin beads	Thermo Fisher Scientific
Permeabilization buffer	0.1% Trion-X100 in PBS

Sodium L-ascorbate	Sigma Aldrich
Washing buffer	10 mM Tris-HCl pH 8.0, 200 mM NaCl, 0.5 mM DTT
Production of recombinant proteins	
D-Tube Dialyzer tubes	Sigma Aldrich
Glutathione Agarose	Thermo Fisher Scientific
Lysis buffer protein purification	50 mM Tris-HCl PH8.0, 300 mM NaCl, 2 mM EDTA, 0.1% TritonX-100, 10% Glycerol, 1 mM DTT (add freshly prepared)
Ni-NTA beads	Qiagen
Rosetta2(DE3)pLysS	Novagen
DNA cloning	
Dh5α	Thermo Fisher Scientific
Dimethyl sulfoxide (DMSO)	Sigma Aldrich
Dpn1	NEB
Electro ligase	NEB
Gateway LR Clonase II Enzyme Mix	Invitrogen
High efficiency DH5α E.coli	NEB
NucleoBond Xtra Midi	Macherey-Nagel
Plasmid Mini Kit	Qiagen
Q5 MasterMix	NEB
QIAquick PCR & Gel Cleanup Kit	Qiagen
Mass spectrometry	
In-gel digestion	
Buffer B	80% ACN, 0.5% acetic acid
Chloroacetamide (CAA)	Sigma Aldrich
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)	Sigma Aldrich
In-solution digestion	
Denaturation buffer	6 M urea, 2 M thiourea in 10 mM HEPES-NaOH pH 8.0
Lysyl endopeptidase (Lys-C)	Wako Chemicals
SepPAK C18 cartridges	Waters
Stage tipping	
Buffer A	0.1% formic acid
Buffer B	80% ACN, 0.1% formic acid
C_{18} elution buffer	50% ACN, 0.1% formic acid
C_{18} 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
Phosphopeptide enrichment	
Phospho binding buffer	50% ACN, 6% TFA

Phospho elution buffer 1	5% NH4OH
Phospho elution buffer 2	10% NH4OH, 25% ACN
Phospho wash buffer	50% ACN, 0.1% TFA
Titansphere TiO bulk material (TiO2) 10 µm	GL Sciences Inc
TMT	
hydroxylamine	Sigma Aldrich
TMT labeling buffer	150 mM HEPES pH 8.5, 30% ACN
TMT10plex Isobaric Label Reagent Set plus TMT11-131C Label Reagent	Thermo Fisher Scientific

Table 2: Antibodies

Antibodies	Product number	Manufacturer	Dilution	Origin
γH2AX	A300-081A-M	Bethyl	1:1000	rabbit
IgG Alexa Fluor 488, 568	A11001, A11004	Life Technology	IF 1:1000	mouse
IgG Alexa Fluor 488, 568	A11008, A11011	Life Technology	IF 1:1000	rabbit
pCHEK1 (S345)	2344	Cell Signaling Technologies	1:1000	rabbit
pCHEK2 (T68)	2661	Cell Signaling Technologies	1:1000	rabbit
pP38 MAPK (Thr180/Tyr182)	9216	Cell Signaling Technologies	1:1000	mouse
pJNK1 (Thr183/Tyr185)	9255	Cell Signaling Technologies	1:1000	mouse
pMAPKAPK-2 (MK2) (Thr334)	3007	Cell Signaling Technologies	1:1000	rabbit
JNK1	1648	Santa Cruz Biotechnology	IF 1:1000	mouse
CPSF6	376228	Santa Cruz Biotechnology	1:500	mouse
CPSF6	ABIN524444	antibodies-online.com	IF 1:200	mouse
CPSF6	HPA039973	Sigma Aldrich	IF 1:1000	rabbit
PSPC1	16714-1-AP	Proteintech	1:500	rabbit
XPC	14768	Cell Signaling Technologies	1:1000	rabbit
RNA Pol II (CTD4H8)	47701	Santa Cruz Biotechnology	IF 1:1000	mouse
RNA Pol II Ser 2 (E1Z3G)	13499	Cell Signaling Technologies	IF 1:1000	rabbit
PCNA	sc-56	Santa Cruz Biotechnology	IF/WB 1:1000	mouse
CPD	CAC-NM-DND-001	Hölzel Diagnostika Handels GmbH	IF 1:500	mouse
6,4-PP	NMDND002	Hölzel Diagnostika Handels GmbH	IF 1:500	mouse
CPSF5	81109	Santa Cruz Biotechnology	IF/WB 1:1000	mouse
CPSF7	393880	Santa Cruz Biotechnology	IF/WB 1:1000	mouse
SC35/ SRSF2	556363	BD Bioscience	IF 1:1000	mouse
pSF3B1 (Thr313)	25009	Cell Signaling Technologies	IF 1:1000	rabbit
U2AF65	37530	Abcam	IF 1:1000	rabbit
S9.6	ENH001	Kerafast	Dotblot 1:10000	mouse
dsDNA	27156	Abcam	Dotblot 1:10000	mouse
HBD/WKK		In house	IF 1:2500	
GFP	sc-9996	Santa Cruz Biotechnology	1:1000	mouse
FLAG M2	F1804	Sigma Aldrich	1:2000	mouse
HA	7392	Santa Cruz Biotechnology	1:1000	mouse
H2A				

AQR	A302-547A	Bethyl	1:2000	rabbit
Streptavidin-HRP	21130	Thermo Scientific	1:1000	
Vinculin	V9264	Sigma Aldrich	1:10000	mouse
β-Actin	A2228	Sigma Aldrich	1:10000	mouse
Secondary antibodies coupled to horseradish peroxidase		Jackson ImmunoResearch Laboratories	IF 1:5000	mouse, rabbit

Table 3: Plasmids

Plasmids	**Source**
pcDNA-DEST47	Life Technologies
pcDNA-DEST53	Life Technologies
pENTR221-CPSF6	IMB
AdML-M3	Addgene
pLX304 CMV FKBP-V5-sTurboID (N)	Addgene
pLX304 CMV FKBP-V5-sTurboID-high affinity (N)	Addgene
pLX304 CMV HA-HaloTag-FRB-sTurboID (C)	Addgene
pLX304 CMV HA-HaloTag-FRB-sTurboID-high affinity (C)	Addgene
Lentivirus production	
pMD2.G	Addgene
psPAX2	Addgene
plix-402	Addgene
plix-Nterm-GFP	This study
plix-Cterm-GFP	This study
Protein production	
pDEST pET-His-MBP-3C	IMB core facility (Protein production)
pMBP-MS2	Addgene 65104

Table 4: Primers

Primers	**Sequence 5'-3'**
PUC19	Fwd – GTACCGAGCTCGAATTCACTGGCCGTCGTT Rev – TGCAGGCATGCAAGCTTGGCGTAATCATGG - [BtnTg]
CPSF6 T404A mutagenesis	Fwd – Phos- CTGGAAGGGAAATGGATGCTGCAAGAACGCCATTG Rev – Phos- CTTCACTCAATGGCGTTCTTGCAGCATCCATTTCC
CPSF6 T407A mutagenesis	Fwd – Phos- CTGGAAGGGAAATGGATACTGCAAGAGCGCCATTG Rev – Phos- CTTCACTCAATGGCGCTCTTGCAGTATCCATTTCC
CPSF6 T404/407A mutagenesis	Fwd – Phos- GGAAATGGATGCTGCAAGAGCGCCATTGAGTGAAG Rev – Phos- CTTCACTCAATGGCGCTCTTGCAGCATCCATTTCC
CPSF6 siRNA1 insensitive mutation	Fwd – Phos- ACGAGAAAAGAGTCGACGACACAAATC Rev – Phos- GTCATGACGGTCTCTACTTCGCGATTT
GAPDH- qPCR	Fwd –TTCTCCATGGTGGTGAAGACG Rev –ATTCCATGGCACCGTCAAGG

ASCC3 long isoform- qPCR	Fwd – CCTCCATCGAGTCCCTTCCT Rev – AACTAAGTCATCCCACGAGCC
ASCC3 short isoform- qPCR	Fwd – TGGGCCTGACATGGAAGAAG Rev – P ATCCAGCCTCCTCTTCTGGA
L3-oligo	Fwd – CCACTTCTTTTTGTCACTTGAAAAACATGTAAAAATAATGTACTAGGAGAC ACTTTCAATAAAGGCAAATGTTTTTATTTGTACACTCTCGGGTGATTATTTACCC Rev – GGGTAAATAATCACCCGAGAGTGTACAAATAAAAACATTTGCCTTTATTGA AAGTGTCTCCTAGTACATTATTTTTACATGTTTTTCAAGTGACAAAAAGAAGTGG
AdML-M3-L3-gibson -fragment	Fwd – aaactcttcgcggtctttccCCACTTCTTTTTGTCACTTGAAAAACATGTAAAAATAATG Rev – gtacggatatcggatccaagGGGTAAATAATCACCCGAGAGTGTACAAA
AdML-M3-L3-gibson -vector	Fwd – TCTCGGGTGATTATTTACCCcttggatccgatatccgtacacc Rev – CAAGTGACAAAAAGAAGTGGggaaagaccgcgaagagtttg
pDEST-47-FKBP-V5-TurboID(N)-Flag-gibson-fragment	Fwd – acaaggacgatgacgacaaatcgaaggagatagaaccatgcagg Rev – gctttttgtacaaacttgtcagaatctgtttagcgttcagcagc
pDEST-47-FKBP-V5-TurboID(N)-Flag-gibson-vector	Fwd – tgaacgctaaacagattctgacaagtttgtacaaaaaagctgaacgag Rev – catggttctatctccttcgatttgtcgtcatcgtccttgtaatcg
pDEST-47-HA-FRB-TurboID (C)-gibson-fragment	Fwd – ggctagacaccatggccagcatgtacccatacgacgttccag Rev – gctttttgtacaaacttgtctttcggcagaccgcagac
pDEST-47-HA-FRB-TurboID (C)- gibson-vector	Fwd – gtctgcggtctgccgaaaagacaagtttgtacaaaaaagctgaacgagaaacgtaaaatga Rev – ggaacgtcgtatgggtacatgctggccatggtgtctagc

Table 5: siRNAs

siRNAs	**Sequence 5'-3'**
CTR	UGGUUUACAUGUUGUGUGA-TT
CTR pool	Dharmacon pool
CPSF6-1	CGUCAUAAAUCCCGUAGUA -TT
CPSF6 pool	Dharmacon pool
XPC	GCAAAUGGCUUCUAUCGAA -TT
AQR	CUGAAUAUGGCGGUGUAGU
JNK1	Dharmacon pool

Table 6: Software

Software
Adobe Illustrator CC2021
Clustal Omega (EMBL-EBI)
Cytoscape version 3.8.2
SnapGene V5
Harmony High-Content Imaging and Analysis Software (PerkinElmer)
IceLogo (University Ghent)
MaxQuant v1.5.2.8 / v1.6.14.0 (Cox lab)
Perseus 1.6.14.0 (Cox lab)

QuikChange Primer Design II tool (NEB)	
R studio v1.2.1335 / R v3.6.1	
Thermo Xcalibur 3.0.63 (Thermo Fisher Scientific)	

Table 7: Machines

Maschines	**Vendor**
AF7000	Leica
Biometra TRIO Thermal Cyclers	Analytikjena
VisiScope Spinning Disk Confocal Microscope	Nikon
ChemiDoc imaging system	BioRad
Easy-LC-1000	Thermo Fisher Scientific
Easy-LC-1200	Thermo Fisher Scientific
GraphPad prism 8	Prism
NanoDrop 2000	Thermo Fisher Scientific
NuPage Novex Gel System	Thermo Fisher Scientific
Opera Phenix High Content Screening System	PerkinElmer
Orbitrap Exploris 480 mass spectrometer	Thermo Fisher Scientific
Plate reader infinite m200	Tecan
Q Exactive Plus	Thermo Fisher Scientific
Sonifier 450	Branson
Thermo Scientific 3311 Forma Steri-Cult CO_2 Incubator	Eppendorf
Thermoshaker	Eppendorf
UV-C irradiator	Inhouse built
Vacufuge Plus	Eppendorf
Xcell II Blot-Modul	Thermo Fisher Scientific

4.2 Cell culture

4.2.1 Cell culture

ATCC provided human osteosarcoma cells (U2OS), human epidermal keratinocyte cells (HaCaT) and HEK293T (human embryonic kidney 293). Cells were cultured in D-MEM medium supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 100 U/ml Penicillin/Streptomycin. Puromycin was also added to stable cell lines at a concentration of 1 g/l. For experiments based on SILAC labeling, Cells were cultured in medium containing either L-arginine and L-lysine (Light), L-arginine [^{13}C6] and L-Lysine [^{2}H4] (Medium), or L-arginine [^{13}C6,^{15}N4] and L-lysine [^{13}C6,^{15}N2] (Heavy). Cells were washed in PBS, detached with 0.05% trypsin, and resuspended in complete D-MEM medium to stop the enzymatic activity. This was followed by spinning down at 250 x g for 5 minutes and re-plated to the desired confluence. All cells were grown in a humidified incubator with 5% CO^2 at 37 °C. In addition, when cells are needed to introduce genotoxic treatment. Inhibitors (e.g., JNK1) were added to cells before UV irradiation, as indicated in the main text. UV light irradiation was performed using the in-house built UV chamber with a detector inside.

4.2.2 Transfection of cells

For siRNA transfection of cells in a 10cm dish, cells were grown to an 60% confluence in 10 ml complete D-MEM before being transfected. Following that, 8 µl of siRNA (10 µM) and 10 µl siRNAMax were diluted in 500 µL Opti-MEM. Similarly, for DNA transfection, 5 µg of plasmid DNA and 15 µL polyethylenimine (PEI) was diluted in 500 µL Opti-MEM. Respectively, The mixtures were combined into a master mix and incubated for 15 minutes at room temperature. They were then added to the cells for 6 hours before the medium was changed to fresh D-MEM. The cells were used for experiments 48 to 72 hours after being transfected with DNA or siRNA, respectively. Transfection volumes in different culture dish sizes were scaled based on surface area depending on the experimental demand.

4.2.3 Stable cell line production by lentivirus transduction

For lentivirus production, the virus is produced in HEK293T cells which were co-transfected in a 4:3:2 w/w/w ratio with an expression plasmid (pliX-GFP-CPSF6 WT, T404A, T407A, or T404A/T407A) and the packaging plasmids psPAX2 and pMD2.G. The supernatant containing the virus was collected 72 hours after transfection, filtered through a 0.45 m filter, mixed with the same volume of fresh medium, and supplemented with polybrene at an 8 µg/L concentration. U2OS or HaCaT cells were transduced by exchanging medium for conditioned supernatant. The virus was removed 24 hours later, and the cells were incubated for another 48 hours before stable cells were selected by adding 1 µg/L puromycin for 7-10 days until all control cells were dead.

4.3 Methods for DNA modification and analysis

4.3.1 Gateway cloning

The pliX-GFP-expression vectors were generated from Destination vectors and Entry / DONR vectors using the Gateway LR Clonase II Enzyme Mix, as directed by the manufacturer. The vectors used in this work are listed in 4.1 (

Table 3).

4.3.2 Site-directed mutagenesis

The QuikChange Primer Design II tool was used to create the primers used for site-directed mutagenesis. Mutagenesis of the pENTR221 CPSF6 was carried out using the PCR conditions listed below. The resulting linear plasmids were column purified using the QIAquick PCR & Gel Cleanup Kit after being treated with 20 U of Dpn1 for 1 hour at 37 °C. The plasmids were ligated using ElectroLigase according to the manufacturer's instructions.

PCR – MIX	
2x Q5 Master Mix	25 µl
Primer Mix	0.5 µM
100 ng Plasmid DNA	100 ng
H_2O	Fill up to 25 µl

Temperature	**Time**	**Cycles**
98°C	30 sec	1
98°C	10 sec	30
55 °C	30 sec	30
72°C	30 sec / kb	30
72°C	5 min	1
12°C	hold	

4.4 Cell-based methods

4.4.1 Colony-forming assay

The respective siRNAs were transfected into the cells as indicated in the main text. Doxycycline was added at a concentration of 1 μg/L to doxycyclin-inducible cell lines. The following day, 4,000 to 12,000 cells were re-seeded in 6-well plates. The cells were irradiated or treated with JNK1 inhibitor for 1 hour before being washed twice with PBS, and the medium was replaced with fresh medium. Cells were grown for 10-14 days after treatment. Then surviving cell colonies were stained with crystal violet solution and counted under the microscope. Each experiment had three technical replicates and three biological replicates.

4.4.2 Cell viability assay

2,500 cells were seeded per well in 96-well plates with a volume of 100 μL. The following day, cells were treated with JNK1 inhibitor for indicated time or irradiated before the medium was replaced. Cell viability was tested after treatment by adding 20 μL of CellTiter-Blue Cell Viability solution for 3 hours. The Tecan Plate Reader was used to measure bioluminescence with the following settings: excitation wavelength 560 nm, emission wavelength 590 nm, number of flashes 25, and integration time 20 μs. Regression analysis in GraphPad Prism (v8.0.2) was used to calculate the LC50 value.

4.4.3 Immunofluorescence and confocal microscopy

For microscopy experiments with Opera Phenix High Content Screening System (PerkinElmer), cells were grown in 10 cm dish and transfected with indicated siRNAs. Then 20,000 were reseeded in 96 well plates (Perkin Elmer 6005550) that were coated with poly-L-lysine (PLL, Sigma-Aldrich). The next day, cells were treated with UV irradiation or pre-treated with JNK1 inhibitor. Cells were pre-extracted with 0.4% NP40 for 30 minutes, then fixed in 4% formaldehyde for 15 minutes, and washed three times with PBS. Cells were treated with blocking buffer (3% BSA in PBS) for 1 hour before incubation with primary antibodies overnight. Cells were then washed three times in PBS containing 0.1% Tween-20 (PBS-T) and were incubated for 1 h with fluorescently labeled secondary antibodies. After three PBS-T washes, DNA was stained with Hoechst 33342 (Sigma-Aldrich).

Images were acquired with Opera Phenix High Content Screening System (PerkinElmer). A 40X water immersion objective (numerical aperture (NA) 1.1) was used. Quantification was analyzed by automated “Harmony High Content Imaging and Analysis Software” (PerkinElmer). Cells in the periphery of the image were excluded from further analysis. Custom-made R scripts (https://www.R-project.org/) and GraphPad Prim (v8.0.2) were used to visualize the data after image analysis in Harmony. Significance values were calculated with one-way ANOVA followed by post-hoc Sidak’s multiple comparisons test in GraphPad Prism (v8.0.2).

4.4.4 Proximity ligation assay

Proximity Ligation Assay was performed according to the manufacturer´s protocol (Duolink®, Sigma-Aldrich). Cells were fixed with 4% paraformaldehyde in PBS and permeabilized with 0.25% Triton X-100. Samples were blocked with Duolink® Blocking Solution for 1 hour at 37°C in a humidity chamber. After removal of the blocking solution, primary antibodies diluted in Duolink® Antibody Diluent were added on the coverslips for 2 hours at room temperature in a humidity chamber. Coverslips were washed 2× with Washing Buffer A. PLA plus and minus probes were put on in a 1:5 dilution in Duolink® Antibody Diluent for 1 hour at 37°C in a humidity chamber. Two washes with Washing Buffer A were followed by Ligase treatment in 1× Ligation Buffer for 30 min at 37°C in a humidity chamber. Ligation buffer was tapped off and coverslips were washed twice with Washing Buffer A. Amplification was achieved by adding the Polymerase in 1× Amplification buffer for 100 min at 37°C in a humidity chamber. After washing the samples 2× with 1× Washing Buffer B and 1× with 0.01× Washing Buffer B, coverslips were stained with 1 µg/ml Hoechst33342 and mounted using Dako mounting medium. Images were taken with a Leica SPE microscope using a 63× 1.4NA oil objective. The number of PLA spots per nucleus was quantified using Fiji/ImageJ (v1.53).

4.4.5 Cell-cycle profiling

Cells were treated with UV (15J/m^2) and left for indicated time before being treated with EdU at a concentration of 10 µM for 1 hour. The cells were harvested and washed with PBS before being fixed for 10 minutes in 4% PFA and the click reaction was performed using Alexa Fluor 647 Azide. Following that, the DNA was stained with either Hoechst (1 µg/mL). Subsequent measurements on the Opera Phenix High Content Screening System were used for cell cycle analysis by microscopy (see 4.4.3). Cell cycle profiles were derived by calculating the Hoechst nuclear intensity sum per nucleu combining with EdU intensity. Custom-made R scripts (https://www.R-project.org/) and GraphPad Prism (v8.0.2) were used to visualize the data after image analysis in Harmony.

4.4.6 Fluorescence recovery after photobleaching (FRAP)

We used established stable cell lines expressing GFP-tagged CPSF6 in HaCaT cells. It is carried out on a confocal microscope (VisiScope Spinning Disk Confocal Microscope) equipped with the necessary laser. Individual GFP-CPSF6 foci at different time points after UV irradiation are photobleached with a laser and then the fluorescence in this region is monitored as it recovers. The fluorescence recovery depends on the molecular dynamics of the phase-separated liquid droplets. If the phase-separated droplets are liquid-like, the fluorescence recovery is quick. Quantification was analyzed in VisView and Imar. GraphPad Prism (v8.0.2) was used to visualize the data after image analysis.

4.4.7 Live imaging

An established stable HaCaT cell line expressing GFP-CPSF6 was plated in a chambered coverslip with 8 individual wells with high glass bottom (ibidi, #80807). Cells were irradiated with UV light (40J/m^2). Immediately, cells were monitored for 24 hours at 37°C with the presence of 5% CO_2,

under a confocal microscope (VisiScope Spinning Disk Confocal Microscope) with the setting: 6 detection views, every 30 seconds of detection frequency, Fluo-confocal-48 and brightfield. Quantification analysis of foci intensity, size and numbers were done in Imaris and the background intensity was removed and visualized in GraphPad Prism (v8.0.2).

4.4.8 1,6-hexanediol and ammonium acetate (NH_4OAc) treatment

Cells were plated on cover slides or 96-well plates 96 well plates (Perkin Elmer 6005550) for 24 hours. The next day, cells were treated with UV irradiation (20J/m^2) and followed by with or without 5% of 1,6-hexanediol treatment for 30 minutes or 50 mM NH_4OAC for 2 minutes. Cells were fixed in 4% formaldehyde for 15 minutes and washed three times with PBS, followed by the permeabilization with 5% of Triton X-100 for 5 minutes. Cells were treated with blocking buffer (3% BSA in PBS) for 1 hour before incubation with respective primary antibodies overnight. Cells were then washed three times in PBS containing 0.1% Tween (PBS-T) and were incubated for 1 h with fluorescently labeled secondary antibodies. After three PBS-T washes, DNA was stained with Hoechst 33342 (Sigma-Aldrich).

Images were acquired with AF7000 Widefield Fluorescence Microscope or Opera Phenix High Content Screening System (PerkinElmer). Quantification was analyzed by Fiji/ImageJ (v1.53) or automated "Harmony High Content Imaging and Analysis Software" (PerkinElmer) as described as above.

4.4.9 Cell lysis

Cells were washed twice with ice-cold PBS and lysed with modified RIPA buffer supplemented with the desired inhibitors (protease inhibitor, phosphatase inhibitors, 10 μM N-ethylmaleimide (NEM)). Additional NaCl was added to samples at a concentration of 500 mM and sonicated. The lysates were centrifuged at 16.000 x g for 15 minutes at 4 °C, and protein concentrations were determined using the QuickStart Bradford Protein assay.

4.4.10 Cell fractionation

Cells were grown in 10cm dish and collected by scraping them in ice-cold PBS and spinning them down at 300 x g and at 4 °C. Pelleted cells were lysed in 500 μL fractionation buffer A, incubated for 15 minutes at 4 °C with rotation, and spun down at 1,300 x g for 5 minutes. The nuclear pellet was washed once with 500 μL buffer A before being resuspended in 500 μL fractionation buffer B supplemented with 1 μM DTT and rotated for 30 minutes at 4 °C. The chromatin fraction was pelleted for 5 minutes at 1,700 x g and washed once with 500 μL buffer B. Finally, the chromatin fraction was resuspended in RIPA buffer (500 mM NaCl), sonicated, and centrifuged at 16,000 x g and at 4 °C for 15 minutes. All buffers were supplemented with inhibitors (protease inhibitor, phosphatase inhibitors, 10 μM N-ethylmaleimide (NEM)).

4.4.11 SDS-PAGE and western blotting

Protein samples in LDS buffer were separated on NuPAGE Bis-Tris gels with a 4-12% gradient at 150 V in MOPS-SDS running buffer. Proteins were transferred onto a 0.45 m nitrocellulose membrane

in transfer buffer for 115 minutes using the Xcell II Blot-Modul at 30~45 V. All subsequent steps were performed on a shaking platform, and the membrane was washed three times with PBS-T buffer in between. PONCEAU staining of the membrane was used first to confirm the transfer quality. The membrane was then incubated in blocking buffer (10% milk in PBST or 5% BSA in PBST) for 1 hour at room temperature (RT) before being incubated with primary antibody diluted in blocking buffer overnight at 4 °C. Finally, the membrane was incubated with a secondary antibody conjugated to horseradish peroxidase (1:5000 diluted in 5% milk in PBST) for 1 hour at RT. Images detection was achieved in the ChemiDoc imaging system in combination with the SuperSignal West Pico Chemiluminescent Substrate.

4.4.12 Expression and purification of recombinant proteins

Rosetta (DE3) cells were transformed with the vectors pMBP-MS2 (Addgene 65104) or pDEST-pET-His-MBP-3C-CPSF6. A single colony was picked for inoculating overnight shaking at 37 °C in LB medium supplemented with the appropriate antibiotic. The bacteria were diluted 1:100 and grown at 37 °C in an antibiotic-containing LB medium while shaking until an optical density of 0.5 was achieved. Protein expression was stimulated with 0.4 mM IPTG, and cells were shaken at 16 °C for 20 hours before being collected at 4000 xg for 10 minutes. After resuspending the pellet in lysis buffer (30 mM Tris-Cl pH 8.0, 500 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 5% glycerol, 100 U/ml Benzonase, EDTA-free cOmplete protease inhibitor cocktail, 15 mM imidazole), it was sonicated for 5 minutes on ice (5 mm tip, 20 cycles, output 5). The lysate was cleared for 30 minutes at 21,000 g. Recombinant proteins were affinity-purified from cleared lysates using a NGC Quest Plus FPLC system (Biorad) and Cytiva columns: HisTrap FF crude (His-MBP-CPSF6) and MBP-Trap (MBP) following the manufacturer's protocols. MBP-MS2 was further subjected to Heparin-based chromatography (HiTrap Heparin HP 5 ml, Cytiva, in 30 mM Na-Hepes, 25 mM NaCl, 5% glycerol) following the manufacturers protocol. All recombinant proteins were concentrated using Amicon spin concentrators (Merck Millipore) and subjected to gel filtration with Superdex 200 16/600 pg (Cytiva) in buffer (25 mM Tris-Cl pH 8.0, 500 mM NaCl,10 % Glycerol). Peak fractions containing the recombinant proteins after gel filtration were pooled. Protein concentration was determined using absorbance spectroscopy and the respective extinction coefficient at 280 nm before aliquots were flash-frozen in liquid nitrogen and stored at -80°C.

4.4.13 Splicing complex enrichment with MBP-MS2 RNA pull-down

Synthesized DNA contains sequences generating pre-mRNA with PAS sites and three MS2 binding sites (hairpins) at the 3'-end. After in vitro transcription with QuantiTect Reverse Transcription Kit (Qiagen), pre-mRNA were mixed with purified MBP-MS2 protein into a form (MBP- MS2-pre-mRNA). We applied HaCaT nuclear extracts to MBP-MS2-pre-mRNA to capture proteins on pre-mRNA. Then this mix is incubated with amylose beads which allow capturing proteins via MBP-MS2 binding for 1.5 hours at 4 °C. Then we eluted the captured proteins with SDS buffer after washing with cell lysis buffer and performed western blotting or LC-MS analysis.

4.4.14 Biotinylated proteins enrichment with NeutrAvidin beads

SILAC-labeled cells were transfected with a construct expressing TurboID-tagged CPSF6 APEX2-tagged CPSF6, or split-TurboID-N/C-tagged HBD. After 48 hours, APEX2-tagged CPSF6 transfected cells were pre-treated with 500 μM biotin phenol (Iris Biochem) for 2 hours at 37°C, followed by incubation with 1 mM H_2O_2 (Sigma-Aldrich) for 2 minutes at room temperature. Cells were washed twice with quenching solution (10 mM sodium azide, 10 mM sodium ascorbate, 5 mM Trolox) and followed by washing twice with PBS. TurboID-tagged CPSF6 transfected cells were treated with 50 μM biotin for 30 minutes. Cells transfected with split-TurboID-N/C-tagged HBD are treated with 50 μM biotin for 12 hours. Cells were lysed on ice with RIPA buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100). Cell lysates were sonicated with Bioruptor to release chromatin-bound proteins with additional NaCl at the final concentration of 500mM. Prior to the pull down, equal amounts of differentially SILAC-labeled cell extracts from three conditions were combined and incubated with pre-equilibrated NeutrAvidin agarose beads (Thermo Scientific) for 2~4 hours at 4°C on a rotation wheel. Beads were washed once with RIPA buffer, three times with 8 M Urea (Sigma) in 1% SDS, and once with 1% SDS in PBS. Eluted proteins were boiled at 95°C for 15 minutes in NuPAGE LDS Sample Buffer (Life Technologies) supplemented with 1 mM DTT. After cooling to room temperature, eluted samples were alkylated for 30 minutes in the dark with 5.5 mM chloroacetamide (CAA) before being loaded onto 4-12% gradient SDS-PAGE gels. Proteins were stained with the Life Technologies Colloidal Blue Staining Kit and digested in-gel with trypsin. Peptides were desalted on reversed-phase C18 StageTips after being extracted from the gel and followed by LC-MS analysis.

4.4.15 Co-immunoprecipitation

In a total volume of 500 μL reaction solution, 0.5~2 mg of protein were immunoprecipitated with 2 μg of antibody (IgG or CPSF6) and 20 μL of Protein-G beads. The samples were rotated for 1.5 hours at 4 °C. 1 ml of cold lysis buffer containing protease inhibitors was used to wash the beads four times. The beads were then boiled at 95°C for 10 minutes in NuPAGE LDS Sample Buffer before being fractionated by SDS-PAGE in 4-12% gradient gels. Proteins were analyzed with immunoblotting or LC-MS analysis, as previously described.

4.4.16 *In vitro* kinase assay

Biotinylated CPSF6 peptide (Biotin-PPGREMDT(p)ART(p)PLSEAEF) was synthesized by GenScript. Peptides were bound to NeutrAvidin agarose beads and subjected to mock-treated or de-phosphorylated for 2 hours at 30°C with 5 μL phosphatase in phosphatase reaction buffer with MnCl2. Following that, 10 μg peptide was incubated with cell lysate for 4 hours at 4 °C with rotation. Beads were washed three times with wash buffer (PBS with 0.1% Triton X-100) containing protease and phosphatase inhibitors.

4.4.17 *In vitro* phase separation assay (*In vitro* droplet formation)

To induce LLPS, MBP-tagged CPSF6 is treated with a specific protease, Human Rhinovirus (HRV) 3C Protease to remove the MBP tag first. Then cleaved proteins are dissolved in LLPS buffer and incubated at the specified temperature and monitored under a microscope.

The cover slides were washed in 2% Hellmanex III solution (Hellma Analytics) for 2 hours before being rinsed with ultrapure water. Cover slides are incubated overnight with methoxy poly(ethylen glycol) (mPEG) silane with an average molecular weight of 5 kDa (25 mg/mL in 95 % EtOH). After that, the cover slides were rinsed with ultrapure water, dried, and sealed with crystal clear tape. 5 µM of purified MBP-CPSF6 was diluted in the *in vitro* droplet formation buffer (25 mM Tris–HCl pH 7.4, 150 mM KCl, 2.5% glycerol, and 0.5 mM DTT) for indicated times (2, 10, 30, 60 minutes) to induce phase separation.

To assess the sensitivity of liquid-like droplets to aliphatic alcohols, 1,6-hexanediol (1,6-Hex) (Sigma Aldrich) was added at a final concentration of 5% was added for 30 minutes. The dynamic of the droplet fusion was monitored under the confocal microscope (VisiScope Spinning Disk Confocal Microscope). In general, at least 5 images were taken in non-overlapping regions that were thought to be representative of the droplet distribution on the slide. FIJI (64V5) was used to process and analyze the images.

4.5 Mass spectrometry-based proteomics

4.5.1 In-gel digestion

Alkylated samples 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. Each lane was divided into 6 to 10 pieces for subsequent processing. Each piece was then divided into 1 mm squares. The gel pieces were destained four times with 1 mL destaining solution for 20 minutes, and dehydrated twice with 1 mL 100% EtOH for 10 minutes.

To digest the proteins into peptdies, 50 µL of trypsin solution (125 ng/ µL trypsin in digestions buffer) was added to the pieces and incubated for 10~30 minutes at 500 rpm. The pieces were covered with digestion buffer and incubated at 37 °C overnight. Then the digestion was stopped by adding 50 µL 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 µL) to cover the gel pieces by sequential incubation with extraction buffer, in-gel buffer B, and ACN. Supernatants from the same gel piece were combined, vacuum centrifuged at 45 °C to a volume of 100 l, and subjected to desalting with StageTip purification (see 4.5.6).

