



HAL
open science

Regulation of alternative splicing and polyadenylation by genotoxic anti-cancer agents

Iris Tanaka

► **To cite this version:**

Iris Tanaka. Regulation of alternative splicing and polyadenylation by genotoxic anti-cancer agents. Biochemistry, Molecular Biology. Université Paris Saclay (COMUE), 2019. English. NNT : 2019SACLS025 . tel-02050405

HAL Id: tel-02050405

<https://theses.hal.science/tel-02050405>

Submitted on 27 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Régulation de l'épissage et de la polyadénylation alternatifs par les agents anticancéreux génotoxiques

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°582

Cancérologie : biologie – médecine – santé (CBMS)
Spécialité de doctorat: aspects moléculaires et cellulaires de la biologie

Iris Tanaka

Composition du Jury :

Robin Farhaeus DR1, Université Paris-Diderot (UMR 1131)	Président
Christiane Branlant DRCE Émerite, Université de Lorraine (UMR7365)	Rapporteuse
Claudio Sette Associate Professor, Università Cattolica del Sacro Cuore	Rapporteur
Ken Olaussen MCU, Université Paris-Sud (UMR981)	Examineur
Martin Dutertre CR1, Institut Curie (UMR3348)	Directeur de thèse



With the support of :



ACKNOWLEDGMENTS - REMERCIEMENTS

First, I would like to thank Pr. Claudio Sette and Dr. Christiane Branlant for kindly accepting to review my PhD thesis manuscript, as well as Dr. Robin Fahraeus and Dr. Ken Olausen for arranging their schedules to attend my thesis defense. Thank you for agreeing to evaluate my PhD thesis work.

Je vais maintenant continuer en français. Je remercie Stéphan de m'avoir accueilli dans son équipe, et Mounira dans son unité. Merci Stéphan pour les échanges toujours constructifs que nous avons eu. Merci Mounira pour ton accueil. Je me suis sentie valorisée en tant que jeune chercheuse dans l'UMR 3348. Je remercie bien sûr Martin, mon directeur de thèse, pour m'avoir soutenue dans ce travail au cours de ces (plus de) quatre années. Tu es un encadrant exceptionnel. J'ai beaucoup aimé travailler avec toi, et les (très) longues discussions sur les projets me manqueront sûrement. J'ai énormément appris à tes côtés. En effet je n'ai pas eu le sentiment de travailler pour toi, mais avec toi. Tu m'as considéré très vite comme une chercheuse à part entière malgré mon manque d'expérience dans le domaine de l'ARN et des dommages à l'ADN, ce qui m'a beaucoup donné confiance en moi. Tu t'es souvent soucié de mon équilibre personnel et tu as aussi su me soutenir discrètement dans les moments difficiles que j'ai traversés. Merci pour tout.

Merci à toute l'équipe Vagner passée et présente, qui m'a permis d'aller au labo chaque jour avec la bonne humeur. Je pense particulièrement à Maricarmen, Michelle, Galina, Ludo qui étaient là au tout début quand je suis arrivée, et avec qui on a partagé quelques fous rires. Merci Sabrina, c'est aussi grâce à toi qu'il y a une bonne ambiance au labo, et que l'on rigole bien. Et puis bien sûr Sylvain, mon voisin de bureau quasi-communicant. Merci pour ta bonne humeur, les rigolades, et ta disponibilité. Merci à l'équipe Amor-Guéret, notamment Sandra, Rosine, Elias, et Hamza qui font partie intégrante de mon expérience dans l'UMR 3348, avec qui j'ai pu discuter science mais pas que. Merci Pascale, tu es devenue petit à petit comme une confidente, nos papotages me manqueront c'est sûr.

Mes copines de l'UMR3348, vous savez qui vous êtes, mais je vais quand même faire quelques spéciale dédicaces. Ma Rymounette, tu dois être la première personne avec qui j'ai parlé en arrivant, et je ne pensais pas trouver une amie si chère pendant ma thèse. Même si tu es Docteure et que tu n'es plus à Orsay (vive la Tunisie), je pense que nos vies sont maintenant liées, et on a encore beaucoup de choses à partager. Hélène, je papote beaucoup plus depuis que tu es dans mon bureau, pour mon plus grand plaisir. Je trouve qu'on est un peu pareil, et ce que j'aime chez toi, c'est qu'on peut parler de tout. Alina my dear, I sure wasn't expecting to meet a lifelong friend like you during my thesis. I really enjoyed going through pregnancy with you. We now have this special mommy connection

and I am so grateful. Mélissa ma chérie, tu es la meilleure voisine de bureau que j'ai eu ! On a partagé beaucoup de choses ensemble, et j'espère que ça va continuer (surtout les pâtisseries). Géraldine, ton côté sensible me touche beaucoup, et je me suis toujours sentie en confiance avec toi. J'espère construire une petite famille qui rayonne d'amour comme la tienne. Virginie, la meuf la plus fit de l'unité ! Je ne me souviens plus comment on s'est rapprochées, mais je crois que c'était dans le RER, tu es venue me taper la discut. Heureusement que tu es plus extravertie que moi, sinon j'aurais peut-être raté une super amitié.

Merci Christine et toute la famille Tencer, ma famille éloignée la plus proche ! Christine, tu sais que tu as une place toute particulière. Merci d'avoir toujours été là, dans tous les bon et mauvais moments. Je sais que je pourrai toujours compter sur toi. Merci Pépé, Mémé, Papa, et Maman. Je suis très fière d'avoir un petit bout de vous, et j'espère que je vous rends fiers aussi.

Merci beaucoup Emmanuelle et Lucien, je ne pouvais pas rêver meilleurs beaux-parents. Merci infiniment pour votre aide ces derniers mois, et votre soutien depuis plus de 5 ans maintenant. C'est grâce à vous aussi que j'ai pu écrire cette thèse.

Et le meilleur pour la fin... Merci Thomas. Quelle aventure... ! Merci de m'avoir soutenu, écouté, rassuré, encouragé toutes ces années. Merci d'avoir été là. Après avoir traversé cette thèse, rythmée par des bas très très bas, et des hauts très très hauts, je suis convaincue que l'on peut tout accomplir ensemble. Je ne pense pas qu'il existe quelqu'un qui me corresponde mieux que toi. Merci d'être un super papa, merci d'illuminer ma vie.

Et enfin ma petite Lana, mon bébé, mon petit monkey, mon précieux, merci d'avoir chamboulé tout ça. Tu as transformé la rédaction de cette thèse en vrai challenge, mais tu m'as rendue plus forte. Tes sourires me font oublier tous mes soucis, merci d'être le meilleur bébé au monde.

Table of contents

ABBREVIATIONS.....	3
PREAMBLE	5
INTRODUCTION	6
I. Alternative splicing (AS) and polyadenylation (APA) in cancer	6
1) AS mechanisms and regulation in cancer.....	6
a. Pre-mRNA splicing	6
b. AS regulation and physiological role	9
c. AS regulation in cancer.....	13
2) APA mechanisms and regulation in cancer	18
a. Pre-mRNA 3' end processing	18
b. APA regulation.....	21
c. Impact of APA.....	23
d. APA regulation in cancer	25
II. Response and resistance to genotoxic anticancer agents: focus on doxorubicin and cisplatin.....	27
1) Genotoxic anticancer agents	27
a. Genotoxic anticancer agents	27
b. DNA topoisomerase II poisons and mechanism of action.....	28
c. Platinum compounds and mechanism of action	32
2) The cell response to genotoxic stress.....	35
a. The DNA damage response (DDR).....	35
b. Repair of doxorubicin and cisplatin DNA lesions.....	38
c. Cellular responses specific to doxorubicin and cisplatin	40
3) Chemotherapy resistance mechanisms	42
a. Resistance to anticancer chemotherapy, general mechanisms	42
b. Resistance mechanisms specific to doxorubicin and cisplatin	45
III. Effects of genotoxic agents on alternative splicing and polyadenylation.....	48
1) AS in genotoxic stress response and resistance.....	48
a. AS response to genotoxic stress	48
b. AS and chemoresistance	52
2) APA in genotoxic stress response and resistance	53
a. Regulation of 3' end processing and APA by genotoxic stress.....	53
b. APA and chemoresistance	54

IV. Thesis aims.....	55
1) Regulation of alternative splicing in breast cancer cell resistance to doxorubicin	55
2) Regulation of alternative polyadenylation in non-small cell lung cancer cells response to cisplatin	55
RESULTS.....	56
I. Paper manuscript 1: “Identification of splicing programs and pathways involved in breast cancer cell resistance to doxorubicin”	56
II. Paper manuscript 2: “Intronic polyadenylation is linked to the antiproliferative effect of specific platinum compounds and generates differentially translated isoforms”	87
DISCUSSION	124
SYNTHÈSE EN FRANÇAIS.....	134
REFERENCES	137

ABBREVIATIONS

3'SS	3' splice site
5'SS	5' splice site
ABC	ATP-binding cassette
ALE	Alternative last exon
ALL	Acute lymphoblastic leukemia
AMO	Antisense morpholino
APA	Alternative polyadenylation
AS	Alternative splicing
ASO	Antisense oligonucleotide
ATM	Ataxia-telangiectasia mutated
ATR	Ataxia-telangiectasia and Rad3-related
BBP	Branch-binding protein
BER	Base excision repair
CFI	Cleavage factor I
CFII	Cleavage factor II
CPSF	Cleavage and polyadenylation specificity factor
CSC	Cancer stem cells
CstF	Cleavage stimulation factor
CTD	Carboxy-terminal domain of Pol II
DACH	Diaminocyclohexane
DDR	DNA damage response
DNA	Desoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
DSB	Double-strand break
dsDNA	Double-stranded DNA
DSE	Downstream sequence element
EMT	Epithelial-to-mesenchymal transition
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
FDA	Food and Drug Administration (USA)
hnRNP	Heterogeneous nuclear ribonucleoprotein
HR	Homologous recombination
ICL	Interstrand crosslink
IPA	Intronic polyadenylation
IR	Ionizing radiation
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
MDR	Multi-drug resistance
MMR	Mismatch repair
mRNA	messenger RNA
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NSCLC	Non-small-cell lung cancer

NTC	Nineteen complex
NTR	Nineteen-related complex
PABPC	Cytosolic poly(A) binding protein
PABPN	Nuclear poly(A) binding protein
PAP	Poly(A) polymerase
PAS	Polyadenylation signal
PDB	Protein-DNA breaks
Pol II	RNA polymerase II
pre-mRNA	pre-messenger RNA
qPCR	Quantitative polymerase chain reaction
RBP	RNA-binding protein
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
ROS	Reactive oxygen species
RPA	Replication protein A
RRM	RNA-recognition motifs
RT-PCR	Reverse transcription and polymerase chain reaction
SCM	Splice-site creating mutations
SF	Splicing factor
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SNV	Single nucleotide variant
SR protein	Serine-Arginine rich protein
SRE	Splicing regulatory element
SSB	Single-strand break
ssDNA	Single-stranded DNA
TCGA	The Cancer Genome Atlas
TDP2	Tyrosyl-DNA-phosphodiesterase 2
TIC	Transcription initiation complex
Top1	Topoisomerase I
Top2	Topoisomerase II
Top2cc	Top2-DNA covalent complex
U2AF	U2 auxiliary factor
UTR	Untranslated region

PREAMBLE

Despite the emergence of targeted therapies, “conventional” DNA-damaging (genotoxic) anti-cancer drugs are still clinically relevant. Despite their important side-effects, for some cancer types it remains the main treatment, sometimes in combination with targeted therapies, and in other cases it is the last resort option. Unfortunately, in many cases chemoresistance to chemotherapy occurs, leading to disease relapse and cancer progression. Understanding various mechanisms of response and resistance to genotoxic anti-cancer drugs is therefore crucial to overcome this major issue.

How does the cancer cell cope with damage to its DNA induced by anticancer drugs? What molecular and cellular mechanisms govern effective (cancer cell death) or ineffective (cancer cell survival) treatment? How can cancer cells that initially responded to treatment become chemoresistant? These are some interrogations researchers in many aspects of biology have been continuously investigating for decades. The DNA damage response signalling and the transcriptional response to genotoxic stress have been intensely investigated. Although our understanding of the cell response to DNA damage is already rich, each year there are new mechanisms and concepts, pathways and molecular machines interplay that are discovered. Post-transcriptional regulations such as alternative splicing and polyadenylation are emerging area of interest in cancer and in the DNA damage response.

During my thesis, I focused on two main aspects: 1) regulation of alternative splicing in breast cancer cell resistance to doxorubicin, and 2) regulation of alternative polyadenylation in response to cisplatin in non-small-cell lung cancer cells. Doxorubicin and cisplatin are two of the oldest DNA-damaging anti-cancer drugs and are still frequently used in clinical settings, but the associated alternative splicing and polyadenylation networks are poorly understood.

In the introduction of this thesis manuscript, I will first expose current knowledge on alternative splicing and polyadenylation mechanisms and regulations in health, disease and stressed conditions. In second part, I will briefly describe generalities on genotoxic anti-cancer drugs, and focus on the specificity of doxorubicin and cisplatin. I will also describe DNA repair and the DNA damage response of the cell, and mechanisms of resistance to chemotherapies, including doxorubicin and cisplatin. Finally, I will synthesize current knowledge on the relationships between alternative splicing, polyadenylation and DNA damage response as well as chemoresistance.

The result section is composed of two paper manuscripts ready for submission: 1) Identification of splicing programs and pathways involved in breast cancer cell resistance to doxorubicin, and 2) Intronic polyadenylation is linked to the antiproliferative effect of specific platinum compounds and generates differentially translated isoforms. In the discussion, I will briefly recapitulate main findings of my thesis projects and discuss possible perspectives, including long-non-coding RNA regulation by intronic polyadenylation, clinically relevant post-transcriptional networks investigation, and RNA therapeutics.

INTRODUCTION

I. Alternative splicing (AS) and polyadenylation (APA) in cancer

The life of a human messenger RNA (mRNA) starts at transcription, and ends at degradation, with multiple complex steps in between: 5' end capping, splicing, 3' end processing (cleavage and polyadenylation), export, and translation. All these steps are crucial, and they can be regulated in a physiological context and altered in diseases. For my thesis, I focused on two of these post-transcriptional steps, pre-messenger RNA (pre-mRNA) splicing and 3' end cleavage/polyadenylation, in the context of cancer and chemotherapy. Indeed, we have known for decades that gene expression is altered in cancer cells, although the complexity of post-transcriptional regulations and their clinical relevance remain partly unexplored.

1) AS mechanisms and regulation in cancer

a. *Pre-mRNA splicing*

The splicing reaction and the spliceosome. Most of human protein-coding genes contain introns that need to be removed during the maturation of the pre-mRNA, by a process called splicing. The splicing reaction is operated by a large ribonucleoprotein (RNP) machinery called the spliceosome. After the discovery of the splicing reaction in 1977 (Berget, Moore and Sharp, 1977; Chow *et al.*, 1977), the spliceosome was identified by several groups in 1985 (Brody and Abelson, 1985; Frendewey and Keller, 1985; Grabowski, Seiler and Sharp, 1985). The splicing reaction requires a large amount of ATP and occurs in two trans-esterification steps catalyzed by five small nuclear ribonucleoprotein (snRNP) complexes: U1, U2, U4, U5, and U6 (Figure 1). Each snRNP is composed of a small nuclear RNA (snRNA), seven core Sm proteins, and several other snRNP-specific core factors. Although spliceosome assembly and catalytic reactions are most often described in yeast, key structure and steps are well conserved in humans (Galej, 2018).

Briefly, U1 snRNP first binds to the 5' splice site (5'SS) junction, while splicing factor 1 (SF1, also called branch-binding protein, BBP) and U2 auxiliary factors (U2AF) recognize the branching point and the 3' splice site (3'SS), respectively (Figure 1b). Next, U2 snRNP displaces SF1 and U2AF to bind itself to the branching point. The first trans-esterification step, commonly known as branching, produces a loop-shaped lariat intermediate of the intron. This is mediated by the recruitment of pre-assembled U4/U6 and U5 snRNPs (tri-snRNP), which forms the pre-catalytic complex (B complex, first full spliceosome assembly) by associating with U1 and U2 snRNP. The spliceosomal complex then undergoes several remodeling to form a catalytically activated (B* complex) complex. During this remodeling, U1 and U4 snRNP are removed from the active complex. In addition, numerous other protein

complexes such as NineTeen (NTC) and NineTeen-related (NTR) complexes are required for spliceosomal activation (Figure 1b). The intron loop formation (the branching reaction, or catalytic step I) occurs by the branching-point adenosine attack to the 5'SS (Figure 1a). In the second step (ligation, or catalytic step II), the now free 3' end of the 5' exon is joined to the 5' end of the following exon by a second trans-esterification (Figure 1a), and the intron lariat is released. The two exons are covalently bound, and the lariat is released with U2, U5, and U6 snRNP attached to it (the intron lariat spliceosome, ILS).

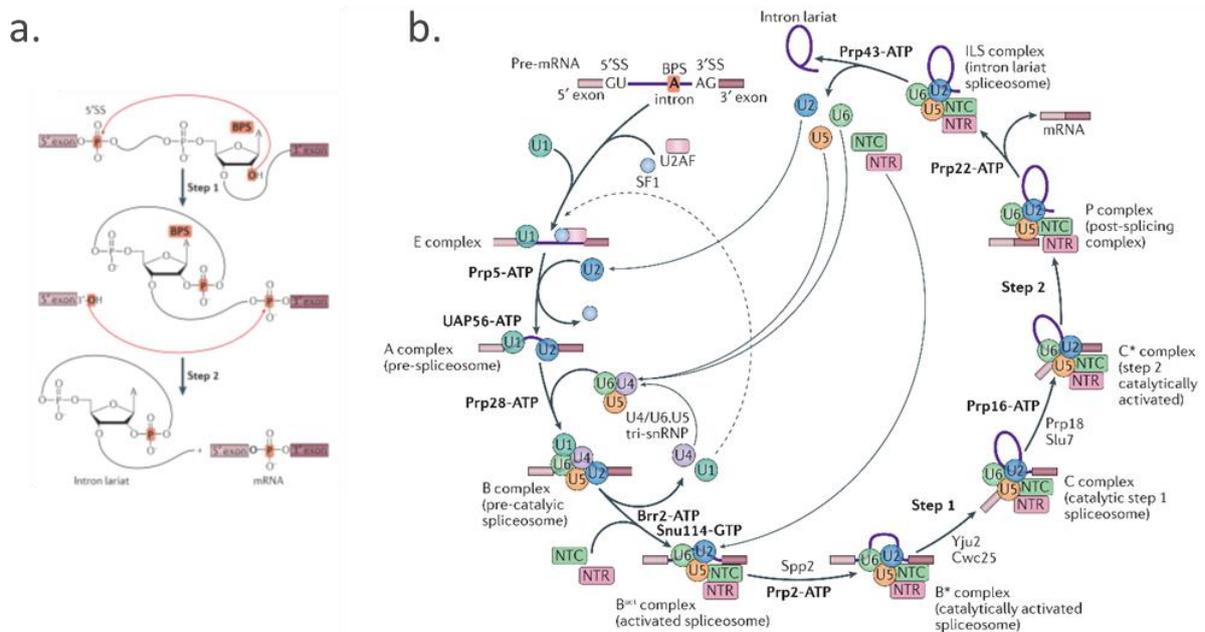


Figure 1: The splicing reaction. a) The two trans-esterification reactions of the splicing reaction. b) The splicing catalytic cycle. (adapted from Shi, 2017)

Co-transcriptional splicing. For mammalian coding genes, splicing is mostly co-transcriptional: the reaction occurs as the pre-mRNA is still attached to the chromatin, and while the RNA polymerase II (Pol II) continues transcription further downstream. In fact, transcription and pre-mRNA maturation process including splicing are mechanistically coupled, and interdependent (Bentley, 2014). It has been shown in yeast that the splicing reaction occurs almost instantly as the pre-mRNA emerges from Pol II (Carrillo Oesterreich *et al.*, 2016). Although we have known that the splicing reaction occurs simultaneously with transcription since 1988, with renowned image by Beyer and Osheim of intron loops and RNP particles on pre-mRNAs still attached to the chromatin (Figure 2), it is only recently that we begun to assess the full extent, mechanisms, interactions, and dynamics of co-transcriptional pre-mRNA processing in humans. Indeed, global analyses demonstrating functional coupling of transcription and mRNA processing in human cells are very recent, and it was made possible by the advance of sequencing techniques, including long-read and nascent RNA sequencing (Nojima *et al.*, 2015; Anvar *et al.*, 2018).

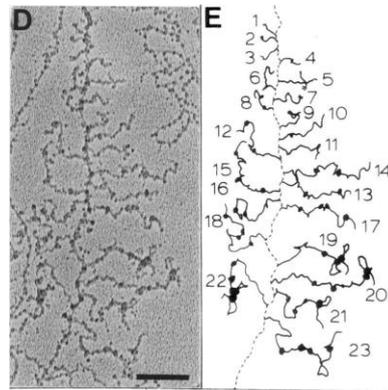


Figure 2: RNP particles at splice junctions on pre-mRNAs (adapted from Beyer and Osheim, 1988). Electron micrograph (left) and its tracing (right) of a *Drosophila* embryo transcription unit.

The carboxy-terminal domain (CTD) of the largest subunit of Pol II act as a “landing pad” to recruit mRNA processing factors to the vicinity of the transcribed pre-mRNA, and its deletion inhibited pre-mRNA processing including splicing (Bentley, 2014). The CTD dynamically orchestrate the action of processing factors throughout the transcription cycle by coordinated phosphorylation events, so-called the “CTD code” (Hsin and Manley, 2012; Bentley, 2014; Nojima *et al.*, 2015). Nojima *et al.* showed that the human CTD code was significantly different from yeast (the main model organism to study splicing), and that co-transcriptional splicing was associated with phosphorylation on the serine 5 position (S5P) of the CTD (Nojima *et al.*, 2015). Interestingly, Pol II pausing is strongly correlated to splice sites and it was suggested that U2AF65 exit from Pol II to the nascent RNA stimulates transcription elongation, but whether pausing is a cause or consequence of splicing is still unclear (Saldi *et al.*, 2016; Herzel *et al.*, 2017). Still, transcription and splicing coupling and interdependent relationship is undeniable.

Roles of introns. Half of the human genome is estimated to be composed of introns (Hubé and Francastel, 2015). Why would the sophisticated machine that is the cell first produce a longer pre-mRNA that will need to be processed by removing introns, using great amount of energy, in order to become a functional and mature mRNA? Most eukaryotes and plants possess “genes in pieces” containing introns and the spliceosome is a highly conserved machinery, suggesting key functions of introns in various organisms although their origin and evolution is still up to debate (Rogozin *et al.*, 2012). In fact, a given intron of a given gene is not necessarily never used to form mRNAs. Due to gene overlap, some introns are exonic parts of other unrelated genes. Also, intron can contain coding or non-coding “nested genes”, meaning that the entirety of a gene is comprised between two exons of another “host gene” (Kumar, 2009).

More interestingly, introns can enhance gene expression by positively affecting transcription rate, polyadenylation, mRNA export and stability, as well as translation efficiency (Le Hir, Nott and Moore, 2003). Indeed, intron-containing transcripts have higher expression profiles compared to intronless versions of the same genes. For example, splicing

signals in introns have been shown to enhance Pol II transcription initiation and elongation, notably via U1 snRNA binding to 5'SS and transcription factors, which stimulates Pol II CTD phosphorylation. Furthermore, it has been shown that some intronic microRNAs ("mirtrons" and "simtrons") biogenesis necessitate the splicing machinery, using the intron lariat as microRNA precursor (Ruby, Jan and Bartel, 2007; Havens *et al.*, 2012; Hubé and Francastel, 2015). Another non coding use of intron is the production of circular RNAs, called circular intronic RNAs (ciRNA)(Zhang *et al.*, 2013). It has also recently been shown in yeast that introns protect the genome against genetic instability by preventing transcription-associated R-loop (DNA-RNA hybrid formed by sequence complementarity) formation (Bonnet *et al.*, 2017). In this study, the authors showed that spliceosome assembly (but not splicing itself) prevented R-loop formation and that intron-rich genes and genomes were best protected against R-loop-induced genetic instability.

b. AS regulation and physiological role

The extent of AS. AS is the inclusion or exclusion of one, several, or parts of exons from a pre-mRNA to form a mature mRNA. A mechanism found from yeast to humans, AS is largely responsible for transcriptome diversity (Keren, Lev-Maor and Ast, 2010). AS is a general mechanism : more than 95% of human multi-exons genes are alternatively spliced and produce several mRNA isoforms potentially contributing to proteomic diversity (Pan *et al.*, 2008; Wang *et al.*, 2008; Mollet *et al.*, 2010). It has been estimated that in average, each gene generates around ten mRNA isoforms (Hu *et al.*, 2015). Given the discrepancy between the number of genes (< 20,000) (Ezkurdia *et al.*, 2014) and proteins (> 100,000), when large-scale studies revealed the global extent of AS in the 2000's, it was considered the main source of proteome diversity (Nilsen and Graveley, 2010).