4.5.2 In-solution digest

For the cell lysates from HaCaT cells, protein concentrations were estimated using the QuickStart Bradford Protein assay. Proteins were precipitated in a 4 fold excess of ice-cold acetone and incubated overnight at -20 °C before being pelleted by centrifugation at 1000 x g for 5 minutes. The

pellets were then redissolved in denaturation buffer at a concentration of 2-8 µg/µl. Cysteines were reduced with 2 mM DTT in the dark for 40 minutes before being alkylated with 10 mM CAA. 1 µg endoproteinase LysC was added to 75 µg of protein for 6 hours to digest it. The mixture was diluted 1:4 in 50 mM Tris buffer, PH8.0, and sequencing grade-modified trypsin was added overnight in a ratio of 1:75. Protease digestion was halted by the addition of 0.5% TFA for 30 minutes incubation at 4 °C. Following that, centrifugation at 4000 x g for 10 minutes removed precipitates and supernatant were kept for further purification and concentration. Reversed-phase Sep-Pak C18 cartridges were prepared by washing them three times with 0.1% TFA and once with 100% ACN. After loading with peptide samples, the columns were washed three times with water, dried, and stored at 4 °C.

4.5.3 TMT labeling

Peptides were eluted from SepPak columns in 2 mL of 50% ACN, thoroughly dried in a vacuum concentrator at room temperature, and resuspended in TMT-labeling buffer (150 mM HEPES pH 8.5, 30% ACN). The Nanodrop at 280 nm was used to measure peptides concentration, which were then adjusted to 5 µg/µl. For each labeling reaction, 100 µg of peptides were mixed with an equal amount of TMT label (20 µg/µl) (TMT:peptide ratio of 16:11 (w/w))and incubated in a thermoshaker for 1 hour at 25 °C and 500 rpm. The labeling reactions were quenched by adding hydroxylamine at a final concentration of 0.4% and incubating for 20 minutes at 25 °C and 1000 rpm in the thermoshaker. Following that, peptides were diluted in 0.1% TFA, lowering the ACN concentration below 3%. Then peptide labeling efficiency was evaluated by running LC-MS analysis. To account for differences in labeled peptides between labels, 5% of the samples were pooled, purified on C18 stage tips, measured with a mass spectrometer, and adjustment factors calculated. The remaining samples were pooled and purified on Sep-Pak C18 columns as previously described (4.6.2).

4.5.4 Phosphopeptide enrichment

TMT-labeled peptides were eluted from Sep-PAK columns in 2 ml of 50% ACN and acidified to a concentration of 6% TFA. 1 mg of peptides were mixed with 2 mg of TiO2 spheres and incubated at room temperature for 1 hour with rotation. Spheres were washed twice with 1 mL of binding buffer and once with 1 mL of wash buffer. They were loaded onto a StageTip made of 1 layer of a C8 47 mm extraction disk and centrifuged at 500 x g to dry. Phosphorylated peptides were eluted by centrifugation at 400 x g with 100 l of elution buffer 1, followed by 100 l of elution buffer 2. To remove NH_4OH, the eluted peptides were vacuum concentrated at 45 °C for 20 minutes. The pH was adjusted to lower than 2 with TFA and peptides were fractionated by Micro-SCX (see 4.5.5).

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

Six disks were cut from 47 mm cation exchange extraction disks with a 17-gauge Hamilton syringe and placed into a 200 µL pipette tip to make micro-SCX tips. By centrifuging at 500 x g, SCX-tips were equilibrated with 50 l of methanol, SCX elution buffer pH 2.5, SCX elution buffer pH 11, and 100 µL of SCX wash buffer. At 400 x g, acidified peptide samples (pH 2) were loaded onto the SCX-

tips. Peptides were then fractionated by eluting them with 100 l of SCX elution buffers at 700 x g from lowest to highest pH. Eluates were vacuum centrifuged for 20 minutes at 45 °C to remove ACN before being desalted by C18 StageTipping (see 4.5.6).

pH of SCX-buffers:

(Chromatin-) proteome: 3.2 / 3.6 / 3.9 / 4.25 / 4.8 / 5.5 / 7 / 11

Phospho-proteome: 2.5 / 2.8 / 3.2 / 3.5 / 3.75 / 4 / 4.25 / 4.5 / 5 / 5.5 / 6.5 / 11

4.5.6 Desalting and concentration of peptides

Peptide purification was carried out with the help of self-made and extremely cost-effective stop-and-go extraction tips (StageTips). StageTips were created by cutting two disks from a 47 mm C18 extraction disk with a 17-gauge Hamilton syringe and inserting them into a 200 µL pipette tip. By centrifuging at 500 x g, stage tips were equilibrated with 25 µL of methanol, followed by 25 µL of stage tip buffer B and 2 x 25 µL of stage tip buffer A. The samples were loaded onto the tip and washed in the same manner with 50 µL Buffer A. StageTips were dried and stored at 4 °C. The peptides were eluted into 96-well sample plates with 50 µL elution buffer and vacuum concentrated in a volume of 2 l. Finally, 3 µL of 0.1% FA was added and followed by introduction into MS.

4.5.7 MS analysis

Peptide fractions were analyzed using a quadrupole Orbitrap mass spectrometer (Q Exactive Plus or Exploris, Thermo Scientific) coupled to a UHPLC system (EASY-nLC 1000 or EASY-nLC 1200, Thermo Scientific). Peptide samples were loaded onto C_{18} reversed-phase columns and eluted with a linear 8 to 40% acetonitrile (ACN) gradient containing 0.1% formic acid. The mass spectrometer was set to data-dependent mode, which switched between MS1 and MS2 acquisition automatically. The Orbitrap was used to collect full survey scan MS spectra (m/z 300-1700). Higher-energy C-trap dissociation (HCD) was used to isolate and fragment the ten most intense ions. Peptides with unassigned charge states, as well as those with charges less than +2, were not fragmented. The Orbitrap mass analyzer was used to collect fragment spectra. The settings for the LC and mass spectrometers are listed below. The initial screen's data was collected entirely on the Exploris mass spectrometer setup.

4.5.8 MS peptide identification

MaxQuant (version 1.5.2.8) was used to analyze raw data files. The Andromeda search engine compared parent ion and MS2 spectra to a database containing 98,566 human protein sequences obtained from the UniProtKB released in 04/2018. Spectra were analyzed with a mass tolerance of 6 ppm in MS mode, 20 ppm in HCD MS2 mode, strict trypsin specificity, and up to two miscleavages allowed. As a fixed modification, cysteine carbamidomethylation was looked for, whereas cysteine modification with NEM, protein N-terminal acetylation, methionine oxidation, and phosphorylation of serine, threonine, and tyrosine were looked for as variable modifications. The dataset was filtered using

posterior error probability (PEP) to achieve a false discovery rate of less than 1% using a target-decoy approach. The match between runs and the re-quantify features were switched on.

TMT samples

Raw files from TMT samples were analyzed with Maxquant version 1.6.14 with the same settings as above. Spectra were compared to a database of 96,788 human protein sequences obtained from the UniProtKB in February 2020. Miscleavages of up to three were permitted, and spectra were filtered for precursor ion fractions (PIF) greater than 0.75.

4.5.9 Data processing and visualization

For the TMT experiments, multiply modified peptides were analyzed separately. Furthermore, the intensities of proteins and peptides were normalized to account for loading errors and batch effects. For each batch, isotope corrected intensities were first divided by the corresponding reference intensity (reference normalization). Then intensity averages for each protein or peptide were calculated per batch. Following that, each intensity was divided by the average mentioned above (per-protein/peptide centering). The R software environment was used for statistical analysis and data visualization (v3.6.1). The limma package was used to calculate moderated t-tests corrected for multiple hypothesis testing (Benjamini Hochberg), with peptides/proteins measured in at least two replicates ("FDR"). Unsupervised hierarchical clustering was used, and heatmaps were created as a result using the pheatmap or the gplots package [784]. The umap package was used for UMAP clustering [785]. Gorilla was used to perform GOterm enrichment analysis [786]. IceLogo was used for sequence motif analysis, and overrepresented motifs were identified using sequence annotation followed by a Fisher's exact test in Perseus (v1.6.14.0) [787]. The STRING database was used to analyze protein interaction networks, which were then visualized with Cytoscape (v3.8.2) [788]. Enrichment analysis for protein clusters was also performed in Cytoscape using "StringApp" or "ClusterViz" [788]. The KSEA algorithm (KSEAapp) was used to predict kinase activities with updated known kinase-substrate relations (PhosphoSitePlus, 04/2021) and a prediction NetworKIN cutoff of 5 [789], [790]. CORAL generated the Kinome analysis with the KSEA predicted kinases' fold change and FDR values (http://phanstiel-lab.med.unc.edu/CORAL/, 11/2021) and the map was redrawn in illustrator [791].

4.6 Genomics

4.6.1 Quantitative PCR (qPCR)

HaCaT cells were transfected with siRNA of CPSF6 or control siRNA. Then cells were treated with UV (15J/m^2) and isolated mRNA with RNeasy Plus Mini Kit (Qiagen) following the manufacturer's protocol. 500 ng of purified RNA was reverse-transcribed into cDNA by using the QuantiTect Reverse Transcription Kit (Qiagen). The quantitative real-time PCR was performed using SYBR Green (ABI) in a reaction system of 2x SYBR Green mix and a 0.5 M final primer mix. It is amplified and monitored in CFX384 BioRad machine.

4.6.2 Quantification of DNA-RNA hybrids using dot blot

Genomic DNA was extracted using the DNeasy mini kit (Qiagen). The isolated gDNA was treated with 1.2 U RNase III (produced in-house) for 2 h at 37°C. After enzyme deactivation at 65°C for 20 min, samples were split in half to digest control samples with 10 U RNaseH1 (NEB) overnight at 37°C. Enzyme deactivation was followed by spotting DNA in a serial dilution on a nitrocellulose membrane (NeoLab Migge GmbH) using a dot blot apparatus (BioRad). DNA was cross-linked to the membrane by UV light and afterwards blocked with 10% skimmed milk solution in PBS supplemented with 0.1% Tween-20. The membrane was incubated overnight at 4°C with the S9.6 antibody (produced in-house). After incubation of secondary antibodies conjugated to horseradish peroxidase (Jackson ImmunoResearch Laboratories), the signal was detected using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific). An antibody against dsDNA was probed as a loading control after stripping the membrane with β-mercaptoethanol (Sigma) and 0.1% SDS in PBS. The detected signal was quantified using Fiji/ImageJ (v1.51), and ratios between the signal resulting from S9.6 and dsDNA staining were calculated to quantify global R-loop levels.

5 Abbreviations

6-4PPs	6'-4' photoproducts
ABC	Ammonium bicarbonate
ACN	Acetonitril
aEJ	Alternative end-joining
AID	Activation-Induced Cytidine Deaminase
alt-NHEJ	Alternative NHEJ
APEX	Ascorbic acid peroxidase
AP-site	Abasic site
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia and Rad3 related
BRCA1	Breast cancer type 1 susceptibility protein
C_{18}	Octadodecyl alkane chains
CAA	Chloroacetamide
CID	Collision-induced dissociation
CPT	Camptothecin
CRL	Cullin-RING E3 ligases
CTR	Control
DDR	DNA damage response
dHJ	Double HJ
D-loop	Displacement loop
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
dNTPs	Deoxyribonucleotide triphosphates
DSB	Double strand break
DTT	Dithiothreitol
EGSEA	Ensemble Gene Set Enrichment Analysis
ESI	Electrospray ionization
EU	5-ethynyl uridine
FA	Formaldehyde
FDR	False discovery rate
GG-NER	Global genome NER
GO	Gene enrichment
GSEA	Gene Set Enrichment Analysis
H_2O_2	Hydrogen peroxide
HCD	Higher-energy c-trap/collision dissociation
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HJ	Holliday junction
HIV	Human immunodeficiency virus
HR	Homologous recombination
ICL	Inter-strand crosslink
IDLs	Insertion-deletion loops
IF	Immunofluorescence
IMB	Institute of Molecular Biology
IP	Immunoprecipitation
IR	Ionizing radiation

KD	Knock-down
KDa	1,000 Dalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knock-out
LC_{50}	Lethal concentration - killing 50% of cells
LC-MS/MS	Liquid chromatography and tandem mass spectrometry
LFQ	Label-free quantification
LysC	Endoproteinase LysC
m/z	Mass to charge
m^6a	N6-methyladensine
Micro-SCX	Micro-tip based strong cation exchange chromatography
MMEJ	Microhomology-mediated end-joining
MMR	Mismatch repair
mRNA	Messenger RNA
MS	Mass spectrometry
MS/MS	Tandem MS
MS1	First stage of mass analysis / precursor spectrum
MS2	Second stage of mass analysis / fragment spectrum
Multiplicity	Number of PTMs on a detected peptide
NEM	N-ethylmaleimide
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NHS	N-hydroxysuccinimide
NLS	Nuclear localization signal
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PEP	Posterior error probability
PTM	Posttranslational modification
RBPs	RNA-binding proteins
RIPA	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RP	Reversed-phase
RP-HPLC	Reversed-phase high-performance chromatography
RPLs	60S ribosomal subunits
RPSs	40S ribosomal subunits
RQC	Ribosome-associated quality control
RT	Room temperature (~22-23°C)
S	Serine
S/TQ motif	Serine or threonine residues followed by glutamine
SCX	Strong cation exchange
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH2/3	Src-homology-2/3
SILAC	Stable isotope labeling with amino acids in cell culture
siRNA	Short interfering RNA

ssDNA	Single stranded DNA
StageTips	Self-made and extremely economical stop-and-go-extraction tips
T (Protein)	Tyrosine
TC-NER	Transcription-coupled NER
TEAB	Triethylammonium bicarbonate
TFs	Transcription factors
TiO_2	Titanium dioxide
TLS	Translesion synthesis
TMT	Tandem mass tag
TMTc	TMTcomplementary
TMT-MS3	TMT-Measuring the reporter ions in the third dimension
U2OS	Human Bone Osteosarcoma Epithelial Cells
UV	Ultraviolet
WB	western blot
WCL	Whole cell lysate
WT	Wild type
Y	Tyrosine
ε-NH2	Epsilon-amino group

6 Appendix

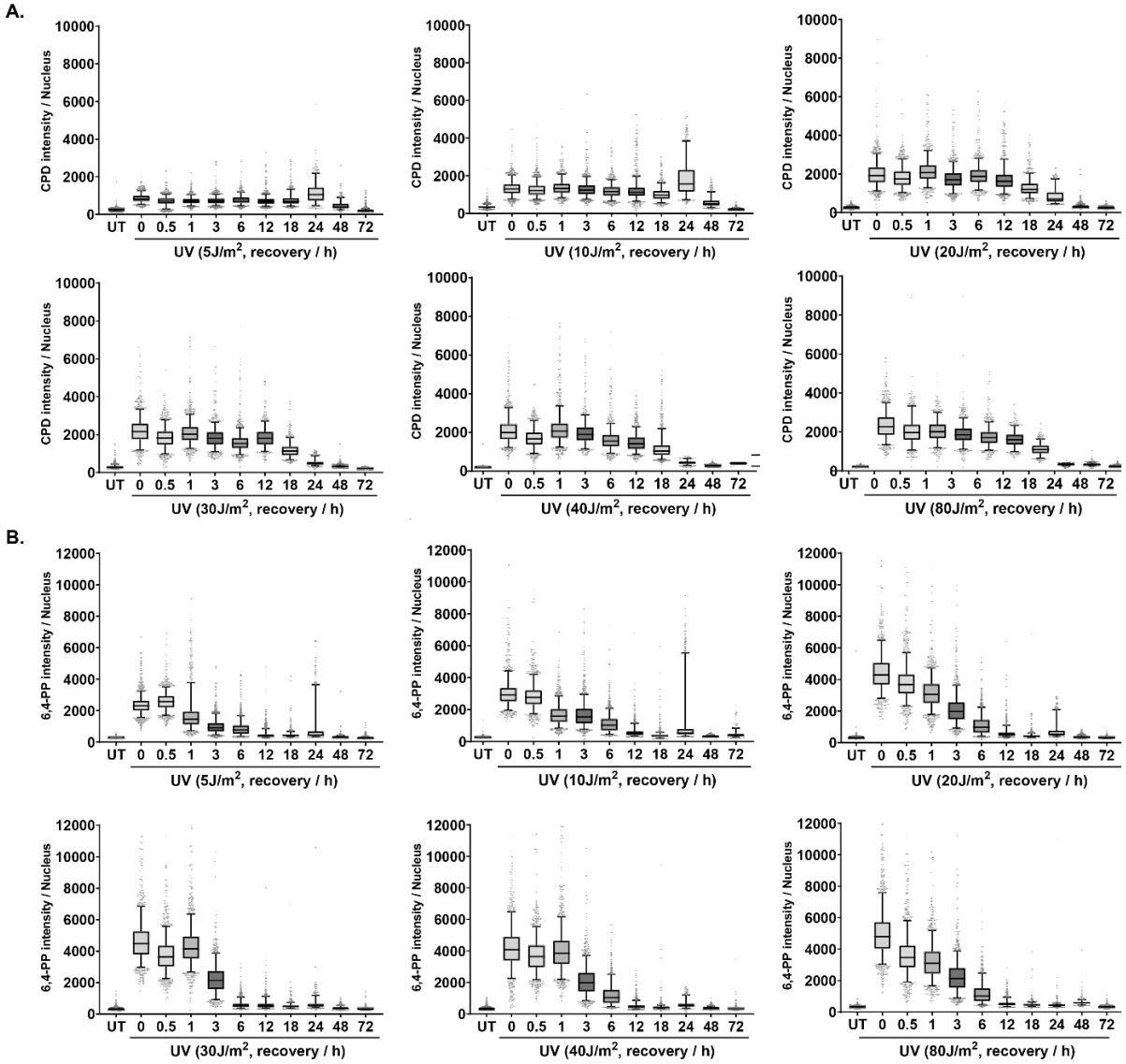

Figure 68: Immunofluorescence analysis of the kinetics of CPD (A) and 6,4-PP (B) dimer removal during the recovery from different dosages of UV irradiation. **See Figure 22**.

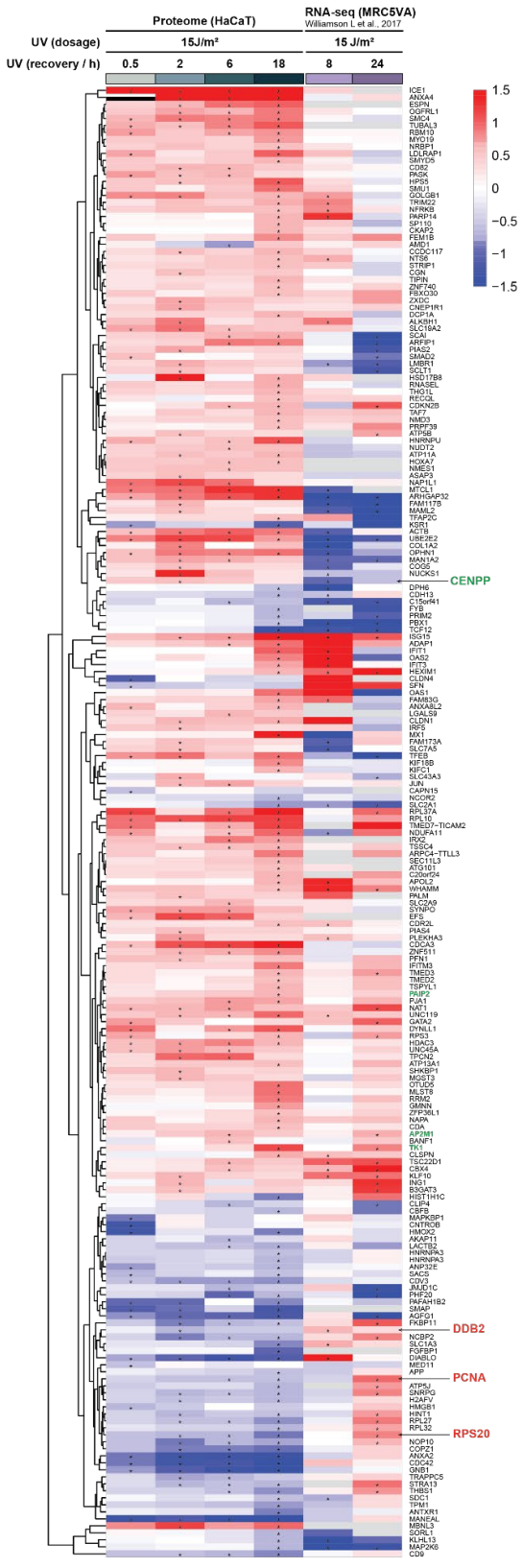

Figure 69: Heatmap of global proteome profiling in HaCaT cell (n= 4) **and the RNA-seq data in MRC5VA cell from the Svejstrup lab upon UV induction (15 J/m², 8h recovery)** (n= 2; fold-change ≥ 1.5, or ≤ 1.5; FDR ≤ 0.01 are indicated with an asterisk). Proteins with upregulated ubiquitylation levels are labeled with red and downregulated ubiquitylation levels are labeled with green after UV treatment from the Svejstrup lab (http://www.biologic-db.org/) [792].

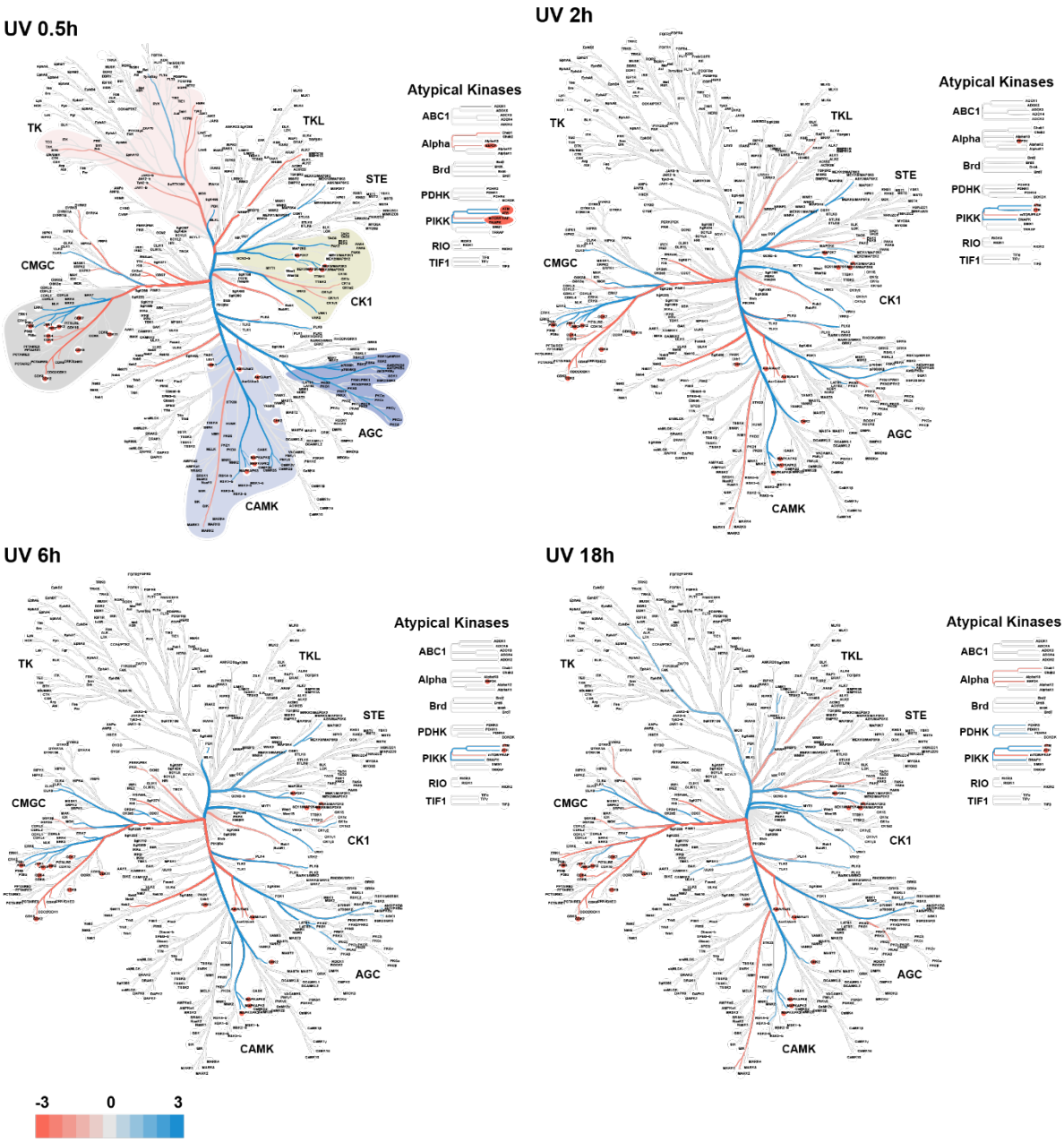

Figure 70: Kinome map for kinase activation for all predicated kinases upon UV irradiation. Kinome map is generated by CORAL with KSEA predicted kinases' fold change and FDR values. Blue indicates down-regulated activity and red indicates upregulated activity.

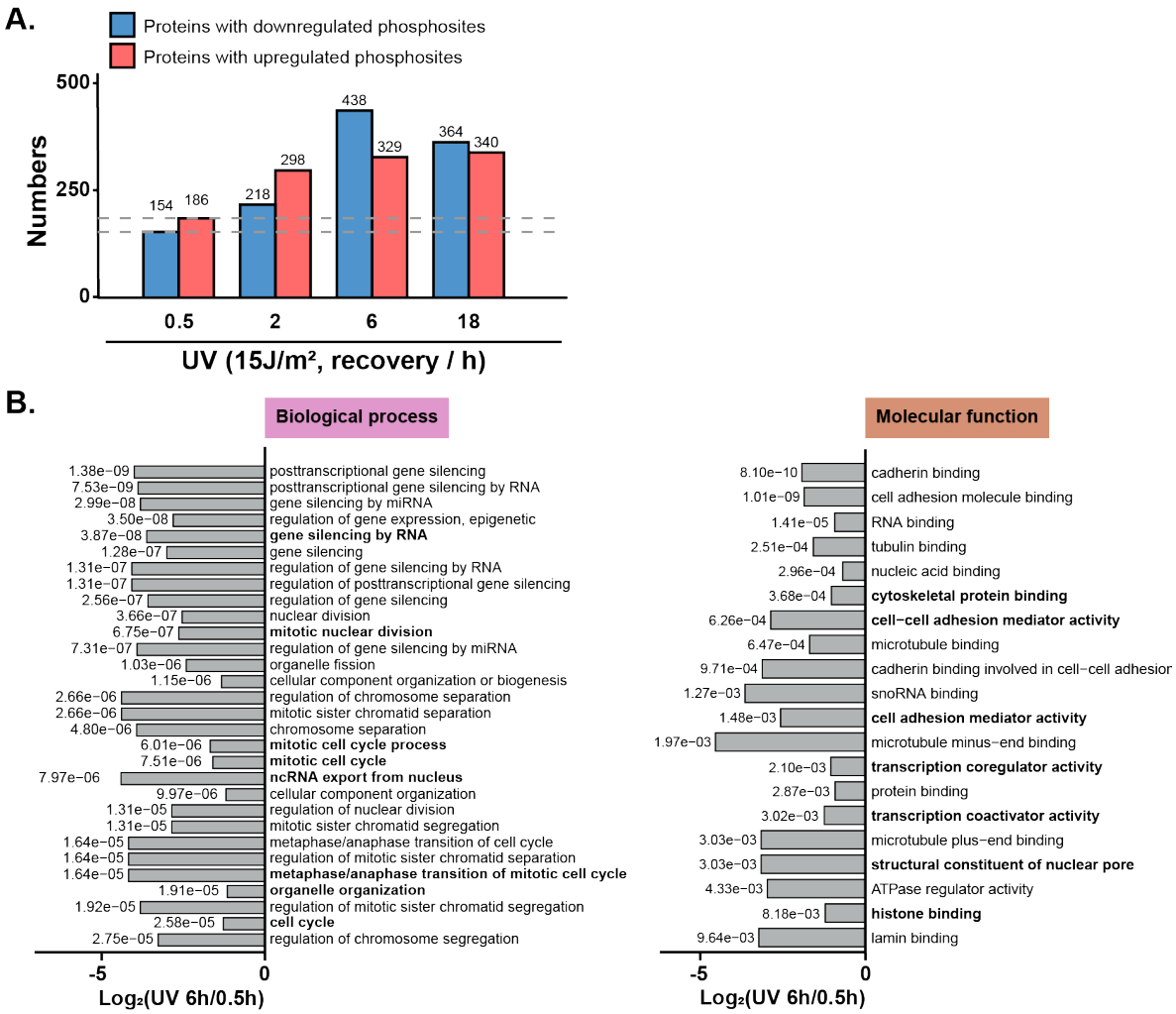

Figure 71: Protein numbers with up- or down-regulated phosphorylation sites (A) and GO enrichment of proteins with downregulated phosphorylation sites while comparing 6h to 0.5h and the highest-scoring GO terms for biological process and molecular function are displayed (B).

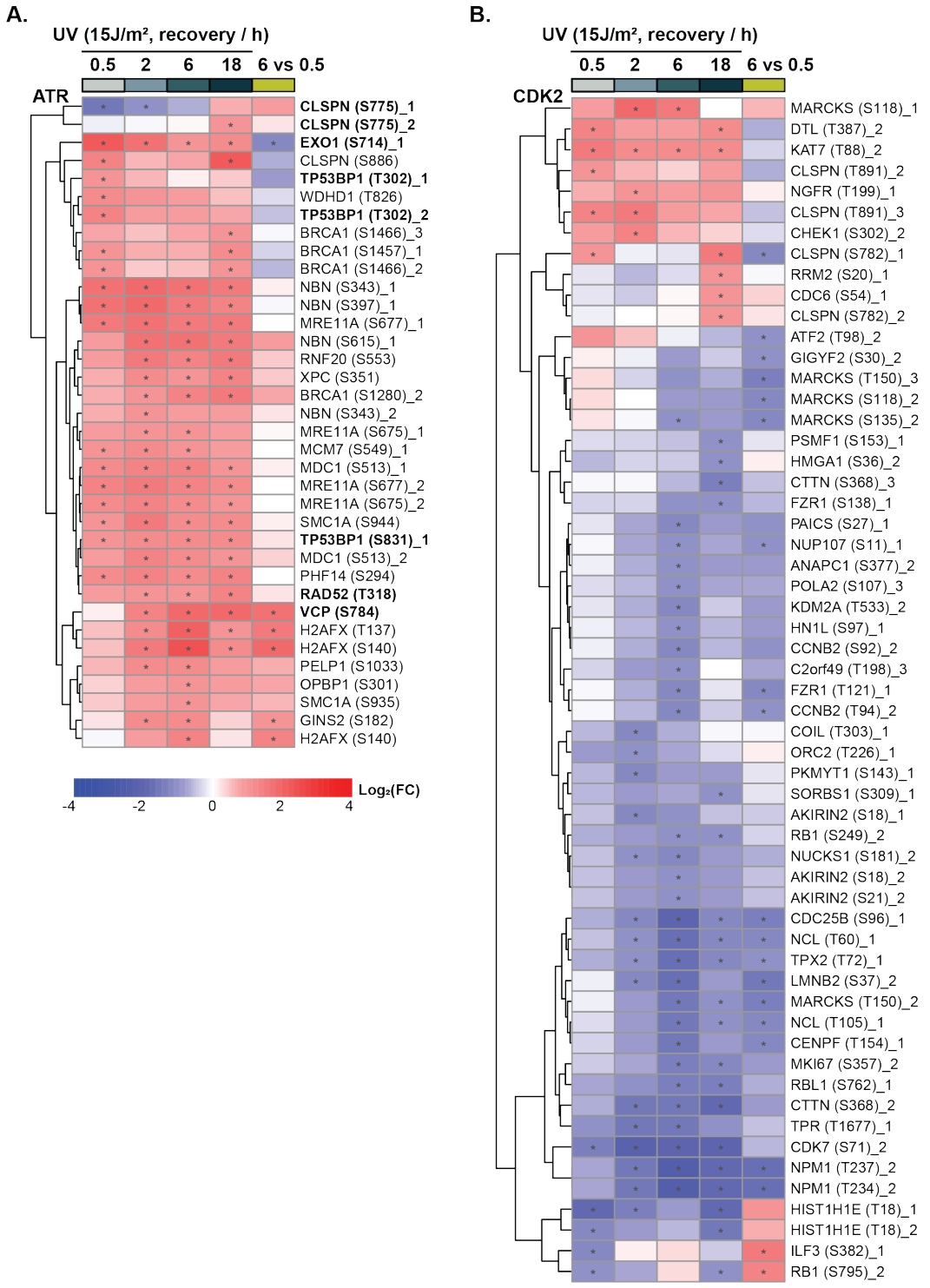

Figure 72: Heatmaps for predicted ATR (A) and CDK2 kinases (B) (peptide fold-change ≥ 2, FDR ≤ 0.05 are indicated with an asterisk).

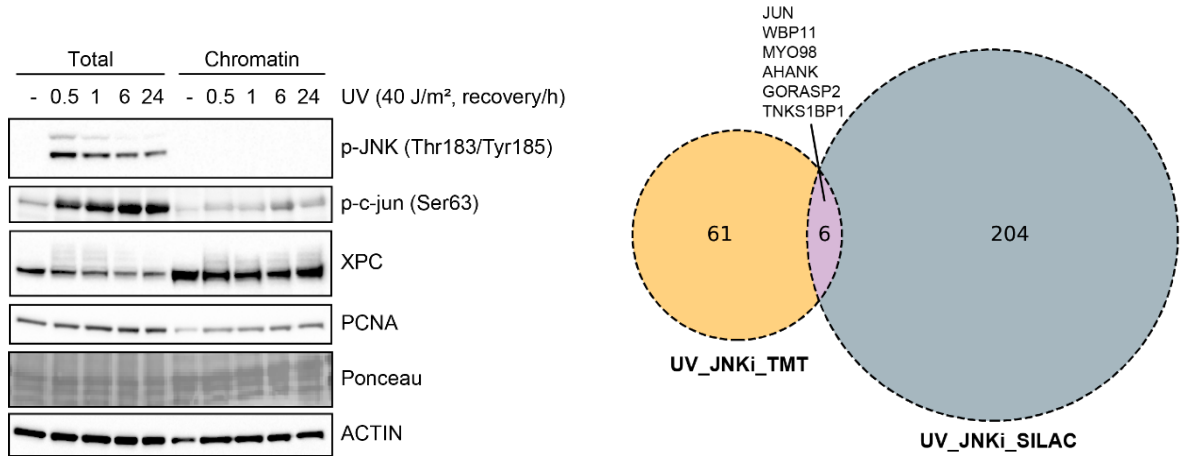

Figure 73: JNK1 activation in different cellular fractions post UV irradiation (left) and overlap of JNK-dependent phosphoproteomics datasets (TMT-labeling and SILAC-labeling) (right).

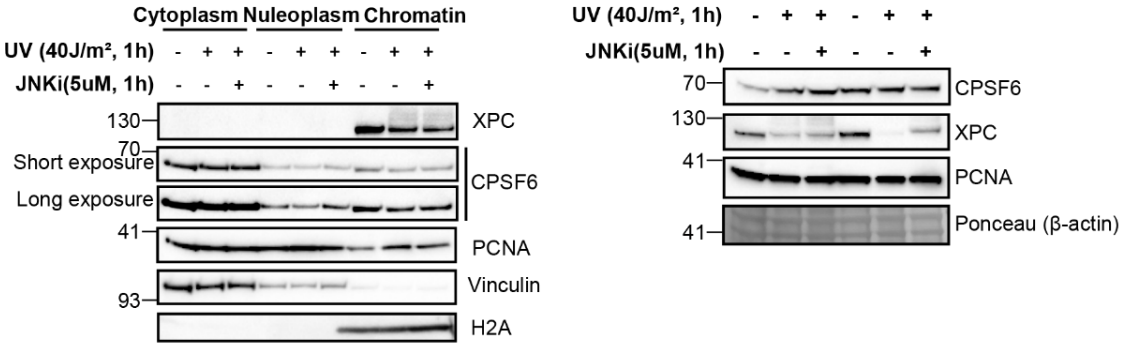

Figure 74: Western blots analysis for different cellular fractions indicates the removal of CPSF6 from chromatin. See Figure 53.

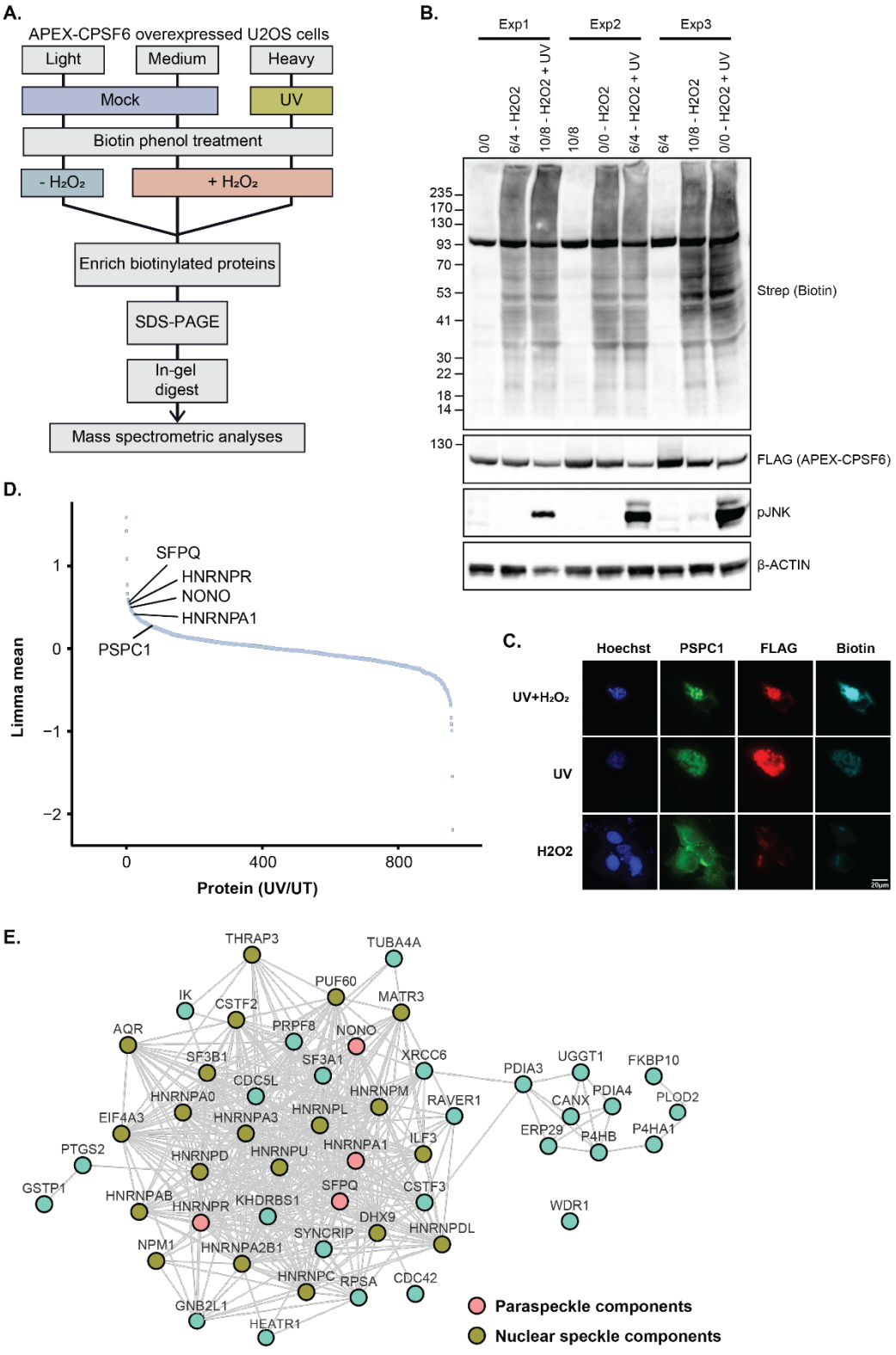

Figure 75: Proximal interactor analysis with APEX-CPSF6 upon UV treatment. A. Experimental workflow for quantitative mass spectrometric analysis of APEX2-tagged CPSF6 proximal complex after UV light treatment. SILAC-labeled cells were transfected with APEX-CPSF6. After 48 hours of transfection, cells were treated with UV light (20J/m^2). After 1h recovery post UV irradiation, biotinylation was induced upon adding 500 µM biotin-phenol for 2 h at 37°C and 1 mM H_2O_2 for 2 min at room temperature. Samples were pooled after cell lysis and biotinylated proteins were purified using NeutrAvidin beads. Immunoprecipitated proteins were resolved on SDS-PAGE and digested in-gel into peptides, followed by LC-MS/MS. **B.** Western blotting analysis for APEX-CPSF6 overexpression and biotinylation in SILAC U2OS cells upon UV treatment from **A. C.** Immunofluorescence analysis of APEX-CPSF6 in U2OS cells after UV light irradiation. Cells were fixed with 4% PFA and nuclei were stained with DAPI. Representative images are shown after Deconvolution with Imaris. **D**. Rank plot of protein groups proximal to APEX-CPSF6. Paraspeckle components are highlighted (first 50) between UV treatment and mock treatment. **E.** Network analysis for proteins enriched upon UV treatment. This experiment was done by Matthias Ostermaier.

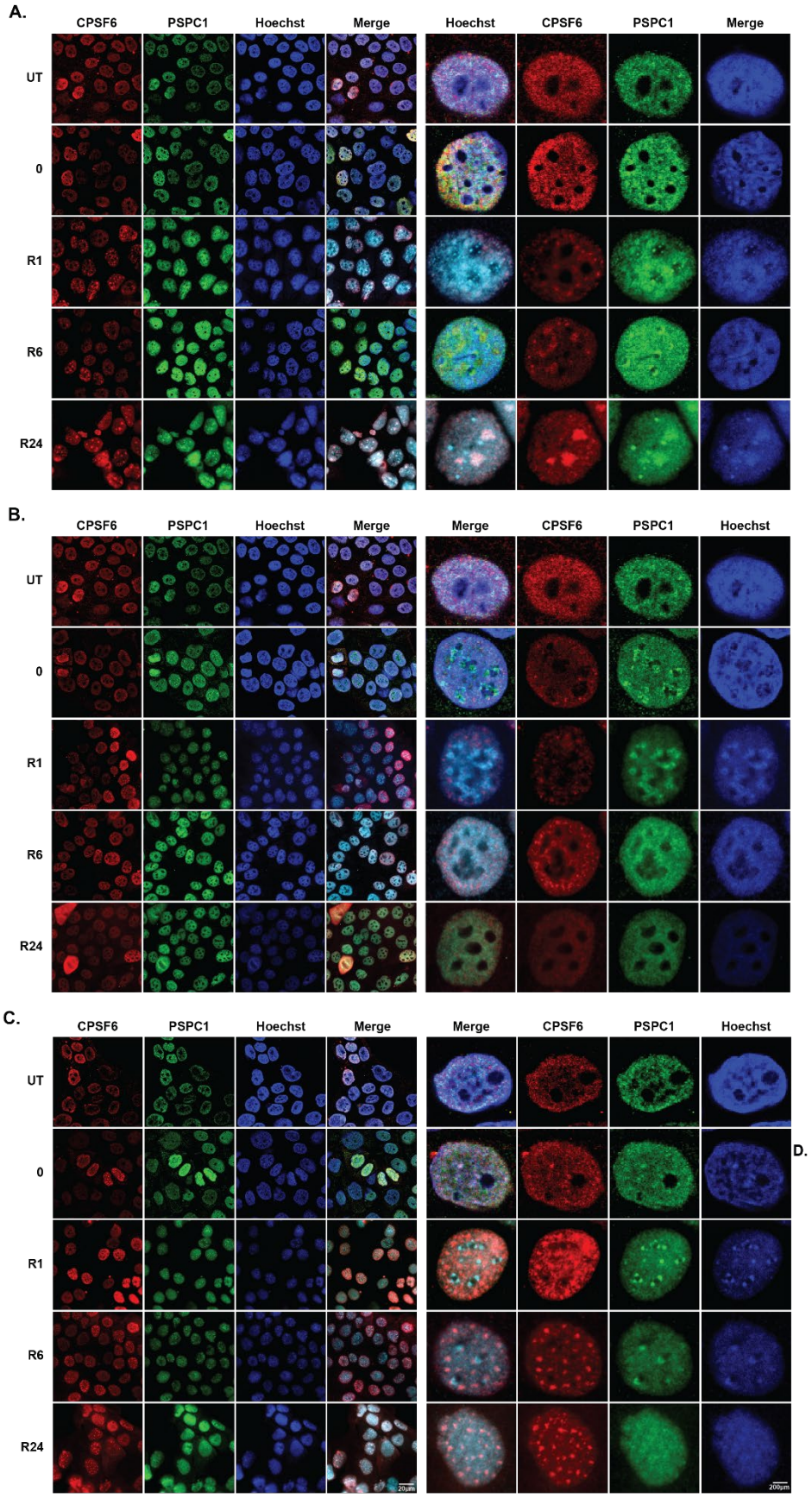

Figure 76: Transcription inhibition induces paraspeckle formation. HaCaT cells were treated with 4NQO (20µM, 1h) (**A**), DRB (10µM, 1h) (**B**) or actinomycin D, ActD (2µg/mL, 1h) (**C**) and washed out with PBS, then cells were left for recovery for the indicated time (R1, R6, R24). Cells were fixed with 4% PFA and stained with CPSF6 and PSPC1 antibodies, and nuclei were stained with Hoechst. Representative images are shown after Deconvolution with Imaris.

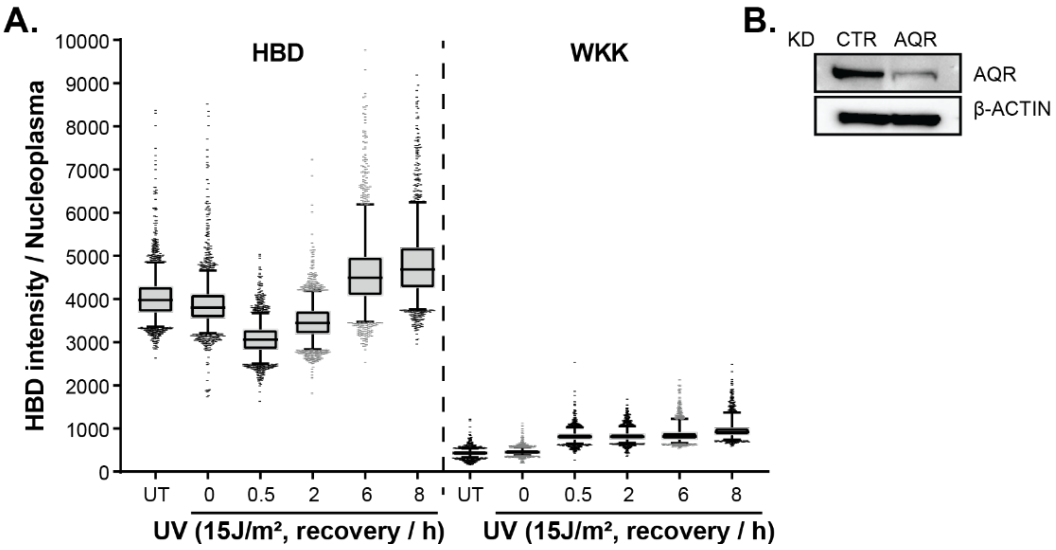

Figure 77: UV induced R-loop formation, see Figure 64

7 References

[1] C. R. Woese, O. Kandler, and M. L. Wheelis, “Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya,” *Proc. Natl. Acad. Sci. U. S. A.*, 1990, doi: 10.1073/pnas.87.12.4576.

[2] J. T. Bonner, “The evolution of evolution,” *J. Exp. Zool. Part B Mol. Dev. Evol.*, 2019, doi: 10.1002/jez.b.22859.

[3] E. Frangedakis *et al.*, “The hornworts: morphology, evolution and development,” *New Phytologist*. 2021, doi: 10.1111/nph.16874.

[4] G. Janusz *et al.*, “Laccase properties, physiological functions, and evolution,” *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21030966.

[5] B. D. Uhal, C. Ramos, I. Joshi, A. Bifero, A. Pardo, and M. Selman, “Cell size, cell cycle, and α-smooth muscle actin expression by primary human lung fibroblasts,” *Am. J. Physiol. - Lung Cell. Mol. Physiol.*, 1998, doi: 10.1152/ajplung.1998.275.5.l998.

[6] A. C. Chien, N. S. Hill, and P. A. Levin, “Cell size control in bacteria,” *Current Biology*. 2012, doi: 10.1016/j.cub.2012.02.032.

[7] M. Westoby *et al.*, “Cell size, genome size, and maximum growth rate are near-independent dimensions of ecological variation across bacteria and archaea,” *Ecol. Evol.*, 2021, doi: 10.1002/ece3.7290.

[8] A. Lazcano, “Cells, Molecules and Evolution: Historical Issues in Molecular Evolution,” *Journal of Molecular Evolution*. 2016, doi: 10.1007/s00239-016-9776-2.

[9] M. Khalili, M. Asadi, H. Kahroba, M. R. Soleyman, H. Andre, and E. Alizadeh, “Corneal endothelium tissue engineering: An evolution of signaling molecules, cells, and scaffolds toward 3D bioprinting and cell sheets,” *Journal of Cellular Physiology*. 2021, doi: 10.1002/jcp.30085.

[10] R. Milo, P. Jorgensen, U. Moran, G. Weber, and M. Springer, “BioNumbers The database of key numbers in molecular and cell biology,” *Nucleic Acids Res.*, 2009, doi: 10.1093/nar/gkp889.

[11] T. Kühn *et al.*, “Protein diffusion in mammalian cell cytoplasm,” *PLoS One*, 2011, doi: 10.1371/journal.pone.0022962.

[12] K. Davies and E. Meimaridou, “Biological basis of child health 1: understanding the cell and genetics,” *Nursing children and young people*. 2020, doi: 10.7748/ncyp.2020.e1047.

[13] M. M. Mysior and J. C. Simpson, “Cell3: A new vision for study of the endomembrane system in mammalian cells,” *Biosci. Rep.*, 2021, doi: 10.1042/bsr20210850.

[14] M. Long, T. Kranjc, M. M. Mysior, and J. C. Simpson, “RNA Interference Screening Identifies Novel Roles for RhoBTB1 and RhoBTB3 in Membrane Trafficking Events in Mammalian Cells,” *Cells*, 2020, doi: 10.3390/cells9051089.

[15] M. Le Vasseur, V. C. Chen, K. Huang, W. A. Vogl, and C. C. Naus, “Pannexin 2 localizes at ER-mitochondria contact sites,” *Cancers (Basel).*, 2019, doi: 10.3390/cancers11030343.

[16] F. Adolf *et al.*, “Proteomic Profiling of Mammalian COPII and COPI Vesicles,” *Cell Rep.*, 2019, doi: 10.1016/j.celrep.2018.12.041.

[17] F. Aniento, V. Sánchez de Medina Hernández, Y. Dagdas, M. Rojas-Pierce, and E. Russinova, “Molecular mechanisms of endomembrane trafficking in plants,” *Plant Cell*, 2021, doi: 10.1093/plcell/koab235.

[18] M. Futai, G. H. Sun-Wada, Y. Wada, N. Matsumoto, and M. Nakanishi-Matsui, “Vacuolar-type ATPase: A proton pump to lysosomal trafficking,” *Proceedings of the Japan Academy Series B: Physical and Biological Sciences*. 2019, doi: 10.2183/pjab.95.018.

[19] A. H. Fox and A. I. Lamond, “Paraspeckles.,” *Cold Spring Harbor perspectives in biology*. 2010,

doi: 10.1101/cshperspect.a000687.

[20] D. L. Spector and A. I. Lamond, “Nuclear speckles,” *Cold Spring Harb. Perspect. Biol.*, vol. 3, no. 2, pp. 1–12, 2011, doi: 10.1101/cshperspect.a000646.

[21] S. P. Shevtsov and M. Dundr, “Nucleation of nuclear bodies by RNA,” *Nat. Cell Biol.*, 2011, doi: 10.1038/ncb2157.

[22] V. Lallemand-Breitenbach and H. de Thé, “PML nuclear bodies.,” *Cold Spring Harbor perspectives in biology*. 2010, doi: 10.1101/cshperspect.a000661.

[23] K. S. Hsu and H. Y. Kao, “PML: Regulation and multifaceted function beyond tumor suppression,” *Cell and Bioscience*. 2018, doi: 10.1186/s13578-018-0204-8.

[24] E. Delbarre and S. M. Janicki, “Modulation of H3.3 chromatin assembly by PML: A way to regulate epigenetic inheritance,” *BioEssays*, 2021, doi: 10.1002/bies.202100038.

[25] M. Tampakaki *et al.*, “PML differentially regulates growth and invasion in brain cancer,” *Int. J. Mol. Sci.*, 2021, doi: 10.3390/ijms22126289.

[26] D. L. J. Lafontaine, J. A. Riback, R. Bascetin, and C. P. Brangwynne, “The nucleolus as a multiphase liquid condensate,” *Nat. Rev. Mol. Cell Biol.*, vol. 22, no. 3, pp. 165–182, Mar. 2021, doi: 10.1038/S41580-020-0272-6.

[27] D. S. W. Protter and R. Parker, “Principles and Properties of Stress Granules,” *Trends Cell Biol.*, vol. 26, no. 9, pp. 668–679, 2016, doi: 10.1016/j.tcb.2016.05.004.

[28] J. Y. Youn *et al.*, “Properties of Stress Granule and P-Body Proteomes,” *Molecular Cell*. 2019, doi: 10.1016/j.molcel.2019.09.014.

[29] G. J. Jang, J. C. Jang, and S. H. Wu, “Dynamics and functions of stress granules and processing bodies in plants,” *Plants*. 2020, doi: 10.3390/plants9091122.

[30] A. A. M. André and E. Spruijt, “Liquid–liquid phase separation in crowded environments,” *Int. J. Mol. Sci.*, 2020, doi: 10.3390/ijms21165908.

[31] W. M. Babinchak *et al.*, “Small molecules as potent biphasic modulators of protein liquid-liquid phase separation,” *Nat. Commun.*, 2020, doi: 10.1038/s41467-020-19211-z.

[32] M. Dang, L. Lim, J. Kang, and J. Song, “ATP biphasically modulates LLPS of TDP-43 PLD by specifically binding arginine residues,” *Commun. Biol.*, 2021, doi: 10.1038/s42003-021-02247-2.

[33] T. Yoshizawa, R. S. Nozawa, T. Z. Jia, T. Saio, and E. Mori, “Biological phase separation: cell biology meets biophysics,” *Biophysical Reviews*. 2020, doi: 10.1007/s12551-020-00680-x.

[34] G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, “Sequence determinants of protein phase behavior from a coarse-grained model,” *PLoS Comput. Biol.*, 2018, doi: 10.1371/journal.pcbi.1005941.

[35] Y. Lin, S. L. Currie, and M. K. Rosen, “Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs,” *J. Biol. Chem.*, vol. 292, no. 46, pp. 19110–19120, 2017, doi: 10.1074/jbc.M117.800466.

[36] Y.-Y. Luo, J.-J. Wu, and Y.-M. Li, “Regulation of liquid–liquid phase separation with focus on post-translational modifications,” *Chem. Commun.*, 2021, doi: 10.1039/d1cc05266g.

[37] S. Farina, F. Esposito, M. Battistoni, G. Biamonti, and S. Francia, “Post-Translational Modifications Modulate Proteinopathies of TDP-43, FUS and hnRNP-A/B in Amyotrophic Lateral Sclerosis,” *Frontiers in Molecular Biosciences*. 2021, doi: 10.3389/fmolb.2021.693325.

[38] R. Henze, M. Wenzel, P. Dittrich, and P. Hemmerich, “Computer modeling of phase separation at PML nuclear bodies,” *Biopolym. Cell*, 2019, doi: 10.7124/bc.0009BE.

[39] C. C. Lin *et al.*, “Receptor tyrosine kinases regulate signal transduction through a liquid-liquid

phase separated state," *bioRxiv*, 2019, doi: 10.1101/783720.

[40] S. Boyko, X. Qi1, T. H. Chen, K. Surewicz, and W. K. Surewicz, "Liquid-liquid phase separation of tau protein: The crucial role of electrostatic interactions," *J. Biol. Chem.*, 2019, doi: 10.1074/jbc.AC119.009198.

[41] S. Boeynaems *et al.*, "Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics," *Mol. Cell*, vol. 65, no. 6, pp. 1044-1055.e5, 2017, doi: 10.1016/j.molcel.2017.02.013.

[42] C. Wang *et al.*, "Stress Induces Dynamic, Cytotoxicity-Antagonizing TDP-43 Nuclear Bodies via Paraspeckle LncRNA NEAT1-Mediated Liquid-Liquid Phase Separation," *Mol. Cell*, 2020, doi: 10.1016/j.molcel.2020.06.019.

[43] J. Woulfe, "Nuclear bodies in neurodegenerative disease," *Biochimica et Biophysica Acta - Molecular Cell Research*. 2008, doi: 10.1016/j.bbamcr.2008.05.005.

[44] M. Dundr and T. Misteli, "Biogenesis of nuclear bodies.," *Cold Spring Harbor perspectives in biology*. 2010, doi: 10.1101/cshperspect.a000711.

[45] E. Taiana *et al.*, "LncRNA NEAT1 in Paraspeckles: A structural scaffold for cellular DNA damage response systems?," *Non-coding RNA*. 2020, doi: 10.3390/NCRNA6030026.

[46] C. S. Bond and A. H. Fox, "Paraspeckles: Nuclear bodies built on long noncoding RNA," *J. Cell Biol.*, vol. 186, no. 5, pp. 637–644, 2009, doi: 10.1083/jcb.200906113.

[47] T. Yamazaki and T. Hirose, "The building process of the functional paraspeckle with long non-coding RNAs," *Frontiers in Bioscience - Elite*. 2015, doi: 10.2741/E715.

[48] Y. Nishimoto *et al.*, "The long non-coding RNA nuclear-enriched abundant transcript 1-2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis," *Mol. Brain*, 2013, doi: 10.1186/1756-6606-6-31.

[49] M. Isobe *et al.*, "Forced isoform switching of Neat1_1 to Neat1_2 leads to the loss of Neat1_1 and the hyperformation of paraspeckles but does not affect the development and growth of mice," *RNA*, 2020, doi: 10.1261/rna.072587.119.

[50] M. Machitani, I. Taniguchi, and M. Ohno, "ARS2 Regulates Nuclear Paraspeckle Formation through 3′-End Processing and Stability of NEAT1 Long Noncoding RNA," *Mol. Cell. Biol.*, 2020, doi: 10.1128/mcb.00269-19.

[51] T. Yamazaki *et al.*, "Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation," *Mol. Cell*, 2018, doi: 10.1016/j.molcel.2018.05.019.

[52] A. Naveed *et al.*, "NEAT1 polyA-modulating antisense oligonucleotides reveal opposing functions for both long non-coding RNA isoforms in neuroblastoma," *Cell. Mol. Life Sci.*, 2021, doi: 10.1007/s00018-020-03632-6.

[53] T. Yamazaki, S. Nakagawa, and T. Hirose, "Architectural RNAs for Membraneless Nuclear Body Formation," *Cold Spring Harb. Symp. Quant. Biol.*, 2020, doi: 10.1101/sqb.2019.84.039404.

[54] E. Knutsen, A. L. Harris, and M. Perander, "Expression and functions of long non-coding RNA NEAT1 and isoforms in breast cancer," *British Journal of Cancer*. 2021, doi: 10.1038/s41416-021-01588-3.

[55] S. Nakagawa, T. Yamazaki, and T. Hirose, "Molecular dissection of nuclear paraspeckles: Towards understanding the emerging world of the RNP milieu," *Open Biol.*, vol. 8, no. 10, 2018, doi: 10.1098/rsob.180150.

[56] S. Souquere, G. Beauclair, F. Harper, A. Fox, and G. Pierron, "Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies," *Mol. Biol. Cell*, 2010, doi: 10.1091/mbc.E10-08-0690.

[57] T. Hirose, T. Yamazaki, and S. Nakagawa, “Molecular anatomy of the architectural NEAT1 noncoding RNA: The domains, interactors, and biogenesis pathway required to build phase-separated nuclear paraspeckles,” *Wiley Interdiscip. Rev. RNA*, 2019, doi: 10.1002/wrna.1545.

[58] † Stefano Cardinale, C. A. Barbara Cisterna,‡ Paolo Bonetti, M. Biggiogera, and and S. M. L. Barabino*, “Subnuclear Localization and Dynamics of the Pre-mRNA 3? End Processing Factor Mammalian Cleavage Factor I 68-kDa Subunit,” *Molecular biology of the cell*, vol. 18, no. 8. pp. 2991–3001, 2007, doi: 10.1091/mbc.E06-09-0846.

[59] C. M. Clemson *et al.*, “An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles,” *Mol. Cell*, 2009, doi: 10.1016/j.molcel.2009.01.026.

[60] A. H. Fox *et al.*, “Paraspeckles: A novel nuclear domain,” *Curr. Biol.*, 2002, doi: 10.1016/S0960-9822(01)00632-7.

[61] G. Pisani and B. Baron, “Nuclear paraspeckles function in mediating gene regulatory and apoptotic pathways,” *Non-coding RNA Research*. 2019, doi: 10.1016/j.ncrna.2019.11.002.

[62] Y. Wang and L. L. Chen, “Organization and function of paraspeckles,” *Essays in Biochemistry*. 2020, doi: 10.1042/EBC20200010.

[63] S. Thankachan, B. K. Bhardwaj, T. Venkatesh, and P. S. Suresh, “Long Non-coding RNA NEAT1 as an Emerging Biomarker in Breast and Gynecologic Cancers: a Systematic Overview,” *Reproductive Sciences*. 2021, doi: 10.1007/s43032-021-00481-x.

[64] P. Li, X. Zhang, L. Lin, G. Chen, and J. Chen, “Silencing of the long non-coding RNA NEAT1 suppresses ovarian cancer cell proliferation, migration and invasion,” *Int. J. Clin. Exp. Pathol.*, 2016.

[65] D. Quan *et al.*, “Identification of lncRNA NEAT1/miR-21/RRM2 axis as a novel biomarker in breast cancer,” *J. Cell. Physiol.*, 2020, doi: 10.1002/jcp.29225.

[66] Y. Shen, X. Wang, L. Lu, and W. Meng, “Aberrant NEAT1 promotes migration in endometrial cancer and as marker of poor prognosis,” *Int. J. Clin. Exp. Pathol.*, 2017.

[67] H. M. Guo, S. H. Yang, S. Z. Zhao, L. Li, M. T. Yan, and M. C. Fan, “LncRNA NEAT1 regulates cervical carcinoma proliferation and invasion by targeting AKT/PI3K,” *Eur. Rev. Med. Pharmacol. Sci.*, 2018, doi: 10.26355/EURREV_201807_15400.

[68] S. Ni, X. Zhao, and L. Ouyang, “Long non-coding RNA expression profile in vulvar squamous cell carcinoma and its clinical significance,” *Oncol. Rep.*, 2016, doi: 10.3892/or.2016.5075.

[69] Q. Guo, X. Shi, and X. Wang, “RNA and liquid-liquid phase separation,” *Non-coding RNA Research*. 2021, doi: 10.1016/j.ncrna.2021.04.003.

[70] W. T. Snead and A. S. Gladfelter, “The Control Centers of Biomolecular Phase Separation: How Membrane Surfaces, PTMs, and Active Processes Regulate Condensation,” *Mol. Cell*, vol. 76, no. 2, pp. 295–305, 2019, doi: 10.1016/j.molcel.2019.09.016.

[71] G. Kabachinski and T. U. Schwartz, “The nuclear pore complex - Structure and function at a glance,” *J. Cell Sci.*, 2015, doi: 10.1242/jcs.083246.

[72] B. Hampoelz, A. Andres-Pons, P. Kastritis, and M. Beck, “Structure and Assembly of the Nuclear Pore Complex,” *Annual Review of Biophysics*. 2019, doi: 10.1146/annurev-biophys-052118-115308.

[73] M. A. D’angelo, “Nuclear pore complexes as hubs for gene regulation,” *Nucleus*, 2018, doi: 10.1080/19491034.2017.1395542.

[74] G. Paci, J. Caria, and E. A. Lemke, “Cargo transport through the nuclear pore complex at a glance,” *J. Cell Sci.*, 2021, doi: 10.1242/jcs.247874.

[75] T. U. Schwartz, “The Structure Inventory of the Nuclear Pore Complex,” *Journal of Molecular Biology*. 2016, doi: 10.1016/j.jmb.2016.03.015.

[76] Z. Hakhverdyan *et al.*, "Dissecting the Structural Dynamics of the Nuclear Pore Complex," *Mol. Cell*, 2021, doi: 10.1016/j.molcel.2020.11.032.

[77] K. L. B. Borden, "The nuclear pore complex and mrna export in cancer," *Cancers*. 2021, doi: 10.3390/cancers13010042.

[78] R. D. Kornberg, "Chromatin structure: A repeating unit of histones and DNA," *Science (80-.).*, 1974, doi: 10.1126/science.184.4139.868.

[79] R. Barth, K. Bystricky, and H. A. Shaban, "Coupling chromatin structure and dynamics by live super-resolution imaging," *Sci. Adv.*, 2020, doi: 10.1126/sciadv.aaz2196.

[80] A. T. Annunziato, "DNA Packaging: Nucleosomes and Chromatin," in *Nature Education*, 2008.

[81] A. T. Annunziato, "DNA Packaging : Nucleosomes and Chromatin The Nucleosome : The Unit of Chromatin," *Nat. Educ.*, 2008.

[82] C. R. Clapier and B. R. Cairns, "The biology of chromatin remodeling complexes," *Annual Review of Biochemistry*. 2009, doi: 10.1146/annurev.biochem.77.062706.153223.

[83] S. Venkatesh and J. L. Workman, "Histone exchange, chromatin structure and the regulation of transcription," *Nature Reviews Molecular Cell Biology*. 2015, doi: 10.1038/nrm3941.

[84] G. Arents and E. N. Moudrianakis, "Topography of the histone octamer surface: Repeating structural motifs utilized in the docking of nucleosomal DNA," *Proc. Natl. Acad. Sci. U. S. A.*, 1993, doi: 10.1073/pnas.90.22.10489.

[85] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, "Crystal structure of the nucleosome core particle at 2.8 Å resolution," *Nature*, 1997, doi: 10.1038/38444.

[86] V. L. Makarov, S. I. Dimitrov, I. R. Tsaneva, and I. G. Pashev, "The role of histone H1 and non-structured domains of core histones in maintaining the orientation of nucleosomes within the chromatin fiber," *Biochem. Biophys. Res. Commun.*, 1984, doi: 10.1016/0006-291X(84)91193-8.

[87] J. O. Thomas, "Histone H1: Location and role," *Current Opinion in Cell Biology*. 1999, doi: 10.1016/S0955-0674(99)80042-8.

[88] S. Jun, "Chromosome, cell cycle, and entropy," *Biophysical Journal*. 2015, doi: 10.1016/j.bpj.2014.12.032.

[89] R. Thadani, J. Kamenz, S. Heeger, S. Muñoz, and F. Uhlmann, "Cell-Cycle Regulation of Dynamic Chromosome Association of the Condensin Complex," *Cell Rep.*, 2018, doi: 10.1016/j.celrep.2018.04.082.

[90] E. Morton Bradbury, "Reversible histone modification and the chromosome cell cycle," *BioEssays*. 1992, doi: 10.1002/bies.950140103.

[91] X. Chu and J. Wang, "Conformational state switching and pathways of chromosome dynamics in cell cycle," *Appl. Phys. Rev.*, 2020, doi: 10.1063/5.0007316.

[92] Y. Bhaud, D. Guillebault, J. F. Lennon, H. Defacque, M. O. Soyer-Gobillard, and H. Moreau, "Morphology and behaviour of dinoflagellate chromosomes during the cell cycle and mitosis," *J. Cell Sci.*, 2000, doi: 10.1242/jcs.113.7.1231.