However, this assumption was recently challenged. It has been argued that large-scale mass-spectrometry experiments found little evidence of protein isoforms resulting from AS, that the majority of the genes had one main protein isoform despite the various transcript isoforms, and that when the AS isoform was translated, the effect on protein function was limited (Tress, Abascal and Valencia, 2017). The authors concluded that most splice variants have little to no function. These conclusions were strongly questioned, arguing back that, on the contrary, numerous studies have shown the link between AS and proteomic complexity, while noting the technical limitation of proteomics studies (Blencowe, 2017). Indeed, Weatheritt *et al.* recently matched RNA-seq and ribosome-profiling datasets of human synchronized cells and concluded that at least 75% of cassette exon variants were ribosome-engaged (Weatheritt, Sterne-Weiler and Blencowe, 2016). Other groups also showed that AS alterations were reflected on the proteome, and that it diversifies protein-protein interactions (Yang *et al.*, 2016; Liu *et al.*, 2017).

Types of AS. There are five main types of AS events: alternative 5'SS and 3'SS, cassette exons, intron retention, and mutually exclusive exons (Figure 3a). Several cassette exons can also be skipped at a time, a phenomenon called multi-skipping, such as it is the case for the cell adhesion gene CD44, which has ten alternative exons between the fifth and sixth

constitutive exons (Figure 3b). Often, alternative exons differ from constitutive exons by their “weaker” splice site sequences, that diverge from the consensus “strong” constitutive splice sites. These “weaker” sites are therefore believed to be less easily recognized by the spliceosome (Fu and Ares, 2014).

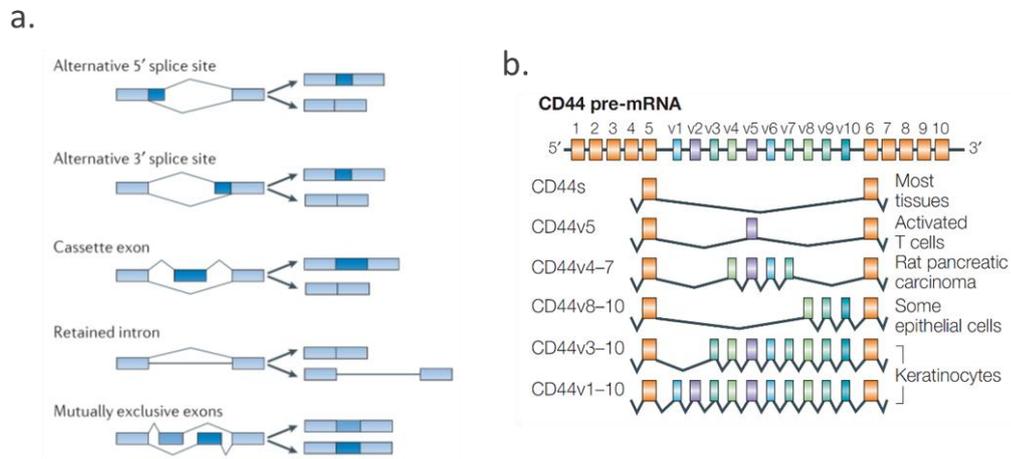


Figure 3: Types of alternative exons. a) Main types of Alternative splicing. Constitutive exons are in light blue and alternative exons in dark blue. (adapted from Dvinge *et al.*, 2016) b) CD44 pre-mRNA multi-skipping exons. (adapted from Ponta, Sherman and Herrlich, 2003)

AS regulation by RBPs. AS regulation is highly contextual and done by the coordination of *cis*-elements (enhancing and silencing sequences on the pre-mRNA) and *trans*-acting factors, which are RNA-binding proteins (RBP) called splicing factors (SF). There are four categories of *cis*-elements, collectively called splicing regulatory elements (SRE): intronic splicing silencer (ISS), exonic splicing silencer (ESS), intronic splicing enhancer (ISE), and exonic splicing enhancer (ESE). SRE tend to have the strongest effect when in close proximity of a splice site (Dvinge *et al.*, 2016). Also, splice sites are quite loosely defined, allowing opportunities of alternative use. Two well-characterized, and often ubiquitously expressed, categories of SF are heterogeneous nuclear RNP (hnRNP) and serine-arginine rich proteins (SR). SR protein usually have one or two RNA recognition motif domains (RRM), which is responsible for the ability to bind RNA, and a carboxy-terminal arginine/serine rich domain used for protein-protein interactions. hnRNP proteins also contain RRM, as well as other functional domains (Singh and Valcárcel, 2005; Dvinge *et al.*, 2016).

As shown in Figure 4, SR proteins are usually presented as splicing activators binding to enhancer sequences (promoting inclusion), and hnRNPs as splicing inhibitors binding to silencer sequences (promoting exclusion) (Kornblihtt *et al.*, 2013). Although this is true in many cases, there is increasing evidence showing the versatility of hnRNP and SR proteins. In fact, depending on the context, it is not rare for SF to have ambivalent, or even antagonistic roles (Fu and Ares, 2014). For example, it has been shown that some hnRNP proteins can also promote the inclusion of alternative exons (Xue *et al.*, 2009; Huelga *et al.*, 2012). Moreover, Pandit *et al.* found that *in vivo*, both SRSF1 and SRSF2 (two SR proteins) equally

promoted inclusion and exclusion of target exons using mouse knockout models and the CLIP-seq technique (UV crosslinking and immunoprecipitation, followed by sequencing) (Pandit *et al.*, 2013). In this study, depletion of one of the SR proteins resulted in binding of other SR proteins, suggesting a compensatory or synergistic relationship between SFs. There are many other RBPs with splicing functions, some of which are tissue-specific and described in Table 1.

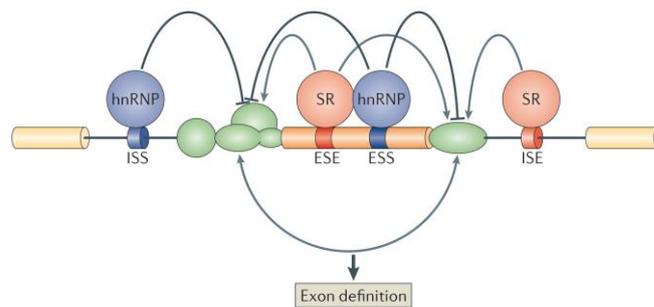


Figure 4: Alternative splicing regulatory elements. Constitutive exons are in light orange, the alternative exon is in dark orange, core spliceosome components are in green. (adapted from Kornblihtt *et al.*, 2013)

Great effort is made to decipher rules of SF binding to RNA in a sequence specific manner, and how it affects the usage of exons (a “splicing code”), but it has been proven extremely complex. It is in part due to the fact that SFs can be redundant, cooperative or competitive, that they can include or exclude an exon depending on where they bind (on exons or introns, near or far from splice site) or which other SF are present, and that their expression can vary from one cell type to another (Fu and Ares, 2014). Complicating the picture, many SFs (for example RBM24 or hnRNP A1) are multi-tasking, participating in gene expression regulation at other levels than AS (mRNA stability, export, translation) (Jean-Philippe, Paz and Caputi, 2013; Lin *et al.*, 2018).

Other mechanisms of AS regulation. In addition to regulatory SF, it has been suggested that the splicing machinery itself can regulate AS too. There are several reports of AS alteration on specific genes after depletion of core spliceosomal proteins such as U2AF35, SF3B1, and PRP8 (from U5 snRNP) (Chen and Manley, 2009; Dvinge, 2018). In addition, several recent large-scale analysis pointed out an extensive role of core spliceosomal components in AS regulation in mammalian cells (Papasaikas *et al.*, 2015; Tejedor, Papasaikas and Valcárcel, 2015; Han *et al.*, 2017). Surprisingly, in Han *et al.* high-throughput study of the effect of trans-acting factors on AS event, it was revealed that a large proportion of AS was mediated by transcription factors (Han *et al.*, 2017). This finding reveals unknown before actors of AS regulation, and versatility of typically DNA-binding proteins to RNA-binding.

The rate of Pol II elongation is also an important regulator of AS, a phenomenon termed “kinetic coupling” model between transcription and splicing. This model proposes that slow elongation promotes the inclusion of alternative exons with weak SRE by allowing a wider

“window of opportunity” for their recognition, and vice-versa (Kornblihtt *et al.*, 2013; Bentley, 2014). There are although exceptions to this general rule of thumb. For example, it has been shown that slow elongation enhanced cystic fibrosis transmembrane conductance regulator (CFTR) exon 9 skipping by favouring the association of a negative SF (Dujardin *et al.*, 2014). Therefore, Pol II elongation speed seems to open wider or narrower the window of opportunity to regulators positively or negatively affecting alternative exon usage. Indeed, a recent study on the genome-wide impact of elongation rate on AS showed that both slow and fast transcription altered inclusion and exclusion of exons (Fong *et al.*, 2014). Interestingly, some exons were observed to be more “rate-sensitive” than others. This study suggested that most exons need an optimal and specific speed for their incorporation. In addition, pre-mRNA secondary structure, as well as epigenetic marks such as chromatin state (histone marks and nucleosome positioning) or DNA methylation have also been reported to affect AS, but the understanding of the mechanisms of these regulations is still emerging (Luco *et al.*, 2011; McManus and Graveley, 2011; Wan *et al.*, 2014; Chen, 2015; Lev Maor, Yearim and Ast, 2015; Saldi *et al.*, 2016; Dvinge, 2018).

Physiological AS. AS has various functions in the healthy cell. For example, AS is believed to correlate with organismal complexity. During the genomic era, it has been discovered that genome size only poorly correlates with phenotypic complexity. However, the prevalence of AS is greater in higher eukaryotes than in lower eukaryote, and in vertebrates than in invertebrates (Harrington *et al.*, 2004; Alekseyenko, Kim and Lee, 2007; Kim, Magen and Ast, 2007; Keren, Lev-Maor and Ast, 2010). It has also been shown in humans that splicing patterns can vary among individuals without causing illness, majorly contributing the high diversity of our phenotypic features (Hull *et al.*, 2007; Park *et al.*, 2018). Furthermore, Braunschweig *et al.* showed in 2014 that intron retention is more prevalent than previously thought in mammals, and that it contributes to transcriptome “fine-tuning” by reducing levels of transcripts irrelevant to the cell of tissue type in which they are detected (Braunschweig *et al.*, 2014).

Another key role of AS in healthy cells is tissue specificity and differentiation. Global analyses have shown that at least 50% of alternative transcripts are differently expressed among tissues, and that specific tissues can be differentiated by their splicing patterns (Yeo *et al.*, 2004; Wang *et al.*, 2008). One of the main reasons for this is tissue-specific expression of splicing factors, as shown in Table 1. The brain is one of the most diverse tissue of the organism and have the highest occurrence of tissue-specific AS (Yeo *et al.*, 2004; Li, Lee and Black, 2007). AS has been proven fundamental for neuronal development and differentiation (Vuong, Black and Zheng, 2016). Notorious brain-specific SFs are PTB and NOVA proteins. An expression switch from PTB to nPTB was revealed fundamental for programming neuronal differentiation, and NOVA2 regulates AS isoforms of genes with synaptic functions (Ule *et al.*, 2005; Boutz *et al.*, 2007). There are also muscle-specific AS patterns, and the SF RBM24 have been shown *in vivo* to be crucial for embryonic cardiac development and skeletal muscle differentiation (Grifone *et al.*, 2014; Yang *et al.*, 2014). Cell-type-specificity of ESRP1 and ESRP2 for epithelial cell-specific splicing programs, and their role for epithelial-to-mesenchymal transition (EMT) have been known for almost a decade, and their functional

relevance was recently demonstrated in knockout mice epidermis, whose development was severely affected (Warzecha *et al.*, 2009; Bebee *et al.*, 2015). Importantly, analysis of the ENCODE Project Consortium datasets revealed that cell-type-specific splicing patterns due to cell-type-specific SF such as ESRP1/2 were maintained independently of tissue of origin (Mallinoud *et al.*, 2014).

Table 1: Tissue-specific splicing factors. (adapted from Chen and Manley, 2009)

Name	Other names	Binding domain	Binding motif	Tissue expression	Target genes
nPTB	brPTB and PTBP2	RRM	CUCUCU	Neurons, myoblasts and testes	<i>BIN1</i> , <i>GLYRA2</i> , <i>ATP2B1</i> , <i>MEF2</i> , <i>NASP</i> , <i>SPAG9</i> and <i>SRC</i>
NOVA1	NA	KH	YCAAY	Neurons of the hindbrain and spinal cord	<i>GABRG2</i> , <i>GLYRA2</i> and <i>NOVA1</i>
NOVA2	NA	KH	YCAAY	Neurons of the cortex, hippocampus and dorsal spinal cord	<i>KCNJ</i> , <i>APLP2</i> , <i>GPHN</i> , <i>JNK2</i> , <i>NEO</i> , <i>GRIN1</i> and <i>PLCB4</i>
FOX1	A2BP1	RRM	(U)GCAUG	Muscle, heart and neurons	<i>ACTN</i> , <i>EWSR1</i> , <i>FGFR2</i> , <i>FN1</i> and <i>SRC</i>
FOX2	RBM9	RRM	(U)GCAUG	Muscle, heart and neurons	<i>EWS</i> , <i>FGFR2</i> , <i>FN1</i> and <i>SRC</i>
RBM35a	ESRP1	RRM	GU rich	Epithelial cells	<i>FGFR2</i> , <i>CD44</i> , <i>CTNND1</i> and <i>ENAH</i>
RBM35b	ESRP2	RRM	GU rich	Epithelial cells	<i>FGFR2</i> , <i>CD44</i> , <i>CTNND1</i> and <i>ENAH</i>
TIA1	mTIA1	RRM	U rich	Brain, spleen and testes	<i>MYPT1</i> , <i>CD95</i> , <i>CALCA</i> , <i>FGFR2</i> , <i>TIAR</i> , <i>IL8</i> , <i>VEGF</i> , <i>NF1</i> and <i>COL2A1</i>
TIAR	TIAL1 and mTIAR	RRM	U rich	Brain, spleen, lung, liver and testes	<i>TIA1</i> , <i>CALCA</i> , <i>TIAR</i> , <i>NF1</i> and <i>CD95</i>
SLM2	KHDRBS3 and TSTAR	KH	UAAA	Brain, tests and heart	<i>CD44</i> and <i>VEGFA</i>
Quaking	QK and QKL	KH	ACUAAAY[...]JUAAY	Brain	<i>MAG</i> and <i>PLP</i>
HUB	HUC, HUD and ELAV2	RRM	AU rich	Neurons	<i>CALCA</i> , <i>CD95</i> and <i>NF1</i>
MBNL	NA	CCCH zinc finger domain	YGCU(U/G)Y	Muscles, uterus and ovaries	<i>TNNT2</i> , <i>INSR</i> , <i>CLCN1</i> and <i>TNNT3</i>
CELF1	BRUNOL2	RRM	U and G rich	Brain	<i>TNNT2</i> and <i>INSR</i>
ETR3	CELF2 and BRUNOL3	RRM	U and G rich	Heart, skeletal muscle and brain	<i>TNNT2</i> , <i>TAU</i> and <i>COX2</i>
CELF4	BRUNOL4	RRM	U and G rich	Muscle	<i>MTMR1</i> and <i>TNNT2</i>
CELF5	BRUNOL5 and NAPOR	RRM	U and G rich	Heart, skeletal muscle and brain	<i>ACTN</i> , <i>TNNT2</i> and <i>GRIN1</i>
CELF6	BRUNOL6	RRM	U and G rich	Kidney, brain and testes	<i>TNNT2</i>

A2BP1, ataxin 2-binding protein 1; *ACTN*, α -actinin; *APLP2*, amyloid- β precursor-like protein 2; *ATP2B1*, ATPase, Ca^{2+} transporting, plasma membrane 1; *BIN1*, bridging integrator 1; *CALCA*, calcitonin-related polypeptide- α ; *CELF*, CUGBP- and ETR3-like factor; *CLCN1*, chloride channel 1; *COL2A1*, collagen, type II, $\alpha 1$; *COX2*, cytochrome c oxidase II; *CTNND1*, catenin $\delta 1$; *EWSR1*, Ewing sarcoma breakpoint region 1; *FGFR2*, fibroblast growth factor receptor 2; *FN1*, fibronectin 1; *GABRG2*, GABA A receptor, $\gamma 2$; *GLYRA2*, glycine receptor, $\alpha 2$ subunit; *GPHN*, gephyrin; *GRIN1*, glutamate receptor, ionotropic, NMDA 3B; *IL8*, interleukin-8; *INSR*, insulin receptor; *JNK2*, Jun N-terminal kinase 2; *KCNJ*, potassium inwardly-rectifying channel, subfamily; *KHDRBS3*, KH domain-containing, RNA-binding, signal transduction-associated protein 3; *MAG*, myelin associated glycoprotein; *MBNL*, muscleblind; *MEF2*, myocyte enhancing factor 2; *MTMR1*, myotubularin-related protein 1; *NASP*, nuclear autoantigenic sperm protein; *NEO*, neogenin; *NF1*, neurofibromin 1; *NOVA*, neuro-oncological ventral antigen; *PLCB4*, phospholipase C $\beta 4$; *PLP*, proteolipid protein; *PTB*, polypyrimidine-tract binding protein; *RBM*, RNA-binding protein; *RRM*, RNA recognition motif; *SLM2*, SAM68-like mammalian protein 2; *SPAG9*, sperm associated antigen 9; *TIA1*, T cell-restricted intracellular antigen 1; *TIAR*, TIA1-related protein; *TNNT2*, troponin T type 2; *VEGF*, vascular endothelial growth factor.

c. AS regulation in cancer

Global alteration of AS. It is well established today that AS is widely altered in numerous diseases including cancer (David and Manley, 2010; Cieply and Carstens, 2015; Daguinet, Dujardin and Valcarcel, 2015; Dvinge *et al.*, 2016). In cancer cells, AS is now known to be altered in a genome-wide manner. In fact, AS in cancer is so important that general AS deregulation is now considered as a “novel hallmark of cancer”, and every common hallmark of cancer, including metastasis, are affected by mis-splicing (Ladomery, 2013; Oltean and Bates, 2014; Pagliarini, Naro and Sette, 2015). Furthermore, a recent large-scale analysis by Climente-González *et al.*, shed light on the functional impact of global AS alteration in cancers (Climente-González *et al.*, 2017). They showed that AS deregulation in cancer (or Cancer-associated splicing changes, CASC) often impacted functional protein domains

frequently mutated in tumours, affected protein-protein interaction in cancer pathways, and proposed that CASC could be oncogenic drivers on their own.

In 2009, Venables *et al.* demonstrated global AS alteration in breast and ovarian cancer compared to normal tissue by high-throughput RT-PCR (Venables *et al.*, 2009). It was one of the first demonstration of the large extent of AS in cancer. In this study, the authors showed that breast and ovarian cancer shared a common master-regulator of AS: the RBP RBFOX2. Numerous other genome-wide studies confirmed cancer-specific global alteration of AS in most cancer types, including leukemias, lung, colorectal, and brain cancers (Germann *et al.*, 2012; Biamonti *et al.*, 2014; Danan-Gotthold *et al.*, 2015; Sveen *et al.*, 2016; Song *et al.*, 2017). Not only cancer and normal tissues can be distinguished by their AS profile, it has also recently been demonstrated using data from The Cancer Genome Atlas (TCGA) project that distinct cancer types could also be separated by their AS pattern specificity (Sebestyén, Zawisza and Eyras, 2015; Tsai *et al.*, 2015; Y. Li *et al.*, 2017). In one study, breast cancer seems to be particularly affected by AS alteration, and AS profiles were able to distinguish the common molecular subtypes (triple negative, non-triple negative and HER2-positive cancers, or basal and luminal) (Sveen *et al.*, 2016), whereas in another, breast cancer had the lowest number of AS events in the 33 cancer types analyzed (Y. Li *et al.*, 2017). Interestingly, widespread intron retention (a rather less studied type of AS) have also recently been shown to be cancer-type-specific and important in tumor-suppressor inactivation (Dvinge and Bradley, 2015; Jung *et al.*, 2015).

Oncogenic AS isoforms. There is a plethora of oncogenic AS isoforms described in the literature. As noted above, every common hallmark of cancer is affected by AS mis-regulation. One of the most notorious and earliest example of oncogenic AS variant are the Bcl-x isoforms (Boise *et al.*, 1993). Interestingly, the two isoform of Bcl-x, Bcl-xL and Bcl-xS, generated by an alternative 5'SS at exon 2, have antagonistic functions. Bcl-xL is anti-apoptotic and Bcl-xS is pro-apoptotic. Bcl-xL/Bcl-xS ratios are higher in many cancer cells compared to normal tissue. Since then, it has been shown that this AS of Bcl-x is regulated by numerous SFs, including Sam 68 (Paronetto *et al.*, 2007), RBM4 (Wang *et al.*, 2014), hnRNPH/F (Garneau *et al.*, 2005), hnRNPK (Revil *et al.*, 2009), SRSF2 (Merdzhanova *et al.*, 2008) and SRSF10 (Shkreta *et al.*, 2016) to name a few. Below is a table summarizing some of the known oncogenic AS isoforms and their function in cancer biology:

Table 2: Validated examples of oncogenic AS variants.

Alternatively spliced gene	Type of AS	exon	Function	Reference
Bcl-x	5'SS	2	apoptosis (Bcl-xL: anti-apoptotic, Bcl-xS: pro-apoptotic)	Boise <i>et al.</i> , 1993
Caspase-2	cassette exon	9	apoptosis (casp-2L: pro-apoptotic, casp-2S: anti-apoptotic)	Wang <i>et al.</i> , 1994
Fas	cassette exon	6	apoptosis (Δ exon6: soluble decoy receptor, anti-apoptotic)	Cheng <i>et al.</i> , 1994
K-RAS	mutually exclusive exons	4a or 4b	apoptosis (K-RAS4A: pro-apoptotic, K-RAS4B: anti-apoptotic)	Plowman <i>et al.</i> , 2006
SYK	cassette exon	9	cell survival (SYK(L): cell growth, survival, SYK(S): apoptosis)	Prinos <i>et al.</i> , 2011
EGFR	cassette exon	4	cell proliferation (Δ exon4: constitutively active receptor)	Wang <i>et al.</i> , 2011
Ron	cassette exon	11	cell invasion (Δ exon11: constitutively active receptor)	Collesi <i>et al.</i> , 1996
CD44	multiple skipping	v1-v10 (between exons 5 and 6)	EMT, metastasis (v4-7 and v6-7)	Günthert <i>et al.</i> , 1991
FGFR2	mutually exclusive exons	3b or 3c	EMT (3b: epithelial, 3c: mesenchymal)	Yan <i>et al.</i> , 1993
Rac1	cassette exon	3b	cell migration, EMT (3b inclusion: constitutively active GTPase signaling)	Jordan <i>et al.</i> , 1999
VEGF-A	3'SS	8	angiogenesis (VEGF-A165: angiogenic, VEGF-A165b: anti-angiogenic)	Bates <i>et al.</i> , 2002
Fibronectin	cassette exon	EDA (between exon 11 and 12) or EDB (between exon 7 and 8)	angiogenesis (EDA: normal vessels, EDB: oncofetal)	Castellani <i>et al.</i> , 1994
MDM2	multiple skipping	4-11	negative regulator of p53 (deletion of p53 binding domain: p53 accumulation)	Sigalas <i>et al.</i> , 1996
BRCA1	cassette exon and 5'SS	11	DNA repair, damage response (ambiguous functions: promoting or repressing growth depending on context)	Tammaro <i>et al.</i> , 2012
PKM	mutually exclusive exons	9 (PKM1) or 10 (PKM2)	glucose metabolism (PKM1: normal differentiated tissues, PKM2: cancer cells, aerobic glycolysis)	Christofk <i>et al.</i> , 2008
BRAF (V600E)	multiple skipping	4-8	drug resistance (deletion of RAS-binding domain)	Poulikakos <i>et al.</i> , 2011

SFs as oncogenes. In addition of AS isoforms with oncogenic potential, SF can sometimes act as oncogenes or tumour-suppressors when deregulated or mutated (Dvinge *et al.*, 2016). An analysis of 261 SF in 21 cancer types showed cancer-specific SF expression, and in general, upregulation (Sveen *et al.*, 2016). The role of the SR protein SRSF1 in cancer has been particularly well-documented. It has been shown back in 2007 that its expression is upregulated by MYC in various tumours, including colon, thyroid, small intestine, kidney and lung (Karni *et al.*, 2007). Also, it was reported that DNA copy-number gain and mRNA overexpression of SRSF1 was correlated with poor survival in a cohort of Chinese small cell lung cancer patients, and its knockdown completely abrogated tumour growth in mice xenografts (Jiang *et al.*, 2016).