[93] D. Gerlich and J. Ellenberg, "Dynamics of chromosome positioning during the cell cycle," *Current Opinion in Cell Biology*. 2003, doi: 10.1016/j.ceb.2003.10.014.

[94] X. Chu and J. Wang, "Conformational state switching and pathways of chromosome dynamics in cell cycle," *bioRxiv*, 2019, doi: 10.1101/2019.12.20.885335.

[95] C. Barrington, D. Pezic, and S. Hadjur, "Chromosome structure dynamics during the cell cycle: a structure to fit every phase," *EMBO J.*, 2017, doi: 10.15252/embj.201798014.

[96] X. Zhou, J. Wang, J. Herrmann, W. E. Moerner, and L. Shapiro, "Asymmetric division yields

progeny cells with distinct modes of regulating cell cycle-dependent chromosome methylation," *Proc. Natl. Acad. Sci. U. S. A.*, 2019, doi: 10.1073/pnas.1906119116.

[97] Z. Zhu and X. Wang, "Roles of cohesin in chromosome architecture and gene expression," *Seminars in Cell and Developmental Biology*. 2019, doi: 10.1016/j.semcdb.2018.08.004.

[98] R. Forey *et al.*, "A Role for the Mre11-Rad50-Xrs2 Complex in Gene Expression and Chromosome Organization," *Mol. Cell*, 2021, doi: 10.1016/j.molcel.2020.11.010.

[99] M. Anger, L. Radonova, A. Horakova, D. Sekach, and M. Charousova, "Impact of global transcriptional silencing on cell cycle regulation and chromosome segregation in early mammalian embryos," *International Journal of Molecular Sciences*. 2021, doi: 10.3390/ijms22169073.

[100] A. M. J. Rattray and B. Müller, "The control of histone gene expression," 2012, doi: 10.1042/BST20120065.

[101] E. Bártová, J. Krejčí, A. Harničarová, G. Galiová, and S. Kozubek, "Histone modifications and nuclear architecture: A review," *Journal of Histochemistry and Cytochemistry*. 2008, doi: 10.1369/jhc.2008.951251.

[102] A. J. Bannister and T. Kouzarides, "Regulation of chromatin by histone modifications," *Cell Research*. 2011, doi: 10.1038/cr.2011.22.

[103] V. K. Gangaraju and B. Bartholomew, "Mechanisms of ATP dependent chromatin remodeling," *Mutat. Res. - Fundam. Mol. Mech. Mutagen.*, 2007, doi: 10.1016/j.mrfmmm.2006.08.015.

[104] A. Travers, "An engine for nucleosome remodeling," *Cell*. 1999, doi: 10.1016/S0092-8674(00)80543-7.

[105] K. Mardinian, J. J. Adashek, G. P. Botta, S. Kato, and R. Kurzrock, " SMARCA4 : Implications of an Altered Chromatin-Remodeling Gene for Cancer Development and Therapy ," *Mol. Cancer Ther.*, 2021, doi: 10.1158/1535-7163.mct-21-0433.

[106] M. Abou El Hassan, T. Yu, L. Song, and R. Bremner, " Polycomb Repressive Complex 2 Confers BRG1 Dependency on the CIITA Locus ," *J. Immunol.*, 2015, doi: 10.4049/jimmunol.1403247.

[107] D. F. V. Corona *et al.*, "ISWI is an ATP-dependent nucleosome remodeling factor," *Mol. Cell*, 1999, doi: 10.1016/S1097-2765(00)80314-7.

[108] M. Yun, J. Wu, J. L. Workman, and B. Li, "Readers of histone modifications," *Cell Research*. 2011, doi: 10.1038/cr.2011.42.

[109] Di. MolinaSerrano, Di. Kyriakou, and A. Kirmizis, "Histone modifications as an intersection between diet and longevity," *Frontiers in Genetics*. 2019, doi: 10.3389/fgene.2019.00192.

[110] J. Füllgrabe, E. Kavanagh, and B. Joseph, "Histone onco-modifications," *Oncogene*. 2011, doi: 10.1038/onc.2011.121.

[111] S. G. Swygert and C. L. Peterson, "Chromatin dynamics: Interplay between remodeling enzymes and histone modifications," *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. 2014, doi: 10.1016/j.bbagrm.2014.02.013.

[112] R. F. Luco, Q. Pan, K. Tominaga, B. J. Blencowe, O. M. Pereira-Smith, and T. Misteli, "Regulation of alternative splicing by histone modifications," *Science (80-.).*, 2010, doi: 10.1126/science.1184208.

[113] L. L. Cao, C. Shen, and W. G. Zhu, "Histone modifications in DNA damage response," *Science China Life Sciences*. 2016, doi: 10.1007/s11427-016-5011-z.

[114] S. Ramazi, A. Allahverdi, and J. Zahiri, "Evaluation of post-translational modifications in histone proteins: A review on histone modification defects in developmental and neurological disorders," *Journal of Biosciences*. 2020, doi: 10.1007/s12038-020-00099-2.

[115] R. Karlić, H. R. Chung, J. Lasserre, K. Vlahoviček, and M. Vingron, "Histone modification

levels are predictive for gene expression," *Proc. Natl. Acad. Sci. U. S. A.*, 2010, doi: 10.1073/pnas.0909344107.

[116] X. Dong and Z. Weng, "The correlation between histone modifications and gene expression," *Epigenomics*. 2013, doi: 10.2217/epi.13.13.

[117] B. Stillman, "Histone Modifications: Insights into Their Influence on Gene Expression," *Cell*. 2018, doi: 10.1016/j.cell.2018.08.032.

[118] A. K. Singh, T. Schauer, L. Pfaller, T. Straub, and F. Mueller-Planitz, "The biogenesis and function of nucleosome arrays," *Nat. Commun.*, 2021, doi: 10.1038/s41467-021-27285-6.

[119] Ž. M. Svedružić, C. Wang, J. V. Kosmoski, and M. J. Smerdon, "Accommodation and repair of a UV photoproduct in DNA at different rotational settings on the nucleosome surface," *J. Biol. Chem.*, 2005, doi: 10.1074/jbc.M509478200.

[120] R. Wang, K. Yang, S. Banerjee, and M. M. Greenberg, "Rotational Effects within Nucleosome Core Particles on Abasic Site Reactivity," *Biochemistry*, 2018, doi: 10.1021/acs.biochem.8b00493.

[121] J. Hapala and E. N. Trifonov, "High resolution positioning of intron ends on the nucleosomes," *Gene*, 2011, doi: 10.1016/j.gene.2011.08.022.

[122] O. Fernandez-Capetillo, A. Lee, M. Nussenzweig, and A. Nussenzweig, "H2AX: The histone guardian of the genome," *DNA Repair*. 2004, doi: 10.1016/j.dnarep.2004.03.024.

[123] S. Wang, X. M. Wu, C. H. Liu, J. Y. Shang, F. Gao, and H. S. Guo, "Verticillium dahliae chromatin remodeling facilitates the DNA damage repair in response to plant ROS stress," *PLoS Pathog.*, 2020, doi: 10.1371/journal.ppat.1008481.

[124] M. Fink, J. S. Thompson, and F. Thoma, "Contributions of histone H3 nucleosome core surface mutations to chromatin structures, silencing and DNA repair," *PLoS One*, 2011, doi: 10.1371/journal.pone.0026210.

[125] Z. Dai, S. J. Mentch, X. Gao, S. N. Nichenametla, and J. W. Locasale, "Methionine metabolism influences genomic architecture and gene expression through H3K4me3 peak width," *Nat. Commun.*, 2018, doi: 10.1038/s41467-018-04426-y.

[126] K. C. Yuen, B. D. Slaughter, and J. L. Gerton, "Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters," *Sci. Adv.*, 2017, doi: 10.1126/sciadv.1700191.

[127] S. S. Dhar *et al.*, "MLL4 Is Required to Maintain Broad H3K4me3 Peaks and Super-Enhancers at Tumor Suppressor Genes," *Mol. Cell*, 2018, doi: 10.1016/j.molcel.2018.04.028.

[128] S. Zhou *et al.*, "The COMPASS-like complex modulates fungal development and pathogenesis by regulating H3K4me3-mediated targeted gene expression in Magnaporthe oryzae," *Mol. Plant Pathol.*, 2021, doi: 10.1111/mpp.13035.

[129] M. Nadal-Ribelles *et al.*, "H3K4 monomethylation dictates nucleosome dynamics and chromatin remodeling at stress-responsive genes," *Nucleic Acids Res.*, 2015, doi: 10.1093/nar/gkv220.

[130] K. M. Tchou-Wong *et al.*, "Effects of nickel treatment on H3K4 trimethylation and gene expression," *PLoS One*, 2011, doi: 10.1371/journal.pone.0017728.

[131] L. M. Soares, P. C. He, Y. Chun, H. Suh, T. S. Kim, and S. Buratowski, "Determinants of Histone H3K4 Methylation Patterns," *Mol. Cell*, 2017, doi: 10.1016/j.molcel.2017.10.013.

[132] Y. Tan *et al.*, "Matrix softness regulates plasticity of tumour-repopulating cells via H3K9 demethylation and Sox2 expression," *Nat. Commun.*, 2014, doi: 10.1038/ncomms5619.

[133] G. Kungulovski, S. Nunna, M. Thomas, U. M. Zanger, R. Reinhardt, and A. Jeltsch, "Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained," *Epigenetics and Chromatin*, 2015, doi: 10.1186/s13072-015-0002-z.

[134] M. Ninova, K. F. Tóth, and A. A. Aravin, “The control of gene expression and cell identity by H3K9 trimethylation,” *Development (Cambridge)*. 2019, doi: 10.1242/dev.181180.

[135] E. Akoury *et al.*, “Disordered region of H3K9 methyltransferase Clr4 binds the nucleosome and contributes to its activity,” *Nucleic Acids Res.*, 2019, doi: 10.1093/nar/gkz480.

[136] Y. Cai *et al.*, “H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions,” *Nat. Commun.*, 2021, doi: 10.1038/s41467-021-20940-y.

[137] W. Lee, J. Kim, J. M. Yun, T. Ohn, and Q. Gong, “MeCP2 regulates gene expression through recognition of H3K27me3,” *Nat. Commun.*, 2020, doi: 10.1038/s41467-020-16907-0.

[138] A. Henckel, K. Nakabayashi, L. A. Sanz, R. Feil, K. Hata, and P. Arnaud, “Histone methylation is mechanistically linked to DNA methylation at imprinting control regions in mammals,” *Hum. Mol. Genet.*, 2009, doi: 10.1093/hmg/ddp277.

[139] M. Pannetier *et al.*, “PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse,” *EMBO Rep.*, 2008, doi: 10.1038/embor.2008.147.

[140] M. D. L. P. Sanchez and C. Gutierrez, “Arabidopsis ORC1 is a PHD-containing H3K4me3 effector that regulates transcription,” *Proc. Natl. Acad. Sci. U. S. A.*, 2009, doi: 10.1073/pnas.0811093106.

[141] J. D. Stender *et al.*, “Control of Proinflammatory Gene Programs by Regulated Trimethylation and Demethylation of Histone H4K20,” *Mol. Cell*, 2012, doi: 10.1016/j.molcel.2012.07.020.

[142] D. M. Nelson *et al.*, “Mapping H4K20me3 onto the chromatin landscape of senescent cells indicates a function in control of cell senescence and tumor suppression through preservation of genetic and epigenetic stability,” *Genome Biol.*, 2016, doi: 10.1186/s13059-016-1017-x.

[143] F. H. C. Crick, L. Barnett, S. Brenner, and R. J. Watts-Tobin, “General nature of the genetic code for proteins,” *Nature*, 1961, doi: 10.1038/1921227a0.

[144] D. A. Hiller and S. A. Strobel, “The chemical versatility of RNA,” *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2011, doi: 10.1098/rstb.2011.0143.

[145] J. E. Kirk, “Nucleotides and nucleic acids.,” *Monogr. Atheroscler.*, 1974, doi: 10.5005/jp/books/12717_8.

[146] J. D. Watson and S. Devons, “ The Double Helix: A Personal Account of the Discovery of the Structure of DNA ,” *Phys. Today*, 1968, doi: 10.1063/1.3035117.

[147] B. Glass, “ The Double Helix. A Personal Account of the Discovery of the Structure of DNA. Text, Commentary, Reviews, Original Papers . James D. Watson , Gunther S. Stent ,” *Q. Rev. Biol.*, 1981, doi: 10.1086/412435.

[148] K. Manchester, “Discovering the DNA double helix and the secret of life,” *South African Journal of Science*. 2003.

[149] J. D. Ferrara, “The annotated and illustrated double helix,” *Crystallogr. Rev.*, 2013, doi: 10.1080/0889311x.2013.812081.

[150] P. A. Wardhani, *OpenStax-Biology*. 2015.

[151] G. Costa dos Santos, M. Renovato-Martins, and N. M. de Brito, “The remodel of the ‘central dogma’: a metabolomics interaction perspective,” *Metabolomics*. 2021, doi: 10.1007/s11306-021-01800-8.

[152] F. Crick, “Central dogma of molecular biology,” *Nature*, 1970, doi: 10.1038/227561a0.

[153] K. Salokas, R. G. Weldatsadik, and M. Varjosalo, “Human transcription factor and protein kinase gene fusions in human cancer,” *Sci. Rep.*, 2020, doi: 10.1038/s41598-020-71040-8.

[154] J. Y. Youn *et al.*, “Functional analysis of kinases and transcription factors in Saccharomyces

cerevisiae using an integrated overexpression library," *G3 Genes, Genomes, Genet.*, 2017, doi: 10.1534/g3.116.038471.

[155] J. K. Rimel *et al.*, "Selective inhibition of CDK7 reveals high-confidence targets and new models for TFIIH function in transcription," *Genes Dev.*, 2020, doi: 10.1101/gad.341545.120.

[156] D. E. Lyons, S. McMahon, and M. Ott, "A combinatorial view of old and new RNA polymerase II modifications," *Transcription*. 2020, doi: 10.1080/21541264.2020.1762468.

[157] S. Hahn, "Structure and mechanism of the RNA polymerase II transcription machinery," *Nature Structural and Molecular Biology*. 2004, doi: 10.1038/nsmb763.

[158] S. Osman and P. Cramer, "Structural Biology of RNA Polymerase II Transcription: 20 Years on," *Annual Review of Cell and Developmental Biology*. 2020, doi: 10.1146/annurev-cellbio-042020-021954.

[159] A. C. Schier and D. J. Taatjes, "Structure and mechanism of the RNA polymerase II transcription machinery," *Genes and Development*. 2020, doi: 10.1101/gad.335679.119.

[160] D. Reines, "Recent advances in understanding RNA polymerase II structure and function," *Fac. Rev.*, 2020, doi: 10.12703/b/9-11.

[161] N. A. Woychik and M. Hampsey, "The RNA polymerase II machinery: Structure illuminates function," *Cell*. 2002, doi: 10.1016/S0092-8674(02)00646-3.

[162] X. Liu, D. A. Bushnell, and R. D. Kornberg, "RNA polymerase II transcription: Structure and mechanism," *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. 2013, doi: 10.1016/j.bbagrm.2012.09.003.

[163] M. Hantsche and P. Cramer, "Conserved RNA polymerase II initiation complex structure," *Current Opinion in Structural Biology*. 2017, doi: 10.1016/j.sbi.2017.03.013.

[164] P. Cramer, "RNA polymerase II structure: From core to functional complexes," *Current Opinion in Genetics and Development*. 2004, doi: 10.1016/j.gde.2004.01.003.

[165] E. Brookes and A. Pombo, "Modifications of RNA polymerase II are pivotal in regulating gene expression states," *EMBO Reports*. 2009, doi: 10.1038/embor.2009.221.

[166] G. Napolitano, L. Lania, and B. Majello, "RNA Polymerase II CTD Modifications: How many tales from a single tail," *J. Cell. Physiol.*, 2014, doi: 10.1002/jcp.24483.

[167] Y. E. Guo *et al.*, "Pol II phosphorylation regulates a switch between transcriptional and splicing condensates," *Nature*. 2019, doi: 10.1038/s41586-019-1464-0.

[168] E. A. Bowman and W. G. Kelly, "RNA Polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation: A tail of two kinases," *Nucleus (United States)*. 2014, doi: 10.4161/nucl.29347.

[169] N. Fong, T. Saldi, R. M. Sheridan, M. A. Cortazar, and D. L. Bentley, "RNA Pol II Dynamics Modulate Co-transcriptional Chromatin Modification, CTD Phosphorylation, and Transcriptional Direction," *Mol. Cell*, vol. 66, no. 4, pp. 546-557.e3, 2017, doi: 10.1016/j.molcel.2017.04.016.

[170] C. Plaschka *et al.*, "Architecture of the RNA polymerase II-Mediator core initiation complex," *Nature*, 2015, doi: 10.1038/nature14229.

[171] M. Boehning *et al.*, "RNA polymerase II clustering through carboxy-terminal domain phase separation," *Nat. Struct. Mol. Biol.*, 2018, doi: 10.1038/s41594-018-0112-y.

[172] J. Y. Chotiner, D. J. Wolgemuth, and P. Jeremy Wang, "Functions of cyclins and CDKs in mammalian gametogenesis," *Biology of Reproduction*. 2019, doi: 10.1093/biolre/ioz070.

[173] J. L. Kohlmeyer, D. J. Gordon, M. R. Tanas, V. Monga, R. D. Dodd, and D. E. Quelle, "CDKs in sarcoma: Mediators of disease and emerging therapeutic targets," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21083018.

[174] J. M. Espinosa, "Transcriptional CDKs in the spotlight," *Transcription*. 2019, doi: 10.1080/21541264.2019.1597479.

[175] J. Chou, D. A. Quigley, T. M. Robinson, F. Y. Feng, and A. Ashworth, "Transcription-associated cyclin-dependent kinases as targets and biomarkers for cancer therapy," *Cancer Discovery*. 2020, doi: 10.1158/2159-8290.CD-19-0528.

[176] C. Engel, S. Neyer, and P. Cramer, "Distinct Mechanisms of Transcription Initiation by RNA Polymerases i and II," *Annual Review of Biophysics*. 2018, doi: 10.1146/annurev-biophys-070317-033058.

[177] V. Haberle and A. Stark, "Eukaryotic core promoters and the functional basis of transcription initiation," *Nature Reviews Molecular Cell Biology*. 2018, doi: 10.1038/s41580-018-0028-8.

[178] H. Liang, J. Du, R. M. Elhassan, X. Hou, and H. Fang, "Recent progress in development of cyclin-dependent kinase 7 inhibitors for cancer therapy," *Expert Opinion on Investigational Drugs*. 2021, doi: 10.1080/13543784.2021.1850693.

[179] S. Egloff, "CDK9 keeps RNA polymerase II on track," *Cellular and Molecular Life Sciences*. 2021, doi: 10.1007/s00018-021-03878-8.

[180] P. K. Parua, S. Kalan, B. Benjamin, M. Sansó, and R. P. Fisher, "Distinct Cdk9-phosphatase switches act at the beginning and end of elongation by RNA polymerase II," *Nat. Commun.*, 2020, doi: 10.1038/s41467-020-18173-6.

[181] S. Nagarkar, R. Wasnik, P. Govada, S. Cohen, and L. S. Shashidhara, "Promoter proximal pausing limits tumorous growth induced by the yki transcription factor in drosophila," *Genetics*, 2020, doi: 10.1534/genetics.120.303419.

[182] J. Zhang, X. Zhang, H. Huang, and Y. Ding, "A review on kinases phosphorylating the carboxyl-terminal domain of RNA polymerase II—Biological functions and inhibitors," *Bioorg. Chem.*, vol. 104, p. 104318, Nov. 2020, doi: 10.1016/J.BIOORG.2020.104318.

[183] K. Burger, M. Schlackow, and M. Gullerova, "Tyrosine kinase c-Abl couples RNA polymerase II transcription to DNA double-strand breaks," *Nucleic Acids Res.*, 2019, doi: 10.1093/nar/gkz024.

[184] M. Kciuk, B. Marciniak, M. Mojzych, and R. Kontek, "Focus on uv-induced dna damage and repair—disease relevance and protective strategies," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21197264.

[185] Y. Jiang, W. Li, L. A. Lindsey-Boltz, Y. Yang, Y. Li, and A. Sancar, "Super hotspots and super coldspots in the repair of UV-induced DNA damage in the human genome," *J. Biol. Chem.*, 2021, doi: 10.1016/j.jbc.2021.100581.

[186] P. S. Thompson and D. Cortez, "New insights into abasic site repair and tolerance," *DNA Repair (Amst).*, 2020, doi: 10.1016/j.dnarep.2020.102866.

[187] I. Tessman and M. A. Kennedy, "DNA polymerase II of Escherichia coli in the bypass of abasic sites in vivo," *Genetics*, 1994, doi: 10.1093/genetics/136.2.439.

[188] C. Walmacq *et al.*, "Mechanism of RNA polymerase II bypass of oxidative cyclopurine DNA lesions," *Proc. Natl. Acad. Sci. U. S. A.*, 2015, doi: 10.1073/pnas.1415186112.

[189] J. H. Shin, L. Xu, and D. Wang, "Mechanism of transcription-coupled DNA modification recognition," *Cell and Bioscience*. 2017, doi: 10.1186/s13578-016-0133-3.

[190] I. Stoimenov, N. Schultz, P. Gottipati, and T. Helleday, "Transcription inhibition by DRB potentiates recombinational repair of UV lesions in mammalian cells," *PLoS One*, 2011, doi: 10.1371/journal.pone.0019492.

[191] D. A. P. Rockx *et al.*, "UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II," *Proc. Natl. Acad. Sci. U. S. A.*, 2000, doi: 10.1073/pnas.180169797.

[192] J. N. Kuehner, J. W. Kaufman, and C. Moore, "Stimulation of RNA Polymerase II ubiquitination and degradation by yeast mRNA 3′-end processing factors is a conserved DNA damage response in eukaryotes," *DNA Repair (Amst).*, vol. 57, pp. 151–160, Sep. 2017, doi: 10.1016/j.dnarep.2017.07.006.

[193] M. T. Pfeifer, P. Koepke, and J. Reuder, "Effects of altitude and aerosol on UV radiation," *J. Geophys. Res. Atmos.*, 2006, doi: 10.1029/2005JD006444.

[194] J. Calbó, D. Pagès, and J. A. González, "Empirical studies of cloud effects on UV radiation: A review," *Reviews of Geophysics*. 2005, doi: 10.1029/2004RG000155.

[195] R. L. McKenzie, L. O. Björn, A. Bais, and M. Ilyasd, "Changes in biologically active ultraviolet radiation reaching the Earth's surface," *Photochem. Photobiol. Sci.*, 2003, doi: 10.1039/b211155c.

[196] A. P. Schuch and C. F. M. Menck, "The genotoxic effects of DNA lesions induced by artificial UV-radiation and sunlight," *J. Photochem. Photobiol. B Biol.*, 2010, doi: 10.1016/j.jphotobiol.2010.03.004.

[197] H. Chen, R. Li, S. Li, J. Andreásson, and J. H. Choi, "Conformational effects of UV light on DNA origami," *J. Am. Chem. Soc.*, 2017, doi: 10.1021/jacs.6b10821.

[198] A. Besaratinia *et al.*, "DNA lesions induced by UV A1 and B radiation in human cells: Comparative analyses in the overall genome and in the p53 tumor suppressor gene," *Proc. Natl. Acad. Sci. U. S. A.*, 2005, doi: 10.1073/pnas.0502311102.

[199] J. Moser *et al.*, "The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions," *DNA Repair (Amst).*, 2005, doi: 10.1016/j.dnarep.2005.01.001.

[200] Q. Sheng *et al.*, "A streamlined solution for processing, elucidating and quality control of cyclobutane pyrimidine dimer sequencing data," *Nature Protocols*. 2021, doi: 10.1038/s41596-021-00496-3.

[201] A. Tufegdžić Vidaković *et al.*, "Regulation of the RNAPII Pool Is Integral to the DNA Damage Response," *Cell*, 2020, doi: 10.1016/j.cell.2020.02.009.

[202] J. Xu *et al.*, "Cockayne syndrome B protein acts as an ATP-dependent processivity factor that helps RNA polymerase II overcome nucleosome barriers," *Proc. Natl. Acad. Sci. U. S. A.*, 2020, doi: 10.1073/pnas.2013379117.

[203] J. C. Weems *et al.*, "Cockayne syndrome B protein regulates recruitment of the elongin a ubiquitin ligase to sites of DNA damage," *J. Biol. Chem.*, 2017, doi: 10.1074/jbc.C117.777946.

[204] B. Ding, D. LeJeune, and S. Li, "The C-terminal repeat domain of Spt5 plays an important role in suppression of Rad26-independent transcription coupled repair," *J. Biol. Chem.*, 2010, doi: 10.1074/jbc.M109.082818.

[205] E. Renaud *et al.*, "Differential contribution of XPC, RAD23A, RAD23B and CENTRIN 2 to the UV-response in human cells," *DNA Repair (Amst).*, 2011, doi: 10.1016/j.dnarep.2011.05.003.

[206] T. van Eeuwen *et al.*, "Cryo-EM structure of TFIIH/Rad4–Rad23–Rad33 in damaged DNA opening in nucleotide excision repair," *Nat. Commun.*, 2021, doi: 10.1038/s41467-021-23684-x.

[207] J. Tang and G. Chu, "Xeroderma pigmentosum complementation group E and UV-damaged DNA-binding protein," *DNA Repair*. 2002, doi: 10.1016/S1568-7864(02)00052-6.

[208] Y. Fujiwara, C. Masutani, T. Mizukoshi, J. Kondo, F. Hanaoka, and S. Iwai, "Characterization of DNA recognition by the human UV-damaged DNA-binding protein," *J. Biol. Chem.*, 1999, doi: 10.1074/jbc.274.28.20027.

[209] M. Noe Gonzalez, D. Blears, and J. Q. Svejstrup, "Causes and consequences of RNA polymerase II stalling during transcript elongation," *Nature Reviews Molecular Cell Biology*. 2021, doi: 10.1038/s41580-020-00308-8.

[210] S. Ito *et al.*, "XPG Stabilizes TFIIH, Allowing Transactivation of Nuclear Receptors: Implications for Cockayne Syndrome in XP-G/CS Patients," *Mol. Cell*, 2007, doi: 10.1016/j.molcel.2007.03.013.

[211] J. O. Fuss and J. A. Tainer, "XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase," *DNA Repair*. 2011, doi: 10.1016/j.dnarep.2011.04.028.

[212] M. Rossignol, I. Kolb-Cheynel, and J. M. Egly, "Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH," *EMBO J.*, 1997, doi: 10.1093/emboj/16.7.1628.

[213] L. Staresincic *et al.*, "Coordination of dual incision and repair synthesis in human nucleotide excision repair," *EMBO J.*, 2009, doi: 10.1038/emboj.2009.49.

[214] J. Zhao *et al.*, "Distinct Mechanisms of Nuclease-Directed DNA-Structure-Induced Genetic Instability in Cancer Genomes," *Cell Rep.*, 2018, doi: 10.1016/j.celrep.2018.01.014.

[215] A. F. Fagbemi, B. Orelli, and O. D. Schärer, "Regulation of endonuclease activity in human nucleotide excision repair," *DNA Repair*. 2011, doi: 10.1016/j.dnarep.2011.04.022.

[216] A. M. Topolska-Woś *et al.*, "A key interaction with RPA orients XPA in NER complexes," *Nucleic Acids Res.*, 2020, doi: 10.1093/nar/gkz1231.

[217] T. Matsuda *et al.*, "DNA repair protein XPA binds replication protein A (RPA)," *J. Biol. Chem.*, 1995, doi: 10.1074/jbc.270.8.4152.

[218] Y. S. Krasikova, N. I. Rechkunova, E. A. Maltseva, and O. I. Lavrik, "RPA and XPA interaction with DNA structures mimicking intermediates of the late stages in nucleotide excision repair," *PLoS One*, vol. 13, no. 1, pp. 1–20, 2018, doi: 10.1371/journal.pone.0190782.

[219] T. Kelly and A. J. Callegari, "Dynamics of DNA replication in a eukaryotic cell," *Proc. Natl. Acad. Sci. U. S. A.*, 2019, doi: 10.1073/pnas.1818680116.

[220] A. Quinet, D. Lemaçon, and A. Vindigni, "Replication Fork Reversal: Players and Guardians," *Molecular Cell*. 2017, doi: 10.1016/j.molcel.2017.11.022.

[221] M. Petropoulos, S. Champeris Tsaniras, S. Taraviras, and Z. Lygerou, "Replication Licensing Aberrations, Replication Stress, and Genomic Instability," *Trends in Biochemical Sciences*. 2019, doi: 10.1016/j.tibs.2019.03.011.

[222] H. Zhang, Y. Xiong, and J. Chen, "DNA-protein cross-link repair: What do we know now?," *Cell and Bioscience*. 2020, doi: 10.1186/s13578-019-0366-z.

[223] X. Wei, Y. Peng, C. Bryan, and K. Yang, "Mechanisms of DNA−protein cross-link formation and repair," *Biochimica et Biophysica Acta - Proteins and Proteomics*. 2021, doi: 10.1016/j.bbapap.2021.140669.

[224] H. Ide, T. Nakano, A. M. H. Salem, and M. I. Shoulkamy, "DNA–protein cross-links: Formidable challenges to maintaining genome integrity," *DNA Repair*. 2018, doi: 10.1016/j.dnarep.2018.08.024.

[225] I. Plo *et al.*, "Association of XRCC1 and tyrosyl DNA phosphodiesterase (Tdp1) for the repair of topoisomerase I-mediated DNA lesions," *DNA Repair (Amst).*, 2003, doi: 10.1016/S1568-7864(03)00116-2.

[226] T. Nakano, X. Xu, A. M. H. Salem, M. I. Shoulkamy, and H. Ide, "Radiation-induced DNA–protein cross-links: Mechanisms and biological significance," *Free Radical Biology and Medicine*. 2017, doi: 10.1016/j.freeradbiomed.2016.11.041.

[227] K. Yang and M. M. Greenberg, "DNA-Protein Cross-Link Formation in Nucleosome Core Particles Treated with Methyl Methanesulfonate," *Chem. Res. Toxicol.*, 2019, doi: 10.1021/acs.chemrestox.9b00314.

[228] B. Vaz, M. Popovic, and K. Ramadan, "DNA–Protein Crosslink Proteolysis Repair," *Trends in Biochemical Sciences*. 2017, doi: 10.1016/j.tibs.2017.03.005.

[229] I. G. Minko, Y. Zou, and R. S. Lloyd, "Incision of DNA-protein crosslinks by UvrABC nuclease suggests a potential repair pathway involving nucleotide excision repair," *Proc. Natl. Acad. Sci. U. S. A.*, 2002, doi: 10.1073/pnas.042700399.

[230] J. Stingele, M. S. Schwarz, N. Bloemeke, P. G. Wolf, and S. Jentsch, "A DNA-dependent protease involved in DNA-protein crosslink repair," *Cell*, 2014, doi: 10.1016/j.cell.2014.04.053.

[231] X. Y. Yan *et al.*, "Insight into the role of p62 in the cisplatin resistant mechanisms of ovarian cancer," *Cancer Cell International*. 2020, doi: 10.1186/s12935-020-01196-w.

[232] C. Y. Sun, Q. Y. Zhang, G. J. Zheng, and B. Feng, "Phytochemicals: Current strategy to sensitize cancer cells to cisplatin," *Biomedicine and Pharmacotherapy*. 2019, doi: 10.1016/j.biopha.2018.12.010.

[233] S. Zhang *et al.*, "Interfering in apoptosis and DNA repair of cancer cells to conquer cisplatin resistance by platinum(iv) prodrugs," *Chem. Sci.*, 2020, doi: 10.1039/d0sc00197j.

[234] X. Zhao, J. Fu, W. Tang, L. Yu, and W. Xu, "Inhibition of serine metabolism promotes resistance to cisplatin in gastric cancer," *Onco. Targets. Ther.*, 2020, doi: 10.2147/OTT.S246430.

[235] Y. Fujikawa, M. Kawanishi, I. Kuraoka, and T. Yagi, "Frequencies of mutagenic translesion DNA synthesis over cisplatin-guanine intra-strand crosslinks in lacZ plasmids propagated in human cells," *Mutat. Res. - Genet. Toxicol. Environ. Mutagen.*, 2014, doi: 10.1016/j.mrgentox.2014.05.006.

[236] N. Chowdhury, T. L. Wood, M. Martínez-Vázquez, R. García-Contreras, and T. K. Wood, "DNA-crosslinker cisplatin eradicates bacterial persister cells," *Biotechnol. Bioeng.*, 2016, doi: 10.1002/bit.25963.

[237] A. Boot *et al.*, "In-depth characterization of the cisplatin mutational signature in human cell lines and in esophageal and liver tumors," *In-depth Charact. cisplatin Mutat. Signat. Hum. cell lines esophageal liver tumors*, 2017, doi: 10.1101/189233.

[238] S. Feuerhahn *et al.*, "XPF-dependent DNA breaks and RNA polymerase II arrest induced by antitumor DNA interstrand crosslinking-mimetic alkaloids," *Chem. Biol.*, 2011, doi: 10.1016/j.chembiol.2011.06.007.

[239] A. Ui, N. Chiba, and A. Yasui, "Relationship among DNA double-strand break (DSB), DSB repair, and transcription prevents genome instability and cancer," *Cancer Sci.*, 2020, doi: 10.1111/cas.14404.

[240] P. Caron, J. van der Linden, and H. van Attikum, "Bon voyage: A transcriptional journey around DNA breaks," *DNA Repair*. 2019, doi: 10.1016/j.dnarep.2019.102686.

[241] A. Cristini, J. H. Park, G. Capranico, G. Legube, G. Favre, and O. Sordet, "DNA-PK triggers histone ubiquitination and signaling in response to DNA double-strand breaks produced during the repair of transcription-blocking topoisomerase I lesions," *Nucleic Acids Res.*, 2016, doi: 10.1093/nar/gkv1196.