As mentioned above, mutations in SF can also be oncogenic. SRSF2, SF3B1, U2AF1, and ZRSR2 are highly mutated in leukemias (Yoshida *et al.*, 2011). Interestingly, all four of these SF mainly target 3'SS and are part of the core spliceosome. SRSF2, SF3B1, and U2AF1 are part of the U2 snRNP complex of the spliceosome responsible of 3'SS recognition. SF3B1 is also reported to be significantly mutated in breast cancer (Ellis *et al.*, 2012) and as much as 20% of uveal melanomas (Harbour *et al.*, 2013; Martin *et al.*, 2013), and is associated with disease progression and poor survival in chronic lymphocytic leukemia patients (Martin *et al.*, 2013). Downregulation or mutation of SF3B1 have been associated with hundreds of AS changes (Dolatshad *et al.*, 2015; L. Wang *et al.*, 2016). The mechanism by which SF3B1 dysregulate AS has been proposed recently by Alsafadi *et al.* using the SF3B1^{R625/K666} mutant. SF3B1^{R625/K666} is a change-of-function mutant and promotes upstream 3'SS by alternative use of branch point usage (Alsafadi *et al.*, 2016). In 2018, a group of researchers from TCGA Research Network analyzed whole-exome sequencing across 33 tumour types and revealed 119 mutated splicing factors (Seiler *et al.*, 2018). Here again, they confirmed high occurrence of recurrent, non-silent hotspot mutations in SRSF2, SF3B1, and U2AF1. In general, mutations in SF are now increasingly considered as cancer drivers (Supek *et al.*, 2014; Hahn *et al.*, 2015; Climente-González *et al.*, 2017; Seiler *et al.*, 2018).

Cis-mutations affecting AS. It has been estimated that mutations in *cis* that disrupt normal splicing could account for a third of all disease causing mutations, therefore they are expected to play a substantial role in cancer as well (Daguenet, Dujardin and Valcarcel, 2015). For example, splice site mutations in the TP53 gene resulting in aberrant transcripts have been known for more than 20 years (Holmila *et al.*, 2003; Leroy, Anderson and Soussi, 2014). In 2014, Supek *et al.* analyzed a subset of cancer-related genes in >3000 cancer exomes and >300 cancer genomes, and showed that synonymous mutations in exonic SREs in oncogenes were associated with abnormal splicing (Supek *et al.*, 2014). In this study, they estimated that half of synonymous drivers alter splicing, shedding light on how silent mutations, originally thought not to alter protein functions, can “make some noise” in human diseases, and especially cancer by activating oncogenes (Supek *et al.*, 2014; Zheng, Kim and Verhaak, 2014). In the study from Jung *et al.* where they showed widespread intron retention as a mechanism of tumour-suppressor inactivation, they actually found such events by analyzing RNA-seq and exome data of 1812 cancer patients for single nucleotide variants (SNV) and identified nearly 900 SNV (including synonymous mutations) causing impaired splicing (Jung

et al., 2015). More recently, Jayasinghe *et al.* analyzed splice-site-creating-mutations (SCM) in the 33 cancer types from the TCGA and discovered nearly 2000 mis-annotated SCMs. They also found recurrent SCMs in known proto-oncogenes such as TP53, BRCA1, GATA3 and PTEN.

Cancer-associated splicing networks. Similarly to transcriptional or signaling networks associated with cancer, AS pathways/networks can contribute to oncogenesis, and their impact is increasingly appreciated. Indeed, it has been demonstrated that oncogenic SF control oncogenic AS event, thus contributing to oncogenesis. As mentioned above, SRSF1 is one of the most well described oncogenic SF. Some of the known oncogenic targets of SRSF1 are listed in Table 3. In addition, a study in breast organotypic three-dimensional culture identified by RNA-seq and validated by RT-PCR hundreds of SRSF1 targets, implicated in cancer-relevant processes, such as cell cycle, proliferation, cell death and survival (Anczuków *et al.*, 2015). It has also been shown that the transcription factor MYC directly activated SRSF1 expression in lung cancer, leading to AS of signaling kinase MKNK2 and transcription factor TEAD1 (Das *et al.*, 2012). In fact MYC, which is overexpressed in many cancers, is the upstream regulator of several other oncogenic SF, such as hnRNPA1 and PTB1 (Cobbold *et al.*, 2010; David *et al.*, 2010; Rauch *et al.*, 2011). In a recent study, MYC was shown to contribute to lymphomagenesis by upregulating key snRNP factors of the core spliceosome, thus maintaining correct splicing and promoting cell survival and proliferation (Koh *et al.*, 2015). The function of SRSF2 mutation in leukemia has also been suggested recently. In blood cells, mutated SRSF2 did not result in a loss of function but in AS misregulation, which led to abnormal splicing of hematopoietic regulators such as EZH2, resulting in impaired hematopoietic differentiation and myelodysplastic syndrome (Kim *et al.*, 2015). Moreover, a systematic analysis of mutation, copy number, and gene expression of 1348 RBP in 11 tumour types recently revealed widespread alteration of RBP expression associated with AS variation in cancer driver genes, and shed light on cancer-associated splicing networks (Sebestyén, Zawisza and Eyras, 2015).

Table 3: Oncogenic splicing targets of SRSF1. (adapted from Das and Krainer, 2014)

Target	Function	SRSF1-induced splicing change	Property of SRSF1-induced isoform
<i>BCL2L1</i>	Apoptosis regulator	Enhanced inclusion of full-length exon 2	Antiapoptotic
<i>BCL2L11</i>	Apoptosis regulator	Increased inclusion of novel alternative 3' exon	Antiapoptotic
<i>BIN1</i>	Apoptosis regulator	Increased exon 12A inclusion	Antiapoptotic
<i>CASP2</i>	Apoptosis regulator	Skipping of a novel 61-bp exon	Proapoptotic
<i>CASP9</i>	Effector of apoptosis	Increased inclusion of cassette exons 3–6	Proapoptotic
<i>CCND1</i>	Cell-cycle regulator	Intron 4 retention	Cell-cycle regulator; pro-oncogenic
<i>CD3ζ</i>	Signaling transducer of T-cell receptor	Suppressed alternative splicing within 3'-UTR	Codes for functional CD3 ζ protein of T-cell receptor
<i>CD44</i>	Cell-cell interaction, cell adhesion, and migration	Increased exon v9 inclusion	Epithelial-specific isoform
<i>DFFA</i>	Inhibitor of caspase-activated DNase	Intron 5 exclusion	CAD inhibitor, no chaperone activity
<i>ENG</i>	Part of TGF β receptor complex; important in angiogenesis	Retention of terminal intron	Promotes endothelial senescence
<i>FN1</i>	Structural component of extracellular matrix, integrin signaling	Inclusion of EDA exon	Promotes cell invasion
<i>MAPT</i>	Stabilization of microtubules	Increased exon 10 inclusion	Microtubule stabilization
<i>MKNK2</i>	Effector in MAPK signaling pathway	Increased inclusion of mutually exclusive exon 13b	Activates p38a-MAPK; pro-oncogenic
<i>MST1R</i>	Receptor for macrophage stimulating protein	Skipping of exon 11	Promotes metastasis
<i>RPS6KB1</i>	Kinase of ribosomal protein S6	Inclusion of novel exons 6a and 6c	Activates mTORC1; pro-oncogenic
<i>SLC39A14</i>	Ion carrier, Zn transport in prostate cells	Increased inclusion of mutually exclusive exon 4b	Biomarker for colorectal cancer
<i>TEAD1</i>	Transcriptional enhancer	Increased exon 5 inclusion	Unknown

2) APA mechanisms and regulation in cancer

a. Pre-mRNA 3' end processing

Pre-mRNA 3' end processing. Pre-mRNA 3'-end processing consists of cleavage at the 3'-end followed by addition of a poly(A) tail composed of approximately 200 adenosine residues. Sites of cleavage are encoded in the DNA sequence of the gene, and for the vast majority it is between the highly conserved polyadenylation signal (PAS) AAUAAA and a downstream sequence element (DSE), usually U or GU rich (Proudfoot, Furger and Dye, 2002). These *cis*-elements are recognized by two protein complexes, cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulation factor (CstF), respectively. The second most common PAS is AUUAAA, and together with AAUAAA accounted for about 60% of all polyadenylation signals (Beaudoing *et al.*, 2000; Gruber *et al.*, 2016). Two factors, cleavage factor I (CFI) and cleavage factor II (CFII), are known to be required for the cleavage reaction, but it is the endonuclease CPSF73 of the CPSF complex that catalyzes the reaction. Next, the cleaved pre-mRNA is polyadenylated by the poly(A) polymerase (PAP). The newly synthesized poly(A) tail is stabilized by the binding of nuclear poly(A) binding proteins (PABPN), which will be replaced by its cytosolic counterpart (PABPC) after exportation from the nucleus (Figure 5).

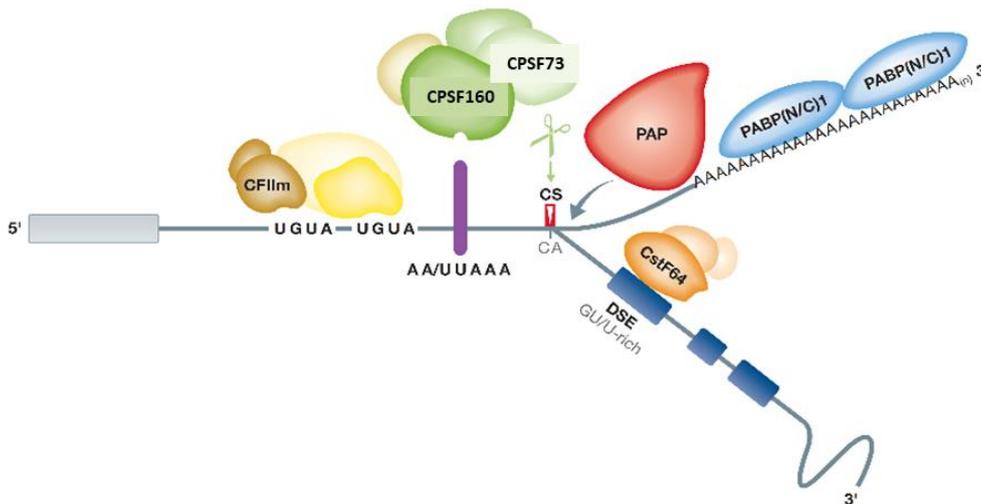


Figure 5 : The pre-mRNA 3'-end processing machinery (Hollerer *et al.*, 2014)

Regulation of 3' end processing. As every other step of gene expression, pre-mRNA 3' end processing can be regulated, and mostly concern PAS recognition and 3' end processing complex assembly (as opposed to regulation after assembly). RBPs, whether they are core 3' end processing factors or other regulatory RBPs, mediate 3' end processing regulation by variation in their RNA binding ability. For example, PTB (a hnRNP) can compete with CstF64 for binding to the DSE, resulting in decreased 3' end processing efficiency and mRNA levels when upregulated (Castelo-Branco *et al.*, 2004). Other RBPs such as HuR and sex-lethal (SXL) also compete with CstF64 (Millevoi and Vagner, 2009). Whether 3' end processing will be efficient depends on the relative amount of the competing RBPs. Unsurprisingly, basal levels

of core 3' end processing factors also affect cleavage and polyadenylation efficiency. When CstF64 is upregulated during B-cell differentiation, it is able to recognize suboptimal polyadenylation site upstream of the canonical site. This allows notably the expression switch of IgM from a membrane-bound to a secreted form (Takagaki *et al.*, 1996). Additionally, post-translational modifications of RBP can affect 3' end processing efficiency. For example, hyperphosphorylation of PAP during mitosis represses its activity and leads to a general inhibition of polyadenylation, a regulation important in normal cell growth (Colgan *et al.*, 1996; Zhao and Manley, 1998). Other post-translational modifications enhance 3' end processing. Sumoylation of CPSF73 and symplekin (another core 3' end processing factor) have been shown to stimulate 3' end processing complex formation (Vethantham, Rao and Manley, 2007). The above mentioned and other mechanisms (some of which are discussed below) of 3' end processing regulation are illustrated in Figure 6.

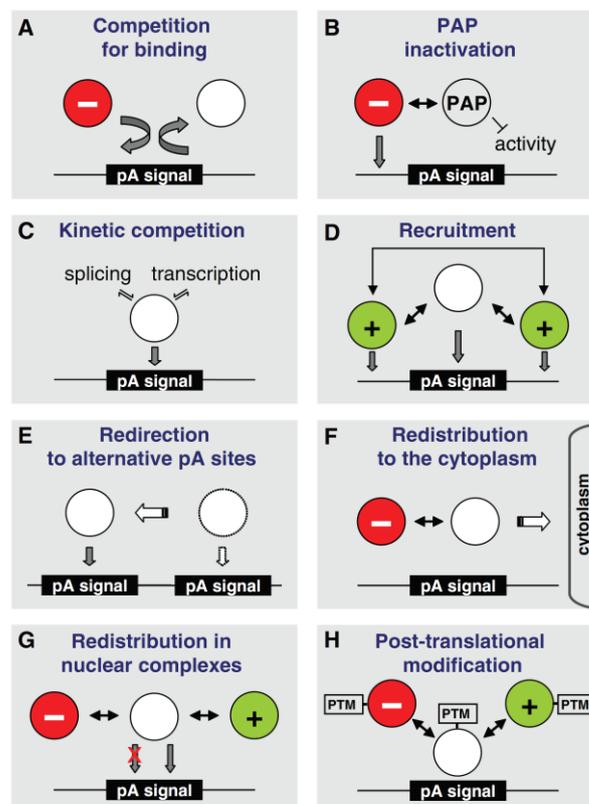


Figure 6 : Mechanisms of 3' end processing regulation. (adapted from Millevoi and Vagner, 2009)

Coupling of 3' end processing and transcription termination. Transcription termination of protein coding genes, resulting in the release of the pre-mRNA from the Pol II and the DNA template, is one of the least understood steps of gene expression. Molecular and mechanistic details on where and how to stop transcription elongation is still a matter of active investigation (Porrúa and Libri, 2015). However, it is known that transcription termination and 3' end processing are intricately linked. Two modes of transcription termination are currently described. In the first one, often referred as the “torpedo model”, transcription termination is dependent on the PAS and triggered by CPSF, CstF, and CFI/II as shown in Figure 7 (Kuehner, Pearson and Moore, 2011). When Pol II transcribes past the PAS,

CPSF bound to the body of Pol II recognizes the PAS on the nascent transcript, which induces Pol II pausing. Upon exposure of the DSE, CstF dislodges CPSF and this remodeling triggers cleavage by CFI/II and release of the pre-mRNA from the Pol II and the DNA. The pre-mRNA is polyadenylated, while the 5'-3' exoribonuclease XRN2 degrades the downstream cleavage product, "chasing" the Pol II by tethering to the RNA exit channel (Kuehner, Pearson and Moore, 2011). This "chasing" is believed to eventually promote Pol II dissociation from the chromatin. In addition, recent *in vitro* studies of HeLa cells extracts demonstrated that the PAS is sufficient and cleavage is not required, whereas conformational changes of the transcription elongation complex was necessary to stimulate termination (Zhang, Rigo and Martinson, 2015). The second mode of termination is an allosteric mechanism dependent on the CFII component PCF11 (Zhang, Fu and Gilmour, 2005). In this model, it is PCF11 that triggers the Pol II dissociation by binding to the Pol II CTD and the nascent pre-mRNA. In this study by Zhang *et al.* in yeast, PCF11 was suggested to form a "bridge" that would induce conformational changes leading to Pol II dissociation. Which mode of termination is used in a given context is not fully understood, and termination is likely to occur via a combination of both modes, as reported by Luo *et al.*'s allosteric/torpedo model where PCF11 recruits the 5'-3' exoribonucleases (Luo, Johnson and Bentley, 2006).

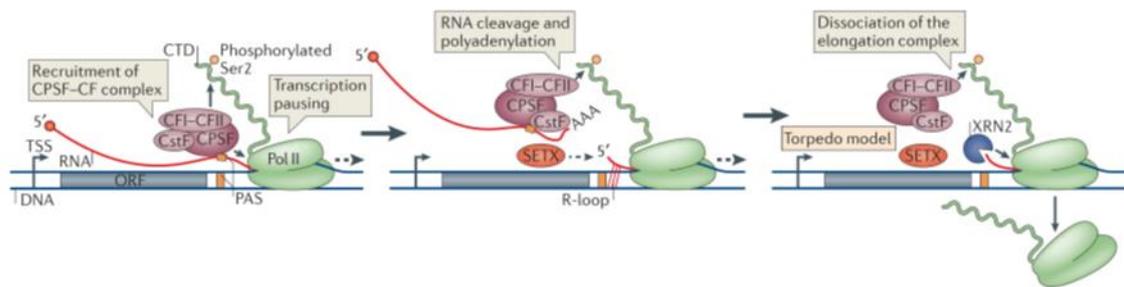


Figure 7: "Torpedo model" of transcription termination and 3' end processing coupling. (adapted from Porrua and Libri, 2015)

Coupling of 3' end processing and splicing. All the pre-mRNA maturation steps are now believed to be coupled and interdependent (Maniatis and Reed, 2002). In this regard, the 3' end processing machinery and the splicing machinery have a particularly special relationship. Splicing of the last intron and 3' end processing of the last exon are functionally coupled. Indeed, the two machineries have a reciprocal regulation: PAS recognition facilitates splicing of the last intron, and recognition of the last intron facilitates cleavage and polyadenylation (Niwa, Rose and Berget, 1990; Niwa and Berget, 1991). Several spliceosome components that play major roles at 3'SS, such as U2AF65 (from U2 snRNP), promote cleavage and polyadenylation by directly binding to 3' end processing core components. At the last exon, U2 snRNP interacts with 3' end processing machinery components instead of U1 snRNP at the 3' end of the exon for exon definition. For example, it has been shown by Millevoi *et al.* that U2AF65 stimulates cleavage and polyadenylation by binding to CFI and recruiting it to the PAS (Millevoi *et al.*, 2002, 2006). U2AF65 is also known to bind PAP and CSTF to regulate 3' end processing (Ko and Gunderson, 2002; Kyburz *et al.*, 2006). U1 snRNP is also involved in 3' end processing, as discussed below. Furthermore, PAP can stimulate

splicing of the last intron through stabilization of U2AF65, leading to spliceosome recruitment (Vagner, Vagner and Mattaj, 2000).

b. APA regulation

Types of APA. APA is the alternative use of non-canonical polyadenylation sites, generating mRNA isoforms that differ at their 3'-end. One of the first evidence of APA was the identification on the immunoglobulin heavy chain, generating both the membrane-bound and the secreted isoform from the same IgM gene in 1980 (Alt *et al.*, 1980). Initially thought as a marginal phenomenon, in 2005 Tian *et al.* showed that at least 32% of human genes were alternatively cleaved (Tian *et al.*, 2005). It is now believed that at least 70% of mammal genes have APA isoforms (Derti *et al.*, 2012; Shi, 2012; Hoque *et al.*, 2013). There are mainly two categories of APA : those which only affect the 3'UTR and those which affect the coding region as well as the 3'UTR (Figure 8). APA that modulate the coding region can be separated into two main types: intronic polyadenylation (IPA) and exonic polyadenylation. Exonic polyadenylation usually generates mRNAs without a stop-codon and are rapidly degraded by the non-stop decay pathway, and will not be discussed in this manuscript (Vasudevan, Peltz and Wilusz, 2002). Last exons of IPA transcripts are also referred to as alternative last exons (ALE). There is currently no consensus nomenclature on the use of IPA, ALE, or other terms to name truncated transcripts generated by APA. IPA and ALE are sometimes used differently because of their differences shown in Figure 8, but are essentially the same. In both case, the truncated transcript is generated by the use of an intronic polyadenylation site, leading to the use of an ALE. In this manuscript, both the terms IPA and ALE will be used and are largely interchangeable. ALEs are an interesting type of exons, at the intersection of AS and APA, where an alternative exon and a corresponding alternative polyadenylation site are used.

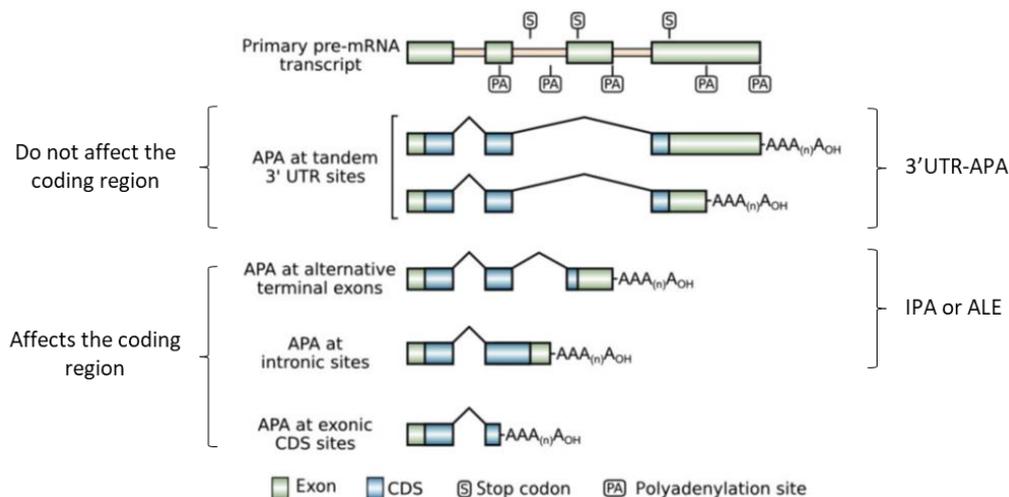


Figure 8: Types of APA. Alternative terminal exons are also called alternative last exons (ALE). (adapted from Gruber, Martin, Keller, *et al.*, 2014)

Regulation of APA by RBPs. Similarly to SF favouring or repressing the use of SS, RBPs also control the use of different polyadenylation sites at the 3'UTR or within the coding region. In fact, many core 3' end processing factors were found to regulate APA as well. It was first demonstrated for CstF64, whose upregulation during B cell maturation results in higher levels of the CSTF complex, and increased usage of the proximal polyadenylation site (IPA) in the IgM heavy chain transcript, associated with the secreted isoform (Takagaki *et al.*, 1996). This role of CstF64 in proximal polyadenylation site usage was also confirmed in cancer cells, where it was described as a “master regulator” of 3'UTR shortening (Xia *et al.*, 2014). PCF11 (from the CF complex) and Fip1 (from the CPSF complex) were also shown to globally promote proximal PAS in 3'UTRs (Lackford *et al.*, 2014; Li *et al.*, 2015). Conversely, two studies using genome-wide approaches described CFI, specifically CFIm25 and CFIm68, as 3'UTR-APA regulators promoting distal polyadenylation site usage (Martin *et al.*, 2012; Li *et al.*, 2015). Other studies confirmed this global trend in CFIm25 knockdown followed by RNA-seq experiments, which increased proximal polyadenylation site selection in coding and 3'UTR regions, in human fibroblasts (Mitra *et al.*, 2018) and glioblastoma tumours (Masamha *et al.*, 2014). The transcription factor E2F involved in cell cycle regulation was shown to regulate expression of 3' end processing genes (Elkon *et al.*, 2012). Furthermore, 3' end processing factors involved in the poly(A) tail formation and maintenance were also unexpectedly found to regulate APA. In 2012, two independent studies showed that knockdown of PAPBPN1, induced global 3'UTR shortening in mouse muscle (De Klerk *et al.*, 2012) and U2OS cells (Jenal *et al.*, 2012). Recently, different PAP isoforms have also been shown to differentially regulate 3'UTR-APA as well as IPA (W. Li *et al.*, 2017). Additionally, other RBPs that are not nuclear 3' end processing factors, such as CPEB1 (Bava *et al.*, 2013), hnRNPC (Gruber *et al.*, 2016), hnRNPK (Yoon *et al.*, 2013), NOVA (Licatalosi *et al.*, 2008) or HuR (Dai, Zhang and Makeyev, 2012; Mansfield and Keene, 2012) were also shown to be involved in APA regulation in various contexts.