[242] F. E. Machour and N. Ayoub, "Transcriptional Regulation at DSBs: Mechanisms and Consequences," *Trends in Genetics*. 2020, doi: 10.1016/j.tig.2020.01.001.

[243] S. W. Awwad, E. R. Abu-Zhayia, N. Guttmann-Raviv, and N. Ayoub, " NELF -E is recruited to DNA double-strand break sites to promote transcriptional repression and repair ," *EMBO Rep.*, 2017, doi: 10.15252/embr.201643191.

[244] L. A. Bishara, F. E. Machour, S. W. Awwad, and N. Ayoub, "NELF complex fosters BRCA1 and RAD51 recruitment to DNA damage sites and modulates sensitivity to PARP inhibition," *DNA Repair (Amst).*, 2021, doi: 10.1016/j.dnarep.2020.103025.

[245] N. M. Shanbhan, I. U. Rafalska-Metcalf, C. Balane- Bolivar, S. M. Janicki, and R. A. Greenberg,

"An ATM-Dependent Transcriptional Silencing Program is Transmitted Through Chromatin in Cis to DNA Double Strand Breaks," *Cell*, 2011.

[246] N. L. Batenburg, E. L. Thompson, E. A. Hendrickson, and X. Zhu, " Cockayne syndrome group B protein regulates DNA double-strand break repair and checkpoint activation ," *EMBO J.*, 2015, doi: 10.15252/embj.201490041.

[247] K. Burger, R. F. Ketley, and M. Gullerova, "Beyond the Trinity of ATM, ATR, and DNA-PK: Multiple Kinases Shape the DNA Damage Response in Concert With RNA Metabolism," *Frontiers in Molecular Biosciences*. 2019, doi: 10.3389/fmolb.2019.00061.

[248] P. Caron *et al.*, "WWP2 ubiquitylates RNA polymerase II for DNA-PK-dependent transcription arrest and repair at DNA breaks," *Genes Dev.*, 2019, doi: 10.1101/gad.321943.118.

[249] D. Munnur *et al.*, "NR4A Nuclear Receptors Target Poly-ADP-Ribosylated DNA-PKcs Protein to Promote DNA Repair," *Cell Rep.*, 2019, doi: 10.1016/j.celrep.2019.01.083.

[250] S. M. Vos, L. Farnung, H. Urlaub, and P. Cramer, "Structure of paused transcription complex Pol II–DSIF–NELF," *Nature*, 2018, doi: 10.1038/s41586-018-0442-2.

[251] M. C. Caron *et al.*, "Poly(ADP-ribose) polymerase-1 antagonizes DNA resection at double-strand breaks," *Nat. Commun.*, 2019, doi: 10.1038/s41467-019-10741-9.

[252] S. Jin and D. T. Weaver, "Double-strand break repair by Ku70 requires heterodimerization with Ku80 and DNA binding functions," *EMBO J.*, 1997, doi: 10.1093/emboj/16.22.6874.

[253] M. Van Der Burg, J. J. M. Van Dongen, and D. C. Van Gent, "DNA-PKcs deficiency in human: Long predicted, finally found," *Current Opinion in Allergy and Clinical Immunology*. 2009, doi: 10.1097/ACI.0b013e3283327e41.

[254] C. Allen, A. Kurimasa, M. A. Brenneman, D. J. Chen, and J. A. Nickoloff, "DNA-dependent protein kinase suppresses double-strand break-induced and spontaneous homologous recombination," *Proc. Natl. Acad. Sci. U. S. A.*, 2002, doi: 10.1073/pnas.052545899.

[255] Y. Liu, E. V. Efimova, A. Ramamurthy, and S. J. Kron, "Repair-independent functions of DNA-PKcs protect irradiated cells from mitotic slippage and accelerated senescence," *J. Cell Sci.*, 2019, doi: 10.1242/jcs.229385.

[256] W. Y. Mansour, K. Borgmann, C. Petersen, E. Dikomey, and J. Dahm-Daphi, "The absence of Ku but not defects in classical non-homologous end-joining is required to trigger PARP1-dependent end-joining," *DNA Repair (Amst).*, 2013, doi: 10.1016/j.dnarep.2013.10.005.

[257] C. Wang *et al.*, "Rational combination therapy for hepatocellular carcinoma with PARP1 and DNA-PK inhibitors," *Proc. Natl. Acad. Sci. U. S. A.*, 2020, doi: 10.1073/pnas.2002917117.

[258] T. Ochi *et al.*, "PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair," *Science (80-.).*, 2015, doi: 10.1126/science.1261971.

[259] M. Bétermier, P. Bertrand, and B. S. Lopez, "Is Non-Homologous End-Joining Really an Inherently Error-Prone Process?," *PLoS Genetics*. 2014, doi: 10.1371/journal.pgen.1004086.

[260] M. Mcvey, S. E. Lee, H. Avenue, and S. Antonio, "MMEJ repair of double-strand breaks: deleted sequences and alternative endings," *Trends Genet.*, 2017.

[261] A. M. Yu and M. McVey, "Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions," *Nucleic Acids Res.*, 2010, doi: 10.1093/nar/gkq379.

[262] B. Pardo, B. Gómez-González, and A. Aguilera, "DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship.," *Cell. Mol. Life Sci.*, 2009, doi: 10.1007/s00018-009-8740-3.

[263] N. Arnoult *et al.*, "Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN," *Nature*, 2017, doi: 10.1038/nature24023.

[264] G. Iliakis, "Backup pathways of NHEJ in cells of higher eukaryotes: Cell cycle dependence,"

Radiotherapy and Oncology. 2009, doi: 10.1016/j.radonc.2009.06.024.

[265] J. Her and S. F. Bunting, "How cells ensure correct repair of DNA double-strand breaks," *Journal of Biological Chemistry*. 2018, doi: 10.1074/jbc.TM118.000371.

[266] Z. Mao, M. Bozzella, A. Seluanov, and V. Gorbunova, "DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells," *Cell Cycle*, 2008, doi: 10.4161/cc.7.18.6679.

[267] X. Zhao *et al.*, "Cell cycle-dependent control of homologous recombination," *Acta Biochimica et Biophysica Sinica*. 2017, doi: 10.1093/abbs/gmx055.

[268] Y. Chen *et al.*, "A PARP1-BRG1-SIRT1 axis promotes HR repair by reducing nucleosome density at DNA damage sites," *Nucleic Acids Res.*, 2019, doi: 10.1093/nar/gkz592.

[269] S. Ying, F. C. Hamdy, and T. Helleday, "Mre11-dependent degradation of stalled DNA replication forks is prevented by BRCA2 and PARP1," *Cancer Res.*, 2012, doi: 10.1158/0008-5472.CAN-11-3417.

[270] N. Schultz, E. Lopez, N. Saleh-Gohari, and T. Helleday, "Poly(ADP-ribose) polymerase (PARP-1) has a controlling role in homologous recombination," *Nucleic Acids Res.*, 2003, doi: 10.1093/nar/gkg703.

[271] K. L. Jensen and P. Russell, "Ctp1-dependent clipping and resection of DNA double-strand breaks by Mre11 endonuclease complex are not genetically separable," *Nucleic Acids Res.*, 2016, doi: 10.1093/nar/gkw557.

[272] H. Wang *et al.*, "CtIP protein dimerization is critical for its recruitment to chromosomal DNA double-stranded breaks," *J. Biol. Chem.*, 2012, doi: 10.1074/jbc.M112.355354.

[273] P. Langerak, E. Mejia-Ramirez, O. Limbo, and P. Russell, "Release of Ku and MRN from DNA ends by Mre11 nuclease activity and Ctp1 is required for homologous recombination repair of double-strand breaks," *PLoS Genet.*, 2011, doi: 10.1371/journal.pgen.1002271.

[274] R. Scully, A. Panday, R. Elango, and N. A. Willis, "DNA double-strand break repair-pathway choice in somatic mammalian cells," *Nature Reviews Molecular Cell Biology*. 2019, doi: 10.1038/s41580-019-0152-0.

[275] S. Qiu and J. Huang, "MRN complex is an essential effector of DNA damage repair," *Journal of Zhejiang University: Science B*. 2021, doi: 10.1631/jzus.B2000289.

[276] K. Czornak, S. Chughtai, and K. H. Chrzanowska, "Mystery of DNA repair: The role of the MRN complex and ATM kinase in DNA damage repair," *Journal of Applied Genetics*. 2008, doi: 10.1007/BF03195638.

[277] R. Sarkar, U. Patra, M. Lo, A. Mukherjee, A. Biswas, and M. Chawla-Sarkar, "Rotavirus activates a noncanonical ATM-Chk2 branch of DNA damage response during infection to positively regulate viroplasm dynamics," *Cell. Microbiol.*, 2020, doi: 10.1111/cmi.13149.

[278] D. R. Whelan and E. Rothenberg, "Super-resolution mapping of cellular double-strand break resection complexes during homologous recombination," *Proc. Natl. Acad. Sci. U. S. A.*, 2021, doi: 10.1073/PNAS.2021963118.

[279] N. Tomimatsu *et al.*, "DNA-damage-induced degradation of EXO1 exonuclease limits DNA end resection to ensure accurate DNA repair," *J. Biol. Chem.*, 2017, doi: 10.1074/jbc.M116.772475.

[280] E. Bolderson *et al.*, "Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks," *Nucleic Acids Res.*, vol. 38, no. 6, pp. 1821–1831, Apr. 2010, doi: 10.1093/nar/gkp1164.

[281] A. Bakr *et al.*, "Involvement of ATM in homologous recombination after end resection and RAD51 nucleofilament formation," *Nucleic Acids Res.*, 2015, doi: 10.1093/nar/gkv160.

[282] I. S. Shkundina, A. A. Gall, A. Dick, S. Cocklin, and A. V. Mazin, "New rad51 inhibitors to

target homologous recombination in human cells," *Genes (Basel).*, 2021, doi: 10.3390/genes12060920.

[283] I. E. Wassing and F. Esashi, "RAD51: Beyond the break," *Seminars in Cell and Developmental Biology*. 2021, doi: 10.1016/j.semcdb.2020.08.010.

[284] O. Belan *et al.*, "Single-molecule analysis reveals cooperative stimulation of Rad51 filament nucleation and growth by mediator proteins," *Mol. Cell*, 2021, doi: 10.1016/j.molcel.2020.12.020.

[285] A. Piazza, S. Shah, W. D. Wright, S. K. Gore, R. Koszul, and W. D. Heyer, "Dynamic Processing of Displacement Loops During Recombinational DNA Repair," *Dyn. Process. Displac. Loops Dur. Recomb. DNA Repair*, 2018, doi: 10.1101/421990.

[286] J. Li, H. Sun, Y. Huang, Y. Wang, Y. Liu, and X. Chen, "Pathways and assays for DNA double-strand break repair by homologous recombination," *Acta Biochimica et Biophysica Sinica*. 2019, doi: 10.1093/abbs/gmz076.

[287] S. S. Jenkins, S. Mukherjee, and W. D. Heyer, "DNA Repair by Homologous Recombination," in *Encyclopedia of Cell Biology*, 2016.

[288] C. Z. Bachrati, R. H. Borts, and I. D. Hickson, "Mobile D-loops are a preferred substrate for the Bloom's syndrome helicase," *Nucleic Acids Res.*, 2006, doi: 10.1093/nar/gkl258.

[289] A. Shibata *et al.*, "DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct MRE11 Nuclease Activities," *Mol. Cell*, 2014, doi: 10.1016/j.molcel.2013.11.003.

[290] M. D. Lavigne, D. Konstantopoulos, K. Z. Ntakou-Zamplara, A. Liakos, and M. Fousteri, "Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate," *Nat. Commun.*, 2017, doi: 10.1038/s41467-017-02145-4.

[291] L. Williamson *et al.*, "UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene," *Cell*, vol. 168, no. 5, pp. 843-855.e13, Feb. 2017, doi: 10.1016/J.CELL.2017.01.019.

[292] Z. Luo, J. Zheng, Y. Lu, and D. B. Bregman, "Ultraviolet radiation alters the phosphorylation of RNA polymerase II large subunit and accelerates its proteasome-dependent degradation," *Mutat. Res. - DNA Repair*, 2001, doi: 10.1016/S0921-8777(01)00097-0.

[293] S. Adar, J. Hu, J. D. Lieb, and A. Sancar, "Genome-wide kinetics of DNA excision repair in relation to chromatin state and mutagenesis," *Proc. Natl. Acad. Sci.*, vol. 113, no. 15, pp. E2124–E2133, 2016, doi: 10.1073/pnas.1603388113.

[294] L. C. Andrade-Lima, A. Veloso, M. T. Paulsen, C. F. M. Menck, and M. Ljungman, "DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes," *Nucleic Acids Res.*, vol. 43, no. 5, pp. 2744–2756, Mar. 2015, doi: 10.1093/NAR/GKV148.

[295] M. E. Borisova *et al.*, "P38-MK2 signaling axis regulates RNA metabolism after UV-light-induced DNA damage," *Nat. Commun.*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-018-03417-3.

[296] Y. Song *et al.*, "DNA Damage Induces Dynamic Associations of BRD4/P-TEFb With Chromatin and Modulates Gene Transcription in a BRD4-Dependent and -Independent Manner," *Front. Mol. Biosci.*, 2020, doi: 10.3389/fmolb.2020.618088.

[297] N. J. Proudfoot, A. Furger, and M. J. Dye, "Integrating mRNA processing with transcription," *Cell*. 2002, doi: 10.1016/S0092-8674(02)00617-7.

[298] R. Adair, G. W. Liebisch, Y. Su, and A. M. Colberg-Poley, "Alteration of cellular RNA splicing and polyadenylation machineries during productive human cytomegalovirus infection," *J. Gen. Virol.*, 2004, doi: 10.1099/vir.0.80450-0.

[299] E. A. Obeng, C. Stewart, and O. Abdel-Wahab, "Altered RNA processing in cancer pathogenesis and therapy," *Cancer Discovery*. 2019, doi: 10.1158/2159-8290.CD-19-0399.

[300] K. Nilsson, C. Wu, and S. Schwartz, "Role of the DNA damage response in human papillomavirus RNA splicing and polyadenylation," *International Journal of Molecular Sciences*. 2018, doi: 10.3390/ijms19061735.

[301] A. R. Kornblihtt, I. E. Schor, M. Alló, G. Dujardin, E. Petrillo, and M. J. Muñoz, "Alternative splicing: A pivotal step between eukaryotic transcription and translation," *Nature Reviews Molecular Cell Biology*. 2013, doi: 10.1038/nrm3525.

[302] Y. H. E. Hsiao *et al.*, "RNA editing in nascent RNA affects pre-mRNA splicing," *Genome Res.*, 2018, doi: 10.1101/gr.231209.117.

[303] B. L. Angarola and O. Anczuków, "Splicing alterations in healthy aging and disease," *Wiley Interdisciplinary Reviews: RNA*. 2021, doi: 10.1002/wrna.1643.

[304] C. H. Su, D. Dhananjaya, and W. Y. Tarn, "Alternative splicing in neurogenesis and brain development," *Frontiers in Molecular Biosciences*. 2018, doi: 10.3389/fmolb.2018.00012.

[305] D. Lipscombe and E. J. Lopez-Soto, "Epigenetic control of ion channel expression and cell-specific splicing in nociceptors: Chronic pain mechanisms and potential therapeutic targets," *Channels*. 2021, doi: 10.1080/19336950.2020.1860383.

[306] Z. Siegfried, S. Bonomi, C. Ghigna, and R. Karni, "Regulation of the Ras-MAPK and PI3K-mTOR signalling pathways by alternative splicing in cancer," *International Journal of Cell Biology*. 2013, doi: 10.1155/2013/568931.

[307] S. C. Bonnal, I. López-Oreja, and J. Valcárcel, "Roles and mechanisms of alternative splicing in cancer — implications for care," *Nature Reviews Clinical Oncology*. 2020, doi: 10.1038/s41571-020-0350-x.

[308] M. Tresini, J. A. Marteijn, and W. Vermeulen, "Bidirectional coupling of splicing and ATM signaling in response to transcription-blocking DNA damage," *RNA Biol.*, vol. 13, no. 3, pp. 272–278, 2016, doi: 10.1080/15476286.2016.1142039.

[309] M. N. Bogdał, B. Hat, M. Kochańczyk, and T. Lipniacki, "Levels of pro-apoptotic regulator Bad and anti-apoptotic regulator Bcl-xL determine the type of the apoptotic logic gate," *BMC Syst. Biol.*, 2013, doi: 10.1186/1752-0509-7-67.

[310] L. Su *et al.*, "Ultraviolet-Ray-Induced Sea Cucumber (Stichopus japonicus) Melting Is Mediated by the Caspase-Dependent Mitochondrial Apoptotic Pathway," *J. Agric. Food Chem.*, 2018, doi: 10.1021/acs.jafc.7b03888.

[311] M. J. Muñoz *et al.*, "Major Roles for Pyrimidine Dimers, Nucleotide Excision Repair, and ATR in the Alternative Splicing Response to UV Irradiation," *Cell Rep.*, vol. 18, no. 12, pp. 2868–2879, 2017, doi: 10.1016/j.celrep.2017.02.066.

[312] M. P. Stokes *et al.*, "Profiling of UV-induced ATM/ATR signaling pathways," *Proc. Natl. Acad. Sci. U. S. A.*, 2007, doi: 10.1073/pnas.0707579104.

[313] I. G. Ustyantsev, J. S. Golubchikova, O. R. Borodulina, and D. A. Kramerov, "Canonical and noncanonical RNA polyadenylation," *Molecular Biology*. 2017, doi: 10.1134/S0026893317010186.

[314] B. Fusby, S. Kim, B. Erickson, H. Kim, M. L. Peterson, and D. L. Bentley, "Coordination of RNA Polymerase II Pausing and 3' end processing factor recruitment with alternative polyadenylation," *Mol. Cell. Biol.*, 2015, doi: 10.1128/mcb.00898-15.

[315] F. Ren, N. Zhang, L. Zhang, E. Miller, and J. J. Pu, "Alternative Polyadenylation: a new frontier in post transcriptional regulation," *Biomark. Res.*, vol. 8, no. 1, pp. 1–10, Dec. 2020, doi: 10.1186/S40364-020-00249-6/FIGURES/4.

[316] J. Zhao, L. Hyman, and C. Moore, "Formation of mRNA 3′ Ends in Eukaryotes: Mechanism, Regulation, and Interrelationships with Other Steps in mRNA Synthesis," *Microbiol. Mol. Biol. Rev.*, 1999, doi: 10.1128/mmbr.63.2.405-445.1999.

[317] Y. Zhu *et al.*, "Molecular Mechanisms for CFIm-Mediated Regulation of mRNA Alternative Polyadenylation," *Mol. Cell*, 2018, doi: 10.1016/j.molcel.2017.11.031.

[318] M. Clerici, M. Faini, L. M. Muckenfuss, R. Aebersold, and M. Jinek, "Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex," *Nat. Struct. Mol. Biol.*, 2018, doi: 10.1038/s41594-017-0020-6.

[319] L. Schönemann *et al.*, "Reconstitution of CPSF active in polyadenylation: Recognition of the polyadenylation signal by WDR33," *Genes Dev.*, 2014, doi: 10.1101/gad.250985.114.

[320] Y. Zhang, Y. Sun, Y. Shi, T. Walz, and L. Tong, "Structural Insights into the Human Pre-mRNA 3′-End Processing Machinery," *Mol. Cell*, 2020, doi: 10.1016/j.molcel.2019.11.005.

[321] Y. Sun, K. Hamilton, and L. Tong, "Recent molecular insights into canonical pre-mRNA 3'-end processing," *Transcription*. 2020, doi: 10.1080/21541264.2020.1777047.

[322] L. Muniz, L. Davidson, and S. West, "Poly(A) Polymerase and the Nuclear Poly(A) Binding Protein, PABPN1, Coordinate the Splicing and Degradation of a Subset of Human Pre-mRNAs," *Mol. Cell. Biol.*, 2015, doi: 10.1128/mcb.00123-15.

[323] I. Vlatkovic *et al.*, "Poly(A) binding protein nuclear 1 regulates the polyadenylation of key synaptic plasticity genes and plays a role in homeostatic plasticity," *bioRxiv*, 2017, doi: 10.1101/121194.

[324] E. De Klerk *et al.*, "Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation," *Nucleic Acids Res.*, 2012, doi: 10.1093/nar/gks655.

[325] T. Abbassi-Daloii *et al.*, "An alanine expanded PABPN1 causes increased utilization of intronic polyadenylation sites," *npj Aging Mech. Dis.*, 2017, doi: 10.1038/s41514-017-0007-x.

[326] M. Simonelig, "PABPN1 shuts down alternative poly(A) sites," *Cell Res.*, 2012, doi: 10.1038/cr.2012.86.

[327] A. R. Gruber, G. Martin, W. Keller, and M. Zavolan, "Cleavage factor Im is a key regulator of 3' UTR length," *RNA Biol.*, vol. 9, no. 12, pp. 1405–1412, 2012, doi: 10.4161/rna.22570.

[328] I. Pereira-Castro and A. Moreira, "On the function and relevance of alternative 3′-UTRs in gene expression regulation," *Wiley Interdisciplinary Reviews: RNA*. 2021, doi: 10.1002/wrna.1653.

[329] Y. Zhang *et al.*, "Alternative polyadenylation: methods, mechanism, function, and role in cancer," *Journal of Experimental and Clinical Cancer Research*. 2021, doi: 10.1186/s13046-021-01852-7.

[330] M. Dutertre, R. Sfaxi, and S. Vagner, "Reciprocal Links between Pre-messenger RNA 3′-End Processing and Genome Stability," *Trends Biochem. Sci.*, vol. 46, no. 7, pp. 579–594, Jul. 2021, doi: 10.1016/J.TIBS.2021.01.009.

[331] B. Tian and J. L. Manley, "Alternative polyadenylation of mRNA precursors," *Nature Reviews Molecular Cell Biology*. 2016, doi: 10.1038/nrm.2016.116.

[332] M. Hoque *et al.*, "Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing," *Nat. Methods*, 2013, doi: 10.1038/nmeth.2288.

[333] M. Hoque, W. Li, and B. Tian, "Accurate mapping of cleavage and polyadenylation sites by 3′ region extraction and deep sequencing," in *Methods in Molecular Biology*, 2014.

[334] A. Derti *et al.*, "A quantitative atlas of polyadenylation in five mammals," *Genome Res.*, 2012, doi: 10.1101/gr.132563.111.

[335] X. Liu *et al.*, "Transcription elongation rate has a tissue-specific impact on alternative cleavage and polyadenylation in Drosophila melanogaster," *RNA*, 2017, doi: 10.1261/rna.062661.117.

[336] J. V. Geisberg, Z. Moqtaderi, and K. Struhl, "The transcriptional elongation rate regulates alternative polyadenylation in yeast," *Elife*, 2020, doi: 10.7554/ELIFE.59810.

[337] E. Whitelaw and N. Proudfoot, "Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene.," *EMBO J.*, 1986, doi: 10.1002/j.1460-2075.1986.tb04587.x.

[338] B. L. Sartini, H. Wang, W. Wang, C. F. Millette, and D. L. Kilpatrick, "Pre-Messenger RNA Cleavage Factor I (CFIm): Potential Role in Alternative Polyadenylation During Spermatogenesis1," *Biol. Reprod.*, vol. 78, no. 3, pp. 472–482, 2008, doi: 10.1095/biolreprod.107.064774.

[339] K. M. Brown and G. M. Gilmartin, "A Mechanism for the Regulation of Pre-mRNA 3′ Processing by Human Cleavage Factor Im," *Mol. Cell*, 2003, doi: 10.1016/S1097-2765(03)00453-2.

[340] Q. Yang, G. M. Gilmartin, and S. Doublié, "The structure of human cleavage factor Im hints at functions beyond UGUA-specific RNA binding: A role in alternative polyadenylation and a potential link to 5′ capping and splicing," *RNA Biology*. 2011, doi: 10.4161/rna.8.5.16040.

[341] Q. Yang, G. M. Gilmartin, and S. Doublié, "Structural basis of UGUA recognition by the Nudix protein CFIm25 and implications for a regulatory role in mRNA 3′ processing," *Proc. Natl. Acad. Sci. U. S. A.*, 2010, doi: 10.1073/pnas.1000848107.

[342] Q. Yang, M. Coseno, G. M. Gilmartin, and S. Doublié, "Crystal structure of a human cleavage factor CFIm25/CFI m68/RNA complex provides an insight into poly(A) site recognition and RNA looping," *Structure*, 2011, doi: 10.1016/j.str.2010.12.021.

[343] Q. Yang, G. M. Gilmartin, and S. Doublié, "The structure of human Cleavage Factor I_m hints at functions beyond UGUA-specific RNA binding," *RNA Biol.*, 2011, doi: 10.4161/rna.8.5.16040.

[344] N. Proudfoot, "New perspectives on connecting messenger RNA 3′ end formation to transcription," *Current Opinion in Cell Biology*. 2004, doi: 10.1016/j.ceb.2004.03.007.

[345] G. Martin, A. R. Gruber, W. Keller, and M. Zavolan, "Genome-wide Analysis of Pre-mRNA 3' End Processing Reveals a Decisive Role of Human Cleavage Factor I in the Regulation of 3' UTR Length," *Cell Reports*, vol. 1, no. 6. pp. 753–763, 2012, doi: 10.1016/j.celrep.2012.05.003.

[346] S. Tan *et al.*, "CPSF6 links alternative polyadenylation to metabolism adaption in hepatocellular carcinoma progression," *J. Exp. Clin. Cancer Res.*, 2021, doi: 10.1186/s13046-021-01884-z.

[347] S. Jang *et al.*, "Differential role for phosphorylation in alternative polyadenylation function versus nuclear import of SR-like protein CPSF6," *Nucleic Acids Res.*, 2019, doi: 10.1093/nar/gkz206.

[348] J. Rappsilber, U. Ryder, A. I. Lamond, and M. Mann, "Large-scale proteomic analysis of the human spliceosome," *Genome Res.*, vol. 12, no. 8, pp. 1231–1245, 2002, doi: 10.1101/gr.473902.

[349] A. Misra, J. Ou, L. J. Zhu, and M. R. Green, "Global Promotion of Alternative Internal Exon Usage by mRNA 3' End Formation Factors," *Mol. Cell*, 2014, doi: 10.1016/j.molcel.2015.03.016.

[350] S. Millevoi *et al.*, "An interaction between U2AF 65 and CF Im links the splicing and 3′ end processing machineries," *EMBO J.*, 2006, doi: 10.1038/sj.emboj.7601331.

[351] C. Cooke, H. Hans, and J. C. Alwine, "Utilization of Splicing Elements and Polyadenylation Signal Elements in the Coupling of Polyadenylation and Last-Intron Removal," *Mol. Cell. Biol.*, 1999, doi: 10.1128/mcb.19.7.4971.

[352] C. Cooke and J. C. Alwine, "Characterization of Specific Protein-RNA Complexes Associated with the Coupling of Polyadenylation and Last-Intron Removal," *Mol. Cell. Biol.*, 2002, doi: 10.1128/mcb.22.13.4579-4586.2002.

[353] D. Nesic and L. E. Maquat, "Upstream introns influence the efficiency of final intron removal and RNA 3'-end formation," *Genes Dev.*, 1994, doi: 10.1101/gad.8.3.363.

[354] C. Cooke and J. C. Alwine, "The cap and the 3' splice site similarly affect polyadenylation efficiency," *Mol. Cell. Biol.*, 1996, doi: 10.1128/mcb.16.6.2579.

[355] N. H. Gehring and J. Y. Roignant, "Anything but Ordinary – Emerging Splicing Mechanisms in Eukaryotic Gene Regulation," *Trends in Genetics*. 2021, doi: 10.1016/j.tig.2020.10.008.

[356] M. Dutertre *et al.*, "Cotranscriptional exon skipping in the genotoxic stress response," *Nat. Struct. Mol. Biol.*, 2010, doi: 10.1038/nsmb.1912.

[357] D. S. Chandler, R. K. Singh, L. C. Caldwell, J. L. Bitler, and G. Lozano, "Genotoxic stress induces coordinately regulated alternative splicing of the p53 modulators MDM2 and MDM4," *Cancer Res.*, 2006, doi: 10.1158/0008-5472.CAN-05-4271.

[358] C. D. Nicholls, M. A. Shields, P. W. K. Lee, S. M. Robbins, and T. L. Beattie, "UV-dependent alternative splicing uncouples p53 activity and PIG3 gene function through rapid proteolytic degradation," *J. Biol. Chem.*, 2004, doi: 10.1074/jbc.M401049200.

[359] S. Dango *et al.*, "DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation," *Mol. Cell*, 2011, doi: 10.1016/j.molcel.2011.08.039.

[360] J. M. Santos-Pereira and A. Aguilera, "R loops: New modulators of genome dynamics and function," *Nature Reviews Genetics*. 2015, doi: 10.1038/nrg3961.

[361] L. Chen *et al.*, "R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters," *Mol. Cell*, vol. 68, no. 4, pp. 745-757.e5, Nov. 2017, doi: 10.1016/J.MOLCEL.2017.10.008.

[362] R. Lin *et al.*, "R-loopBase: a knowledgebase for genome-wide R-loop formation and regulation," *Nucleic Acids Res.*, 2021, doi: 10.1093/nar/gkab1103.

[363] C. Grunseich *et al.*, "Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters," *Mol. Cell*, 2018, doi: 10.1016/j.molcel.2017.12.030.

[364] B. P. Belotserkovskii, J. H. S. Shin, and P. C. Hanawalt, "Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter," *Nucleic Acids Res.*, vol. 45, no. 11, pp. 6589–6599, Jun. 2017, doi: 10.1093/NAR/GKX403.

[365] R. Boque-Sastre *et al.*, "Head-to-head antisense transcription and R-loop formation promotes transcriptional activation," *Proc. Natl. Acad. Sci. U. S. A.*, 2015, doi: 10.1073/pnas.1421197112.

[366] J. C. Morales *et al.*, "XRN2 Links Transcription Termination to DNA Damage and Replication Stress," *PLoS Genet.*, 2016, doi: 10.1371/journal.pgen.1006107.

[367] A. Cristini, M. Groh, M. S. Kristiansen, and N. Gromak, "RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage," *Cell Rep.*, vol. 23, no. 6, pp. 1891–1905, May 2018.

[368] K. Skourti-Stathaki, N. J. Proudfoot, and N. Gromak, "Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination," *Mol. Cell*, 2011, doi: 10.1016/j.molcel.2011.04.026.

[369] X. Zhang *et al.*, "Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis," *Nat. Commun.*, vol. 8, Jun. 2017, doi: 10.1038/NCOMMS15908.

[370] M. Tresini *et al.*, "The core spliceosome as target and effector of non-canonical ATM signalling," *Nature*, 2015, doi: 10.1038/nature14512.

[371] F. C. Lam *et al.*, "BRD4 prevents the accumulation of R-loops and protects against transcription–replication collision events and DNA damage," *Nat. Commun.*, 2020, doi: 10.1038/s41467-020-17503-y.

[372] C. St Germain, H. Zhao, and J. H. Barlow, "Transcription-replication collisions—a series of unfortunate events," *Biomolecules*. 2021, doi: 10.3390/biom11081249.

[373] A. L. de Septenville, S. Duigou, H. Boubakri, and B. Michel, "Replication fork reversal after replication-transcription collision," *PLoS Genet.*, 2012, doi: 10.1371/journal.pgen.1002622.

[374] Y. Liu *et al.*, "Transcription shapes DNA replication initiation to preserve genome integrity," *Genome Biol.*, 2021, doi: 10.1186/s13059-021-02390-3.

[375] A. Helmrich, M. Ballarino, and L. Tora, "Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes," *Mol. Cell*, 2011, doi: 10.1016/j.molcel.2011.10.013.

[376] J. P. Matson and L. Zou, "A genome-wide and cotranscriptional suppressor of R loops," *Genes Dev.*, 2020, doi: 10.1101/GAD.339861.120.

[377] K. S. Lang *et al.*, "Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis," *Cell*, vol. 170, no. 4, pp. 787-799.e18, Aug. 2017, doi: 10.1016/J.CELL.2017.07.044.

[378] S. Hamperl, M. J. Bocek, J. C. Saldivar, T. Swigut, and K. A. Cimprich, "Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses," *Cell*, vol. 170, no. 4, pp. 774-786.e19, Aug. 2017.

[379] B. Gómez-González *et al.*, "Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles," *EMBO J.*, 2011, doi: 10.1038/emboj.2011.206.

[380] J. A. Aoki and J. L. Manley, "The Role of Cotranscriptional Recruitment of RNA-Binding Proteins in the Maintenance of Genomic Stability," in *Post-Transcriptional Gene Regulation: RNA Processing in Eukaryotes*, 2013.

[381] P. Shah, B. Zhao, L. Qiang, and Y. Y. He, "Phosphorylation of xeroderma pigmentosum group C regulates ultraviolet-induced DNA damage repair," *Nucleic Acids Res.*, 2018, doi: 10.1093/nar/gky239.

[382] K. Paszkowska-Szczur *et al.*, "Xeroderma pigmentosum genes and melanoma risk," *Int. J. Cancer*, 2013, doi: 10.1002/ijc.28123.

[383] S. E. Tsutakawa *et al.*, "Human XPG nuclease structure, assembly, and activities with insights for neurodegeneration and cancer from pathogenic mutations," *Proc. Natl. Acad. Sci. U. S. A.*, 2020, doi: 10.1073/pnas.1921311117.

[384] R. N. Rainey, S. Y. Ng, J. Llamas, G. T. J. van der Horst, and N. Segil, "Mutations in cockayne syndrome-associated genes (Csa and Csb) predispose to cisplatin-induced hearing loss in mice," *J. Neurosci.*, 2016, doi: 10.1523/JNEUROSCI.3890-15.2016.