IPA repression by U1 snRNP. U1 snRNP role in 3' end processing was already speculated in the early 1990s (Wassarman and Steitz, 1993). One of the first evidence of U1 snRNP role in 3' end processing regulation was its autoregulation by binding of two U1A proteins to U1A pre-mRNA 3'UTR and PAP C-terminal domain, inhibiting its polyadenylation in a negative feedback loop (Boelens *et al.*, 1993; van Gelder *et al.*, 1993; Gunderson *et al.*, 1997; Vagner *et al.*, 2000). More recently, the role of U1 snRNP in 3' end processing regulation was further detailed thanks to high-throughput sequencing technologies. When U1 snRNP interaction with 5'SS was knocked down using an antisense morpholino oligo (AMO), in addition to the expected accumulation of unspliced pre-mRNA, numerous pre-mRNA failed to complete full-length transcription and were prematurely cleaved and polyadenylated, within 5 kilobases from the transcription start site (Kaida *et al.*, 2010). This was not due to an inhibition of splicing, as U2 snRNP did not recapitulate this phenomenon, and importantly, these short pre-mRNAs were correctly processed at their 3' end. It was demonstrated in this study that U1 snRNP protected the transcriptome from premature cleavage and polyadenylation by binding to cryptic 5'SS upstream of PAS along the pre-mRNA (Kaida *et al.*, 2010). In addition, the same group showed that moderate depletion of functional U1 snRNP (no enough to inhibit splicing) resulted in a dose-dependent proximal last exon shifting, as shown in Figure

9 (Berg *et al.*, 2012). Interestingly, Vorlová *et al.* showed that depletion of U1 snRNP induced IPA, resulting in the expression of truncated mRNAs coding for dominant-negative soluble decoy versions of numerous receptor tyrosine-kinases such as EGFR and VEGFR (Vorlová *et al.*, 2011). Thus, U1 snRNP has important extra-splicing functions: it is crucial in IPA repression and gene isoform expression regulation.

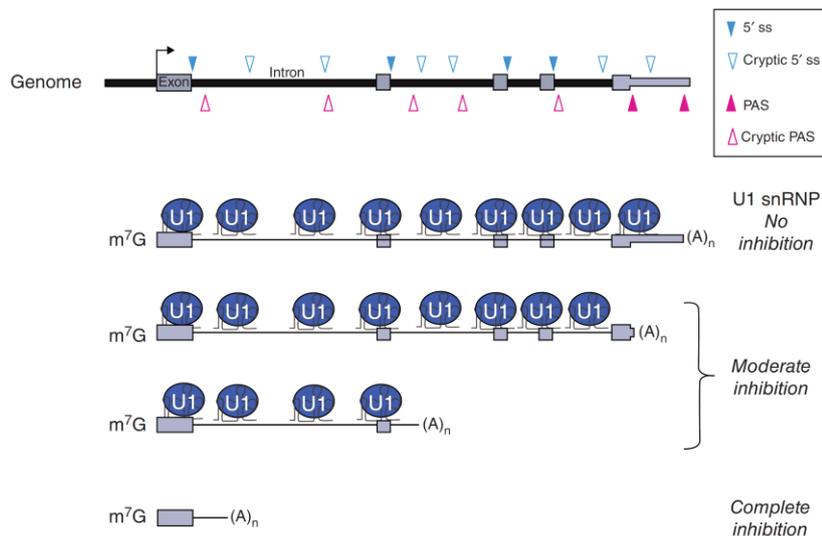


Figure 9: U1 snRNP expression level-dependent protection of intronic PAS. (adapted from Kaida, 2016)

c. Impact of APA

Impact of 3'UTR-APA. APA only affecting the 3'UTR, also called tandem APA, shorten or lengthen the 3'UTR and therefore its “targetability” by RBPs, microRNAs, or other *trans*-factors (Figure 10a). The most studied 3'UTR-APA regulation is 3'UTR shortening. As 3'UTR contain regulatory sequences involved in various RNA metabolism mechanisms, the stability, export, translation, and cellular localization of the mRNA, can be greatly affected by 3'UTR-APA. For example, as shown in Figure 10b, cancer cells expressing shorter 3'UTR on cyclin D2 and IMP-1 lacked let-7 and/or miR-15/16 binding sites and were more proliferative and oncogenic, respectively, than those which expressed the longer and canonical 3'UTR (Mayr and Bartel, 2009). In this study by Mayr and Bartel, it was also shown that the mRNAs with shorter 3'UTR were also generally more efficiently translated by escaping microRNA translation repression.

This 3'UTR shortening effect on translation efficiency was also confirmed by several groups (Sterne-Weiler *et al.*, 2013; Masamha *et al.*, 2014; Chang *et al.*, 2015; Floor and Doudna, 2016; Hoffman *et al.*, 2016). One groundbreaking report was of Sandberg *et al.* in 2008 when they discovered that general 3'UTR shortening during T-cell activation led to increased protein expression (Sandberg *et al.*, 2008). However, in some cases the shorter 3'UTR isoform is less translated (Pinto *et al.*, 2011), and the impact of 3'UTR shortening on

translation efficiency is still debated. Indeed, some studies showed at most limited effect, or even the opposite: increased translation efficiency or protein abundance with longer 3'UTR (Spies, Burge and Bartel, 2013; Gruber, Martin, Müller, *et al.*, 2014). Indeed, the relationship between 3'UTR length, microRNA binding sites, and translation efficiency seem to be more complex than initially appreciated. Fu *et al.* recently demonstrated that 3'UTR shortening is associated with better translational efficiency by escaping microRNA repression in six cell lines. They did confirm that in the NIH3T3 cell line used in Spies *et al.* study, transcripts with longer 3'UTR tended to be associated with polysomes meaning increased translation efficiency, but demonstrated that this was due to contact inhibition and senescence. Additionally, Hoffman *et al.* showed that not only the presence or absence of microRNA binding sites, but also their relative distance to APA sites is important for gene expression regulation (Hoffman *et al.*, 2016). They found that “stronger”, conserved microRNA binding sites were enriched immediately upstream of APA sites. In this case, 3'UTR shortening is associated with stronger microRNA repression. In proliferating cells, 3'UTR shortening therefore not only increase expression of pro-proliferating genes, but also represses it for anti-proliferation genes by this mechanism (Hoffman *et al.*, 2016). Finally, 3'UTR also modulates mRNA stability, their localization, cell type specific expression, and was even recently shown to be able to act as a scaffold for protein-protein interactions (Mayr, 2016).

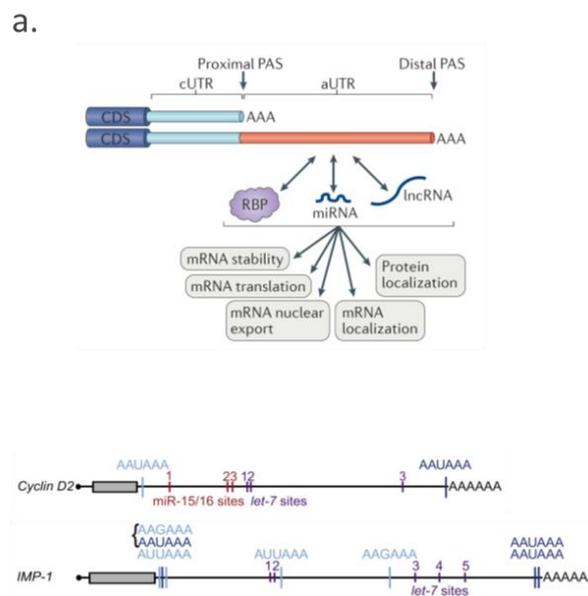


Figure 10: 3'UTR-APA. a) Impact of 3'UTR-APA regulation. (adapted from Tian and Manley, 2016) b) PAS and microRNA targeting sites of Cyclin D2 and IMP-1 genes. (adapted from Mayr and Bartel, 2009)

Impact of IPA. Since ALE and IPA also affect the coding region, we can expect ALE or IPA isoforms to have distinct functions, or no function at all, from the canonical isoform (Figure 11). The example of IgM cited above are actually ALE isoforms with distinct functions. As mentioned above, Vorlová *et al.* identified dominant negative receptor tyrosine kinases acting as soluble decoy receptor as a result of IPA (Vorlová *et al.*, 2011). In mouse muscle,

IPA of platelet-derived growth factor receptor α (PDGFR α) produced a truncated isoform also acting as a soluble decoy receptor and preventing fibrosis (Mueller *et al.*, 2016). Recently, Singh *et al.* analyzed 46 3'-seq and RNA-seq datasets from human normal tissues and various cell types and identified widespread IPA isoforms expressed in specific tissues or cell types, especially in immune cells (Singh *et al.*, 2018). Furthermore, it has also been shown with the example of the ASCC3 gene, that an IPA isoform can in fact be a long-non-coding RNA (lncRNA) with antagonistic functions to the full length mRNA (Williamson *et al.*, 2017).

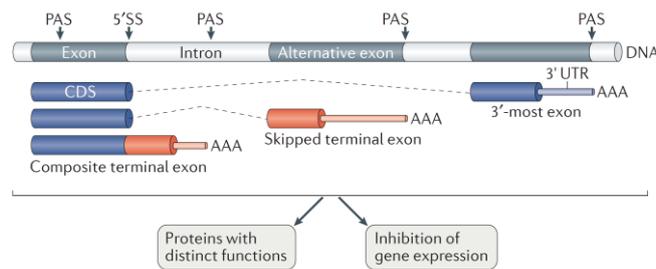


Figure 11: Impact of IPA/ALE. (adapted from Tian and Manley, 2016)

d. APA regulation in cancer

Regulation of 3'UTR length in cancer. One of the major characteristics of cancer cells is their unlimited, high proliferative state (Hanahan and Weinberg, 2011). Global 3'UTR shortening was associated with proliferating cells (Sandberg *et al.*, 2008; Elkon *et al.*, 2012; Beisang, Reilly and Bohjanen, 2014; Hoffman *et al.*, 2016; Mitra *et al.*, 2018). Therefore, it would be expected that 3'UTR shortening, or APA modulation in general would be associated with cancer as well. In 2009, Mayr and Bartel showed widespread shortening of 3'UTR in cancer cell lines compared to non-transformed cell lines, resulting in oncogene activation. This was the first evidence of global APA modulation in cancer cells. Following this study, others have shown consistent results (Bava *et al.*, 2013; Masamha *et al.*, 2014; Xia *et al.*, 2014). In 2014, Masamha *et al.* showed that CFIm25 from the 3'end processing machinery repressed proximal poly(A) sites, and when its expression was low, cells had shorter 3'UTR and more tumorigenic properties (Masamha *et al.*, 2014). The same year, Xia *et al.* developed a bioinformatic algorithm to identify APA from RNA-seq datasets. They could identify global APA in cancer versus normal cells pairs from the TCGA, mostly consisting of 3'UTR shortening (Xia *et al.*, 2014). Here again, a protein from the 3'-end processing machinery (CstF64) was shown to be a master regulator of 3'UTR shortening.

Another interesting role of 3'UTR shortening is on stem cell differentiation and renewal, embryonic cellular processes frequently hijacked by cancer cells. In 2009, Ji and Tian observed 3'UTR shortening during somatic cell reprogramming into induced pluripotent cells (Ji and Tian, 2009). Another study showed that Fip1, a 3'end processing factor interacting with CPSF, promotes somatic cell reprogramming and embryonic stem cell self-renewal by favouring 3'UTR shortening in the majority of its target genes, which are enriched in stem cell related functions (Lackford *et al.*, 2014). In line with this, Brumbaugh *et al.* also recently

showed CFIm25 role in 3'UTR modulation in stem cell reprogramming and further linked it to chromatin regulation (Brumbaugh *et al.*, 2018).

3'UTR-APA regulation in cancer is not always about 3'UTR shortening. For example, it has been shown that compared to non-transformed breast cell line MCF10A, estrogen receptor positive MCF-7 and triple-negative MDA-MB231 breast cancer cell lines displayed shortened and lengthened 3'UTRs, respectively (Fu *et al.*, 2011). In a colorectal cancer development study using patient samples, Morris *et al.* showed that depending on the functions of the genes, 3'UTR shortening or lengthening was observed (Morris *et al.*, 2012). A recent study pointed out that previous studies assumed that 3'UTR-APA concerned two polyadenylation sites, a proximal and a distal one. In fact, it is clear today that most human genes with several poly(A) sites have at least three, rendering the "proximal" and "distal" terms inappropriate. By developing a better RNA-seq algorithm, they showed that a vast amount of tandem APA events could not simply be defined by 3'UTR shortening or lengthening, adding to the complexity of these regulations (Xue *et al.*, 2018).

IPAs/ALEs in cancer. Over the past ten years, most global studies of APA regulation relationship to cancer studies were focused on 3'UTR-APA, and few studies looked at the other type of APA altering the coding region. However, some ALE isoforms of oncogenes have been known for decades, but were initially described as AS variants. It is the case for cyclin D1 (Betticher *et al.*, 1995), caspase 8 (Himeji *et al.*, 2002), and p53 (Arai *et al.*, 1986) for example. More recently, Ni and Kupperwasser described a truncated MAGI3 isoform generated by IPA, which acts as a dominant-negative of the full-length isoform and promotes tumorigenesis of breast cancer (Ni and Kuperwasser, 2016). Two studies have also noted widespread IPA regulation in proliferation, with a trend for transcript shortening in one case (Elkon *et al.*, 2012), but not in the other case (Sandberg *et al.*, 2008). In 2014, my supervisor showed that ALEs were the main type of exons induced by doxorubicin treatment in the breast cancer cell line MCF-7 (Dutertre, Chakrama, *et al.*, 2014). Finally, two very recent studies from the same group showed widespread IPA induction in blood diseases. Lee *et al.* showed widespread upregulation of truncated mRNAs by IPA, predominantly tumour-suppressor genes, in chronic lymphocytic leukemia (Lee *et al.*, 2018), and Singh *et al.* showed widespread repression of physiologic immune cell-specific IPA isoforms in multiple myelomas (Singh *et al.*, 2018).

II. Response and resistance to genotoxic anticancer agents: focus on doxorubicin and cisplatin

Despite the promising emergence of targeted therapies during the last decade, “conventional” genotoxic anticancer agents such as doxorubicin and cisplatin are still widely used in the clinic. Often, cancer patients are treated with both targeted therapies and genotoxic chemotherapy, in an effort to combine therapeutic strategies. One of the main limitations of cancer therapy, genotoxic or not, is the development of resistance resulting in tumour recurrence, metastasis or even death. It is therefore important to understand mechanisms of response and resistance to various compounds, in order to be able in the future to offer personalized, better treatment options to cancer patients. In this part, I will especially focus on two widely used types of genotoxic anticancer agents: the anthracycline and DNA topoisomerase II poison doxorubicin, and the platinum compound cisplatin.

1) Genotoxic anticancer agents

a. Genotoxic anticancer agents

History of DNA-damaging agents in cancer therapy. First “successful” treatment of cancer begun in the 1940’s, with the discovery of nitrogen mustards for non-Hodgkin’s lymphoma patients (Cheung-Ong, Giaever and Nislow, 2013). Louis Goodman and Alfred Gilman from Yale School of Medicine observed that soldiers who died from sulphur mustard gas exposure during World War I had severe lymphoid hypoplasia and myelosuppression. After promising mice experiments (Gilman, Philips and Hedgpeth, 1946), they convinced their collaborator and thoracic surgeon Gustav Lindskog to inject a related molecule, nitrogen mustard (or β -chloroethyl amine), in patients with advanced non-Hodgkin’s lymphoma (Chabner and Roberts, 2005). Supporting their predictions, tumours regressed. Even though remission lasted only a few weeks, this was the first evidence of cancer management with chemotherapy. Around the same time, Sidney Farber at Harvard Medical School was investigating the effects of folic acid on leukemia patients. When he found that folic acid increased acute lymphoblastic leukemia (ALL) cells proliferation, he developed folate analogues (folic acid antagonists) and treated children with ALL (Chabner and Roberts, 2005). This led to success in decreasing cancerous cells, but yet again, remissions were brief (Farber *et al.*, 1948). A decade later, in the 1950’s, it was identified that nitrogen mustard and anti-folate were both effective due to their DNA-damaging properties, by alkylation of DNA and inhibition of DNA synthesis, respectively. These findings and other successful use of genotoxic agents in leukemia prompted researchers to actively investigate on DNA-damaging agents for cancer therapy (Cheung-Ong, Giaever and Nislow, 2013). In 1955, the National Cancer Institute (NCI) in the United-States made a bold move to create the National Cancer Chemotherapy Service Center (NCCSC) and a drug screening programme (Chabner and Roberts, 2005). Until the 1960’s medical oncologists and standardized chemotherapy didn’t exist, and were even frowned upon, until first successful treatments of childhood

leukemias and Hodgkin's disease by combination chemotherapy (DeVita and Chu, 2008). Since the 1970's genotoxic anticancer drugs are widely used and in constant development, which participated in the decrease of cancer mortality in the United-States and Europe in the 1990's (DeVita and Chu, 2008; Bosetti *et al.*, 2013).

Exploiting genomic instability in cancer therapy. Genotoxic agents are effective at killing cancer cells compared to normal cells because of their rapid cell cycle and their genomic instability. Cancer cells usually have a less efficient DNA repair system, so damaged DNA is often not properly repaired and ignored, while the cell continues to cycle rapidly. Cancer cells then accumulate DNA lesions and mutations at high rate, eventually leading to cell death (Cheung-Ong, Giaever and Nislow, 2013). Derivatives of alkylating agents and inhibitors of DNA synthesis discovered in the 1940's are still used today to treat several types of cancers (Jones, 2016). For example, cyclophosphamide (an alkylating agent) is one of the most commonly used anticancer agent, most of the time in combination with others. Cyclophosphamide and doxorubicin are often used together to treat breast or blood cancers. Other DNA-damaging anticancer agents are platinum drugs and topoisomerase inhibitors, which will be discussed in detail in the following sections. Below is a table recapitulating most common type of genotoxic anticancer agents:

Table 3: Major categories of genotoxic anticancer agents. ALL=Acute Lymphoblastic Leukemia, AML=Acute Myeloid Leukemia, CLL=Chronic Lymphoblastic Leukemia, CML=Chronic Lymphoblastic Leukemia, HL=Hodgkin's Lymphoma, NHL=Non-Hodgkin's Lymphoma. (source: <https://www.cancer.gov/about-cancer/treatment/drugs>, NIH-NCI)

Category	Mode of action	Exemples	Used to treat
Alkylating agents	Covalently bind to DNA bases	Cyclophosphamide	Breast cancer, Ovarian cancer, ALL, AML, CLL, CML, HL, NHL, Neuroblastoma
		Mitomycin C	Gastric and pancreatic adenocarcinoma
		Nitrosourea	HL, NHL, Brain tumours
		Bendamustine	NHL, CLL
		Temozolomide	Glioblastoma
Platinum drugs	Form intrastrand and interstrand crosslinks on DNA	Cisplatin	Testicular cancer, Ovarian cancer, Cervical cancer, Bladder cancer, Gastric cancer, Head and neck cancer, Lung cancer, Neuroblastoma, Brain tumours
		Carboplatin	Ovarian cancer, Neuroblastoma, Brain tumours, Lung cancer, Bladder cancer
		Oxaliplatin	Colorectal cancer
Antimetabolite	Analogues of DNA and RNA synthesis metabolites, impair DNA synthesis	Methotrexate	Breast cancer, ALL, NHL, Lung cancer, Osteosarcoma
		5-Fluorouracil	Breast cancer, Colorectal cancer, Gastric cancer, Pancreatic cancer
		Gemcitabine	Breast cancer, Lung cancer, Ovarian cancer, Pancreatic cancer
		Mercaptopurine	ALL
		Hydroxyurea	CML, Head and neck cancer
Topoisomerase inhibitors/poisons	Inhibit DNA topoisomerases enzymatic activity	Topotecan	Ovarian cancer, Cervical cancer, Lung cancer
		Irinotecan	Colorectal cancer
		Doxorubicin	Breast cancer, Ovarian cancer, ALL, AML, NHL, Neuroblastoma, Lung cancer, Sarcomas, Thyroid cancer
		Etoposide	Lung cancer, Testicular cancer
Radiomimetic drugs	Induce DSB	Bleomycin	HL, NHL, Cervical cancer, Testicular cancer, Head and neck cancer

b. DNA topoisomerase II poisons and mechanism of action

Human DNA topoisomerases. DNA topoisomerases are universal key enzymes that relax topological constraints on double-stranded DNA (dsDNA) during transcription and replication. During transcription and replication, dsDNA needs to be separated to allow passing of the

RNA or DNA polymerases. This opening of the strands inevitably results in DNA supercoiling in the flanking regions, that needs to be uncoiled to avoid stalling of the polymerases or the formation of abnormal structures (Pommier *et al.*, 2010). Humans have six topoisomerases classified in three types: type IA, type IB, and type IIA (Figure 12a). Type I topoisomerases break one strand of DNA to allow passage of the other strand (IA) (Figure 12b), or to let the downstream coiled dsDNA rotate (IB), and reseal the break (Figure 12c). Type II topoisomerases create double-strand breaks (DSB) of the DNA and then actively pass another dsDNA through the break before re-ligating both strands (Figure 12d). Moreover, type IIA topoisomerases are separated into two isoenzymes, Top2 α and Top2 β . Top2 α is linked to cell proliferation and increases during G2/M phases, whereas Top2 β is expressed ubiquitously, including in non-dividing cells (Pommier *et al.*, 2010).