[385] S. Moriwaki, "Hereditary Disorders with Defective Repair of UV-Induced DNA Damage," *Japanese Clin. Med.*, 2013, doi: 10.4137/jcm.s10730.

[386] T. Nardo *et al.*, "A UV-sensitive syndrome patient with a specific CSA mutation reveals separable roles for CSA in response to UV and oxidative DNA damage," *Proc. Natl. Acad. Sci. U. S. A.*, 2009, doi: 10.1073/pnas.0902113106.

[387] Y. Nakazawa *et al.*, "Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair," *Nat. Genet.*, 2012, doi: 10.1038/ng.2229.

[388] N. Sonenberg and A. G. Hinnebusch, "Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets," *Cell*. 2009, doi: 10.1016/j.cell.2009.01.042.

[389] K. E. Deigan and A. R. Ferré-D'Amaré, "Riboswitches: Discovery of drugs that target bacterial gene-regulatory RNAs," *Acc. Chem. Res.*, 2011, doi: 10.1021/ar200039b.

[390] W. C. Lin, F. T. Lin, and J. R. Nevins, "Selective induction of E2F1 in response to DNA damage,

mediated by ATM-dependent phosphorylation," *Genes Dev.*, 2001.

[391] R. M. Augusto Da Costa, L. Riou, A. Paquola, C. F. Martins Menck, and A. Sarasin, "Transcriptional profiles of unirradiated or UV-irradiated human cells expressing either the cancer-prone XPB/CS allele or the noncancer-prone XPB/TTD allele," *Oncogene*, 2005, doi: 10.1038/sj.onc.1208288.

[392] K. E. Rieger and G. Chu, "Portrait of transcriptional responses to ultraviolet and ionizing radiation in human cells," *Nucleic Acids Res.*, 2004, doi: 10.1093/nar/gkh783.

[393] K. A. Spriggs, M. Stoneley, M. Bushell, and A. E. Willis, "Re-programming of translation following cell stress allows IRES-mediated translation to predominate," *Biol. Cell*, 2008, doi: 10.1042/bc20070098.

[394] J. D. Thomas and G. J. Johannes, "Identification of mRNAs that continue to associate with polysomes during hypoxia," *RNA*, 2007, doi: 10.1261/rna.534807.

[395] M. Bushell *et al.*, "Polypyrimidine Tract Binding Protein Regulates IRES-Mediated Gene Expression during Apoptosis," *Mol. Cell*, 2006, doi: 10.1016/j.molcel.2006.06.012.

[396] R. Grover, P. S. Ray, and S. Das, "Polypyrimidine tract binding protein regulates IRES-mediated translation of p53 isoforms," *Cell Cycle*, 2008, doi: 10.4161/cc.7.14.6271.

[397] J. D. Blais *et al.*, "Activating Transcription Factor 4 Is Translationally Regulated by Hypoxic Stress," *Mol. Cell. Biol.*, 2004, doi: 10.1128/mcb.24.17.7469-7482.2004.

[398] A. O. Ibegbu, L. Fyfe, D. McBean, and I. Mullaney, "The effects of hypoxia and opioid receptor agonists treatment in cortical B50 neuronal cells in culture," *J. Biol. Environ. Sci.*, 2012.

[399] I. R. Powley *et al.*, "Translational reprogramming following UVB irradiation is mediated by DNA-PKcs and allows selective recruitment to the polysomes of mRNAs encoding DNA repair enzymes," *Genes Dev.*, 2009, doi: 10.1101/gad.516509.

[400] P. Llabata *et al.*, "Involvement of the eIF2α Kinase GCN2 in UV-B Responses," *Front. Plant Sci.*, 2019, doi: 10.3389/fpls.2019.01492.

[401] A. V. Kondrashov, K. A. Spriggs, M. Bushell, and A. E. Willis, "Co-ordinated regulation of translation following DNA damage," *Cell Cycle*. 2009, doi: 10.4161/cc.8.19.9467.

[402] E. Thomson, S. Ferreira-Cerca, and E. Hurt, "Eukaryotic ribosome biogenesis at a glance," *J. Cell Sci.*, 2013, doi: 10.1242/jcs.111948.

[403] I. C. Kos-Braun and M. Koš, "Post-transcriptional regulation of ribosome biogenesis in yeast," *Microbial Cell*. 2017, doi: 10.15698/mic2017.05.575.

[404] L. Qian and H. Zhang, "Research progress of ribosome biogenesis and cancer," *Cancer Res. Prev. Treat.*, 2020, doi: 10.3971/j.issn.1000-8578.2020.19.1138.

[405] K. F. Chau *et al.*, "Downregulation of ribosome biogenesis during early forebrain development," *Elife*, 2018, doi: 10.7554/eLife.36998.

[406] B. W. Tye *et al.*, "Proteotoxicity from aberrant ribosome biogenesis compromises cell fitness," *Elife*, 2019, doi: 10.7554/eLife.43002.

[407] K. J. Abraham *et al.*, "Nucleolar RNA polymerase II drives ribosome biogenesis," *Nature*, no. December 2018, 2020, doi: 10.1038/s41586-020-2497-0.

[408] M. Penzo, L. Montanaro, D. Treré, and M. Derenzini, "The Ribosome Biogenesis—Cancer Connection," *Cells*, 2019, doi: 10.3390/cells8010055.

[409] V. Prakash *et al.*, "Ribosome biogenesis during cell cycle arrest fuels EMT in development and disease," *Nat. Commun.*, 2019, doi: 10.1038/s41467-019-10100-8.

[410] V. A. Halim *et al.*, "Doxorubicin-induced DNA damage causes extensive ubiquitination of ribosomal proteins associated with a decrease in protein translation," *Mol. Cell. Proteomics*,

2018, doi: 10.1074/mcp.RA118.000652.

[411] F. J. LaRiviere, S. E. Cole, D. J. Ferullo, and M. J. Moore, "A late-acting quality control process for mature eukaryotic rRNAs," *Mol. Cell*, 2006, doi: 10.1016/j.molcel.2006.10.008.

[412] S. Juszkiewicz, V. Chandrasekaran, Z. Lin, S. Kraatz, V. Ramakrishnan, and R. S. Hegde, "ZNF598 Is a Quality Control Sensor of Collided Ribosomes," *Mol. Cell*, 2018, doi: 10.1016/j.molcel.2018.08.037.

[413] A. Garzia *et al.*, "The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs," *Nat. Commun.*, 2017, doi: 10.1038/ncomms16056.

[414] L. M. Tam, J. Jiang, P. Wang, and Y. Wang, "Arsenite Binds to ZNF598 to Perturb Ribosome-Associated Protein Quality Control," *Chem. Res. Toxicol.*, 2020, doi: 10.1021/acs.chemrestox.9b00412.

[415] K. L. Hickey *et al.*, "GIGYF2 and 4EHP inhibit translation initiation of defective messenger rnas to assist ribosome-associated quality control," *bioRxiv*, 2019, doi: 10.1101/792994.

[416] M. Morita *et al.*, "A Novel 4EHP-GIGYF2 Translational Repressor Complex Is Essential for Mammalian Development," *Mol. Cell. Biol.*, 2012, doi: 10.1128/mcb.00455-12.

[417] E. Sundaramoorthy, M. Leonard, R. Mak, J. Liao, A. Fulzele, and E. J. Bennett, "ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation," *Mol. Cell*, 2017, doi: 10.1016/j.molcel.2016.12.026.

[418] S. Juszkiewicz and R. S. Hegde, "Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination," *Mol. Cell*, 2017, doi: 10.1016/j.molcel.2016.11.039.

[419] S. Juszkiewicz, S. H. Speldewinde, L. Wan, J. Q. Svejstrup, and R. S. Hegde, "The ASC-1 Complex Disassembles Collided Ribosomes," *Mol. Cell*, 2020, doi: 10.1016/j.molcel.2020.06.006.

[420] E. A. Ponomarenko *et al.*, "The Size of the Human Proteome: The Width and Depth," *International Journal of Analytical Chemistry*. 2016, doi: 10.1155/2016/7436849.

[421] S. H. White and R. E. Jacobs, "Statistical distribution of hydrophobic residues along the length of protein chains. Implications for protein folding and evolution," *Biophys. J.*, 1990, doi: 10.1016/S0006-3495(90)82611-4.

[422] C. B. Anfinsen, "Principles that govern the folding of protein chains," *Science*. 1973, doi: 10.1126/science.181.4096.223.

[423] F. U. Hartl and M. Hayer-Hartl, "Protein folding. Molecular chaperones in the cytosol: From nascent chain to folded protein," *Science*. 2002, doi: 10.1126/science.1068408.

[424] C. Linghu *et al.*, "Recording of cellular physiological histories along optically readable self-assembling protein chains," *bioRxiv*, 2021.

[425] O. Krishnadev, K. V. Brinda, and S. Vishveshwara, "A graph spectral analysis of the structural similarity network of protein chains," *Proteins Struct. Funct. Genet.*, 2005, doi: 10.1002/prot.20532.

[426] S. Piana, P. Robustelli, D. Tan, S. Chen, and D. E. Shaw, "Development of a Force Field for the Simulation of Single-Chain Proteins and Protein-Protein Complexes," *J. Chem. Theory Comput.*, 2020, doi: 10.1021/acs.jctc.9b00251.

[427] M. Patel and H. Shah, "Protein secondary structure prediction using support vector machines (SVMs)," 2014, doi: 10.1109/ICMIRA.2013.124.

[428] K. Haglund and I. Dikic, "Ubiquitylation and cell signaling," *EMBO Journal*. 2005, doi: 10.1038/sj.emboj.7600808.

[429] A. H. Phillips and J. E. Corn, “Using protein motion to read, write, and erase ubiquitin signals,” *Journal of Biological Chemistry*. 2015, doi: 10.1074/jbc.R115.653675.

[430] L. Spasser and A. Brik, “Chemistry and biology of the ubiquitin signal,” *Angewandte Chemie - International Edition*. 2012, doi: 10.1002/anie.201200020.

[431] M. Akutsu, I. Dikic, and A. Bremm, “Ubiquitin chain diversity at a glance,” *J. Cell Sci.*, 2016, doi: 10.1242/jcs.183954.

[432] L. M. Fennell, S. Rahighi, and F. Ikeda, “Linear ubiquitin chain-binding domains,” *FEBS Journal*. 2018, doi: 10.1111/febs.14478.

[433] M. E. French, C. F. Koehler, and T. Hunter, “Emerging functions of branched ubiquitin chains,” *Cell Discovery*. 2021, doi: 10.1038/s41421-020-00237-y.

[434] K. Rittinger and F. Ikeda, “Linear ubiquitin chains: Enzymes, mechanisms and biology,” *Open Biology*. 2017, doi: 10.1098/rsob.170026.

[435] Y. N. WANG and B. ZHAO, “Mechanisms of ubiquitin chain formation,” *Progress in Biochemistry and Biophysics*. 2021, doi: 10.16476/j.pibb.2020.0304.

[436] P. Xu *et al.*, “Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation,” *Cell*, 2009, doi: 10.1016/j.cell.2009.01.041.

[437] M. D. Petroski and R. J. Deshaies, “Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34,” *Cell*, 2005, doi: 10.1016/j.cell.2005.09.033.

[438] H. Kim, P. Vick, J. Hedtke, D. Ploper, and E. M. De Robertis, “Wnt Signaling Translocates Lys48-Linked Polyubiquitinated Proteins to the Lysosomal Pathway,” *Cell Rep.*, 2015, doi: 10.1016/j.celrep.2015.04.048.

[439] 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.*, 2018, doi: 10.1073/pnas.1716673115.

[440] S. J. L. van Wijk *et al.*, “Fluorescence-Based Sensors to Monitor Localization and Functions of Linear and K63-Linked Ubiquitin Chains in Cells,” *Mol. Cell*, 2012, doi: 10.1016/j.molcel.2012.06.017.

[441] J. A. Nathan, H. Tae Kim, L. Ting, S. P. Gygi, and A. L. Goldberg, “Why do cellular proteins linked to K63-polyubiquitin chains not associate with proteasomes?,” *EMBO J.*, 2013, doi: 10.1038/emboj.2012.354.

[442] F. Ohtake, Y. Saeki, S. Ishido, J. Kanno, and K. Tanaka, “The K48-K63 Branched Ubiquitin Chain Regulates NF-κB Signaling,” *Mol. Cell*, 2016, doi: 10.1016/j.molcel.2016.09.014.

[443] N. A. Snyder and G. M. Silva, “Deubiquitinating enzymes (DUBs): Regulation, homeostasis, and oxidative stress response,” *Journal of Biological Chemistry*. 2021, doi: 10.1016/j.jbc.2021.101077.

[444] V. de Cesare, D. C. Lopez, P. D. Mabbitt, O. Antico, N. T. Wood, and S. Virdee, “Deubiquitinating enzyme amino acid profiling reveals a class of ubiquitin esterases,” *bioRxiv*, 2020, doi: 10.1101/2020.05.11.087965.

[445] T. Hermanns and K. Hofmann, “Bacterial dubs: Deubiquitination beyond the seven classes,” *Biochemical Society Transactions*. 2019, doi: 10.1042/BST20190526.

[446] T. Hermanns, I. Woiwode, R. F. M. Guerreiro, R. Vogt, M. Lammers, and K. Hofmann, “An evolutionary approach to systematic discovery of novel deubiquitinases, applied to Legionella,” *Life Sci. Alliance*, 2020, doi: 10.26508/LSA.202000838.

[447] P. Paudel *et al.*, “Crystal structure and activity-based labeling reveal the mechanisms for linkage-specific substrate recognition by deubiquitinase USP9X,” *Proc. Natl. Acad. Sci. U. S. A.*, 2019,

doi: 10.1073/pnas.1815027116.

[448] F. E. Reyes-Turcu, K. H. Ventii, and K. D. Wilkinson, "Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes," *Annual Review of Biochemistry*. 2009, doi: 10.1146/annurev.biochem.78.082307.091526.

[449] A. S. Schober and E. Berra, "DUBs, new members in the hypoxia signaling clUb," *Frontiers in Oncology*. 2016, doi: 10.3389/fonc.2016.00053.

[450] R. K. Meray and P. T. Lansbury, "Reversible monoubiquitination regulates the Parkinson disease-associated ubiquitin hydrolase UCH-L1," *J. Biol. Chem.*, 2007, doi: 10.1074/jbc.M611153200.

[451] D. Komander, M. J. Clague, and S. Urbé, "Breaking the chains: Structure and function of the deubiquitinases," *Nature Reviews Molecular Cell Biology*. 2009, doi: 10.1038/nrm2731.

[452] P. Roos-Mattjus and L. Sistonen, "The ubiquitin-proteasome pathway," *Annals of Medicine*. 2004, doi: 10.1080/07853890310016324.

[453] S. Haq and S. Ramakrishna, "Deubiquitylation of deubiquitylases," *Open Biology*. 2017, doi: 10.1098/rsob.170016.

[454] Y. van der Weegen *et al.*, "The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II," *Nat. Commun.*, 2020, doi: 10.1038/s41467-020-15903-8.

[455] M. Fousteri, W. Vermeulen, A. A. van Zeeland, and L. H. F. Mullenders, "Cockayne Syndrome A and B Proteins Differentially Regulate Recruitment of Chromatin Remodeling and Repair Factors to Stalled RNA Polymerase II In Vivo," *Molecular Cell*. 2006, doi: 10.1016/j.molcel.2006.06.029.

[456] E. Dubois, S. Gerber, A. Kisselev, A. Harel-Bellan, and R. Groisman, "UV-dependent phosphorylation of COP9/signalosome in UV-induced apoptosis," *Oncol. Rep.*, 2016, doi: 10.3892/or.2016.4671.

[457] R. Groisman *et al.*, "The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage," *Cell*, 2003, doi: 10.1016/S0092-8674(03)00316-7.

[458] E. S. Fischer *et al.*, "The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation," *Cell*, 2011, doi: 10.1016/j.cell.2011.10.035.

[459] Y. Nakazawa *et al.*, "Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair," *Cell*, 2020, doi: 10.1016/j.cell.2020.02.010.

[460] G. Kokic, F. R. Wagner, A. Chernev, H. Urlaub, and P. Cramer, "Structural basis of human transcription–DNA repair coupling," *Nature*, vol. 598, no. 7880, pp. 368–372, Oct. 2021, doi: 10.1038/s41586-021-03906-4.

[461] A. Tufegd Zi C Vidakovi *et al.*, "Regulation of the RNAPII Pool Is Integral to the DNA Damage Response Article Regulation of the RNAPII Pool Is Integral to the DNA Damage Response," *Cell*, 2020, doi: 10.1016/j.cell.2020.02.009.

[462] J. He, Q. Zhu, G. Wani, and A. A. Wani, "UV-induced proteolysis of RNA polymerase II is mediated by VCP/p97 segregase and timely orchestration by Cockayne syndrome B protein," *Oncotarget*, 2017, doi: 10.18632/oncotarget.14205.

[463] M. E. Geijer *et al.*, "Elongation factor ELOF1 drives transcription-coupled repair and prevents genome instability," *Nat. Cell Biol.*, 2021, doi: 10.1038/s41556-021-00692-z.

[464] Y. van der Weegen *et al.*, "ELOF1 is a transcription-coupled DNA repair factor that directs RNA polymerase II ubiquitylation," *Nat. Cell Biol.*, 2021, doi: 10.1038/s41556-021-00688-9.

[465] Q. Zhu *et al.*, "USP7-mediated deubiquitination differentially regulates CSB but not UVSSA

upon UV radiation-induced DNA damage," *Cell Cycle*, 2020, doi: 10.1080/15384101.2019.1695996.

[466] E. G. KREBS, D. J. GRAVES, and E. H. FISCHER, "Factors affecting the activity of muscle phosphorylase b kinase.," *J. Biol. Chem.*, 1959, doi: 10.1016/s0021-9258(18)69685-1.

[467] C. Smal *et al.*, "Identification of in Vivo Phosphorylation Sites on Human Deoxycytidine Kinase," *J. Biol. Chem.*, 2006, doi: 10.1074/jbc.m512129200.

[468] W. Xu and Y. Wang, "Post-translational Modifications of Serine/Threonine and Histidine Kinases and Their Roles in Signal Transductions in Synechocystis Sp. PCC 6803," *Applied Biochemistry and Biotechnology*. 2021, doi: 10.1007/s12010-020-03435-2.

[469] S. Manuse, A. Fleurie, L. Zucchini, C. Lesterlin, and C. Grangeasse, "Role of eukaryotic-like serine/threonine kinases in bacterial cell division and morphogenesis," *FEMS Microbiology Reviews*. 2015, doi: 10.1093/femsre/fuv041.

[470] D. W. Hommes, M. P. Peppelenbosch, and S. J. H. Van Deventer, "Mitogen activated protein (MAP) kinase signal transduction pathways and novel anti-inflammatory targets," *Gut*. 2003, doi: 10.1136/gut.52.1.144.

[471] R. Heinrich, B. G. Neel, and T. A. Rapoport, "Mathematical models of protein kinase signal transduction," *Mol. Cell*, 2002, doi: 10.1016/S1097-2765(02)00528-2.

[472] R. J. Davis, "The mitogen-activated protein kinase signal transduction pathway," *Journal of Biological Chemistry*. 1993, doi: 10.1016/s0021-9258(18)82362-6.

[473] S. Ghorbani, P. D. Szigetvari, J. Haavik, and R. Kleppe, "Serine 19 phosphorylation and 14-3-3 binding regulate phosphorylation and dephosphorylation of tyrosine hydroxylase on serine 31 and serine 40," *J. Neurochem.*, 2020, doi: 10.1111/jnc.14872.

[474] M. Boyce *et al.*, "A selective inhibitor of elF2α dephosphorylation protects cells from ER stress," *Science (80-.).*, 2005, doi: 10.1126/science.1101902.

[475] O. M. Power, M. A. Fenelon, J. A. O'Mahony, and N. A. McCarthy, "Dephosphorylation of caseins in milk protein concentrate alters their interactions with sodium hexametaphosphate," *Food Chem.*, 2019, doi: 10.1016/j.foodchem.2018.07.086.

[476] A. N. Blackford and S. P. Jackson, "ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response," *Molecular Cell*. 2017, doi: 10.1016/j.molcel.2017.05.015.

[477] D. Durocher and S. P. Jackson, "DNA-PK, ATM and ATR as sensors of DNA damage: Variations on a theme?," *Current Opinion in Cell Biology*. 2001, doi: 10.1016/S0955-0674(00)00201-5.

[478] B. P. C. Chen *et al.*, "Ataxia telangiectasia mutated (ATM) is essential for DNA-PKcs phosphorylations at the Thr-2609 cluster upon DNA double strand break," *J. Biol. Chem.*, 2007, doi: 10.1074/jbc.M611605200.

[479] H. Yamaguchi, S. R. Durell, D. K. Chatterjee, C. W. Anderson, and E. Appella, "The Wip1 phosphatase PPM1D dephosphorylates SQ/TQ motifs in checkpoint substrates phosphorylated by PI3K-like kinases," *Biochemistry*, 2007, doi: 10.1021/bi701096s.

[480] D. Chowdhury, M. C. Keogh, H. Ishii, C. L. Peterson, S. Buratowski, and J. Lieberman, "γ-H2AX dephosphorylation by protein phosphatase 2A facilitates DNA double-strand break repair," *Mol. Cell*, 2005, doi: 10.1016/j.molcel.2005.10.003.

[481] J. Risso-Ballester and R. Sanjuán, "High fidelity deep sequencing reveals no effect of ATM, ATR, and DNA-PK cellular DNA damage response pathways on adenovirus mutation rate," *Viruses*, 2019, doi: 10.3390/v11100938.

[482] A. E. Cambindo Botto *et al.*, "Reciprocal regulation between alternative splicing and the DNA damage response," *Genetics and Molecular Biology*. 2020, doi: 10.1590/1678-4685-GMB-2019-0111.

[483] L. Ding *et al.*, "The roles of cyclin-dependent kinases in cell-cycle progression and therapeutic strategies in human breast cancer," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21061960.

[484] Y. Keshet and R. Seger, "The MAP kinase signaling cascades: a system of hundreds of components regulates a diverse array of physiological functions.," *Methods in molecular biology (Clifton, N.J.)*. 2010, doi: 10.1007/978-1-60761-795-2_1.

[485] S. H. Yang, A. D. Sharrocks, and A. J. Whitmarsh, "MAP kinase signalling cascades and transcriptional regulation," *Gene*. 2013, doi: 10.1016/j.gene.2012.10.033.

[486] L. Chang and M. Karin, "Mammalian MAP kinase signalling cascades," *Nature*. 2001, doi: 10.1038/35065000.

[487] J. B. Hoek, "MAP KINASE SIGNALING PROTOCOLS," *Shock*, 2004, doi: 10.1097/01.shk.0000137186.99856.40.

[488] S. T. Eblen, "Extracellular-Regulated Kinases: Signaling From Ras to ERK Substrates to Control Biological Outcomes," in *Advances in Cancer Research*, 2018.

[489] J. D. Thatcher, "The Ras-MAPK signal transduction pathway," *Science Signaling*. 2010, doi: 10.1126/scisignal.3119tr1.

[490] Y. Ou, Z. Zheng, B. Niu, J. Su, and H. Su, "Different MAPK signal transduction pathways play different roles in the impairment of glucose-stimulated insulin secretion in response to IL-1β," *Mol. Med. Rep.*, 2020, doi: 10.3892/mmr.2020.11366.

[491] J. M. Kyriakis and J. Avruch, "Mammalian MAPK signal transduction pathways activated by stress and inflammation: A 10-year update," *Physiological Reviews*. 2012, doi: 10.1152/physrev.00028.2011.

[492] C. Y. Xiong, J. Zhang, and D. W. Guan, "P38MAPK signal transduction pathway and tissue lesion," *Chinese J. Forensic Med.*, 2008.

[493] A. De Luca, M. R. Maiello, A. D'Alessio, M. Pergameno, and N. Normanno, "The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: Role in cancer pathogenesis and implications for therapeutic approaches," *Expert Opinion on Therapeutic Targets*. 2012, doi: 10.1517/14728222.2011.639361.

[494] O. M. Seternes, A. M. Kidger, and S. M. Keyse, "Dual-specificity MAP kinase phosphatases in health and disease," *Biochim. Biophys. Acta - Mol. Cell Res.*, 2019, doi: 10.1016/j.bbamcr.2018.09.002.

[495] H. B. Low and Y. Zhang, "Regulatory roles of MAPK phosphatases in cancer," *Immune Network*. 2016, doi: 10.4110/in.2016.16.2.85.

[496] D. V. Bulavin *et al.*, "Inactivation of the Wip1 phosphatase inhibits mammary tumorigenesis through p38 MAPK-mediated activation of the p16Ink4a-p19 Arf pathway," *Nat. Genet.*, 2004, doi: 10.1038/ng1317.

[497] M. Takekawa *et al.*, "p53-inducible Wip1 phosphatase mediates a negative feedback regulation of p38 MAPK-p53 signaling in response to UV radiation," *EMBO J.*, 2000, doi: 10.1093/emboj/19.23.6517.

[498] D. D. Hirsch and P. J. S. Stork, "Mitogen-activated protein kinase phosphatases inactivate stress-activated protein kinase pathways in vivo," *J. Biol. Chem.*, 1997, doi: 10.1074/jbc.272.7.4568.

[499] K. Ono and J. Han, "The p38 signal transduction pathway Activation and function," *Cellular Signalling*. 2000, doi: 10.1016/S0898-6568(99)00071-6.

[500] E. Herlaar and Z. Brown, "p38 MAPK signalling cascades in inflammatory disease," *Molecular Medicine Today*. 1999, doi: 10.1016/S1357-4310(99)01544-0.

[501] J. Raingeaud *et al.*, "Pro-inflammatory cytokines and environmental stress cause p38 mitogen-

activated protein kinase activation by dual phosphorylation on tyrosine and threonine," *J. Biol. Chem.*, 1995, doi: 10.1074/jbc.270.13.7420.

[502] A. Cuenda and S. Rousseau, "p38 MAP-Kinases pathway regulation, function and role in human diseases," *Biochimica et Biophysica Acta - Molecular Cell Research*. 2007, doi: 10.1016/j.bbamcr.2007.03.010.

[503] A. Cuadrado and A. R. Nebreda, "Mechanisms and functions of p38 MAPK signalling," *Biochemical Journal*. 2010, doi: 10.1042/BJ20100323.

[504] C. Falcicchia, F. Tozzi, O. Arancio, D. M. Watterson, and N. Origlia, "Involvement of p38 mapk in synaptic function and dysfunction," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21165624.

[505] A. Pranteda, V. Piastra, L. Stramucci, D. Fratantonio, and G. Bossi, "The p38 mapk signaling activation in colorectal cancer upon therapeutic treatments," *Int. J. Mol. Sci.*, 2020, doi: 10.3390/ijms21082773.

[506] D. V. Bulavin, O. Kovalsky, M. C. Hollander, and A. J. Fornace, " Loss of Oncogenic H-ras-Induced Cell Cycle Arrest and p38 Mitogen-Activated Protein Kinase Activation by Disruption of Gadd45a ," *Mol. Cell. Biol.*, 2003, doi: 10.1128/mcb.23.11.3859-3871.2003.

[507] M. S. Sheikh, M. C. Hollander, and A. J. Fornace, "Role of Gadd45 in apoptosis," 2000, doi: 10.1016/S0006-2952(99)00291-9.

[508] A. Cuadrado *et al.*, "A new p38 MAP kinase-regulated transcriptional coactivator that stimulates p53-dependent apoptosis," *EMBO J.*, 2007, doi: 10.1038/sj.emboj.7601657.

[509] A. R. Goloudina and V. A. Pospelov, "The role of stress-kinase p38 in cell cycle blocking delay under osmotic stress," *Tsitologiia*, 2004.

[510] H. Li *et al.*, "P38 mAPK-MK2 pathway regulates the heat-stress-induced accumulation of reactive oxygen species that mediates apoptotic cell death in glial cells," *Oncol. Lett.*, 2018, doi: 10.3892/ol.2017.7360.

[511] S. Soni, P. Anand, and Y. S. Padwad, "MAPKAPK2: The master regulator of RNA-binding proteins modulates transcript stability and tumor progression," *Journal of Experimental and Clinical Cancer Research*. 2019, doi: 10.1186/s13046-019-1115-1.

[512] P. Gupta, B. Saha, S. Chattopadhyay, and B. S. Patro, "Pharmacological targeting of differential DNA repair, radio-sensitizes WRN-deficient cancer cells in vitro and in vivo," *Biochem. Pharmacol.*, 2021, doi: 10.1016/j.bcp.2021.114450.

[513] B. Cánovas *et al.*, "Targeting p38α Increases DNA Damage, Chromosome Instability, and the Anti-tumoral Response to Taxanes in Breast Cancer Cells," *Cancer Cell*, 2018, doi: 10.1016/j.ccell.2018.04.010.

[514] M. M. L. De Sousa, K. Ø. Bjørås, A. Hanssen-Bauer, K. Solvang-Garten, and M. Otterlei, "P38 MAPK signaling and phosphorylations in the BRCT1 domain regulate XRCC1 recruitment to sites of DNA damage," *Sci. Rep.*, 2017, doi: 10.1038/s41598-017-06770-3.

[515] A. Martínez-Limón, M. Joaquin, M. Caballero, F. Posas, and E. de Nadal, "The p38 pathway: From biology to cancer therapy," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21061913.

[516] M. Loesch and G. Chen, "The p38 MAPK stress pathway as a tumor suppressor or more?," *Frontiers in Bioscience*. 2008, doi: 10.2741/2951.

[517] S. Gupta *et al.*, "Selective interaction of JNK protein kinase isoforms with transcription factors," *EMBO J.*, 1996, doi: 10.1002/j.1460-2075.1996.tb00636.x.

[518] Y. Fleming, C. G. Armstrong, N. Morrice, A. Paterson, M. Goedert, and P. Cohen, "Synergistic activation of stress-activated protein kinase 1/c-Jun N-terminal kinase (SAPK1/JNK) isoforms by mitogen-activated protein kinase kinase 4 (MKK4) and MKK7," *Biochem. J.*, 2000, doi:

10.1042/0264-6021:3520145.

[519] J. Yang and S. Yao, "JNK-Bcl-2/Bcl-xL-Bax/Bak pathway mediates the crosstalk between matrine-induced autophagy and apoptosis via interplay with Beclin 1," *Int. J. Mol. Sci.*, 2015, doi: 10.3390/ijms161025744.

[520] K. Sabapathy, K. Hochedlinger, S. Y. Nam, A. Bauer, M. Karin, and E. F. Wagner, "Distinct roles for JNK1 and JNK2 in regulating JNK activity and c-Jun-dependent cell proliferation," *Mol. Cell*, 2004, doi: 10.1016/j.molcel.2004.08.028.

[521] K. Sabapathy *et al.*, "Distinct Roles for JNK1 and JNK2 in Regulating JNK Activity and c-Jun-Dependent Cell Proliferation University of California at San Diego," *Mol. Cell*, 2004.

[522] S. Y. Fuchs, L. Dolan, R. J. Davis, and Z. Ronai, "Phosphorylation-dependent targeting of c-Jun ubiquitination by Jun N-kinase," *Oncogene*, 1996.

[523] F. Xiao, B. Liu, and Q. X. Zhu, "c-Jun N-terminal kinase is required for thermotherapyinduced apoptosis in human gastric cancer cells," *World J. Gastroenterol.*, 2012, doi: 10.3748/wjg.v18.i48.7348.

[524] Z. Lu, H. Chen, X. M. Zheng, and M. L. Chen, "Experimental study on the apoptosis of cervical cancer Hela cells induced by juglone through c-Jun N-terminal kinase/c-Jun pathway," *Asian Pac. J. Trop. Med.*, 2017, doi: 10.1016/j.apjtm.2017.06.005.

[525] S. Wen *et al.*, "Induction of mitochondrial apoptosis pathway mediated through caspase-8 and c-Jun N-terminal kinase by cadmium-activated Fas in rat cortical neurons," *Metallomics*, 2021, doi: 10.1093/mtomcs/mfab042.

[526] C. Tournier *et al.*, "Requirement of JNK for stress-induced activation of the cytochrome c-mediated death pathway," *Science (80-.).*, 2000, doi: 10.1126/science.288.5467.870.

[527] X. Mao, C. Rong Yu, W. Hua Li, and W. Xin Li, "Induction of apoptosis by shikonin through a ROS/JNK-mediated process in Bcr/Abl-positive chronic myelogenous leukemia (CML) cells," *Cell Res.*, 2008, doi: 10.1038/cr.2008.86.

[528] A. Malki and E. S. El Ashry, "Quinuclidinone derivative 6 induced apoptosis in human breast cancer cells via sphingomyelinase and JNK signaling," *J. Chemother.*, 2012, doi: 10.1179/1973947812Y.0000000035.

[529] R. Eferl *et al.*, "Liver tumor development: c-Jun antagonizes the proapoptotic activity of p53," *Cell*, 2003, doi: 10.1016/S0092-8674(03)00042-4.

[530] T. Sakurai, S. Maeda, L. Chang, and M. Karin, "Loss of hepatic NF-κB activity enhances chemical hepatocarcinogenesis through sustained c-Jun N-terminal kinase 1 activation," *Proc. Natl. Acad. Sci. U. S. A.*, 2006, doi: 10.1073/pnas.0603499103.

[531] L. Hui, K. Zatloukal, H. Scheuch, E. Stepniak, and E. F. Wagner, "Proliferation of human HCC cells and chemically induced mouse liver cancers requires JNK1-dependent p21 downregulation," *J. Clin. Invest.*, 2008, doi: 10.1172/JCI37156.

[532] F. Chen and V. Castranova, "Beyond apoptosis of JNK1 in liver cancer," *Cell Cycle*. 2009, doi: 10.4161/cc.8.8.8200.

[533] Q. B. She, N. Chen, A. M. Bode, R. A. Flavell, and Z. Dong, "Deficiency of c-Jun-NH2-terminal kinase-1 in mice enhances skin tumor development by 12-O-tetradecanoylphorbol-13-acetate," *Cancer Res.*, 2002.

[534] N. Chen *et al.*, "Suppression of skin tumorigenesis in c-Jun NH2-terminal kinase-2-deficient mice," *Cancer Res.*, 2001.

[535] M. H. Cheng and S. J. Kim, "Inhibitory effect of probenecid on osteoclast formation via JNK, ROS and COX-2," *Biomol. Ther.*, 2020, doi: 10.4062/biomolther.2019.047.

[536] S. Wang *et al.*, "Andrographolide induces apoptosis in human osteosarcoma cells via the

ROS/JNK pathway," *Int. J. Oncol.*, 2020, doi: 10.3892/ijo.2020.5032.

[537] Z. L. Sun, J. L. Dong, and J. Wu, "Juglanin induces apoptosis and autophagy in human breast cancer progression via ROS/JNK promotion," *Biomed. Pharmacother.*, 2017, doi: 10.1016/j.biopha.2016.11.030.

[538] P. C. Calses *et al.*, "DGCR8 Mediates Repair of UV-Induced DNA Damage Independently of RNA Processing," *Cell Rep.*, 2017, doi: 10.1016/j.celrep.2017.03.021.

[539] J. C. Ferreon *et al.*, "Acetylation disfavors tau phase separation," *Int. J. Mol. Sci.*, 2018, doi: 10.3390/ijms19051360.

[540] C. Xia, Y. Tao, M. Li, T. Che, and J. Qu, "Protein acetylation and deacetylation: An important regulatory modification in gene transcription (Review)," *Exp. Ther. Med.*, 2020, doi: 10.3892/etm.2020.9073.

[541] W. Fan and J. Luo, "SIRT1 regulates UV-induced DNA repair through deacetylating XPA," *Mol. Cell*, 2010, doi: 10.1016/j.molcel.2010.07.006.

[542] R. Guo, J. Chen, D. L. Mitchell, and D. G. Johnson, "GCN5 and E2F1 stimulate nucleotide excision repair by promoting H3K9 acetylation at sites of damage," *Nucleic Acids Res.*, 2011, doi: 10.1093/nar/gkq983.

[543] R. Guo *et al.*, "E2F1 localizes to sites of UV-induced DNA damage to enhance nucleotide excision repair," *J. Biol. Chem.*, 2010, doi: 10.1074/jbc.M110.121939.

[544] A. Di Lorenzo and M. T. Bedford, "Histone arginine methylation," *FEBS Letters*. 2011, doi: 10.1016/j.febslet.2010.11.010.

[545] A. E. Raposo and S. C. Piller, "Protein arginine methylation: An emerging regulator of the cell cycle," *Cell Division*. 2018, doi: 10.1186/s13008-018-0036-2.

[546] J. P. Bryant, J. Heiss, and Y. K. Banasavadi-Siddegowda, "Arginine methylation in brain tumors: Tumor biology and therapeutic strategies," *Cells*. 2021, doi: 10.3390/cells10010124.

[547] Z. Guo *et al.*, "Methylation of FEN1 suppresses nearby phosphorylation and facilitates PCNA binding," *Nat. Chem. Biol.*, 2010, doi: 10.1038/nchembio.422.

[548] B. N. Borsos, H. Majoros, and T. Pankotai, "Emerging Roles of Post-Translational Modifications in Nucleotide Excision Repair," *Cells*. 2020, doi: 10.3390/cells9061466.

[549] F. Liebelt *et al.*, "Transcription-coupled nucleotide excision repair is coordinated by ubiquitin and SUMO in response to ultraviolet irradiation," *Nucleic Acids Res.*, 2020, doi: 10.1093/nar/gkz977.

[550] Q. E. Wang, Q. Zhu, G. Wani, M. A. El-Mahdy, J. Li, and A. A. Wani, "DNA repair factor XPC is modified by SUMO-1 and ubiquitin following UV irradiation," *Nucleic Acids Res.*, 2005, doi: 10.1093/nar/gki684.

[551] A. Min and S. A. Im, "PARP inhibitors as therapeutics: Beyond modulation of parylation," *Cancers*. 2020, doi: 10.3390/cancers12020394.

[552] H. Wei and X. Yu, "Functions of PARylation in DNA Damage Repair Pathways," *Genomics, Proteomics and Bioinformatics*. 2016, doi: 10.1016/j.gpb.2016.05.001.

[553] Z. Li *et al.*, "Destabilization of linker histone H1.2 is essential for ATM activation and DNA damage repair," *Cell Res.*, 2018, doi: 10.1038/s41422-018-0048-0.

[554] B. S. King, K. L. Cooper, K. J. Liu, and L. G. Hudson, "Poly(ADP-ribose) contributes to an association between Poly(ADP-ribose) polymerase-1 and xeroderma pigmentosum complementation group A in nucleotide excision repair," *J. Biol. Chem.*, 2012, doi: 10.1074/jbc.M112.393504.

[555] R. G. Cooks, K. L. Busch, and G. L. Glish, "Mass spectrometry: Analytical capabilities and potentials," *Science*. 1983, doi: 10.1126/science.6353576.

[556] A. C. Leney and A. J. R. Heck, "Native Mass Spectrometry: What is in the Name?," *J. Am. Soc. Mass Spectrom.*, 2017, doi: 10.1007/s13361-016-1545-3.

[557] A. W. S. Fung, V. Sugumar, A. H. Ren, and V. Kulasingam, "Emerging role of clinical mass spectrometry in pathology," *Journal of Clinical Pathology*. 2020, doi: 10.1136/jclinpath-2019-206269.

[558] T. Herl and F. M. Matysik, "Recent Developments in Electrochemistry–Mass Spectrometry," *ChemElectroChem*. 2020, doi: 10.1002/celc.202000442.

[559] B. Bogdanov and R. D. Smith, "Proteomics by fticr mass spectrometry: TOP down and bottom up," *Mass Spectrometry Reviews*. 2005, doi: 10.1002/mas.20015.

[560] T. Dau, G. Bartolomucci, and J. Rappsilber, "Proteomics Using Protease Alternatives to Trypsin Benefits from Sequential Digestion with Trypsin," *Anal. Chem.*, 2020, doi: 10.1021/acs.analchem.0c00478.

[561] B. Manadas, V. M. Mendes, J. English, and M. J. Dunn, "Peptide fractionation in proteomics approaches," *Expert Review of Proteomics*. 2010, doi: 10.1586/epr.10.46.

[562] B. Lombardi, N. Rendell, M. Edwards, M. Katan, and J. G. Zimmermann, "Evaluation of phosphopeptide enrichment strategies for quantitative TMT analysis of complex network dynamics in cancer-associated cell signalling," *EuPA Open Proteomics*, 2015, doi: 10.1016/j.euprot.2015.01.002.

[563] B. Cañas Montalvo, D. López-Ferrer, A. Ramos-Fernández, E. Camafeita, and E. Calvo, "Mass spectrometry technologies for proteomics," *Briefings in Functional Genomics and Proteomics*. 2006, doi: 10.1093/bfgp/eli002.

[564] M. Dilillo, E. L. De Graaf, A. Yadav, M. E. Belov, and L. A. McDonnell, "Ultraviolet Photodissociation of ESI- and MALDI-Generated Protein Ions on a Q-Exactive Mass Spectrometer," *J. Proteome Res.*, 2019, doi: 10.1021/acs.jproteome.8b00896.

[565] A. Michalski *et al.*, "Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer," *Mol. Cell. Proteomics*, 2011, doi: 10.1074/mcp.M111.011015.

[566] A. Michalski *et al.*, "Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer," *Mol. Cell. Proteomics*, vol. 10, no. 9, p. M111.011015, Sep. 2011, doi: 10.1074/MCP.M111.011015.

[567] M. A. van Agthoven, Y. P. Y. Lam, P. B. O'Connor, C. Rolando, and M. A. Delsuc, "Two-dimensional mass spectrometry: new perspectives for tandem mass spectrometry," *Eur. Biophys. J.*, 2019, doi: 10.1007/s00249-019-01348-5.

[568] M. Jora *et al.*, "Differentiating Positional Isomers of Nucleoside Modifications by Higher-Energy Collisional Dissociation Mass Spectrometry (HCD MS)," *J. Am. Soc. Mass Spectrom.*, 2018, doi: 10.1007/s13361-018-1999-6.

[569] J. E. Elias and S. P. Gygi, "Target-decoy search strategy for mass spectrometry-based proteomics.," *Methods Mol. Biol.*, 2010, doi: 10.1007/978-1-60761-444-9_5.

[570] F. Yu, S. E. Haynes, and A. I. Nesvizhskii, "IonQuant enables accurate and sensitive label-free quantification with FDR-controlled match-between-runs," *Mol. Cell. Proteomics*, 2021, doi: 10.1016/J.MCPRO.2021.100077.

[571] T. J. Hedl *et al.*, "Proteomics approaches for biomarker and drug target discovery in als and ftd," *Frontiers in Neuroscience*. 2019, doi: 10.3389/fnins.2019.00548.

[572] A. Hogrebe, L. Von Stechow, D. B. Bekker-Jensen, B. T. Weinert, C. D. Kelstrup, and J. V. Olsen, "Benchmarking common quantification strategies for large-scale phosphoproteomics," *Nat. Commun.*, 2018, doi: 10.1038/s41467-018-03309-6.

[573] J. Li *et al.*, "TMTpro-18plex: The Expanded and Complete Set of TMTpro Reagents for Sample

Multiplexing," *J. Proteome Res.*, 2021, doi: 10.1021/acs.jproteome.1c00168.

[574] L. Williamson *et al.*, "UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene," *Cell*, vol. 168, no. 5, pp. 843-855.e13, Feb. 2017, doi: 10.1016/J.CELL.2017.01.019.

[575] S. Boeing *et al.*, "Multiomic Analysis of the UV-Induced DNA Damage Response," *Cell Rep.*, vol. 15, no. 7, pp. 1597–1610, May 2016, doi: 10.1016/J.CELREP.2016.04.047.

[576] E. Martinez-Salas, A. Embarc-Buh, and R. Francisco-Velilla, "Emerging roles of Gemin5: From snRNPs assembly to translation control," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21113868.

[577] A. B. Prusty, R. Meduri, B. K. Prusty, J. Vanselow, A. Schlosser, and U. Fischer, "Impaired spliceosomal UsnRNP assembly leads to Sm mRNA down-regulation and Sm protein degradation," *J. Cell Biol.*, 2017, doi: 10.1083/jcb.201611108.

[578] S. Otter, M. Grimmler, N. Neuenkirchen, A. Chari, A. Sickmann, and U. Fischer, "A comprehensive interaction map of the human survival of motor neuron (SMN) complex," *J. Biol. Chem.*, 2007, doi: 10.1074/jbc.M608528200.

[579] M. Moreno-Morcillo, R. Francisco-Velilla, A. Embarc-Buh, J. Fernández-Chamorro, S. Ramón-Maiques, and E. Martinez-Salas, "Structural basis for the dimerization of Gemin5 and its role in protein recruitment and translation control," *Nucleic Acids Res.*, 2020, doi: 10.1093/nar/gkz1126.

[580] S. Hofer *et al.*, "Contradictory effects of chemical filters in UV/ROS-stressed human keratinocyte and fibroblast cells," *ALTEX*, 2019, doi: 10.14573/altex.1808201.

[581] É. Hideg, M. A. K. Jansen, and Å. Strid, "UV-B exposure, ROS, and stress: Inseparable companions or loosely linked associates?," *Trends in Plant Science*. 2013, doi: 10.1016/j.tplants.2012.09.003.

[582] M. Alhamdoosh, C. W. Law, L. Tian, J. M. Sheridan, M. Ng, and M. E. Ritchie, "Easy and efficient ensemble gene set testing with EGSEA," *F1000Research*, 2017, doi: 10.12688/f1000research.12544.1.

[583] M. Gatei *et al.*, "Role for ATM in DNA damage-induced phosphorylation of BRCA1," *Cancer Res.*, 2000.

[584] D. Cortez, Y. Wang, J. Qin, and S. J. Elledge, "Requirement of ATM-dependent phosphorylation of Brca1 in the DNA damage response to double-strand breaks," *Science (80-.).*, 1999, doi: 10.1126/science.286.5442.1162.

[585] N. Tomimatsu *et al.*, "Exo1 plays a major role in DNA end resection in humans and influences double-strand break repair and damage signaling decisions," *DNA Repair (Amst).*, 2012, doi: 10.1016/j.dnarep.2012.01.006.

[586] D. R. Matson *et al.*, "High nuclear TPX2 expression correlates with TP53 mutation and poor clinical behavior in a large breast cancer cohort, but is not an independent predictor of chromosomal instability," *BMC Cancer*, 2021, doi: 10.1186/s12885-021-07893-7.

[587] M. Ognibene, M. Podestà, A. Garaventa, and A. Pezzolo, "Role of GOLPH3 and TPX2 in neuroblastoma DNA damage response and cell resistance to chemotherapy," *Int. J. Mol. Sci.*, 2019, doi: 10.3390/ijms20194764.

[588] N. Hanagata, F. Zhuang, S. Connolly, J. Li, N. Ogawa, and M. Xu, "Molecular responses of human lung epithelial cells to the toxicity of copper oxide nanoparticles inferred from whole genome expression analysis," *ACS Nano*, 2011, doi: 10.1021/nn202966t.

[589] Z. Deng, S. Zhang, S. Gu, X. Ni, W. Zeng, and X. Li, "Useful bicistronic reporter system for studying Poly(A) site-defining cis elements and regulation of alternative polyadenylation," *Int. J. Mol. Sci.*, 2018, doi: 10.3390/ijms19010279.

[590] K. M. Brown and G. M. Gilmartin, "A Mechanism for the Regulation of Pre-mRNA 3′

Processing by Human Cleavage Factor Im," *Mol. Cell*, 2003, doi: 10.1016/S1097-2765(03)00453-2.

[591] L. Williamson *et al.*, "UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene," *Cell*, vol. 168, no. 5, pp. 843-855.e13, 2017, doi: 10.1016/j.cell.2017.01.019.

[592] S. Boeing *et al.*, "Multiomic Analysis of the UV-Induced DNA Damage Response," *Cell Rep.*, vol. 15, no. 7, pp. 1597–1610, 2016, doi: 10.1016/j.celrep.2016.04.047.

[593] F. Yuan, Y. Zhang, L. Ma, Q. Cheng, G. Li, and T. Tong, "Enhanced NOLC1 promotes cell senescence and represses hepatocellular carcinoma cell proliferation by disturbing the organization of nucleolus," *Aging Cell*, 2017, doi: 10.1111/acel.12602.

[594] E. S. Johnson, P. C. M. Ma, I. M. Ota, and A. Varshavsky, "A proteolytic pathway that recognizes ubiquitin as a degradation signal," *J. Biol. Chem.*, vol. 270, no. 29, pp. 17442–17456, 1995, doi: 10.1074/jbc.270.29.17442.

[595] O. Cazzalini, P. Perucca, R. Mocchi, S. Sommatis, E. Prosperi, and L. A. Stivala, "DDB2 association with PCNA is required for its degradation after UV-induced DNA damage," *Cell Cycle*, 2014, doi: 10.4161/cc.26987.

[596] X. Wu *et al.*, "DDB2 regulates DNA replication through PCNA-independent degradation of CDT2," *Cell Biosci.*, 2021, doi: 10.1186/s13578-021-00540-5.

[597] S. Li *et al.*, "Developing Gender-Specific Gene Expression Biodosimetry Using a Panel of Radiation-Responsive Genes for Determining Radiation Dose in Human Peripheral Blood," *Radiat. Res.*, 2019, doi: 10.1667/RR15355.1.

[598] D. O. Warmerdam *et al.*, "PHF6 promotes non-homologous end joining and G2 checkpoint recovery," *EMBO Rep.*, 2020, doi: 10.15252/embr.201948460.

[599] L.-M. Appel *et al.*, "PHF3 regulates neuronal gene expression through the Pol II CTD reader domain SPOC," *Nat. Commun.*, 2021, doi: 10.1038/s41467-021-26360-2.

[600] Q. Wu, W. Wu, V. Jacevic, T. C. C. Franca, X. Wang, and K. Kuca, "Selective inhibitors for JNK signalling: a potential targeted therapy in cancer," *Journal of Enzyme Inhibition and Medicinal Chemistry*. 2020, doi: 10.1080/14756366.2020.1720013.

[601] X. Xie *et al.*, "Abstract 750: JNK-IN-8: a novel covalent inhibitor targeting JNK signaling in triple-negative breast cancer," 2014, doi: 10.1158/1538-7445.am2014-750.

[602] S. Gupta, D. Campbell, B. Dérijard, and R. J. Davis, "Transcription factor ATF2 regulation by the JNK signal transduction pathway," *Science (80-.).*, 1995, doi: 10.1126/science.7824938.

[603] H. Okazawa and S. Estus, "The JNK/c-Jun cascade and Alzheimer's disease," *Am. J. Alzheimers. Dis. Other Demen.*, 2002, doi: 10.1177/153331750201700209.

[604] S. Bagchi, R. Fredriksson, and Å. Wallén-Mackenzie, "In Situ Proximity Ligation Assay (PLA)," *Methods Mol. Biol.*, 2015, doi: 10.1007/978-1-4939-2742-5_15.

[605] T. C. Branon *et al.*, "Efficient proximity labeling in living cells and organisms with TurboID," *Nat. Biotechnol.*, vol. 36, no. 9, pp. 880–898, 2018, doi: 10.1038/nbt.4201.

[606] H. Li, A. M. Frankenfield, R. Houston, S. Sekine, and L. Hao, "Thiol-Cleavable Biotin for Chemical and Enzymatic Biotinylation and Its Application to Mitochondrial TurboID Proteomics," *J. Am. Soc. Mass Spectrom.*, 2021, doi: 10.1021/jasms.1c00079.

[607] T. C. Branon *et al.*, "Directed evolution of TurboID for efficient proximity labeling in living cells and organisms," *bioRxiv*, 2017, doi: 10.1101/196980.

[608] X. Bao *et al.*, "Capturing the interactome of newly transcribed RNA," *Nat. Methods*, 2018, doi: 10.1038/nmeth.4595.

[609] N. Binothman, I. Y. Hachim, J. J. Lebrun, and S. Ali, "CPSF6 is a Clinically Relevant Breast

Cancer Vulnerability Target: Role of CPSF6 in Breast Cancer," *EBioMedicine*, 2017, doi: 10.1016/j.ebiom.2017.06.023.

[610] C. M. Borini Etichetti, A. Tenaglia, M. N. Arroyo, and J. E. Girardini, "Expression of zebrafish cpsf6 in embryogenesis and role of protein domains on subcellular localization," *Gene Expr. Patterns*, 2020, doi: 10.1016/j.gep.2020.119114.

[611] "Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles inline-supplementary-material-4." .

[612] T. Naganuma and T. Hirose, "Paraspeckle formation during the biogenesis of long non-coding RNAs," *RNA Biology*. 2013, doi: 10.4161/rna.23547.

[613] D. M. Mitrea and R. W. Kriwacki, "Phase separation in biology; Functional organization of a higher order Short linear motifs - The unexplored frontier of the eukaryotic proteome," *Cell Commun. Signal.*, vol. 14, no. 1, pp. 1–20, 2016, doi: 10.1186/s12964-015-0125-7.

[614] A. A. Hyman, C. A. Weber, and F. Jülicher, "Liquid-liquid phase separation in biology," *Annu. Rev. Cell Dev. Biol.*, vol. 30, pp. 39–58, 2014, doi: 10.1146/ANNUREV-CELLBIO-100913-013325.

[615] B. A. Gibson *et al.*, "Organization of Chromatin by Intrinsic and Regulated Phase Separation," *Cell*, 2019, doi: 10.1016/j.cell.2019.08.037.

[616] Y. Lin, D. S. W. Protter, M. K. Rosen, and R. Parker, "Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins," *Mol. Cell*, vol. 60, no. 2, pp. 208–219, 2015, doi: 10.1016/j.molcel.2015.08.018.

[617] S. Kroschwald, S. Maharana, and A. Simon, "Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments," *Matters*, pp. 1–7, 2017, doi: 10.19185/matters.201702000010.

[618] A. Jain and R. D. Vale, "RNA phase transitions in repeat expansion disorders," *Nature*, 2017, doi: 10.1038/nature22386.

[619] F. Pessina *et al.*, "Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors," *Nat. Cell Biol. 2019 2110*, vol. 21, no. 10, pp. 1286–1299, Sep. 2019, doi: 10.1038/s41556-019-0392-4.

[620] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, "Biomolecular condensates: Organizers of cellular biochemistry," *Nat. Rev. Mol. Cell Biol.*, vol. 18, no. 5, pp. 285–298, May 2017, doi: 10.1038/NRM.2017.7.

[621] M. Tari, V. Manceau, J. Matha Salone, A. Kobayashi, D. Pastré, and A. Maucuer, " U2 AF 65 assemblies drive sequence-specific splice site recognition ," *EMBO Rep.*, vol. 20, no. 8, pp. 1–18, 2019, doi: 10.15252/embr.201847604.

[622] P. Quintero-Cadena, T. L. Lenstra, and P. W. Sternberg, "RNA Pol II Length and Disorder Enable Cooperative Scaling of Transcriptional Bursting," *Mol. Cell*, 2020, doi: 10.1016/j.molcel.2020.05.030.

[623] A. Pancholi *et al.*, "RNA polymerase II clusters form in line with surface condensation on regulatory chromatin," *Mol. Syst. Biol.*, 2021, doi: 10.15252/msb.202110272.

[624] K. M. Harlen and L. S. Churchman, "The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain," *Nature Reviews Molecular Cell Biology*. 2017, doi: 10.1038/nrm.2017.10.

[625] Z. Zhou, J. Sim, J. Griffith, and R. Reed, "Purification and electron microscopic visualization of functional human spliceosomes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 19, pp. 12203–12207, 2002, doi: 10.1073/pnas.182427099.

[626] L. Wan *et al.*, "Translation stress and collided ribosomes are co-activators of cGAS," *Mol. Cell*, 2021, doi: 10.1016/j.molcel.2021.05.018.

[627] S. Kim *et al.*, "ATAD5 restricts R-loop formation through PCNA unloading and RNA helicase maintenance at the replication fork," *Nucleic Acids Res.*, 2020, doi: 10.1093/nar/gkaa501.

[628] M. Groh, L. O. Albulescu, A. Cristini, and N. Gromak, "Senataxin: Genome Guardian at the Interface of Transcription and Neurodegeneration," *Journal of Molecular Biology*. 2017, doi: 10.1016/j.jmb.2016.10.021.

[629] V. Pfeiffer, J. Crittin, L. Grolimund, and J. Lingner, "The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening," *EMBO J.*, 2013, doi: 10.1038/emboj.2013.217.

[630] R. Luna, A. G. Rondón, C. Pérez-Calero, I. Salas-Armenteros, and A. Aguilera, "The THO complex as a paradigm for the prevention of cotranscriptional R-Loops," *Cold Spring Harb. Symp. Quant. Biol.*, 2019, doi: 10.1101/sqb.2019.84.039594.

[631] P. Chakraborty, J. T. J. Huang, and K. Hiom, "DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing," *Nat. Commun.*, 2018, doi: 10.1038/s41467-018-06677-1.

[632] A. Chakraborty *et al.*, "FMRP bridges R-loops and DHX9 through direct interactions," *bioRxiv*, 2021.

[633] H. D. Nguyen *et al.*, "Spliceosome mutations induce R loop-associated sensitivity to ATR inhibition in myelodysplastic syndromes," *Cancer Res.*, 2018, doi: 10.1158/0008-5472.CAN-17-3970.

[634] "Myelodysplastic Syndrome Splicing Factor Mutations Induce R-Loops," *Cancer discovery*. 2018, doi: 10.1158/2159-8290.CD-RW2018-025.

[635] M. Tresini *et al.*, "The core spliceosome as target and effector of non-canonical ATM signalling," *Nature*, 2015, doi: 10.1038/nature14512.

[636] M. S. Domínguez-Sánchez, S. Barroso, B. Gómez-González, R. Luna, and A. Aguilera, "Genome instability and transcription elongation impairment in human cells depleted of THO/TREX," *PLoS Genet.*, 2011, doi: 10.1371/journal.pgen.1002386.

[637] T. Mosler *et al.*, "R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability," *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–17, Dec. 2021, doi: 10.1038/s41467-021-27530-y.

[638] L. A. Sanz and F. Chédin, "High-resolution, strand-specific R-loop mapping via S9.6-based DNA–RNA immunoprecipitation and high-throughput sequencing," *Nat. Protoc.*, 2019, doi: 10.1038/s41596-019-0159-1.

[639] M. Malig, S. R. Hartono, J. M. Giafaglione, L. A. Sanz, and F. Chedin, "High-Throughput Single-Molecule R-loop Footprinting Reveals Principles of R-loop Formation," *bioRxiv*, 2019, doi: 10.1101/640094.

[640] J. Y. Chen, X. Zhang, X. D. Fu, and L. Chen, "R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1," *Nat. Protoc.*, 2019, doi: 10.1038/s41596-019-0154-6.

[641] M. Graf *et al.*, "Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle," *Cell*, vol. 170, no. 1, pp. 72-85.e14, Jun. 2017.

[642] Y. Okamoto *et al.*, "FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress," *FEBS J.*, 2019, doi: 10.1111/febs.14700.

[643] J. Li and D. F. Stern, "DNA damage regulates Chk2 association with chromatin," *J. Biol. Chem.*, 2005, doi: 10.1074/jbc.M509299200.

[644] S. Britton, J. Coates, and S. P. Jackson, "A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair," *J. Cell Biol.*, 2013, doi: 10.1083/jcb.201303073.

[645] S. A. Gaidamakov *et al.*, "Eukaryotic RNases H1 act processively by interactions through the

duplex RNA-binding domain," *Nucleic Acids Res.*, 2005, doi: 10.1093/nar/gki510.

[646] K. F. Cho *et al.*, "Split-TurboID enables contact-dependent proximity labeling in cells," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 22, 2020, doi: 10.1073/pnas.1919528117.

[647] P. Dent, A. Yacoub, P. B. Fisher, M. P. Hagan, and S. Grant, "MAPK pathways in radiation responses," *Oncogene*. 2003, doi: 10.1038/sj.onc.1206701.

[648] M. Burotto, V. L. Chiou, J. M. Lee, and E. C. Kohn, "The MAPK pathway across different malignancies: A new perspective," *Cancer*. 2014, doi: 10.1002/cncr.28864.

[649] Y. Guo, W. Pan, S. Liu, Z. Shen, Y. Xu, and L. Hu, "ERK/MAPK signalling pathway and tumorigenesis (Review)," *Exp. Ther. Med.*, 2020, doi: 10.3892/etm.2020.8454.

[650] X. W. LI, R. Z. HE, Y. LI, and Z. F. RUAN, "Tizoxanide mitigates inflammatory response in LPS-induced neuroinflammation in microglia via restraining p38/MAPK pathway," *Eur. Rev. Med. Pharmacol. Sci.*, 2020, doi: 10.26355/eurrev_202006_21543.

[651] K. V. Krishna, S. K. Dubey, G. Singhvi, G. Gupta, and P. Kesharwani, "MAPK pathway: Potential role in glioblastoma multiforme," *Interdisciplinary Neurosurgery: Advanced Techniques and Case Management*. 2021, doi: 10.1016/j.inat.2020.100901.

[652] O. Atay and J. M. Skotheim, "Spatial and temporal signal processing and decision making by MAPK pathways," *Journal of Cell Biology*. 2017, doi: 10.1083/jcb.201609124.

[653] Y. A. E. Elghobashy, M. F. Assar, A. A. Mahmoud, A. Monem A Eltorgoman, and S. Elmasry, "The relation between mitogen activated protein kinase (MAPK) pathway and different genes expression in patients with beta Thalassemia," *Biochem. Biophys. Reports*, 2020, doi: 10.1016/j.bbrep.2020.100836.

[654] P. Botwina *et al.*, "Berberine hampers influenza a replication through inhibition of MAPK/ERK pathway," *Viruses*, 2020, doi: 10.3390/v12030344.

[655] J. S. Zhao *et al.*, "Loss of PR55α promotes proliferation and metastasis by activating MAPK/AKT signaling in hepatocellular carcinoma," *Cancer Cell Int.*, 2021, doi: 10.1186/s12935-021-01796-0.

[656] N. A. Carne, S. Bell, A. P. Brown, A. Määttä, M. J. Flagler, and A. M. Benham, "Reductive Stress Selectively Disrupts Collagen Homeostasis and Modifies Growth Factor-independent Signaling through the MAPK/Akt Pathway in Human Dermal Fibroblasts," *Mol. Cell. Proteomics*, 2019, doi: 10.1074/mcp.RA118.001140.

[657] H. H. Yu *et al.*, "Benzoylaconitine inhibits production of IL-6 and IL-8 via MAPK, Akt, NF-κB Signaling in IL-1β-induced human synovial cells," *Biol. Pharm. Bull.*, 2020, doi: 10.1248/bpb.b19-00719.

[658] Y. H. Yuan *et al.*, "Epigenetic inactivation of HOXD10 is associated with human colon cancer via inhibiting the RHOC/AKT/MAPK signaling pathway," *Cell Commun. Signal.*, 2019, doi: 10.1186/s12964-018-0316-0.

[659] J. Mitra, P. M. Hegde, and M. L. Hegde, "Loss of endosomal recycling factor RAB11 coupled with complex regulation of MAPK/ERK/AKT signaling in postmortem spinal cord specimens of sporadic amyotrophic lateral sclerosis patients," *Mol. Brain*, 2019, doi: 10.1186/s13041-019-0475-y.

[660] J. Yang, M. Y. Xie, X. L. Yang, B. H. Liu, and H. H. Lin, "Phosphoproteomic Profiling Reveals the Importance of CK2, MAPKs and CDPKs in Response to Phosphate Starvation in Rice," *Plant Cell Physiol.*, 2019, doi: 10.1093/pcp/pcz167.

[661] P. Wińska, O. Karatsai, M. Staniszewska, M. Koronkiewicz, K. Chojnacki, and M. J. Rędowicz, "Synergistic interactions of 5-fluorouracil with inhibitors of protein kinase ck2 correlate with p38 mapk activation and fak inhibition in the triple-negative breast cancer cell line," *Int. J. Mol. Sci.*, 2020, doi: 10.3390/ijms21176234.

[662] F. Stark, J. Pfannstiel, I. Klaiber, and T. Raabe, "Protein kinase CK2 links polyamine metabolism to MAPK signalling in Drosophila," *Cell. Signal.*, 2011, doi: 10.1016/j.cellsig.2011.01.013.

[663] E. Chevet *et al.*, "Phosphorylation by CK2 and MAPK enhances calnexin association with ribosomes," *EMBO J.*, 1999, doi: 10.1093/emboj/18.13.3655.

[664] G. Cozza *et al.*, "Quinalizarin as a potent, selective and cell-permeable inhibitor of protein kinase CK2," *Biochem. J.*, 2009, doi: 10.1042/BJ20090069.

[665] K. Husain, T. T. Williamson, N. Nelson, and T. Ghansah, "Protein kinase 2 (CK2): a potential regulator of immune cell development and function in cancer," *Immunological Medicine*. 2021, doi: 10.1080/25785826.2020.1843267.

[666] M. Montenarh, "Protein kinase CK2 in DNA damage and repair," *Translational Cancer Research*. 2016, doi: 10.3978/j.issn.2218-676X.2016.01.09.

[667] X. Shiyang *et al.*, "Casein kinase 2 modulates the spindle assembly checkpoint to orchestrate porcine oocyte meiotic progression," *J. Anim. Sci. Biotechnol.*, 2020, doi: 10.1186/s40104-020-00438-1.

[668] Y. Wang, X. Wang, G. Xu, and S. Gou, "Novel CK2-Specific Pt(II) Compound Reverses Cisplatin-Induced Resistance by Inhibiting Cancer Cell Stemness and Suppressing DNA Damage Repair in Non-small Cell Lung Cancer Treatments," *J. Med. Chem.*, 2021, doi: 10.1021/acs.jmedchem.1c00079.

[669] V. Hammoudi, "DNA damage response priming by CK2: A matter of life or cell death in root apical meristems," *Plant Cell*. 2021, doi: 10.1093/plcell/koab018.

[670] I. M. Johnston, S. J. Allison, J. P. Morton, L. Schramm, P. H. Scott, and R. J. White, "CK2 Forms a Stable Complex with TFIIIB and Activates RNA Polymerase III Transcription in Human Cells," *Mol. Cell. Biol.*, 2002, doi: 10.1128/mcb.22.11.3757-3768.2002.

[671] D. Graczyk, J. Dębski, G. Muszyńska, M. Bretner, O. Lefebvre, and M. Boguta, "Casein kinase II-mediated phosphorylation of general repressor Maf1 triggers RNA polymerase III activation," *Proc. Natl. Acad. Sci. U. S. A.*, 2011, doi: 10.1073/pnas.1010010108.

[672] J. Liu, L. Liu, J. He, Y. Xu, and Y. Wang, "Multi-omic analysis of altered transcriptome and epigenetic signatures in the UV-induced DNA damage response," *DNA Repair (Amst).*, vol. 106, p. 103172, Oct. 2021, doi: 10.1016/J.DNAREP.2021.103172.

[673] C. L. Yang, J. Y. Meng, M. S. Yao, and C. Y. Zhang, "Transcriptome Analysis of Myzus persicae to UV-B Stress," *J. Insect Sci.*, 2021, doi: 10.1093/jisesa/ieab033.

[674] D. Yamada *et al.*, "The human protein kinase HIPK2 phosphorylates and downregulates the methyl-binding transcription factor ZBTB4," *Oncogene*, 2009, doi: 10.1038/onc.2009.109.

[675] A. Weber *et al.*, "Zbtb4 represses transcription of P21CIP1 and controls the cellular response to p53 activation," *EMBO J.*, 2008, doi: 10.1038/emboj.2008.85.

[676] H. Ikehata and T. Ono, "Significance of CpG Methylation for Solar UV-induced Mutagenesis and Carcinogenesis in Skin," *Photochem. Photobiol.*, 2006, doi: 10.1562/2006-02-28-ir-822.