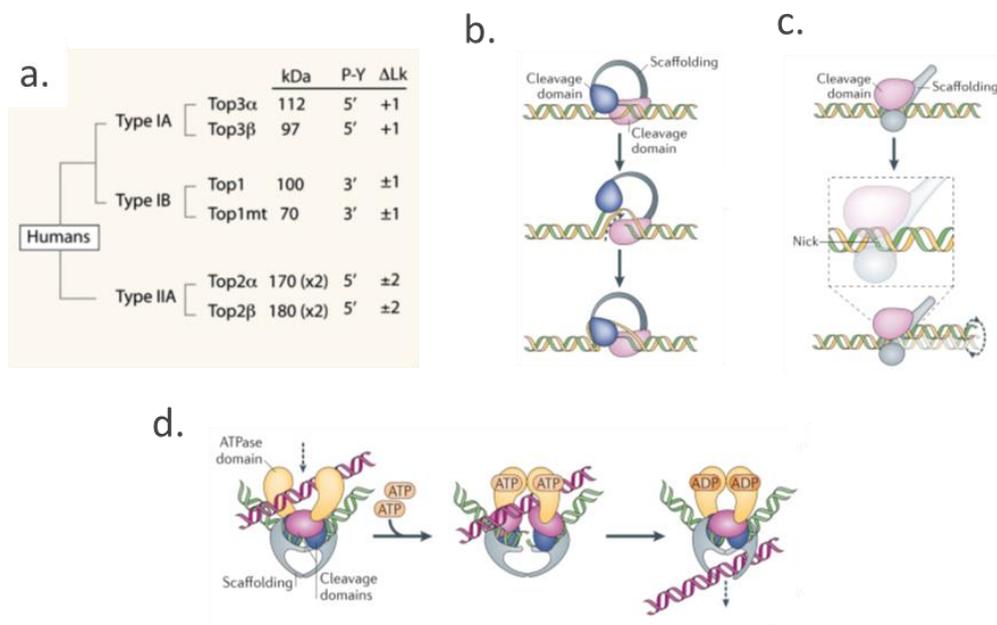


Figure 12: Human DNA topoisomerases. a) Human DNA Topoisomerase categories. (adapted from Pommier *et al.*, 2010). b) Type IA topoisomerases activity. c) Type IB topoisomerases activity. d) Type IIA topoisomerases activity. (b,c,d) adapted from Vos *et al.*, 2011)

A common feature of all topoisomerases is their nucleophilic tyrosine residue. This lets the topoisomerases break the DNA by transesterification reactions, using the active site tyrosine to attack the DNA phosphodiester backbone, and then form an enzyme-DNA covalent link to protect the cleaved DNA from degradation (Pommier, 2013). Top3 enzymes from the topoisomerase IA family only relax negatively supercoiled (under-twisted) DNA, restoring the lacking twists for proper dsDNA conformation. Type IB topoisomerases do not pass DNA strands through the breaks they create, but rather let the dsDNA molecule to rotate. Top1 can restore positively or negatively supercoiled DNA, is required for transcription activation, and is mechanistically coupled to RNA Pol II in transcription elongation (Delgado *et al.*, 2018). Type IIA topoisomerases differ from type IA topoisomerases by cleaving both strands of the DNA instead of one, allowing passage of a

dsDNA instead of the opposite strand, and doing so using ATP. The specificity of Top2 is also their ability to disentangle long intertwined chromosomes and DNA catenanes (Vos *et al.*, 2011).

Poisoning Top2. Topoisomerases currently targeted by anticancer agents are Top1 and Top2 (Pommier *et al.*, 2010; Delgado *et al.*, 2018). The only clinically approved Top1 inhibitors are the plant alkaloid, camptothecin derivatives: Topotecan and Irinotecan. (Figure 13). Camptothecin was isolated from the stem and bark of *Camptotheca acuminata* (“Chinese happy tree”), a tree traditionally used in Chinese medicine (Martino *et al.*, 2017). Topotecan and Irinotecan are used to treat various cancers, such as ovarian, lung, colorectal, glioblastomas, sarcomas, and cervix cancers (Pommier, 2013). Their detailed chemistry and mechanism of action in the context of cancer treatment is reviewed in Pommier, 2009.

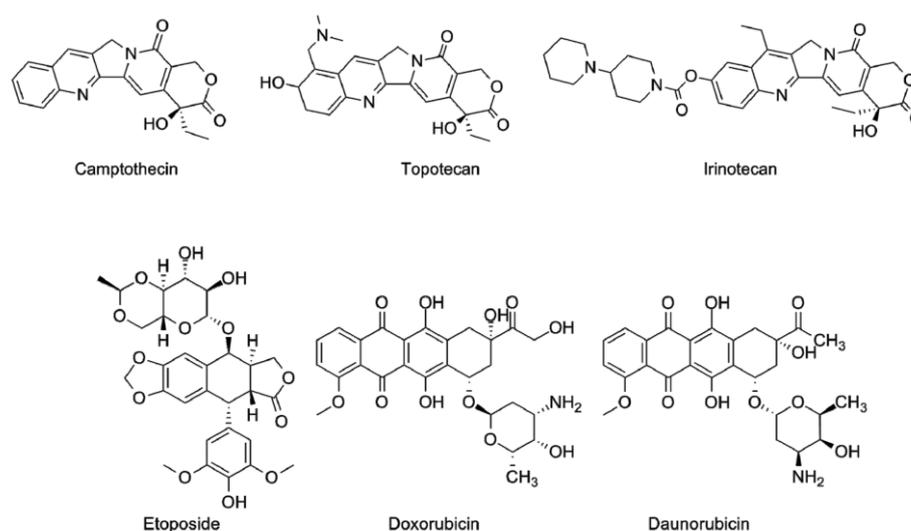


Figure 13: Topoisomerase inhibitors chemical structure. (adapted from Delgado *et al.*, 2018)

Top2 activity can be inhibited in two ways: catalytic inhibition or enzyme poisoning (Figure 14). So far, only poisoning Top2 has shown anticancer effects and is used in the clinic. Catalytic inhibitors decrease the efficacy of the enzymatic reaction without generating Top2-DNA covalent complexes (Top2cc), while Top2 poisons such as doxorubicin trap Top2 to the DNA after cleavage of the dsDNA and prevent its religation, resulting in high amount of toxic Top2cc. By blocking the specific step of DNA religation, not only enzymatic activity of Top2 is inhibited, but it also induces cytotoxic damage (DSB) to the cell. One of the reasons why Top2 poisons proved greater efficacy as anticancer agents is likely because cancer cells frequently overexpress Top2, reducing the efficacy of catalytic inhibitors. An interesting example is HER2-positive breast cancer. Doxorubicin is especially effective in HER2-positive breast cancer, because Top2 is overexpressed due to amplification of the locus on chromosome 17 containing TOP2A and ERBB2 genes (O’Malley *et al.*, 2009).

Anthracyclines are the family of Top2 poisons derived from daunorubicin. Daunorubicin and doxorubicin were initially discovered in the 1950s from *Streptomyces* soil bacteria

(Pommier, 2013). Anthracyclines act by interfacial inhibition. By positioning between -1 and +1 of the DSB, and at the interface between the protein (Top2) and the DNA, they disrupt the special organization necessary for religation to occur. This mechanism is the main mode of action of anthracyclines, although it is not their only cytotoxic effect (Nitiss, 2009; Pommier, 2013). They can also intercalate DNA, interfere with DNA helicase activity and DNA strand separation, and generate reactive oxygen species (ROS) (Henninger and Fritz, 2017). Etoposide (Figure 13) is another Top2 poison used in the clinic that prevents dsDNA religation after strand passage, while it differs from anthracyclines in many ways. Etoposide generates Top2cc at higher frequency, but in a reversible manner, mainly generates SSB as opposed to DSB, and does not intercalate DNA (Pommier, 2013). Interestingly, etoposide has lower long-term side-effects, but also a narrower range of anticancer activity (Hande, 1998).

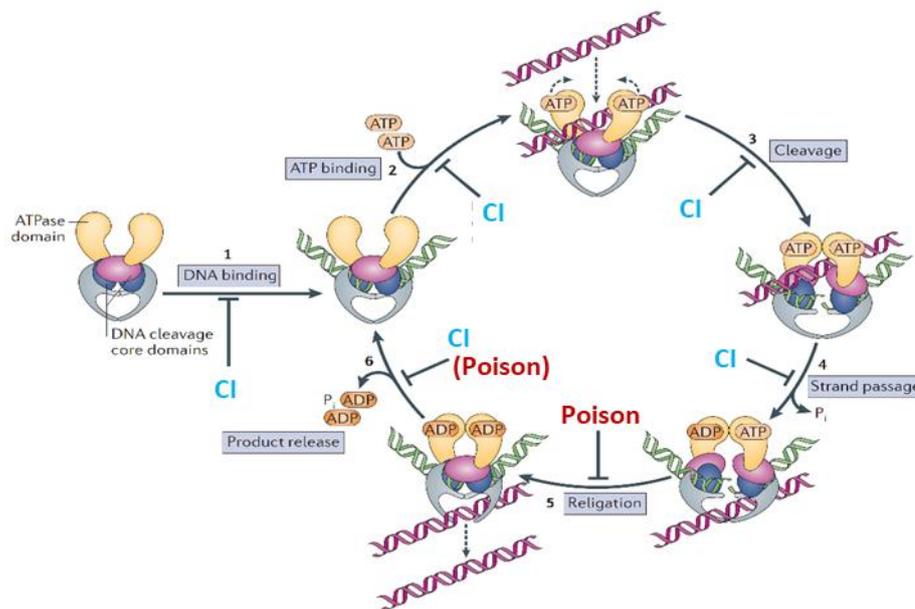


Figure 14: Top2 catalytic cycle and Top2 inhibitors. CI: catalytic inhibitors. (adapted from Vos *et al.*, 2011)

Clinical limitations of doxorubicin. Doxorubicin is one of the most widely used genotoxic anticancer agent, but its use is limited by resistance and dose-limiting side-effects: cardiotoxicity and development of secondary cancers (Turner, Biganzoli and Di Leo, 2015). Doxorubicin-induced congestive heart failure, characterized by reduced left ventricular ejection fraction, often arises years or decades after initial chemotherapy, and is largely dependent on the cumulative dose received by the patient (Henninger and Fritz, 2017). To prevent congestive heart failure, use of cardioprotective agents such as dexrazoxane together with chemotherapy used to be common practice until recently. However, the FDA limited the use of dexrazoxane in 2011. Alternative cardioprotective care such as liposome encapsulation of doxorubicin, or the use of drugs traditionally used in cardiology (β -blockers or statins for example), are therefore of great clinical interest (Vejpongsa and Yeh, 2014; Henninger and Fritz, 2017). Similarly to congestive heart failure, the risk of doxorubicin-induced leukemias is dependent on the cumulative dose, but is low enough (less than 1%)

that the benefit of an anthracycline-based chemotherapy is considered to outweighs the risks (Turner, Biganzoli and Di Leo, 2015).

c. Platinum compounds and mechanism of action

Cisplatin as an anti-cancer drug. *Cis*-diamminetetrachloroplatinum(II) (Figure 15a), or cisplatin, was initially synthesized for the first time in 1844 by Michele Peyrone, and was known as Peyrone's chloride. It was (re)discovered by accident in 1965 by Rosenberg *et al.*, when they realized that a magnetic field generated by platinum electrodes blocked *E. coli* cell division (Figure 15b) (Rosenberg, Van Camp and Krigas, 1965). The team rapidly tested the compound on transplanted tumours in mice and discovered its potent anti-tumour effect (Rosenberg *et al.*, 1969). Clinical trials were shortly started, and FDA-approval for metastatic testicular and ovarian cancer was given in 1978, less than ten years after the initial biophysics experiments of Rosenberg (Kelland, 2007). Its use has now been extended to several other types of solid tumours, including lung, cervical, breast, colorectal, bladder, and head and neck cancers (Johnstone, Suntharalingam and Lippard, 2016).

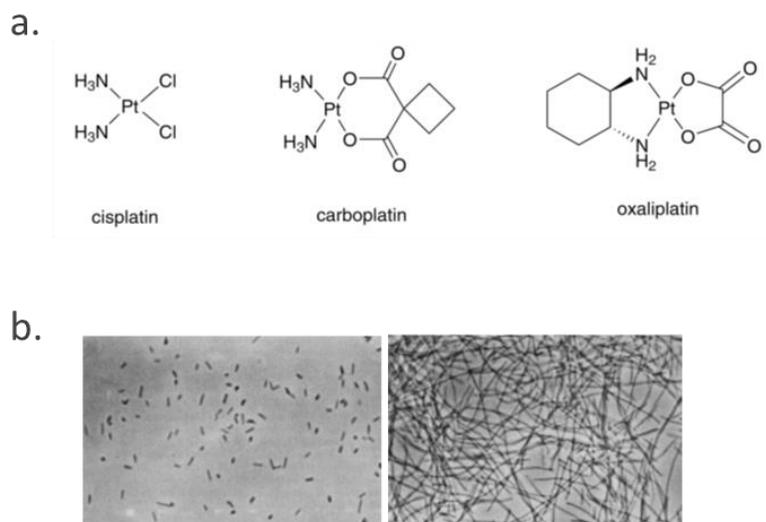


Figure 15: Cisplatin discovery. a) Cisplatin and its analogues (adapted from Dilruba and Kalayda, 2016) b) Unstressed (left) and filament-shaped (right) *E. coli* subject to magnetic field generated from platinum electrode (adapted from Rosenberg, Van Camp and Krigas, 1965)

Cisplatin acts by forming intra-strand or inter-strand platinum crosslink, as well as monoadducts on the DNA (Figure 16). After entering the cell, cisplatin is hydrolyzed by spontaneous aquation reactions. Due to lower chloride concentration (4-20 mM) compared to outside of the cell (100 mM), the chloride leaving groups are replaced with water molecules, allowing the platinum ion to covalently bind guanine residues, and to a lesser extent, adenine residues, at N7 (Kelland, 2007; Dilruba and Kalayda, 2016). When two adjacent purines are bound with cisplatin, they form an intra-strand adduct, resulting in important distortion of the double-helix. This DNA distortion prevents replication and

transcription, which ultimately leads to apoptosis (Kelland 2007). Intra-strand adduct account for the vast majority of damages caused by cisplatin on the DNA, but inter-strand adducts can also be formed when two purines bases from opposite strands are linked (Jung and Lippard, 2007; Kelland, 2007). Cisplatin is also known to bind to RNA and interfere with cellular RNA processing, which probably assist in the action of the drug (see II. 2) c.).

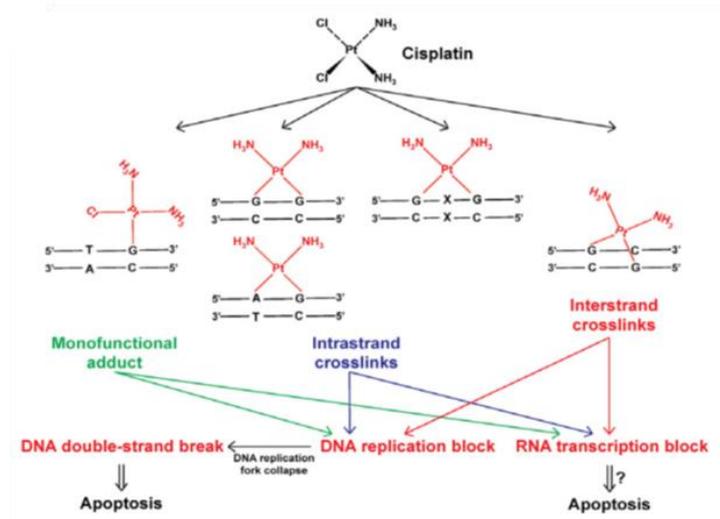


Figure 16: Types of cisplatin-adducts on DNA and their cellular effects. (adapted from Roos and Kaina, 2013)

Clinical limitations of cisplatin. Cisplatin was proven very efficient in a wide range of solid tumours, even showing a nearly 100% response rate in testicular cancer if detected early (Wang and Lippard, 2005). However, its clinical use is limited due to its notorious dose-limiting nephrotoxicity, neurotoxicity, ototoxicity (hearing loss due to damage to hair cells of the cochlea), myelosuppression (decreased blood cell and platelet production), as well as nausea and vomiting (Florea and Büsselberg, 2011; Dilruba and Kalayda, 2016). Nephrotoxicity can be alleviated to some extent by prehydration (hyperhydration with a saline solution) of the patient to increase drug evacuation by the kidneys, whereas hearing loss is permanent (Kelland, 2007; Karasawa and Steyger, 2015; Breglio *et al.*, 2017). In addition to severe dose-limiting toxicity to healthy organs, the principal obstacle to cisplatin use in the clinic is the occurrence of intrinsic (colorectal, prostate, lung, or breast cancer) or acquired (ovarian cancer) resistance after initial response (Kelland, 2007).

Cisplatin analogues. To overcome toxic side-effects and chemoresistance, continuous efforts have been made to design new platinum drugs over the last 40 years. Although thousands of cisplatin analogues have been tested only a handful entered clinical trials, the clinical situation of platinum-based anticancer drugs have not changed in almost 20 years: there are still only three platinum compounds, namely cisplatin, carboplatin, and oxaliplatin, that are globally approved drugs (Johnstone, Suntharalingam and Lippard, 2016).

Carboplatin was developed in the 1980's shortly after the FDA approval of cisplatin. Carboplatin has a more stable leaving group than chloride (100-fold lower aquation rate), rendering it less reactive once inside the cell by its chelating nature while keeping good

aqueous solubility (Figure 15a) (Wang and Lippard, 2005; Wheate *et al.*, 2010; Johnstone, Suntharalingam and Lippard, 2016), and is practically devoid of nephrotoxicity (Kelland, 2007). Due to its lower toxicity, it is suitable for a more “aggressive” therapy regimen, and widely used in ovarian cancer combination therapy where resistance to cisplatin often occurs (Dilruba and Kalayda, 2016). In fact, patients treated with cisplatin or carboplatin show similar survival rate (Aabo *et al.*, 1998). The main dose-limiting constraint of carboplatin is its myelosuppressive side-effect. Interestingly, carboplatin generates the same DNA adducts as cisplatin, but at a slower rates: it has been shown that up to 40 fold higher concentration of carboplatin is necessary to obtain similar effects (Knox *et al.*, 1986).

Oxaliplatin is described as third-generation platinum drug, as it was developed to overcome carboplatin resistance (Dilruba and Kalayda, 2016). It is structurally different from cisplatin and carboplatin by its *R,R*-diaminocyclohexane (DACH) chelating groups replacing the non-leaving amine groups (Figure 15a). This was thought to confer oxaliplatin the ability to be more efficient in DNA replication blocking, and therefore a higher toxicity, by different distortion of the DNA strands and DNA repair than cisplatin and carboplatin. Although like cisplatin, the main type of adducts generated are intra-strand crosslinks on adjacent guanines, oxaliplatin creates less DNA adducts whilst being more cytotoxic (Raymond *et al.*, 2002). Additionally, oxaliplatin induces increased replicative bypass (Dilruba and Kalayda, 2016). Another peculiarity of oxaliplatin is its exclusive FDA-approval for treatment of advanced colorectal cancers, likely due to their high expression of organic cation transporters (OCT) involved in oxaliplatin uptake by the cell (Johnstone, Suntharalingam and Lippard, 2016). Importantly, oxaliplatin is able to maintain its cytotoxic effect in cisplatin and carboplatin primary resistant colon cancers (Raymond *et al.*, 2002). This could be due to the fact that the main cytotoxic effect of oxaliplatin was shown not to be DNA-damage-induced apoptosis, but by ribosome biogenesis stress, explaining the high cytotoxicity with less DNA-damage than cisplatin or carboplatin (Bruno *et al.*, 2017).

Since carboplatin and oxaliplatin, a plethora of platinum-containing molecules have been investigated in the hopes of finding an anticancer drug with even less side-effects and to overcome various resistance mechanisms, including trans-platinum (II) and platinum(IV) pro-drugs, polynuclear complexes, as well as new modes of delivery (nanoparticles, polymers, proteins) (Johnstone, Suntharalingam and Lippard, 2016). So far, none have reached the clinical success of cisplatin, carboplatin, and oxaliplatin, and sometimes trials have been stopped at stage III (Dilruba and Kalayda, 2016). In fact, clinical trial regulations have evolved since the discovery of anticancer properties of cisplatin, which in fact, might not pass approval with current authorities due to its severe side-effects.

2) The cell response to genotoxic stress

a. The DNA damage response (DDR)

Specific responses for specific damages. The DNA in our cells is continuously exposed to DNA damage, whether it is from environmental origin, chemotherapeutic agents, or products from normal cellular metabolism (Jackson and Bartek, 2010). Exogenous DNA damage include UV light and cosmic radiation from the atmosphere, genotoxic compounds found in cigarette smoke or pollution, and agents used in cancer chemotherapy. Endogenous DNA damage are mostly reactive oxygen species (ROS) from oxidative respiration. Different agents induce different types of lesions, and the cells have developed a panel of intricate mechanisms, collectively called the DNA-damage response (DDR), to address each type of issue (Table 4) (Giglia-Mari, Zotter and Vermeulen, 2011; Lord and Ashworth, 2012; Roos and Kaina, 2013).

Table 4: DNA lesions and their repair pathways. (adapted from Giglia-Mari, Zotter and Vermeulen, 2011)

Lesion	Cause	Repair process(es)
CPD, 6-4PP ⁽¹⁾	Sunlight	NER
Bulky adducts ⁽²⁾	Food, cigarette smoke	NER
Intrastrand crosslinks	Chemotherapy (e.g., Cis-Pt)	NER
8-oxo-dG ⁽³⁾	ROS ⁽⁴⁾ , respiration	BER
Thymineglycol ⁽³⁾	ROS ⁽⁴⁾ , respiration	BER
N ⁷ -Alkyl-dG, N ³ -Alkyl-dA	Food, pollutants	BER
O ⁶ -Alkyl-dG	Food, pollutants	DR ⁽⁵⁾ , BER?
5-methyl-dC	DNMT ⁽⁶⁾	BER/AID-BER/NER?
Uracil, (Hypo)Xanthine	Spontaneous deamination	BER
Abasic site	Spontaneous hydrolysis	BER
Single-strand breaks	Ionizing radiation, ROS	Ligation, BER
Double-strand breaks	Ionizing radiation, ROS, VDJ-rec	HR, NHEJ
Tyrosyl-3'DNA	Topo-I inhibition, ROS	SSBR
Mismatches	Replication errors	MMR
Small insertion/deletions	Replication slippage	MMR
Interstrand crosslinks	Chemotherapy	ICLR/ HR?

1. CPD: cyclobutane pyrimidine dimer; 6-4 PP: 6-4 pyrimidine-pyrimidone photo-product.

2. A large group of chemicals conjugated to bases that cause DNA helix destabilization such as: Benzo[α]pyrene (a polycyclic aromatic hydrocarbon); Aflatoxins (present in fungal food contaminations); and Nitrosamines (tobacco smoke).

3. A large group of different oxidation products affecting either the base or the phosphate-sugar backbone of which 8-oxo-dG is the most abundant.

4. ROS: reactive oxygen species, produced as side-product of respiration/metabolism and ionizing radiation.

5. DR: direct reversal, involving the suicide enzyme MGMT.

6. DNMT: DNA methyltransferase, functions in epigenetic gene-expression control (e.g., at CpG islands).

The most complex to repair, and potentially the most toxic DNA lesions are double-strand breaks (DSB). DSBs can be directly induced by topoisomerase inhibitors and ionizing radiation (IR) used in cancer radiotherapy, X-rays, or indirectly occurring as a result of unrepaired single-strand breaks (SSB) or crosslinks during replication. There are two major mechanisms to repair DSB : homologous recombination (HR) and non-homologous end-joining (NHEJ) (Figure 17a). For HR, part of the DNA surrounding the break is removed by a process called resection, then the sister chromatid is used as a template to replace the missing DNA sequence. On the other hand, NHEJ is a mostly error-prone repair process

which directly ligates the DNA ends together. It is a faster repair, but subject to loss or changes in nucleotides, thus possibly mutagenic. NHEJ is also the DSB repair pathway most used in mammalian cells (Goldstein and Kastan, 2015). The probability of which DSB repair will be preferentially used mostly depends on the cell cycle phase. Since HR requires a homologous sister chromatid template, it will act during S and G2 phases, while NHEJ can happen throughout the cell cycle (Giglia-Mari, Zotter and Vermeulen, 2011; Goldstein and Kastan, 2015).

When the break affects only one strand, it is repaired by base excision repair (BER) (Figure 17a). This type of damage is often of endogenous origin, but are also generated by IR in cancer therapy (Hoeijmakers, 2001; Goldstein and Kastan, 2015). In BER, the opposite strand can serve as template. A single nucleotide (short-patch BER) or up to 13 nucleotides (long-patch BER) are removed and replaced with newly synthesized DNA. BER is also used to repair rather subtle modifications on nucleotides, such as oxidative lesions (8-oxoguanines) and alkylation products. In this case, the damaged bases are first removed from the double-helix DNA by specific glycolases, creating a SSB that will be repaired in the same manner as the direct SSB. Bulkier single-strand lesions that provoke distortion of the DNA helix, such as lesions induced by UV light and intrastrand crosslinks, are repaired by nucleotide excision repair (NER). NER is subclassified in two categories: the global genome NER (GG-NER) that surveys the genome for distortion injury, and the transcription-coupled NER (TC-NER) that occurs when lesions blocks RNA polymerases elongation. In NER, after recognition of the lesion, DNA is unwinded, DNA portions containing the lesions are excised, and the gap is filled by the replication machinery (Hoeijmakers, 2001). Lastly, another important repair mechanism is mismatch repair (MMR), which corrects insertions and deletions, and mismatches in base pairing resulting from replication errors. Base mismatching creates distortion on the double-helical structure of DNA, and so is recognized as a DNA lesion (Lord and Ashworth, 2012). During replication, the newly synthesized and error-containing DNA is excised and resynthesized. This versatile machinery is reviewed in Jiricny, 2006.