[677] C. Rao *et al.*, "Ksr1-and erk-dependent translational regulation of the epithelial-to-mesenchymal transition," *Elife*, 2021, doi: 10.7554/eLife.66608.

[678] Y. Okina *et al.*, "TGF-β1-driven reduction of cytoglobin leads to oxidative DNA damage in stellate cells during non-alcoholic steatohepatitis," *J. Hepatol.*, 2020, doi: 10.1016/j.jhep.2020.03.051.

[679] K. Liu, J. Wang, X. Gao, and W. Ren, "C1q/TNF-Related Protein 9 Inhibits Coxsackievirus B3-Induced Injury in Cardiomyocytes through NF-κ B and TGF-β 1/Smad2/3 by Modulating THBS1," *Mediators Inflamm.*, 2020, doi: 10.1155/2020/2540687.

[680] J. M. Soll, J. R. Brickner, M. C. Mudge, and N. Mosammaparast, "RNA ligase-like domain in

activating signal cointegrator 1 complex subunit 1 (ASCC1) regulates ASCC complex function during alkylation damage," *J. Biol. Chem.*, 2018, doi: 10.1074/jbc.RA117.000114.

[681] H. W. Cho, H. S. Jin, and Y. Bin Eom, "Association between non-Caucasian-specific ASCC1 gene polymorphism and osteoporosis and obesity in Korean postmenopausal women," *J. Bone Miner. Metab.*, 2020, doi: 10.1007/s00774-020-01120-2.

[682] P. A. Madureira, R. Hill, P. W. K. Lee, and D. M. Waisman, "Genotoxic Agents Promote the Nuclear Accumulation of Annexin A2: Role of Annexin A2 in Mitigating DNA Damage," *PLoS One*, 2012, doi: 10.1371/journal.pone.0050591.

[683] Y. Huang, C. H. Yan, and S. Bin Fu, "The cloning and expression of apoptosis associated gene ANNEXIN A2 induced by p53 gene," *Chinese J. Med. Genet.*, 2005.

[684] K. M. Waters *et al.*, "Annexin A2 modulates radiation-sensitive transcriptional programming and cell fate," *Radiat. Res.*, 2013, doi: 10.1667/RR3056.1.

[685] P. A. Madureira, R. Hill, V. A. Miller, C. Giacomantonio, P. W. K. Lee, and D. M. Waisman, "Annexin A2 is a novel cellular redox regulatory protein involved in tumorigenesis," *Oncotarget*, 2011, doi: 10.18632/oncotarget.375.

[686] Z. Xing, X. Wang, J. Liu, M. Zhang, K. Feng, and X. Wang, "Expression and prognostic value of CDK1, CCNA2, and CCNB1 gene clusters in human breast cancer," *J. Int. Med. Res.*, 2021, doi: 10.1177/0300060520980647.

[687] N. Y. Ji *et al.*, "Evaluation of annexin II as a potential serum marker for hepatocellular carcinoma using a developed sandwich ELISA method," *Int. J. Mol. Med.*, 2009, doi: 10.3892/ijmm_00000290.

[688] C. COMéRA, B. ROTHHUT, and F. RUSSO-MARIE, "Identification and characterization of phospholipase A2 inhibitory proteins in human mononuclear cells," *Eur. J. Biochem.*, 1990, doi: 10.1111/j.1432-1033.1990.tb15381.x.

[689] H. Tsukamoto *et al.*, "Annexin A2 regulates a disintegrin and metalloproteinase 17-mediated ectodomain shedding of pro-tumor necrosis factor-α in monocytes and colon epithelial cells," *Inflamm. Bowel Dis.*, 2013, doi: 10.1097/MIB.0b013e318281f43a.

[690] D. C. Yánez, S. Ross, and T. Crompton, "The IFITM protein family in adaptive immunity," *Immunology*. 2020, doi: 10.1111/imm.13163.

[691] R. Dowran *et al.*, "Various interferon (IFN)-inducible transmembrane (IFITM) proteins for COVID-19, is there a role for the combination of mycophenolic acid and interferon?," *Biochimie*, 2020, doi: 10.1016/j.biochi.2020.08.006.

[692] C. Prelli Bozzo *et al.*, "IFITM proteins promote SARS-CoV-2 infection and are targets for virus inhibition in vitro," *Nat. Commun.*, 2021, doi: 10.1038/s41467-021-24817-y.

[693] H. P. Li *et al.*, "Characterization and anti-inflammation role of swine IFITM3 gene," *Oncotarget*, 2017, doi: 10.18632/oncotarget.20568.

[694] A. Takashima and P. R. Bergstresser, "Impact of UVB radiation on the epidermal cytokine network," *Photochem. Photobiol.*, 1996, doi: 10.1111/j.1751-1097.1996.tb03054.x.

[695] T. Geuens, D. Bouhy, and V. Timmerman, "The hnRNP family: insights into their role in health and disease," *Human Genetics*. 2016, doi: 10.1007/s00439-016-1683-5.

[696] M. Fioramonte, G. Reis-de-Oliveira, C. Brandão-Teles, and D. Martins-de-Souza, "A glimpse on the architecture of hnRNP C1/C2 interaction network in cultured oligodendrocytes," *Biochim. Biophys. Acta - Proteins Proteomics*, 2021, doi: 10.1016/j.bbapap.2021.140711.

[697] K. Mori *et al.*, " Reduced hn RNPA 3 increases C9orf72 repeat RNA levels and dipeptide-repeat protein deposition ," *EMBO Rep.*, 2016, doi: 10.15252/embr.201541724.

[698] Y. Kawabe, K. Mori, T. Yamashita, S. Gotoh, and M. Ikeda, " The RNA exosome complex

degrades expanded hexanucleotide repeat RNA in C9orf72 FTLD / ALS ," *EMBO J.*, 2020, doi: 10.15252/embj.2019102700.

[699] J. Jiang *et al.*, "Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs," *Neuron*, 2016, doi: 10.1016/j.neuron.2016.04.006.

[700] P. McGoldrick *et al.*, "Unaffected mosaic C9orf72 case: RNA foci, dipeptide proteins, but upregulated C9orf72 expression," *Neurology*, 2018, doi: 10.1212/WNL.0000000000004865.

[701] Z. T. McEachin, J. Parameswaran, N. Raj, G. J. Bassell, and J. Jiang, "RNA-mediated toxicity in C9orf72 ALS and FTD," *Neurobiology of Disease*. 2020, doi: 10.1016/j.nbd.2020.105055.

[702] R. Simone *et al.*, " G-quadruplex-binding small molecules ameliorate C9orf72 FTD / ALS pathology in vitro and in vivo ," *EMBO Mol. Med.*, 2018, doi: 10.15252/emmm.201707850.

[703] M. H. Schludi and D. Edbauer, " Targeting RNA G-quadruplexes as new treatment strategy for C9orf72 ALS / FTD ," *EMBO Mol. Med.*, 2018, doi: 10.15252/emmm.201708572.

[704] E. G. Conlon *et al.*, "The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains," *Elife*, 2016, doi: 10.7554/eLife.17820.

[705] G. Y. Liu *et al.*, "Nucleophosmin Regulates Intracellular Oxidative Stress Homeostasis via Antioxidant PRDX6," *J. Cell. Biochem.*, 2017, doi: 10.1002/jcb.26135.

[706] X. Hua, W. Chi, L. Su, J. Li, Z. Zhang, and X. Yuan, "ROS-induced Oxidative Injury involved in Pathogenesis of Fungal Keratitis via p38 MAPK Activation," *Sci. Rep.*, 2017, doi: 10.1038/s41598-017-09636-w.

[707] S. Wiebe *et al.*, "The eIF4E homolog 4EHP (eIF4E2) regulates hippocampal long-term depression and impacts social behavior," *Mol. Autism*, 2020, doi: 10.1186/s13229-020-00394-7.

[708] B. Zinshteyn, N. K. Sinha, S. U. Enam, B. Koleske, and R. Green, "Translational repression of NMD targets by GIGYF2 and EIF4E2," *PLoS Genet.*, 2021, doi: 10.1371/journal.pgen.1009813.

[709] N. K. Sinha *et al.*, "EDF1 coordinates cellular responses to ribosome collisions," *Elife*, 2020, doi: 10.7554/ELIFE.58828.

[710] R. Abdella, A. Talyzina, S. Chen, C. J. Inouye, R. Tjian, and Y. He, "Structure of the human Mediator-bound transcription preinitiation complex," *Science (80-.).*, 2021, doi: 10.1126/science.abg3074.

[711] R. P. Fisher, "Cdk7: a kinase at the core of transcription and in the crosshairs of cancer drug discovery," *Transcription*. 2019, doi: 10.1080/21541264.2018.1553483.

[712] K. Glover-Cutter *et al.*, "TFIIH-Associated Cdk7 Kinase Functions in Phosphorylation of C-Terminal Domain Ser7 Residues, Promoter-Proximal Pausing, and Termination by RNA Polymerase II," *Mol. Cell. Biol.*, 2009, doi: 10.1128/mcb.00637-09.

[713] C. C. Ebmeier *et al.*, "Human TFIIH Kinase CDK7 Regulates Transcription-Associated Chromatin Modifications," *Cell Rep.*, 2017, doi: 10.1016/j.celrep.2017.07.021.

[714] V. Álvarez, C. Frattini, M. P. Sacristán, A. Gallego-Sánchez, R. Bermejo, and A. Bueno, "PCNA Deubiquitylases Control DNA Damage Bypass at Replication Forks," *Cell Rep.*, 2019, doi: 10.1016/j.celrep.2019.09.054.

[715] B. M. Ripley, M. S. Gildenberg, and M. Todd Washington, "Control of DNA damage bypass by ubiquitylation of PCNA," *Genes*. 2020, doi: 10.3390/genes11020138.

[716] L. Fan, T. Bi, L. Wang, and W. Xiao, "DNA-damage tolerance through PCNA ubiquitination and sumoylation," *Biochemical Journal*. 2020, doi: 10.1042/BCJ20190579.

[717] J. Lykke-Andersen and E. J. Bennett, "Protecting the proteome: Eukaryotic cotranslational quality control pathways," *Journal of Cell Biology*. 2014, doi: 10.1083/jcb.201311103.

[718] L. Daftuar, Y. Zhu, X. Jacq, and C. Prives, “Ribosomal Proteins RPL37, RPS15 and RPS20 Regulate the Mdm2-p53-MdmX Network,” *PLoS One*, 2013, doi: 10.1371/journal.pone.0068667.

[719] G. M. Cooper and R. E. Hausman, *The Cell: A Molecular Approach 2nd Edition*. 2007.

[720] J. Méndez and B. Stillman, “Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis,” *Mol. Cell. Biol.*, 2000, doi: 10.1128/mcb.20.22.8602-8612.2000.

[721] G. Kustatscher *et al.*, “Proteomics of a fuzzy organelle: Interphase chromatin,” *EMBO J.*, 2014, doi: 10.1002/embj.201387614.

[722] G. Kustatscher, K. L. H. Wills, C. Furlan, and J. Rappsilber, “Chromatin enrichment for proteomics,” *Nature Protocols*. 2014, doi: 10.1038/nprot.2014.142.

[723] G. Batugedara *et al.*, “The chromatin bound proteome of the human malaria parasite,” *Microb. Genomics*, 2020, doi: 10.1099/mgen.0.000327.

[724] G. van Mierlo, R. A. Wester, and H. Marks, “A Mass Spectrometry Survey of Chromatin-Associated Proteins in Pluripotency and Early Lineage Commitment,” *Proteomics*, 2019, doi: 10.1002/pmic.201900047.

[725] Y. Kito *et al.*, “Cell cycle–dependent localization of the proteasome to chromatin,” *Sci. Rep.*, 2020, doi: 10.1038/s41598-020-62697-2.

[726] G. Kustatscher, P. Grabowski, and J. Rappsilber, “Multiclassifier combinatorial proteomics of organelle shadows at the example of mitochondria in chromatin data,” *Proteomics*, 2016, doi: 10.1002/pmic.201500267.

[727] P. A. Ginno, L. Burger, J. Seebacher, V. Iesmantavicius, and D. Schübeler, “Cell cycle-resolved chromatin proteomics reveals the extent of mitotic preservation of the genomic regulatory landscape,” *Nat. Commun.*, 2018, doi: 10.1038/s41467-018-06007-5.

[728] C. S. Hughes, S. Foehr, D. A. Garfield, E. E. Furlong, L. M. Steinmetz, and J. Krijgsveld, “Ultrasensitive proteome analysis using paramagnetic bead technology,” *Mol. Syst. Biol.*, 2014, doi: 10.15252/msb.20145625.

[729] S. Guan, P. P. Taylor, Z. Han, M. F. Moran, and B. Ma, “DDIA: Data dependent-independent acquisition proteomics - DDA and DIA in a single LC-MS/MS run,” *bioRxiv*, 2019, doi: 10.1101/802231.

[730] F. Klont *et al.*, “Assessment of Sample Preparation Bias in Mass Spectrometry-Based Proteomics,” *Anal. Chem.*, 2018, doi: 10.1021/acs.analchem.8b00600.

[731] V. Demichev, C. B. Messner, S. I. Vernardis, K. S. Lilley, and M. Ralser, “DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput,” *Nat. Methods*, 2020, doi: 10.1038/s41592-019-0638-x.

[732] J. Cicenas, “The Aurora kinase inhibitors in cancer research and therapy,” *Journal of Cancer Research and Clinical Oncology*. 2016, doi: 10.1007/s00432-016-2136-1.

[733] T. M. Pitts, S. L. Davis, S. G. Eckhardt, and E. L. Bradshaw-Pierce, “Targeting nuclear kinases in cancer: Development of cell cycle kinase inhibitors,” *Pharmacology and Therapeutics*. 2014, doi: 10.1016/j.pharmthera.2013.12.010.

[734] J. Qi *et al.*, “Selective inhibition of Aurora A and B kinases effectively induces cell cycle arrest in t(8;21) acute myeloid leukemia,” *Biomed. Pharmacother.*, 2019, doi: 10.1016/j.biopha.2019.109113.

[735] J. A. Pérez-Fidalgo, V. Gambardella, B. Pineda, O. Burgues, O. Piñero, and A. Cervantes, “Aurora kinases in ovarian cancer,” *ESMO Open*. 2020, doi: 10.1136/esmoopen-2020-000718.

[736] X. W. Xiao, R. Liu, Q. J. Shun, Y. F. Fei, and M. Z. Qi, "Overexpression of Aurora-A kinase promotes tumor cell proliferation and inhibits apoptosis in esophageal squamous cell carcinoma cell line," *Cell Res.*, 2006, doi: 10.1038/sj.cr.7310046.

[737] J. A. Puig-Butille *et al.*, "AURKA Overexpression Is Driven by FOXM1 and MAPK/ERK Activation in Melanoma Cells Harboring BRAF or NRAS Mutations: Impact on Melanoma Prognosis and Therapy," *J. Invest. Dermatol.*, 2017, doi: 10.1016/j.jid.2017.01.021.

[738] C. R. Elbæk, V. Petrosius, and C. S. Sørensen, "WEE1 kinase limits CDK activities to safeguard DNA replication and mitotic entry," *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. 2020, doi: 10.1016/j.mrfmmm.2020.111694.

[739] F. Bassermann, D. Frescas, D. Guardavaccaro, L. Busino, A. Peschiaroli, and M. Pagano, "The Cdc14B-Cdh1-Plk1 Axis Controls the G2 DNA-Damage-Response Checkpoint," *Cell*, 2008, doi: 10.1016/j.cell.2008.05.043.

[740] J. A. Riback *et al.*, "Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response," *Cell*, vol. 168, no. 6, pp. 1028-1040.e19, 2017, doi: 10.1016/j.cell.2017.02.027.

[741] Y. Hirose, R. Tacke, and J. L. Manley, "Phosphorylated RNA polymerase II stimulates pre-mRNA splicing," *Genes Dev.*, 1999, doi: 10.1101/gad.13.10.1234.

[742] C. J. David, A. R. Boyne, S. R. Millhouse, and J. L. Manley, "The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex," *Genes Dev.*, 2011, doi: 10.1101/gad.2038011.

[743] M. M. Keshwani *et al.*, "Conserved proline-directed phosphorylation regulates SR protein conformation and splicing function," *Biochem. J.*, 2015, doi: 10.1042/BJ20141373.

[744] V. F. Thompson *et al.*, "Transcription-Dependent Formation of Nuclear Granules Containing FUS and RNA Pol II," *Biochemistry*, 2018, doi: 10.1021/acs.biochem.8b01097.

[745] J. C. Schwartz, E. R. Podell, S. S. W. Han, J. D. Berry, K. C. Eggan, and T. R. Cech, "FUS is sequestered in nuclear aggregates in ALS patient fibroblasts," *Mol. Biol. Cell*, 2014, doi: 10.1091/mbc.E14-05-1007.

[746] A. M. Janke *et al.*, "Lysines in the RNA Polymerase II C-Terminal Domain Contribute to TAF15 Fibril Recruitment," *Biochemistry*, 2018, doi: 10.1021/acs.biochem.7b00310.

[747] M. I. Martinez-Macias *et al.*, "FUS (fused in sarcoma) is a component of the cellular response to topoisomerase I–induced DNA breakage and transcriptional stress," *Life Sci. Alliance*, 2019, doi: 10.26508/lsa.201800222.

[748] A. Bertolotti, Y. Lutz, D. J. Heard, P. Chambon, and L. Tora, "hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II," *EMBO J.*, 1996, doi: 10.1002/j.1460-2075.1996.tb00882.x.

[749] M. Watson and K. Stott, "Disordered domains in chromatin-binding proteins," *Essays in Biochemistry*. 2019, doi: 10.1042/EBC20180068.

[750] A. S. Singatulina *et al.*, "PARP-1 Activation Directs FUS to DNA Damage Sites to Form PARG-Reversible Compartments Enriched in Damaged DNA," *Cell Rep.*, 2019, doi: 10.1016/j.celrep.2019.04.031.

[751] V. Spegg and M. Altmeyer, "Biomolecular condensates at sites of DNA damage: More than just a phase," *DNA Repair (Amst).*, vol. 106, p. 103179, Oct. 2021, doi: 10.1016/J.DNAREP.2021.103179.

[752] F. Pessina *et al.*, "DNA Damage Triggers a New Phase in Neurodegeneration," *Trends in Genetics*. 2021, doi: 10.1016/j.tig.2020.09.006.

[753] X.-J. Fan *et al.*, "Original Article NONO phase separation enhances DNA damage repair by accelerating nuclear EGFR-induced DNA-PK activation," 2021.

[754] G. M. Harami *et al.*, “Phase separation by ssDNA binding protein controlled via protein-protein and protein-DNA interactions,” *Proc. Natl. Acad. Sci. U. S. A.*, 2020, doi: 10.1073/pnas.2000761117.

[755] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, “Biomolecular condensates: Organizers of cellular biochemistry,” *Nature Reviews Molecular Cell Biology*. 2017, doi: 10.1038/nrm.2017.7.

[756] F. Michelini *et al.*, “Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks,” *Nat. Cell Biol.*, 2017, doi: 10.1038/ncb3643.

[757] A. J. Price *et al.*, “CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication,” *PLoS Pathog.*, 2012, doi: 10.1371/journal.ppat.1002896.

[758] T. Fricke *et al.*, “The ability of TNPO3-depleted cells to inhibit HIV-1 infection requires CPSF6,” *Retrovirology*, 2013, doi: 10.1186/1742-4690-10-46.

[759] X. Wang *et al.*, “Cellular Cleavage and Polyadenylation Specificity Factor 6 (CPSF6) Mediates Nuclear Import of Human Bocavirus 1 NP1 Protein and Modulates Viral Capsid Protein Expression,” *J. Virol.*, 2020, doi: 10.1128/jvi.01444-19.

[760] S. Rasheedi *et al.*, “The Cleavage and polyadenylation specificity factor 6 (CPSF6) subunit of the capsid-recruited pre-messenger RNA cleavage factor I (CFIm) complex mediates HIV-1 integration into genes,” *J. Biol. Chem.*, 2016, doi: 10.1074/jbc.M116.721647.

[761] S. Jang *et al.*, “Differential role for phosphorylation in alternative polyadenylation function versus nuclear import of SR-like protein CPSF6,” *Nucleic Acids Res.*, vol. 47, no. 9, pp. 4663–4683, 2019, doi: 10.1093/nar/gkz206.

[762] J. Wang *et al.*, “A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins,” *Cell*, 2018, doi: 10.1016/j.cell.2018.06.006.

[763] K. Lasker *et al.*, “A modular platform for engineering function of natural and synthetic biomolecular condensates,” *bioRxiv*, 2021.

[764] K. Kamagata *et al.*, “Molecular principles of recruitment and dynamics of guest proteins in liquid droplets,” *Sci. Rep.*, 2021, doi: 10.1038/s41598-021-98955-0.

[765] K. M. Ruff *et al.*, “Sequence Grammar Underlying Unfolding and Phase Separation of Globular Proteins,” *SSRN Electron. J.*, 2021, doi: 10.2139/ssrn.3929009.

[766] Y. Shen *et al.*, “Biomolecular condensates undergo a generic shear-mediated liquid-to-solid transition,” *bioRxiv*, 2020, doi: 10.1101/2020.01.21.912857.

[767] S. Chong and M. Mir, “Towards Decoding the Sequence-Based Grammar Governing the Functions of Intrinsically Disordered Protein Regions,” *Journal of Molecular Biology*. 2021, doi: 10.1016/j.jmb.2020.11.023.

[768] G. Krainer *et al.*, “Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions,” *bioRxiv*, 2020, doi: 10.1101/2020.05.04.076299.

[769] K. L. Saar *et al.*, “Learning the molecular grammar of protein condensates from sequence determinants and embeddings,” *Proc. Natl. Acad. Sci. U. S. A.*, 2021, doi: 10.1073/pnas.2019053118.

[770] P. K. Jha, P. S. Desai, J. Li, and R. G. Larson, “pH and salt effects on the associative phase separation of oppositely charged polyelectrolytes,” *Polymers (Basel).*, 2014, doi: 10.3390/polym6051414.

[771] S. Alberti, A. Gladfelter, and T. Mittag, “Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates,” *Cell*, vol. 176, no. 3, pp. 419–434, 2019, doi: 10.1016/j.cell.2018.12.035.

[772] Y. Zhu *et al.*, "Molecular Mechanisms for CFIm-Mediated Regulation of mRNA Alternative Polyadenylation," *Mol. Cell*, vol. 69, no. 1, pp. 62-74.e4, 2018, doi: 10.1016/j.molcel.2017.11.031.

[773] Y. Zheng *et al.*, "Cleavage and Polyadenylation Specificity Factor 6 Is Required for Efficient HIV-1 Latency Reversal," *MBio*, 2021, doi: 10.1128/mBio.01098-21.

[774] Z. Fan *et al.*, "CDK13 cooperates with CDK12 to control global RNA polymerase II processivity," *Sci. Adv.*, 2020, doi: 10.1126/sciadv.aaz5041.

[775] T. Sasado, H. Kondoh, M. Furutani-Seiki, and K. Naruse, "Mutation in cpsf6/CFIm68 (Cleavage and Polyadenylation Specificity Factor Subunit 6) causes short 3'UTRs and disturbs gene expression in developing embryos, as revealed by an analysis of primordial germ cell migration using the medaka mutant naruto," *PLoS One*, 2017, doi: 10.1371/journal.pone.0172467.

[776] S. D. Kasowitz *et al.*, "Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development," *PLoS Genet.*, 2018, doi: 10.1371/journal.pgen.1007412.

[777] M. Das, D. S. Garlick, D. L. Greiner, and R. J. Davis, "The role of JNK in the development of hepatocellular carcinoma," *Genes Dev.*, 2011, doi: 10.1101/gad.1989311.

[778] S. Khan and H. Zaki, "Crosstalk between NLRP12 and JNK during hepatocellular carcinoma," *International Journal of Molecular Sciences*. 2020, doi: 10.3390/ijms21020496.

[779] E. Desideri and M. R. Ciriolo, "Inhibition of JNK increases the sensitivity of hepatocellular carcinoma cells to lysosomotropic drugs via LAMP2A destabilization," *Cell Death Discov.*, 2021, doi: 10.1038/s41420-021-00408-0.

[780] A. Olazabal-Herrero, A. M. Green, X. Chen, P. Sung, M. M. Pillai, and G. M. Kupfer, "Binding of FANCD2 to SRSF1 Splicing Factor Prevents Genomic Instability through R Loop Regulation," *Blood*, 2020, doi: 10.1182/blood-2020-141369.

[781] 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, pp. 412-425.e6, 2018, doi: 10.1016/j.molcel.2017.12.029.

[782] M. Jiménez *et al.*, "Splicing events in the control of genome integrity: Role of SLU7 and truncated SRSF3 proteins," *Nucleic Acids Res.*, 2019, doi: 10.1093/nar/gkz014.

[783] L. Zardoni *et al.*, "Elongating RNA polymerase II and RNA:DNA hybrids hinder fork progression and gene expression at sites of head-on replication-transcription collisions," *Nucleic Acids Res.*, vol. 1, no. 1256879, pp. 13–14, 2013, doi: 10.1093/NAR/GKAB1146.

[784] R. Kolde, "Pretty Heatmaps R package," *Version 1.0.12*, 2019. .

[785] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform Manifold Approximation and Projection," *J. Open Source Softw.*, 2018, doi: 10.21105/joss.00861.

[786] R. Navon, E. Eden, and I. Steinfeld, "GOrilla - a tool for identifying enriched GO terms," *cbl-gorilla.cs.technion.ac.il*. 2009.

[787] M. Salvi, L. Cesaro, and L. A. Pinna, "Variable contribution of protein kinases to the generation of the human phosphoproteome: A global weblogo analysis," *Biomol. Concepts*, 2010, doi: 10.1515/bmc.2010.013.

[788] N. T. Doncheva, J. H. Morris, J. Gorodkin, and L. J. Jensen, "Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data," *J. Proteome Res.*, 2019, doi: 10.1021/acs.jproteome.8b00702.

[789] D. D. Wiredja, M. Koyutürk, and M. R. Chance, "The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics," *Bioinformatics*, 2017, doi: 10.1093/bioinformatics/btx415.

[790] H. Fu, M. Zhong, and P. Tong, "The effects of KSEA interaction on the ground-state properties of spin chains in a transverse field," *Eur. Phys. J. B*, 2020, doi: 10.1140/epjb/e2020-100542-1.

[791] K. S. Metz *et al.*, "Coral: Clear and Customizable Visualization of Human Kinome Data," *Cell Syst.*, 2018, doi: 10.1016/j.cels.2018.07.001.

[792] S. Boeing *et al.*, "Multiomic Analysis of the UV-Induced DNA Damage Response," *Cell Rep.*, vol. 15, no. 7, pp. 1597–1610, May 2016, doi: 10.1016/J.CELREP.2016.04.047.

[793] E. C. Hills, "The Degree of Doctor of Philosophy," *Bull. Am. Assoc. Univ. Profr.*, 1927, doi: 10.2307/40217738.

[794] "Doctor of Philosophy," in *Definitions*, 2020.

8 Acknowledgments

Finally, you are at the last part of my thesis. Thank you very much for reading through! You may also be interested in who contributed to all this work and how I managed to be here.

I come from an impoverished and remote rural village in northern China, where there was no electricity, no running water, and no modern tools when I was born. When I was young, there were only fields when looking outside of our door. Because it is in the north of China, it is very dry. I often went with my family to the springs in the mountains to find water, and collected rain and snow to store water on rainy and snowy days. To this day, we still have a huge tank at home to store water. I saw electric lights for the first time in my primary school classroom and learned how to turn on the light. It was turned on with a rope but no button. I still remember that the rope was often broken, and our teacher needed to step on the desk to repair the light bulb or connect the rope. Although the power went out every night, it didn't stop us from studying. Writing and reading by candlelight was a "scene" every night. When I was about to graduate from high school (2007), running water finally came to our village, and I remember our village chief said, "Finally, we don't have to pick water from the valley anymore! We will be happy from now on!"

I had never left our village until going to high school, and my only source of outside information was the school books. I loved reading all books given out at school, reading them repeatedly. I got to know Beijing's Tiananmen Square, the Great Wall, and traffic lights. When I was in first grade, there was a song about traffic lights in the book: "Stop on red, go on green, wait for the yellow light!". However, in the whole village, no one had ever seen a traffic light. My mother drew one light on the ground for us, and my sisters and I spent that autumn evening imagining what we would do when the light came on. My mom's "Time to eat!" took us back to the warmth of the night.

When I was in junior high school, most of the children in our village abandoned school, either because they couldn't afford the tuition or because their parents didn't want to support them. Fortunately, my parents did not give up and still insisted on paying for my schooling. Special thanks to my class teacher (Ms. Liu), at that time she allowed me to pay the tuition six months late when the harvest was done and my family had money again. With her understanding and support, I could continue to study in school. Surprisingly, I was rewarded with 10 Yuan for my outstanding performance. This 10 Yuan reward gave me great encouragement and even influenced my study for the next years. Another special thanks need to be given to the principal at that time (Principal Huang), who waived my tuition fees for the last two years of junior high school to help me with my studies.

When I was going to high school, I left our village. Although I learned later that our county was tiny, it was so big and strange at that time. All the other students were from the city. They wore new and clean clothes, and they were polite and well-mannered. In contrast to them, I was afraid to talk to others

because of my low self-esteem and timid character. Fortunately, the school waived all my fees, considering my good grades. More fortunately, I met a respectable and loving Chinese teacher (Mr. Gu). He led me step by step through all the struggles, trained me to speak in public, and often encouraged me with high fives! He often said to me, “It’s okay! You can do it! “

My mother always told my sisters and me about the courage and independence of women that she had read about in books. She wished all of her children have the courage to stand on their own feet. However, such a strong woman was not favored by fate. The year I graduated from high school (2008), she was diagnosed with breast cancer. Thanks to the government’s support for rural health insurance and the financial support from the company of Wu Long Group, my mom had the opportunity to undergo surgery, chemotherapy, and other treatments, and I could continue my education in college. Because of my mom’s illness, I choose biology as a subject. Fortunately, my mom has recovered and has been in complete remission until now.

Finally, when college came around, I went to the big city, Beijing! I saw a traffic light for the first time and made use of the song I learned when I was 7 years old. To this day, I still sing that song silently in my mind every time I go through a traffic light. My dad accompanied me, and together we learned how to ride the elevator, the bus, the train, and many other things. Here I met my academic mentors, Dr. Professor. Zhang and Dr. Professor. Xiao. I still remember the joy and excitement when I first saw PCR bands in my gel and the first extracted mRNA precipitate from stem cells, and the satisfaction of seeing the embryonic stem cells growing happily every day. Especially in Dr. Professor Zhang, I saw a different kind of confidence and determination in her. The many experiences they told me about studying abroad deepened my courage and determination to go out and broaden my view again and again!

Finally, with the help and encouragement of everyone, I went from Nao Yao Zhuang Primary School (which is one minute away from home by walking), to Bai Di Cun Primary School (which is one hour away from home by walking), to Ying Yan Junior High School (which is two hours away from home by walking), and then to Xing Da Middle School and Yuan Qu Middle School (which are two hours away from home by bus), to Capital Normal University (which is 23 hours by train) and finally to Germany (8,000 kilometers away from home).

Although I met many difficulties during this journey, the help from others and the excitement about all the scientific results firmed my belief to insist on studying and leaving the village to see more!

This journey brought me here, writing this thesis. The beginning of these five and half years was slow and long for me, but it went faster and faster with the help of all people along the way. I really like to say: Thank you (谢谢, Dankeschön, Спасибо, Grazie, HVALA, Teşekkürler, Obrigada, Terima kasih, شكرًا)! Without you guys, your support, and your encouragement, the adventure would not have been possible.

In this more than 30-year life journey, although there have been many things that have not been perfect, many wonderful people and events have made my poor and unsatisfactory life better and better day by day. Thanks to all that has happened to me! Of note, I am lucky to grow up in China's fast-developing period, giving me all opportunities to explore and freedom to develop myself.

I still remember how fantastic it was when I read that story from "The Feynman Lectures on Physics". It was a story told by the great astronomer Arthur Eddington, who had just figured out that the stars get their power from burning hydrogen in a nuclear reaction producing helium. The story goes as follows: He recounted how, on the night after his discovery, he was sitting on a bench with his girlfriend. She said, "Look how pretty the stars shine." To which he replied,"Yes, and right now, I'm the only man in the world who knows how they shine." Although he was describing a kind of wonderful loneliness you have when you make a discovery, I was deeply moved by his achievement and the loneliness. This excitement came along with me and always encouraged me during scientific research.

Three poems always encouraged me along the journey, and I want to share them with you. First, "Let life be beautiful like summer flowers And Death like autumn leaves." ("生如夏花之绚烂，死如秋叶之静美。") from the <Stray Birds> by Rabindranath Tagore. The second one is the poem <If> from Rudyard Kipling.

Lastly, another inspirational poem <Life Is> by Mother Teresa.

Life Is – Mother Teresa

Life is an opportunity, benefit from it.
Life is beauty, admire it.
Life is a dream, realize it.
Life is a challenge, meet it.
Life is a duty, complete it.
Life is a game, play it.
Life is a promise, fulfill it.
Life is sorrow, overcome it.
Life is a song, sing it.
Life is a struggle, accept it.
Life is a tragedy, confront it.
Life is an adventure, dare it.
Life is luck, make it.
Life is too precious, do not destroy it.
Life is life, fight for it.

9 Curriculum Vitae

PhD is *doctor philosophiae* in Latin. In the context of Doctor of Philosophy, philosophy was applied with its original Greek meaning of "Love of wisdom" [793], [794]. I wish we were all equipped with love and wisdom simultaneously during our PhD time. And I wish I have shown you that I am still in love with acquiring knowledge and wisdom to apply some methods and methodology to do effective research to broaden the knowledge base.

Thank you very much again for reading through all chapters!

Tomorrow will be another good day!