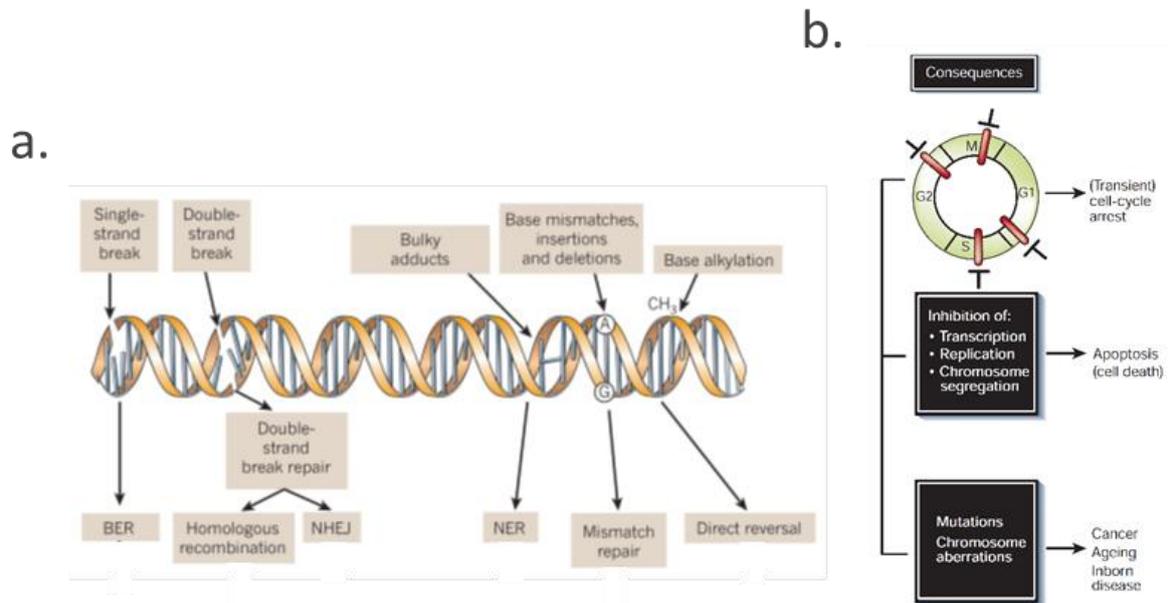


Figure 17: The DNA-damage response. a) DNA repair mechanisms for different types of DNA lesions. (adapted from Hoeijmakers, 2001) b) Cellular consequences of DNA damage. (adapted from Lord and Ashworth, 2012)

DDR signaling. After a DNA break, the cell activates intricate signaling cascades leading to cell cycle arrest, which allows a window of opportunity to repair damaged DNA before replication. In healthy cells, three outcomes are possible: (1) DNA is repaired and cell cycle resumes, (2) DNA repair was not successful and the cell goes into senescence, (3) damage to the DNA is too important and the cell undergo apoptosis (Figure 17b). In cancer cells, mitosis can occur even if DNA is not properly repaired after damage. Depending on the type of lesion and the cell cycle phase, cell cycle will be arrested at G1/S, during S phase, or at G2/M (Giglia-Mari, Zotter and Vermeulen, 2011). The general DNA damage signaling is done by damage-sensing protein complexes which recruit apical kinases of the phosphatidylinositol 3-kinase-like kinase (PIKK) family (ATM, ATR, and DNA-PK), which transduce signal to mediators and effectors. This initial signaling and recruitment of repair factors can be referred to as the “early DDR” (Figure 18). This results in activation of necessary cellular processes (DNA repair, replication and transcription arrest, cell cycle arrest, apoptosis, senescence, etc.) by a transcriptional response due to activation of key transcription factors such as p53, and subsequent expression of their target genes, a process that can be referred to as the “late DDR” (Figure 18).

Following a DSB, the lesion is typically recognized by the MRN (MRE11-RAD50-NBS1) complex, and the ataxia-telangiectasia mutated (ATM) kinase is recruited and activated by phosphorylation. Once ATM is at the site of damage and is activated it phosphorylates a number of targets important for the cell response to DNA damage, which will amplify the ATM signaling, as well as the downstream kinase Chk2. Chk2 phosphorylation activates p53 stabilization, which will upregulate p21 and induce cell cycle arrest in G1/S through cyclin E and cyclin-dependent kinase 2 (CDK2) inhibition (Giglia-Mari, Zotter and Vermeulen, 2011).

Another sensor of DSB not pictured in Figure 18 is DNA-dependent protein kinase (DNA-PK), a crucial enzyme in NHEJ. It is activated when it is recruited to the site of DSB by the Ku70/80 complex, and directly promotes NHEJ by recruiting necessary factors such as XRCC4 and DNA ligase IV (Blackford and Jackson, 2017).

When there is a SSB, from direct damage, as an intermediate of repair pathways mentioned above, or from replication fork stalling, it is ataxia-telangiectasia and Rad3-related (ATR) that senses the lesion. ATR recognizes replication protein A (RPA) coated ssDNA, via association with ATR-interacting protein (ATRIP), and phosphorylates Chk1, which in turn inhibits the cell cycle-dependent 25 (CDC25) proteins, resulting in cell cycle arrest in S and G2/M. Loading of Rad9-Rad1-Hus1 (or the 9-1-1 complex) to RPA coated ssDNA also triggers ATR activation (Giglia-Mari, Zotter and Vermeulen, 2011). Although ATM and ATR are activated by different types of DNA lesions, they converge to similar downstream signaling and cellular outcomes (Figure 18).

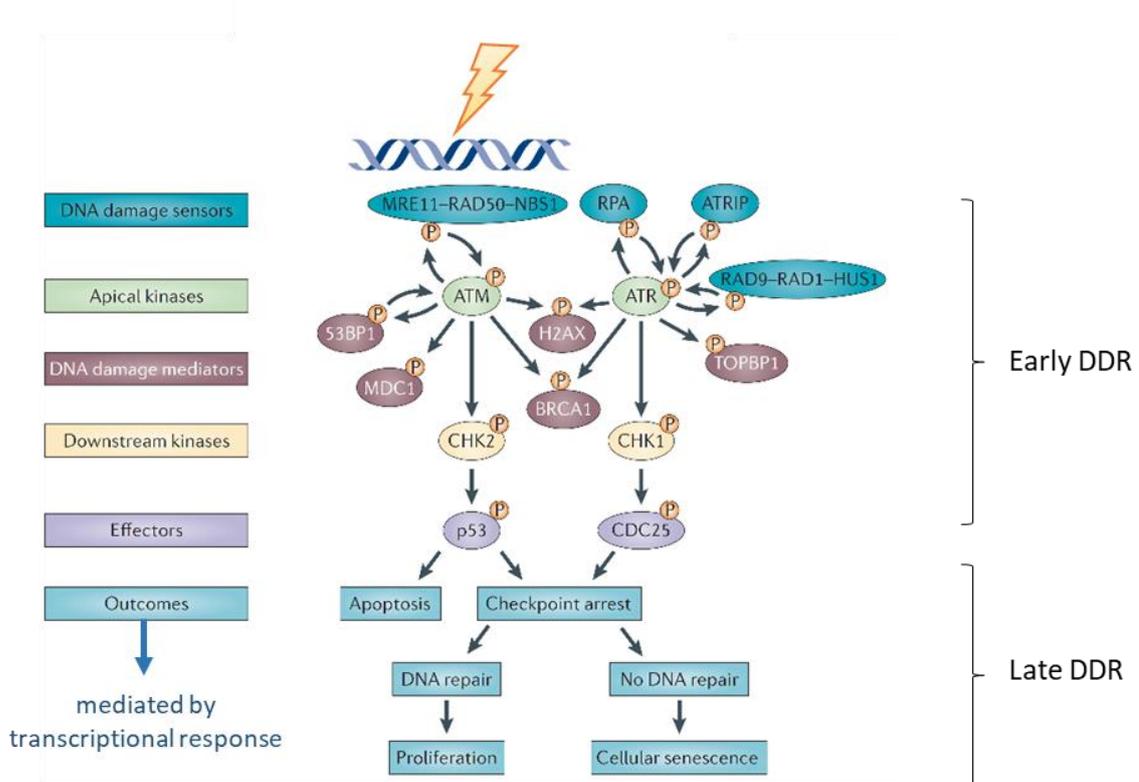


Figure 18: The DDR signaling. (adapted from Sulli, Di Micco and Di Fagagna, 2012)

b. Repair of doxorubicin and cisplatin DNA lesions

Repair of doxorubicin-induced DNA lesions. The main type of DNA lesion induced by doxorubicin is DSB, as detailed in II. 1) a., or more specifically Top2cc resulting in protein-linked DNA breaks (PDB). The process by which Top2 poison-induced PDB are repaired is only beginning to be elucidated. In order to repair such breaks, the cell must recognize PDB as genuine damage, rather than a transitional state of an enzymatic activity (Nitiss, 2009).

Therefore, PDB must be transformed into “regular” DSB by removal of Top2 covalently bound to the DNA. This first requires proteolytic degradation of stalled Top2, as shown by Gao *et al.* (Gao *et al.*, 2014). Once Top2 of the Top2cc is partially degraded, it is precisely disjoined from DNA, without cleaving DNA, by tyrosyl-DNA-phosphodiesterase 2 (TDP2), leaving a DSB substrate for NHEJ (Ashour, Atteya and El-Khamisy, 2015). Consistently, depletion of TDP2 in cell lines and in mice increased their sensitivity to etoposide (Ledesma *et al.*, 2009; Gómez-Herreros *et al.*, 2013). Notably, it has been shown that TDP2 generates “ligation-ready” DNA termini (5′ phosphate and 3′ hydroxyl) resulting in an error-free NHEJ repair, because Top2 is bound to DNA by its 5′-ends (while Top1 binds 3′-ends) (Gómez-Herreros *et al.*, 2013). This specific Top2cc related PDB repair is most likely coordinated by ATM activation, but details about TDP2 recruitment to PDB, its relationship with DNA repair and other cellular pathways, and its regulators still need further investigation (Pommier, 2013; Ashour, Atteya and El-Khamisy, 2015).

Repair of cisplatin-induced DNA lesions. As mentioned above, cisplatin mainly induces intrastrand crosslink, and to a lesser extent, interstrand crosslinks (ICL). Intrastrand crosslinks are classically repaired by the NER pathway alone (Figure 17), whereas ICL are trickier to repair and the mechanism is not fully elucidated (Goldstein and Kastan, 2015). It has been reported that ICL repair calls for a coordination of NER, MMR, HR, and a DNA damage tolerance mechanism called translesion synthesis, as well as a specific repair pathway called the Fanconi anaemia pathway (Roos and Kaina, 2013; Goldstein and Kastan, 2015; Ceccaldi, Sarangi and D’Andrea, 2016; Dilruba and Kalayda, 2016). How to repair ICL depends on the cycling state of the cell. In cycling cells and senescent cell, replication-dependent and independent repair mechanisms are activated, respectively. In cycling cells, ICL repair is triggered by the replication fork collision with the lesion, since it prevents the separation of the DNA strands (Clouston, Schärer and Niedernhofer, 2013; Ceccaldi, Sarangi and D’Andrea, 2016). The Fanconi anaemia pathway is named after a rare genetic disorder of DNA repair deficiency (predisposing to cancer) due to inactivation of one of the genes in the pathway. It is activated in the S phase during replication, and the subsequent steps are: lesion recognition, DNA incision, lesion bypass, and lesion repair (Figure 19a) (Ceccaldi, Sarangi and D’Andrea, 2016). When two replication forks collide at the ICL site, the replicative helicase CMG complex is removed by BRCA1 (another function of BRCA1 besides HR) (Long *et al.*, 2014; Ceccaldi, Sarangi and D’Andrea, 2016). The FANCM protein, phosphorylated by ATR, senses the lesion and binds to ICL which recruits the Fanconi anaemia core complex composed of 14 proteins. This will activate FANCD2-FANCI heterodimer by a reversible monoubiquitylation, that seem to serve as a “landing pad” for NER repair proteins, which will do the DNA nucleolytic incision to release the ICL from one of the two strands (unhooking). Lesion bypass on the ICL tethered strand is then done by translesion synthesis, a DNA damage tolerance process which allows the replication machinery to bypass DNA lesion by switching to error-prone DNA translesion polymerases (REV1 and Pol ζ in humans) (Ceccaldi, Sarangi and D’Andrea, 2016). Meanwhile, the DSB generated by the nucleolytic incision on the other strand is mainly repaired by HR, although HR-independent mechanisms have also recently been reported (Wang *et al.*, 2015). Although there has been a considerable expansion of ICL repair pathways in the recent years, the exact role, the timing

of recruitment, post-translational regulations, and mechanistic insights of the majority of Fanconi anaemia proteins and other ICL repair-related proteins is far from being fully understood (Ceccaldi, Sarangi and D'Andrea, 2016).

Senescent cells also need to repair DNA damages outside of the S phase (in G₀/G₁) when there are ICLs. First studies of replication-independent cisplatin-induced ICL repair mechanisms are quite recent, and mainly based on cell-free extract and reporter plasmid experiments. They showed that replication-independent cisplatin-induced ICL were dependent on the replication-coupled NER pathway, and involved the translesion DNA polymerase Pol κ (Enoiu, Jiricny and Schärer, 2012; Williams, Gottesman and Gautier, 2012). However, implication of MMR proteins in ICL repair and cisplatin resistance have been investigated for more than 20 years (Aebi *et al.*, 1996; Jung and Lippard, 2007). In a 2017 study by Kato *et al.*, it has finally been demonstrated in *Xenopus* cell-free extracts supporting neither replication nor transcription, that the MMR proteins (MutS α and MutL α) are required for ICL sensing and activation of repair (Figure 19b) (Kato *et al.*, 2017).

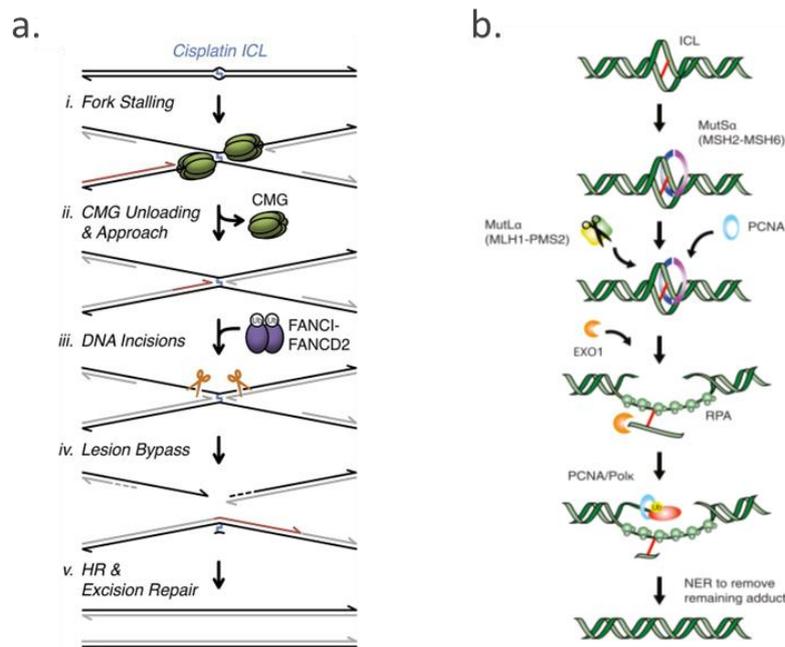


Figure 19: Repair of cisplatin-induced ICL. a) Replication-dependent ICL repair in cycling cells. (adapted from Long *et al.*, 2014) b) Replication-independent ICL repair in senescent cells. (adapted from Kato *et al.*, 2017)

c. Cellular responses specific to doxorubicin and cisplatin

Cellular responses to doxorubicin. As mentioned in II. 1) a., anthracyclines such as doxorubicin have a wide range of cellular actions in addition to their Top2 poisoning function. By intercalating DNA, it blocks transcription and replication. In addition, it has been reported that doxorubicin induces histone eviction from the chromatin, including the histone variant

H2AX crucial for DDR signaling, deregulating the transcriptome and driving apoptosis (Pang *et al.*, 2013). Moreover, doxorubicin is reported to induce cell death in many ways: apoptosis, necrosis, and mitotic catastrophe (inability of the cell to complete mitosis and ending up in a large cell with multiple micronuclei) (Eom *et al.*, 2005; Tacar, Sriamornsak and Dass, 2013). Surprisingly, it was also reported that doxorubicin induces a viral infection-like immune response, via type I interferon (Sistigu *et al.*, 2014).

The generation of ROS by doxorubicin is well documented, and it has been thought to be the link with cardiotoxicity. The quinone moiety of doxorubicin is known to reduce oxygen molecules, but more importantly doxorubicin forms complexes with intracellular iron in a redox cycle that convert superoxide anion ($O_2^{\cdot-}$) and hydrogen peroxide (H_2O_2) into more potent hydroxyl radicals ($\cdot OH$) (Minotti, 2004). It was hypothesized that this accumulation of ROS in oxidative stress-sensitive cardiomyocytes was responsible for the onset of congestive heart failure in patients (Henninger and Fritz, 2017). It has been shown that instead, it is most likely due to cardiomyocyte-specific Top2 β poisoning (cardiomyocytes express Top2 β but not Top2 α , as they are quiescent) and ROS resulting from mitochondrial dysfunction rather than redox cycling (Zhang *et al.*, 2012).

Cellular responses to cisplatin. Following DNA damage induced by cisplatin, cells mainly die by p53-induced apoptosis, involving various signaling cascades and cell cycle arrest (Figure 20). In addition to the NER, Fanconi anaemia, and MMR proteins, high-mobility group proteins (HMG) also recognize platinated DNA (Siddik, 2003). HMG1 preferentially binds to distorted DNA, and has been shown to inhibit NER of cisplatin crosslinks by shielding the damage site from repair, therefore increasing cisplatin cytotoxicity (Jung and Lippard, 2007). Also, HMG1 enhance p53 binding to DNA, and induction of apoptosis.

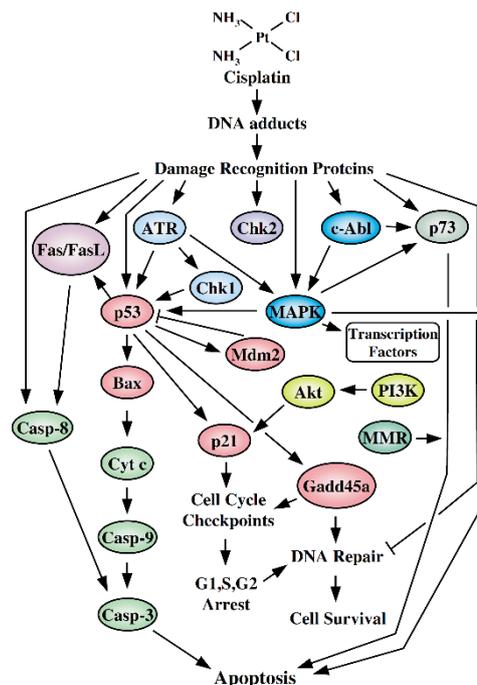


Figure 20: The cisplatin-induced DDR. (adapted from Siddik, 2003)

Although main cytotoxic effect of cisplatin is almost always attributed to its DNA binding, it turns out that only about 1% of cisplatin molecules are bound to nuclear DNA (Gonzalez *et al.*, 2001). This suggests that a large part of administered cisplatin stays outside of the cells or in the cytoplasm. As mentioned above, once in the cytoplasm, cisplatin is hydrolyzed and is transformed into highly reactive mono-aqua and di-aqua complexes. In this form, cisplatin can easily bind any nucleophilic substrate that it encounters, whether it is DNA, RNA, proteins (especially cysteine-rich), and any thiol-containing molecule (Cepeda *et al.*, 2007). For example, it is known to bind with high affinity with reduced glutathione (GSH), inducing oxidative stress by sequestering antioxidants such as GSH and shifting the redox status in the cell (Cepeda *et al.*, 2007; Karasawa and Steyger, 2015). In addition, cisplatin can also directly generate ROS. It has been shown that by binding to mitochondrial DNA, which lacks NER, cisplatin blocked mitochondrial DNA transcription, protein synthesis, and eventually impaired mitochondrial respiration (Marullo *et al.*, 2013). This oxidative stress could in part explain toxicity to non-dividing healthy tissues (nephrotoxicity, ototoxicity), as they are unlikely to die from DDR and apoptosis induced by cisplatin crosslinks during replication (Karasawa and Steyger, 2015).

DNA and RNA molecules are chemically extremely similar. It is therefore not surprising that cisplatin interacts with RNA too, but these relationships have been less studied even though it has been reported that cisplatin interacts with various RNA species, including ribosomal and telomeric RNAs. Recently, Melnikov *et al.* gave further insight into translation inhibition by cisplatin by showing in bacterial ribosomes that cisplatin intercalates between the ribosome and the mRNA, impairing effective mRNA translocation for translation (Melnikov *et al.*, 2016).

3) Chemotherapy resistance mechanisms

a. Resistance to anticancer chemotherapy, general mechanisms

The chemotherapy resistance problem. Drug resistance, whether intrinsic or acquired during treatment, is the main hurdle to clinical success of cancer chemotherapy. The cancer cell constantly and impressively adapts to its environment to keep on thriving. There are multiple mechanisms described for genotoxic anticancer drug resistance, from drug efflux, modification of the drug target and cell signaling pathways, to the cancer stem cell (CSC) phenotype and EMT. The reason why chemoresistance is so difficult to predict and to overcome is because the cell uses a combination of mechanisms to resist to a single agent (while in practice patients are almost always treated with a combination of drugs), and even this combination can evolve over time. Moreover, the multi-drug resistance phenomenon and intra-tumour heterogeneity makes the situation even more complex. Although we are at the beginning of the era of personalized medicine, prognostic markers of chemoresistance are not yet to the point of being used to design a specific therapy regimen for a cancer patient. It is therefore of crucial importance to understand the molecular basis of each

mechanism and their interactions, to attempt to depict a more precise picture of the resistance networks involved for a given drug (or a combination of drugs) at a given time (or through time).

Intracellular mechanisms. An easy way for the cancer cell to avoid exposure to a drug is to simply increase its evacuation or to restrict its import through transporters, limiting intracellular accumulation (Wijdeven *et al.*, 2016). Notably, the ATP-binding cassette (ABC) transporter superfamily responsible for multidrug resistance (MDR) is particularly well documented (Szakács *et al.*, 2006; Li *et al.*, 2016). There are 49 members in the ABC transporter superfamily, but only three have been most studied in the context of cancer chemoresistance : ABCB1 (also named P-glycoprotein (P-gp) and multidrug resistance protein 1 (MDR1)), ABCC1 (also named MDR-associated protein 1 (MDP1)), and ABCG2 (also named breast cancer resistance protein (BCRP)) (Holohan *et al.*, 2013). In normal cells, ABCB1 is especially important for excretory functions of epithelial cells of the digestive tract (Holohan *et al.*, 2013). It is overexpressed in many cancers, and regulate transport of a wide range of drugs used in anticancer therapy, such as anthracyclines, antimetabolites, and taxanes (Wijdeven *et al.*, 2016).

Alterations of the drug target is also common, whether it is up- or down-regulation, or mutation, to compensate for the direct effect of the anticancer agent. For example, 5-Fluorouracil inhibits thymidylate synthase, a key enzyme for synthesis of thymidine and therefore DNA, as its name suggests. Expression of thymidylate synthase is increased in 5-Fluorouracil treated cancer cells through a negative feedback loop mechanism, and higher levels of thymidylate synthase were correlated to poor response to chemotherapy in gastric and colorectal cancers (Kachalaki *et al.*, 2016). Down-regulation and mutation of drug targets cause chemoresistance too, as it results in the inability for the drug to bind its target. Classical hallmarks of cancer are the ability of cancer cells to avoid apoptosis, unrestricted cell growth, and inefficient DDR (Hanahan and Weinberg, 2011). It is therefore logical to use these advantages to resist genotoxic stress induced by chemotherapy agents at multiple levels. Cells adapt by deregulating apoptotic response (for example, by over-expressing anti-apoptotic BCL-2 family proteins) and by activating pro-survival signaling (for example, EGFR and NF- κ B signaling) (Holohan *et al.*, 2013). Faulty DDR is a “mixed blessing” in the context of chemotherapy, as it confers sensitivity to genotoxic drugs and some DDR inhibitors, but also causes genomic instability potentially responsible for the acquisition of genetic alterations leading to chemoresistance. To resist genotoxic chemotherapy, the cancer cell can modulate its DDR by three main mechanisms : 1) genetic reversion, which restores the drug target function by mutation or internal deletions; 2) mutation of DDR genes that impairs the function of the drug target; and 3) DDR alteration, by which the cells uses an alternative DNA repair pathway than the one targeted by chemotherapy (Bouwman and Jonkers, 2012).

Intracellular resistance mechanisms discussed above are summarized in Figure 21:

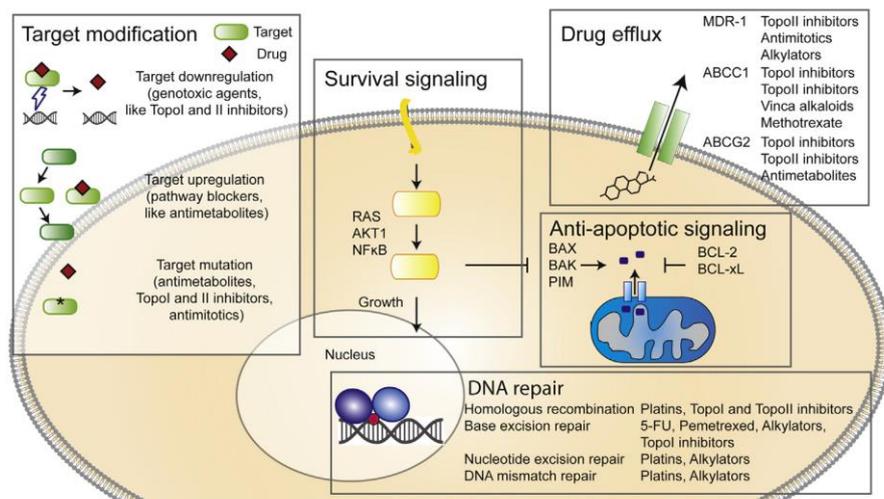


Figure 21: Intracellular resistance mechanisms (adapted from Wijdeven *et al.*, 2016)

Systemic mechanisms. In addition to intracellular mechanisms, the cancer cell can acquire chemoresistance as a population and by interacting with its micro-environment. Activation of survival signaling mentioned above can be of cell-extrinsic origin, notably by cytokines and growth factors secreted by cancer-associated fibroblasts (CAF) and tumour-associated macrophages (TAM) (Rebucci and Michiels, 2013). In a cancer cell lines and stromal cells co-culture experiment, it was shown by Straussman *et al.* that dermal fibroblasts were able to confer complete innate resistance to gemcitabine in colon and pancreatic cancers (Straussman *et al.*, 2012). Moreover, DNA damage induced by genotoxic chemotherapy in prostate cancer patients was shown to promote EMT and acquired resistance through paracrine signaling of the Wnt pathway from the tumour micro-environment (Sun *et al.*, 2012).

During EMT, epithelial cells lose their polarized organization and switch to a more invasive mesenchymal phenotype. This transition was initially suggested to be the key initiator mechanism of metastasis, but two independent groups showed recently that EMT was dispensable for metastasis but an important factor for chemoresistance (to cyclophosphamide and gemcitabine) in mouse models (Fischer *et al.*, 2015; Zheng *et al.*, 2015). In addition, the link between cancer stem cells (CSCs) and chemoresistance have been increasingly acknowledged over the past few years. The CSC theory states that a small number of quiescent, self-renewing cells within the tumour are responsible and capable of tumour initiation. CSC harbor intrinsic chemoresistance features, such as ABC transporters expression, activation of anti-apoptotic and pro-survival signaling, as well as enhanced DNA damage repair and EMT phenotype (Holohan *et al.*, 2013). More importantly, they are hardly sensitive to genotoxic anticancer agents effective to fast-cycling cells, due to their quiescent nature. Confirming this, enrichment in CSC population after chemotherapy in patients have

been observed in several cancers, such as breast, colorectal, pancreatic, lung, prostate cancers as well as AML and CML (Vidal *et al.*, 2014).

The contribution of the immune system, and specifically immunogenic cell death, to the efficacy of chemotherapy is increasingly recognized (Krysko *et al.*, 2012; Kroemer *et al.*, 2013). DNA damage can induce immunogenic cell death, and an immunosuppressive environment can promote resistance to this particular type of cell death. For example, it was recently demonstrated that in oxaliplatin-treated prostate cancer, immunogenic cell death by cytotoxic CD8⁺ T-cells was dependent on B cells plasmocytes, highly present in resistant tumours (Shalapour *et al.*, 2015).

Finally, autophagy can also promote chemoresistance by facilitating cancer cell survival during metabolic stresses caused by anticancer agents (Holohan *et al.*, 2013). Systemic resistance mechanisms discussed here are summarized in Figure 22:

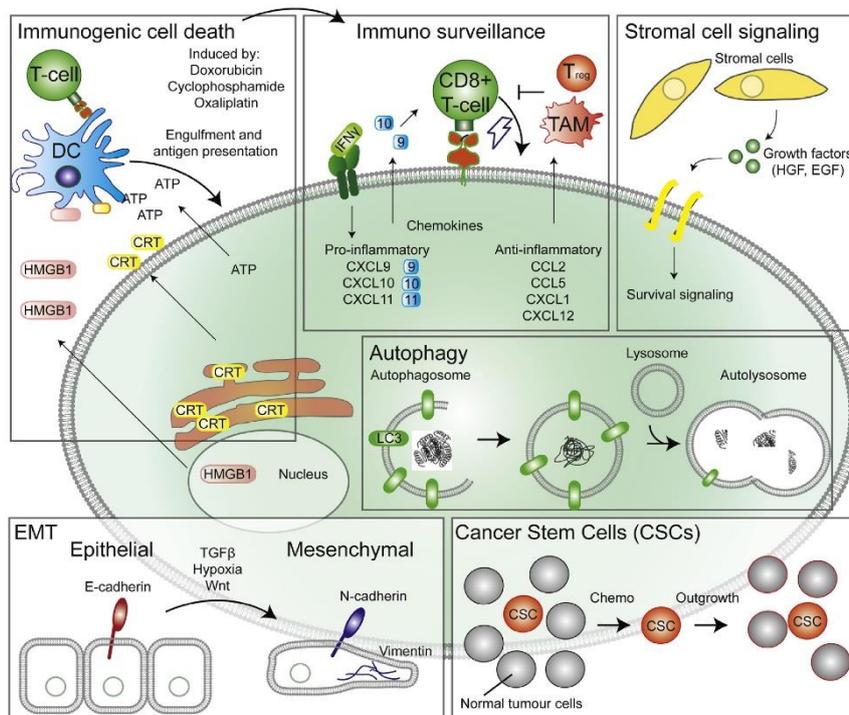


Figure 22: Systemic resistance mechanisms (adapted from Wijdeven *et al.*, 2016)

b. Resistance mechanisms specific to doxorubicin and cisplatin

Resistance to doxorubicin. Resistance mechanisms to doxorubicin are as various as its modes of action. One of the most described doxorubicin resistance mechanism is drug efflux via ABCB1. Although ABCB1 and ABC transporters in general have been associated with chemoresistance in a large panel of cancers, use of ABC transporters inhibitors in patients have failed to ameliorate doxorubicin sensitivity, suggesting that MDR via ABC transporters is not the most prominent resistance mechanism, and that ABC transporters might have a high degree of redundancy (Holohan *et al.*, 2013; Wijdeven *et al.*, 2016). Indeed, several

clinical trials of chemotherapy-ABC transporter inhibitor combination failed to show any improvement in patient outcome (Szakács *et al.*, 2006). This being acknowledged, further investigations on the precise role of ABCB1 in doxorubicin resistance, could give better insights and therapeutical opportunities. For example, it has been shown that ABCB1 expression was mediated by micro-RNAs, chromosomal amplification, and epigenetic modulations (Genovese *et al.*, 2017), as well as the transcription factor oestrogen receptor α (ER α) (Chen *et al.*, 2018). Doxorubicin can also be sequestered in lysosomes to prevent nuclear drug exposure, and have been shown to trigger lysosome biogenesis by nuclear translocation of the transcription factor TFEB, further enhancing lysosomal drug entrapment and MDR (Zhitomirsky and Assaraf, 2015). Interestingly, it was also shown that doxorubicin lysosomal sequestration was dependent on lysosomal ABCB1 and the ability of doxorubicin to be ionized at lysosomal pH (Yamagishi *et al.*, 2013).

Other common doxorubicin resistance mechanisms are mutation or downregulation of the target Top2, increased anti-oxidant defense of the cell, decreased DNA damage or increased DNA repair, and anti-apoptotic signaling due to p53 mutation (Chien and Moasser, 2008; Wijdeven *et al.*, 2016). A recent genome-wide screen identified novel factors contributing to doxorubicin resistance all converging to DSB repair, namely Keap1 and SWI/SNF that attenuated Top2 activity and therefore DSB generation, and C9orf82/CAAP1 which when depleted accelerated DSB repair (Wijdeven *et al.*, 2015). An oncogenic kinase, lemur tyrosine kinase 3 (LMTK3), involved in endocrine therapy resistance was also shown mediate doxorubicin resistance by delaying formation of DSB and inducing expression of DNA repair genes (Stebbing *et al.*, 2018). Because doxorubicin also induces immunogenic cell death, overexpression of CD73, an important enzyme in the catalysis of AMP to adenosine (a potent immunosuppressor), was shown to be associated with triple-negative breast cancer patients doxorubicin resistance. CD73 impaired immunogenic cell death after doxorubicin treatment by suppressing CD8⁺ T-cell activation via A2A adenosine receptor signaling (Loi *et al.*, 2013).

Additionally, actors of chemoresistance can have multiple functions. For example, it has been shown recently that the p53 negative regulator MDM2, has p53-independent functions in resistance to doxorubicin, but not other genotoxic drugs: p53 mutant cells overexpressing MDM2 were specifically resistant to doxorubicin by reducing DSB (Senturk, Bohlman and Manfredi, 2017). Another example is the histone deacetylase 10 (HDAC10) was recently shown to be important for doxorubicin containing lysosome exocytosis as well as DSB repair in chemotherapy-resistant neuroblastoma cell lines, and was not related to lysosomal drug efflux via ABCB1 (Ridinger *et al.*, 2018).

Resistance to cisplatin. Resistance to cisplatin is mostly a result of two broad mechanisms: failure to reach DNA, and failure to achieve cell death after DNA damage (Kelland, 2007). The ability of cisplatin to reach DNA can be disturbed in several ways, such as increased drug efflux, decreased cellular uptake, and cytoplasmic inactivation (Amable, 2016). Although increased drug efflux is believed not be the main cause of limited intracellular accumulation of cisplatin (as opposed to doxorubicin), some transporters such as ATP7A/B and MRP2 were shown to be upregulated in cisplatin resistant cancer cell lines

(Galluzzi *et al.*, 2012). The copper transporter protein 1 (CTR1) is one of the primary cisplatin transporters involved in cisplatin uptake (Ishida *et al.*, 2002). Cisplatin resistant ovarian cancer cell lines displayed low levels of CTR1, and high expression of CTR1 in ovarian cancer patients correlated with increased survival, suggesting a rationale for the use of copper chelating agents to overcome cisplatin resistance in patients (Liang *et al.*, 2012). Similar observations were found in stage III non-small-cell lung cancer patients (Chen *et al.*, 2012). As mentioned in II. 2) c., cisplatin binds with high affinity to intracellular GSH. On one hand, this results in oxidative stress participating in cisplatin cytotoxicity, but on the other hand GSH can act as a cisplatin scavenger and limits drug availability (drug inactivation) (Galluzzi *et al.*, 2012). Indeed, high expression of GSH was observed in patient-derived ovarian cancer cell lines after the onset of chemoresistance (Wolf *et al.*, 1987). However, Chen *et al.* demonstrated that overexpression of GSH sensitized small-cell lung cancer cells to cisplatin by upregulating CTR1 expression (Chen *et al.*, 2008). Therefore, the role of GSH in cisplatin resistance is still largely unclear and needs further investigation.

Failure to achieve cell death after DNA damage is mostly due to increased DNA repair and DDR. The most notorious example is increased expression of ERCC1 from the NER machinery crucial for cisplatin-induced DNA crosslinks (see II. 2) b.). It is widely described in the literature for its negative correlation with cisplatin response in patients with various cancer types (Galluzzi *et al.*, 2012; Amable, 2016). Indeed, low ERCC1 expression was correlated with better survival in cisplatin treated NSCLC patients, and have proved to be a useful biomarker to predict response to cisplatin in an adjuvant setting (Bowden, 2014). However, it is important to note that expression levels of ERCC1 do not seem to correlate with NER proficiency (Galluzzi *et al.*, 2012), and that methods of ERCC1 protein expression by immune-histochemistry have been questioned recently (different batches of the same antibody yielded contradictory results) (Friboulet *et al.*, 2013). Therefore, the role of ERCC1 in cisplatin resistance has to be taken carefully, as it has not been shown how ERCC1 mechanistically contributes to cisplatin resistance, and even associated with better response to cisplatin in some contexts (Bowden, 2014). DNA crosslinks induced by cisplatin can also be repaired by MMR and HR (see II. 2) b.). Accordingly, downregulation of several actors of these pathways have been linked to cisplatin resistance in cell lines and patients, such as MSH2 and MLH1 (MMR), and BRCA1/2 (HR) (Galluzzi *et al.*, 2012; Wijdeven *et al.*, 2016). In addition, hyperactivation of TLS (important in ICL repair during replication) is also reported to participate in cisplatin resistance (Haynes *et al.*, 2015). In conclusion, any, or a combination of, the DNA repair pathways involved in the various types of adducts generated by cisplatin throughout the cell cycle may be involved in its chemoresistance.

III. Effects of genotoxic agents on alternative splicing and polyadenylation

In the past two parts, I exposed fundamental and current knowledge on AS and APA function and regulation in cancer, and on cell response and resistance to genotoxic anticancer agents doxorubicin and cisplatin. Transcriptional responses to DNA damage and in chemoresistance are widely described. The relationship between genotoxic stress and post-transcriptional regulations is also increasingly appreciated, especially in an acute response context. Interestingly, genome-wide screens and proteomic studies on the DDR identified factors involved in mRNA processing (Matsuoka *et al.*, 2007; Paulsen *et al.*, 2009; Solier *et al.*, 2010). In fact, post-transcriptional regulations seem to be an integrative part of the DDR. However, how the cell uses AS and APA to respond to genotoxic stress or to become resistant to anticancer agents remains largely unexplored.

1) AS in genotoxic stress response and resistance

a. AS response to genotoxic stress

Global AS alteration in response to genotoxic stress. Isolate examples of splice variants expressed in response to genotoxic stress have been reported since the 1990's, but first global analyses of the extent of AS after DNA damage are less than 10 years old. These first studies were done on splice-sensitive micro-arrays, which depend on the number and types of probes included in the array design. In 2009, Muñoz *et al.* analyzed AS changes in 482 genes involved in cancer, cell cycle, proliferation and death after UV irradiation and observed AS alteration in 22% of these genes (Muñoz *et al.*, 2009). In this study, they showed that AS of apoptosis-related genes such as Bcl-X and caspase 9 was mediated by inhibition of Pol II elongation by hyperphosphorylation of the CTD induced by UV (Muñoz *et al.*, 2009). In 2010 and 2011, several studies from my supervisor and others showed genome-wide AS changes in response to camptothecin (Dutertre *et al.*, 2010; Solier *et al.*, 2010; Ip *et al.*, 2011). My supervisor also found 248 exons regulated in response to doxorubicin treatment in MCF-7 cells by pan-genomic exon-array (Dutertre, Chakrama, *et al.*, 2014), as mentioned in I. 2) d.. Many of the AS isoforms induced by cisplatin were found by Gabriel *et al.* in 2015 (Gabriel *et al.*, 2015). After observing that MDM2 and VEGF were alternatively spliced after cisplatin treatment, they carried out an RNA-seq on cisplatin treated MCF-7 cells. They found 717 AS events in 619 transcripts, and only five genes were regulated in AS as well as transcriptionally. These findings illustrate that the splicing response to genotoxic stress is specific and has a distinct role from transcriptional regulation.

Moreover, it was recently shown that the DDR and AS are mechanistically linked. In 2015, Tresini *et al.* showed that UV induces displacement of spliceosome complex U2/U5/U6 from the chromatin via a non-canonical ATM signaling in a positive feedback loop involving Pol II arrest and R-loop formation, resulting in increased intron retention and exon skipping.

Importantly, they demonstrated reciprocal regulation of the spliceosome and ATM: spliceosome organization modulation activated ATM, and ATM was capable of influencing AS (Tresini *et al.*, 2015).

AS variants after DNA damage. Interestingly, many of the genes affected by AS following genotoxic stress are SF and genes involved in the DDR as shown in Table 5 (Giono *et al.*, 2016), corroborating the strong link between splicing and the DDR.

Table 5: Validated AS variants of SF and DDR genes. (adapted from Giono *et al.*, 2016) (IR = Ionizing radiation, CIS = cisplatin, DOX = doxorubicin, ETO = Etoposide, MMS = Methyl methanesulfonate)

Gene-AS Event	Stimulus	Effect
<i>(a) AS of genes known to be involved in the DDR</i>		
Mdm2	UV, IR, CIS	Exons 4–11 skipping. p53 stabilization
	CPT, DOX, CIS	Exon(s) skipping. p53 stabilization
	DOX, actinomycin D	Intron 10 retention, truncated inactive variant. p53 stabilization
Mdm4	UV	Exon(s) skipping. p53 stabilization
Fas receptor (CD95)	mitomycin C	Increased soluble Fas isoform (exclusion of E6)
Caspase 2	ETO, CPT, DOX, mitoxantrone	Increased anti-apoptotic 2S form (inclusion of E9)
Caspase 8	CIS	Reduced anti-apoptotic form (inclusion of E8), via SRSF2
Caspase 9	UV	Increased pro-apoptotic caspase-9a form (inclusion of E3–4-5-6 cassette)
	multiple drugs, cyclophosphamide	Increased pro-apoptotic caspase-9a form via E2F1/SRSF2
Bcl-x	UV, CIS	Increased pro-apoptotic Bcl-x _S form
	CIS, oxaliplatin	Increased pro-apoptotic Bcl-x _S form via ATM/Chk2/p53
	cyclophosphamide	Increased pro-apoptotic Bcl-x _S form via E2F1/SRSF2
Bax, BRCA1	multiple drugs	Increased pro-apoptotic form
survivin	multiple drugs	Increased pro-apoptotic form
GADD45G	IR	Reduced inclusion of E2
PIG3 (TP53I3)	UV, CIS	Increased exon E4 skipping (PIG3AS). Increased degradation
Smac/DIABLO	UV, CIS	Increased exon E4 skipping (Smac3). Proapoptotic
NBS1	IR, MMS	DSB sensor complex. Decreased inclusion of exon with premature termination codon
Mcl-1	IR	Increased anti-apoptotic Mcl-1 L form. Radioresistance
Abl1, Chk2, MAP4K2	UV	Increased exon skipping via relocalization of EWS
DHX9	UV	Increased inclusion of exon 6 A (NMD+ form)
<i>(b) AS of genes involved in mRNA metabolism</i>		
SRSF1	UV	Splicing factor, reduced intron inclusion (NMD+ form)
SF1, RBM8A, ZRANB2, BAT1	CPT	Splicing factors, increased exon skipping
HNRPDL, HNRPLL, SRSF2,	CPT, UV	Splicing factors, increased exon inclusion
SF3B3,		
SNRNPB, MBNL2, PRPF3,		
PRPF38B		
FUBP1	UV	DNA and RNA binding protein, increased exon inclusion
TIA-1, TIAL1	UV	RNA binding proteins, increased exon inclusion
TIA-1, TIAF1	CIS, topotecan	RNA binding proteins

For example, EWS, an RNA and DNA binding protein that interacts with Pol II, is involved in the AS of DDR genes such as MDM2, Abl1, Chk2 and MAP4K2, and its effects on its target genes is impaired following various types of genotoxic stress (camptothecin, doxorubicin, cisplatin, UV) inducing exon skipping (Dutertre *et al.*, 2010; Paronetto, Miñana and Valcárcel, 2011). My supervisor showed that after genotoxic stress induced by camptothecin, the interaction between EWS and YB-1, a spliceosome-associated factor, is altered and leads to skipping of multiple MDM2 alternative exons. This results in the decrease of the full length MDM2 and p53 stabilization, a mechanism that could explain the necessity of EWS for cell survival in HeLa after UV irradiation (Paronetto, Miñana and Valcárcel, 2011). SRSF1 has also

been shown in minigene experiments to promote MDM2 alternative multi-exon skipping in response to cisplatin and UV, by binding to exon 11 of MDM2 (Comiskey *et al.*, 2015). In response to cisplatin, Gabriel *et al.* showed that DNA damage-induced AS changes and cell death were partly mediated by the SF SRSF4. Post-translational modifications of SF are also involved in the DDR-related AS changes. Indeed, Edmond *et al.* showed that in response to cisplatin, SRSF2 protein is accumulated by hypoacetylation and phosphorylation that prevent its degradation by the proteasome, and regulates AS of caspase 8 to decrease the anti-apoptotic isoform (Edmond *et al.*, 2011). In addition, Shkreta *et al.* showed that after oxaliplatin treatment in HEK-293 cells, SRSF10 was dephosphorylated, which abrogates its interaction with splicing activators hnRNPH/F and removes splicing repressor hnRNPK from Bcl-x 5'SS, allowing the expression of the pro-apoptotic isoform (Bcl-xS) (Shkreta *et al.*, 2016).

AS response to doxorubicin and cisplatin. As discussed above, there has been an increasing interest in shedding light on the AS response to genotoxic stress in the past decade, but few studies focused on Top2 poisons including doxorubicin. RT-PCR validated AS isoforms described in the literature as induced by doxorubicin are summarized below in Table 6:

Table 6: Validated AS variants regulation in response to doxorubicin treatment

Alternatively spliced gene	Exon	Type of AS	Up/Down	Cells	Reference
APIP	2	cassette exon	↓	human iPSC-derived cardiomyocytes	Knowles <i>et al.</i> , 2018
Bcl-x	2	alternative 5'SS	↑	HEK-293, HeLa, MCF-7, PC-3	Shkreta <i>et al.</i> , 2008
CASP2	9	cassette exon	↑	U-937	Solier <i>et al.</i> , 2004
CDC25C	3+5+6	cassette exon	↓	MCF-7, MDA-MB-231	Albert <i>et al.</i> , 2012
MDM2	110	cryptic exon	↑	NIH3T3	Lents <i>et al.</i> , 2008
	4-11	cassette exon, multi-exon	↓	MCF-7	Dutertre <i>et al.</i> , 2010
Survivin	12	cryptic exon	↑	EU-3	Zhu <i>et al.</i> , 2004
	3	cassette exon	↓		

Examples of AS isoforms regulated by cisplatin are listed in Table 7. Most of the AS changes induced by doxorubicin or cisplatin are related to the p53 pathway. As discussed above, DDR genes are now increasingly recognized in the AS response to DNA damage. However, AS variants in Table 6 and 7 were mostly found on a candidate-gene approach, and further investigation on the AS response to genotoxic stress is needed.

Table 7: Validated AS variants in response to cisplatin treatment.

Alternatively spliced gene	Exon	Type of AS	Up/Down	Cells	Reference
Bcl-x	2	alternative 5'SS	↑	HEK-293, HeLa, MCF-7, PA-1, PC-3, SKOV-3	Shkreta et al., 2008
				Hep3B	Muñoz et al., 2009
CASP8	8	alternative 5'SS	↑	H358	Edmond et al., 2011
CDC25C	3+5+6	cassette exon	↓	MCF-7	Albert et al., 2012
EIF4A2	4	cassette exon	↓	MCF-7	Gabriel et al., 2015
hnRNPD	8	cassette exon	↑	MCF-7, AT5BIVA, MO59J	Gabriel et al., 2015
	6	cassette exon	↓		
H-RAS	5	cassette exon	↑	SH-SY5Y	Vivarelli <i>et al.</i> , 2013
MAD2	2-3	multi-exon	↓	HCT116	López-Saavedra <i>et al.</i> , 2016
MAGOH	3	cassette exon	↓	MCF-7	Gabriel et al., 2015
MDM2	4-11	multi-exon	↓	Ishikawa, HT1080, RD, MG63, MSU-1, HDF-1, HDF-2, BT-549, MDA-MB-231	Gabriel et al., 2015
				H1299, U2OS	Chandler <i>et al.</i> , 2006
				MCF-7	Chandler et al., 2006; Dutertre et al., 2010; Gabriel et al., 2015
MDM4	4-9	multi-exon	↓	HCT116, H1299, MCF-7, IMR90	Markey and Berberich, 2008
TMPO	6-8	multi-exon	↓	MCF-7	Gabriel et al., 2015
VEGF	6	cassette exon	↓	MCF-7, Ishikawa	Gabriel et al., 2015

SFs involved in DDR beyond splicing. Several examples of SFs contribution to DNA repair and the DDR at multiple levels have been described (Dutertre, Lambert, *et al.*, 2014; Naro *et al.*, 2015). For example, some SFs are recruited to the site of DNA damage and act directly on DNA repair. In a genome-wide siRNA based screen to identify components of the HR machinery, it was unexpectedly found that RNA post-transcriptional modification was the most enriched function in the positive regulators of HR (Adamson *et al.*, 2012). In this study, Adamson *et al.* showed that the SF RBMX localized to sites of damage and promoted HR in a PARP1-dependent manner. Another splicing factor, FUS, which was recently proposed to also participate in co-transcriptional splicing by mediating U1 snRNP and Pol II interaction (Yu and Reed, 2015), was also shown to be recruited to sites of DNA damage to induce HR via PARP1 (Mastrocola *et al.*, 2013). Another example is PRP19, a component of the core spliceosome and E3 ubiquitin-ligase, is ubiquitylated following DNA damage, recognizes RPA-coated SSBs, ubiquitylates RPA and activates ATR-mediated DDR (Maréchal *et al.*, 2014). Depletion of other SFs such as SRSF1, SRSF2, and several hnRNPs have been shown to promote genomic instability, through formation of R-loops. Furthermore, DNA damage induces relocalization of and post-translational modifications on SF, adding additional layers of AS and DDR interplay (Dutertre, Lambert, *et al.*, 2014; Naro *et al.*, 2015; Shkreta and Chabot, 2015).

DDR proteins involved in AS. Surprisingly, some proteins traditionally involved in DNA repair and DDR such as BRCA1, 53BP1, and Ku70/80, turned out to be able to interact directly or indirectly with RNA (Dutertre and Vagner, 2017). For example, BRCA1 was shown to form a splicing complex with Bcl2-associated transcription factor 1 (BCLAF1) that interacts with members of the core spliceosome (Prp8, U2AF65, U2AF35, SF3B1). Following DNA damage induced by etoposide, BRCA1 is phosphorylated by ATM, and forms a complex with BCLAF1 that relocalizes to promoter regions of genes involved in the DDR to ensure their proper expression and splicing (Savage *et al.*, 2014).

b. AS and chemoresistance

In the context of chemoresistance, AS regulations are mainly isolate observations, and there is a very limited number of global analyses. In an interesting recent study, it has been shown that BRCA1 mutated cancer cells could become resistant to PARP inhibitors and cisplatin by partially rescuing their HR functions by using an alternative 5'SS of exon 11, producing a shorter isoform (BRCA1-Δq) with residual activity (Y. Wang *et al.*, 2016). The exon 11 of BRCA1, with over 3kb in length, is one of the longest of the human genome, and this AS results in the use of only the first 117 bp of exon 11 (Raponi *et al.*, 2014).

Some SF have also been described in the resistance to genotoxic anticancer agents. One of the first examples is the role of the SF SPF45 (or RBM17), a component of the spliceosome, in the multidrug resistance phenotype. Indeed, SPF45 is overexpressed in many solid tumours such as bladder, breast, colon, lung, ovarian, pancreas, and prostate cancer (Sampath *et al.*, 2003), and has been observed to confer resistance to carboplatin, vinorelbine and vincristine (microtubule poisons), doxorubicin, etoposide, mitoxantrone (another Top2 poison), gemcitabine, and pemetrexed (an antimetabolite) in a stably transfected ovarian cancer cell line (Perry *et al.*, 2005). Its target exons have not been elucidated in these early studies, but we know today that one of its targets is the exon 6 of the apoptosis gene Fas receptor. SPF45 promotes Fas exon 6 skipping by directly interacting with U2AF65, which produces a soluble protein lacking the transmembrane domain with antagonistic functions to the inclusion isoform (Corsini *et al.*, 2007), and it has been recently suggested that this regulation by SPF45 involved the Fas-antisense lncRNA and was responsible for apoptosis resistance (Villamizar *et al.*, 2016). Another SF involved in chemoresistance is the polypyrimidine tract binding protein PTBP1 that is involved in incurable disease pancreatic ductal adenocarcinoma (PDAC) gemcitabine resistance by modulating the AS of the pyruvate kinase gene PKM's mutually exclusive exon 9 (included in isoform PKM1) and 10 (included in isoform PKM2) (Calabretta *et al.*, 2016). PKM2 isoform was induced by PTBP1 binding to exon 8, was associated with poor prognostic PDAC, and promoted gemcitabine resistance in cells. Indeed, when PKM2 isoform was specifically depleted by an antisense oligonucleotide (ASO), thus restored sensitivity to gemcitabine as well as cisplatin in cells (Calabretta *et al.*, 2016).

One of the very few global analyses of AS in chemoresistance found thousands of AS events regulated in daunorubicin (a Top2 poison molecularly very close to doxorubicin)

resistant acute myeloid leukemia (AML) cells and chemotherapy resistant patient samples, and these regulations occurred in mostly different genes from those regulated at the transcriptional level (Mohamed *et al.*, 2016). However this study did not go further with mechanistical and functional analysis to demonstrate AS isoforms or SF involved in chemoresistance.

2) APA in genotoxic stress response and resistance

a. Regulation of 3' end processing and APA by genotoxic stress

Effect of genotoxic stress on 3' end processing. Following DNA damage such as UV irradiation, global cellular levels of mRNA are transiently decreased. It is suggested that in addition to a general decrease in transcription, inhibition of 3'-end processing is participating in this decrease (Cevher and Kleiman, 2010). Indeed, it has been shown that 3' end processing is strongly but transiently inhibited following UV irradiation (Kleiman and Manley, 2001). After UV-induced DNA damage, CstF (that otherwise promote 3' end cleavage) forms a complex with DDR proteins BRCA1-associated RING domain protein 1 (BARD1) and BRCA1 and this inhibits cleavage and induces deadenylation of transcripts by poly(A)-specific ribonuclease (PARN) which leads to their degradation (Cevher *et al.*, 2010). Moreover, the same group later showed that p53 directly participated in 3' end processing inhibition following DNA damage (Nazeer *et al.*, 2011). They also recently showed that another layer of the CstF-BARD1-BRCA1 complex 3' end processing inhibition is promoting ubiquitination of RNA Pol II and histones (Fonseca *et al.*, 2017). Ubiquitination and degradation of Pol II is a "last resort" DDR mechanism (Wilson, Harreman and Svejstrup, 2013), and is conserved between yeast and humans (Kuehner, Kaufman and Moore, 2017). It is particularly efficient for the cell to decrease its mRNA production while maintaining it for crucial DDR genes after a genotoxic injury. Some insights into this regulation lie in the G-quadruplex structure downstream of p53's PAS. This structure is stabilized by the RNA helicase DHX36 (Newman *et al.*, 2017), and its 3' end processing is maintained following UV damage via an interaction with hnRNPH/F (Decorsière *et al.*, 2011). Also, a non-canonical PAP, Star-PAP, has been shown to promote 3' end processing of the mitochondrial apoptosis gene BCL2-interacting killer (BIK) expression after etoposide-induced DNA damage (Li *et al.*, 2012). Additional mechanisms of 3' end processing and DDR interplay are reviewed in Cevher and Kleiman, 2010.

APA regulation by DNA damage. Investigations on the effect of genotoxic stress on APA are quite recent, and some of the first studies were done in yeast. In 2013, Graber *et al.* analyzed genome-wide APA after *Saccharomyces cerevisiae* exposure to a UV-mimetic drug and found 2031 genes (about 35% of all yeast genes) whose polyadenylation site usage was altered, giving the first evidence of DNA-damage induced APA regulation (Graber *et al.*, 2013). In this study, DNA damage repressed proximal and favoured distal polyadenylation sites of elongated isoforms. Interestingly, a similar effect was observed in response to doxorubicin in mammalian cells. When treated with doxorubicin, 248 exons were alternatively spliced, as mentioned above. The majority of these exons were ALEs, and doxorubicin favoured the use of the evolutionarily recent distal ALE in DDR and cell cycle genes (Dutertre, Chakrama, *et al.*, 2014). Two recent large-scale studies also showed APA

regulations in response to UV, but the effects were in the opposite direction. In 2016, Devany *et al.* demonstrated by performing 3' region extraction and deep-sequencing (3'READS) that UV induced widespread APA including IPA as well as 3'UTR shortening and lengthening in colon carcinoma cells (Devany *et al.*, 2016). IPA was biased towards the 5' end of genes enriched in DDR functions and was mediated by a U1 snRNP decrease, consistent with the role of U1 snRNP in IPA repression discussed in I. 2) a.. In HEK-293 cells, UV also induced 5' proximal IPA/ALEs. This shift from long to short isoform resulted in the expression of a non-coding RNA isoform of ASCC3, which opposed the function of the full-length protein involved in transcription regulation (Williamson *et al.*, 2017).

So far, very few studies on APA regulation by genotoxic stress focused on anticancer agents. In a study where the role of the three PAP isoforms in APA choice was investigated, the authors found that after epirubicin exposure the mRNA and protein levels of the tumour suppressor gene PTEN was increased, and this was abrogated by knocking down Star-PAP, suggesting their role in APA in response to DNA damage (W. Li *et al.*, 2017). Additionally, the 3'UTR of the transcription factor ZEB1, which was found to be associated with EMT in pancreatic ductal adenocarcinoma, is shortened upon gemcitabine treatment, resulting in its increased translation efficiency by escape from microRNA repression (Passacantilli *et al.*, 2017).

b. APA and chemoresistance

As investigations on APA response to genotoxic stress are quite recent, it is not surprising that there is very limited literature on APA and anticancer drug resistance. To my knowledge, there is no study on APA regulation in the context of resistance to cancer chemotherapy, except two examples initially described as AS events. In PDAC cells, gemcitabine induces the expression of SRSF1, promoting the expression of the MAPK-interacting kinase 2 (MNK2) b isoform (Adesso *et al.*, 2013). This isoform is in fact the result of the use of the 13b ALE lacking the MAPK binding site instead of the canonical last exon 13a. The expression of MNK2b isoform attenuated gemcitabine-induced cell death, and authors suggested that PDAC cells activate a MNK2-dependent pro-survival pathway in response to gemcitabine treatment. The other example is IPA of Top2 α discovered in an etoposide resistance context, described as an intron retention event (Kanagasabai *et al.*, 2016). In this recent study, authors identified a possibly dominant-negative 90 kDa truncated isoform of Top2 α (full length is 170 kDa), which was the product of an IPA in intron 19, in etoposide resistant K562 leukemia cells. When transfected into sensitive cells, this isoform reduced Top2cc and DNA damage in response to etoposide.

IV. Thesis aims

Many AS events and several SF have been involved in cell response to genotoxic agents including anticancer drugs, and various AS variants have been described in chemoresistance. However, in the context of chemoresistance, there are very few studies of genome-wide AS regulation and of AS pathways linking RBPs, transcript isoforms, and cellular processes. In addition, the regulation of IPA by genotoxic anticancer drugs also remains poorly understood.

Given the undeniable role of AS and APA in cancer, and its increasingly well-described regulation in response to some types of DNA damage, I hypothesized the existence of AS and APA regulation pathways/networks controlling gene expression which modulate cell response and sensitivity to genotoxic chemotherapeutic agents. To address this, my PhD project had two main aims detailed below:

1) Regulation of alternative splicing in breast cancer cell resistance to doxorubicin

Doxorubicin is one of the most commonly used anti-cancer agent in breast cancer treatment. Previous studies from my supervisor demonstrated involvement of AS in doxorubicin response in breast cancer cell lines. However, there is no description so far of an “AS pathway” (SF → AS variant → cellular process → resistance) involved in doxorubicin sensitivity or resistance.

To better understand the role of AS in doxorubicin resistance, I used a sensitive and resistant MCF-7 breast cancer cell line model to decipher:

1. The extent of AS regulation in doxorubicin resistance
2. SF and AS variants modulating doxorubicin sensitivity
3. Cellular processes affected by AS leading to doxorubicin resistance

2) Regulation of alternative polyadenylation in non-small cell lung cancer cells response to cisplatin

Cisplatin is one of the oldest anticancer drugs, and is frequently used to treat various cancers, including NSCLC. The study of ALEs in cancer and therapy response is an emerging area and very little is currently known. To decipher the extent and cellular impact of ALE regulations in cisplatin response of NSCLC cells, I investigated the following aspects:

1. The extent of ALE regulation in NSCLC in response to platinum drugs
2. Translation efficiency of cisplatin-induced ALEs

RESULTS

- I. Paper manuscript 1: “Identification of splicing programs and pathways involved in breast cancer cell resistance to doxorubicin”

Identification of splicing programs and pathways involved in breast cancer cell resistance to doxorubicin

Iris TANAKA¹, Olivier SAULNIER², Clara BENOIT³, Alina CHAKRABORTY¹, Sophie VACHER⁴, Eric W. LAM⁵, Ivan BIECHE⁴, Olivier DELATTRE², Stéphan VAGNER¹, Didier AUBOEUF³, Martin DUTERTRE¹

¹CNRS UMR 3348, Institut Curie, Orsay, France; Equipe Labellisée Ligue; Paris-Sciences-Lettres Research University; Université Paris Saclay

²Institut Curie Research Center, SIREDO Oncology Center, Paris-Sciences-Lettres Research University, INSERM U830, Laboratory of Biology and Genetics of Cancers, Paris, France; Université Paris Diderot, Sorbonne Paris Cité, France

³CNRS UMR 5239, Ecole Normale Supérieure de Lyon, Lyon, France

⁴Unité de Pharmacogénomique, Service de génétique, Institut Curie, Paris, France ; Université Paris Descartes, Paris, France

⁵Imperial College London, London, United Kingdom

ABSTRACT

Alternative splicing (AS) networks involved in oncogenesis are increasingly known. With respect to anticancer drug resistance, various AS events and a few splicing factors have been involved, but little is known about regulatory pathways and genome-wide programs. Doxorubicin is widely used in chemotherapy of ER+ breast cancer. Here, using genome-wide analyses, we identified over 1700 AS events and 40 splicing factors regulated in a breast cancer cell model of acquired resistance to doxorubicin (comparing DoxoR to parental cells). An RNAi screen identified two poorly characterized splicing factors, ZRANB2 and SYF2, whose depletion partially reversed doxorubicin resistance. Using RNAi and RNA-seq, we found that ZRANB2 and SYF2 mainly regulate cassette exons and alternative 3' splice sites, respectively, that were enriched in cytoskeleton- and proliferation-related genes. Both ZRANB2 and SYF2 directly promoted the inclusion of exon 5 in ECT2 gene transcripts,

resulting in its higher inclusion in resistant *versus* sensitive cells. In resistant cells, ZRANB2, SYF2 and the ECT2-Ex5+ variant promote cell cycle arrest in S phase in response to doxorubicin. Targeting ECT2 exon 5, which encodes a BRCT domain, partially sensitized DoxoR cells to doxorubicin *in vitro*. ECT2 exon 5 inclusion was associated with bad prognosis in breast cancer, specifically in ER+ tumors treated with chemotherapy. Altogether, our data identify the splicing programs controlled by two splicing factors and converging on the same AS event, that participate to breast cancer cell resistance to doxorubicin.

Keywords

Alternative splicing; genotoxic agents; breast cancer; chemotherapy; resistance.

INTRODUCTION

A major problem in anticancer therapy, either conventional or targeted, is the frequent acquisition of resistance to treatment. One of the main classes of anticancer agents are genotoxic agents. Resistance can involve various processes (often in combination), such as drug efflux or metabolism, drug target regulation, DNA-damage response, cell survival and death pathways, epithelial-mesenchymal transition, and cancer stem cell phenotype¹. Acquired resistance is associated with mutation or expression regulation of genes that are either involved in these processes, or in the expression regulation of such genes. Transcriptomic analyses have found many protein-coding genes, microRNAs and long non-coding RNAs that are differentially expressed in resistant *versus* sensitive cells. While most of these alterations are likely passenger rather than driver events, studies have identified resistance-associated gene regulatory pathways connecting altered regulators and target genes that play a role in resistance. These regulatory pathways have been mainly limited to quantitative gene expression regulation at the levels of transcription, RNA stability, and translation^{1,2}.

In addition to quantitative regulation, human gene expression is also regulated qualitatively, in a large part through alternative splicing (AS) that generates alternative transcripts in more than 90% of protein-coding genes. AS is controlled in a large part by more than 300 splicing factors that bind specific RNA motifs in pre-messenger RNAs and/or are part of the core

spliceosome machinery³. In various cancers, hundreds of AS regulations are found in tumors *versus* healthy tissues, and several splicing factors are recurrently mutated or overexpressed in specific cancers and have been shown to have oncogenic properties^{4–6}. Recent studies on oncogenic splicing factors have identified the genome-wide AS programs they control, as well as target splice variants that are phenotypically relevant, suggesting AS regulatory pathways involved in oncogenesis^{7–10}.

For various anticancer agents, studies on candidate genes have identified splice variants mediating resistance in cellular models or associated with resistance in patients, and a few splicing factors have been involved in resistance^{11–14}. However, the AS regulatory pathways connecting splicing factors and AS events involved in anticancer drug resistance, are usually unknown. In two studies, the splicing factors PTBP1 and TRA2A were up-regulated in resistant cells and promoted resistance to gemcitabine in pancreatic cancer through AS regulation of the PKM gene, and to paclitaxel in triple-negative breast cancer through AS of RSRC2, respectively^{15,16}. In addition, very few studies identified genome-wide AS programs in resistant *versus* sensitive cells^{17,18}, and their role and upstream regulators were not identified. Thus, while AS regulation can play a role in anticancer drug resistance^{11–14}, AS regulatory pathways and genome-wide programs involved in anticancer drug resistance remain poorly understood.

To address this question, we studied ER+ breast cancer cell resistance to doxorubicin (Doxo), which is commonly used in chemotherapy for this cancer type. AS regulation by Doxo treatment in ER+ breast cancer cells has been previously analyzed in the context of acute response¹⁹, but not in the context of resistance. The classical model of acquired Doxo resistance in ER+ breast cancer is MCF7-Doxo-R cells²⁰. In this study, using this model, we identified two splicing factors that mediate Doxo resistance, as well as their genome-wide AS programs that converge on a common AS target involved in the resistance phenotype, thus identifying AS programs and pathways involved in breast cancer cell resistance to Doxo.

RESULTS

Acquired resistance to Doxo is accompanied by widespread AS regulation

To identify splicing variants regulated in a breast cancer cell model of acquired resistance to Doxo, we used the previously described MCF7-Doxo-R cells that are about 200 times less sensitive to Doxo when compared to parental MCF-7 cells²⁰. Stringent genome-wide analyses using exon-junction arrays and RNA-seq identified 639 and 1723 exonic events, respectively, that are regulated between MCF7-Doxo-R cells and parental MCF-7 cells (Tables S1 and S2). The most prominent type of exonic regulations identified were single-exon skipping (ASE), but many regulations of multiple-exon skipping, alternative 3' and 5' splice sites, retained introns, mutually exclusive exons, alternative last exons, and alternative first exons were also identified (Fig. 1A,B). Exonic regulations were enriched in several functions, mainly related to cytoskeleton, transcription, cell cycle, DNA damage response (DDR), and cell death (Fig. 1C). There was little overlap with the previously identified set of exons regulated in acute response to Doxo in MCF-7 cells¹⁹ (data not shown). We validated many AS regulations by RT-PCR (Fig. 1D and see below).

An EMT-related splicing switch accompanies, but does not directly explain Doxo resistance

Given the large number of splice variants that were regulated in Doxo resistance, we looked for upstream regulators. For this, we first looked for known RNA motifs corresponding to known RNA-binding proteins (RBPs) and enriched in our dataset of ASEs regulated in resistance. The most enriched motif was an RBFOX binding motif (Fig. 2A). Interestingly, a switch between RBFOX2 and ESRP1/2 splicing factors has been involved in EMT²¹, and both EMT and the related CSC phenotype were previously found associated with Doxo resistance in breast cancer cell models and tumors^{1,22,23}. Consistently, in comparison to MCF-7 cells, MCF7-Doxo-R cells displayed a strong decrease of several epithelial markers including CDH1 and ESRP1/2, and an increase in several mesenchymal markers including CDH2, vimentin and RBFOX2 (Fig. 2B). MCF7-Doxo-R cells also displayed CSC markers (CD44^{high}/CD24^{low}; data not shown). This was accompanied by a splicing switch of several known EMT-related splice variants (Fig. 2C). Thus, the switch between RBFOX2 and ESRP1/2 likely explains in part the widespread AS regulations in genes related to cytoskeleton and EMT-related functions that we noted in MCF7-Doxo-R *versus* MCF-7 cells (Fig. 1C, Fig. 2C and data not shown). However, depletion of either RBFOX2 in MCF7-Doxo-R cells or ESRP1 and 2 (alone or in combination) in MCF-7 cells did not affect Doxo sensitivity (Fig. 2D,E). While we cannot exclude that

depletion of RBFOX2 or ESRP1/2 for longer time than 5 days might affect Doxo sensitivity, these data suggest that these factors do not control AS programs directly involved in Doxo sensitivity.

Identification of the splicing factors ZRANB2 and SYF2 as mediators of Doxo resistance

To further investigate on splicing factors involved in Doxo resistance, we then identified by microarray analysis, about 40 splicing-related factors whose expression was either increased or decreased in MCF7-Doxo-R *versus* MCF7 cells (Table 1). We then carried out an siRNA screen on these factors. Among 22 factors down-regulated in resistant cells, the depletion of only one factor (SNRPA1) increased MCF-7 cell survival to Doxo, but a strong effect was also seen in the absence of Doxo, thus lending little credence to this potential hit (Fig. 3A and Suppl. Fig. S1A). Meanwhile, among 19 factors up-regulated in resistant cells, the depletion of four factors (ZRANB2, SYF2, IGF2BP1 and QKI) decreased MCF7-Doxo-R cell survival to Doxo, with little or no impact on cell survival in the absence of Doxo (Fig. 3B and Suppl. Fig. S1B). We did not pursue on IGF2BP1, whose primary function is the regulation of cytoplasmic mRNA fate and translation²⁴, and we could not confirm a robust effect of QKI in further experiments (data not shown). Further analysis with independent siRNAs confirmed that depletion of ZRANB2 and SYF2 in MCF7-Doxo-R cells decreased their survival to Doxo (Fig. 3C,D). In addition, both RT-qPCR and Western blot analyses confirmed ZRANB2 overexpression at both RNA and protein levels in MCF7-Doxo-R cells compared to MCF-7 cells (Fig. 3E). In the case of SYF2, its over-expression in MCF7-Doxo-R cells could be confirmed by RT-qPCR, but not by Western blot, which is less quantitative (Fig. 3E).

Identification of exons regulated by ZRANB2

ZRANB2 is well known to bind RNA and was shown to regulate AS of several genes²⁵, but little is known about its biological function and its AS targets genome-wide. To identify the AS events it regulates in MCF7-Doxo-R cells, we analyzed the effects of its depletion using two independent siRNAs. This analysis identified 78 exons regulated by both siRNAs, 51% of which were ASEs (Fig. 4A and Table S3). Remarkably, 55 (71%) of the 78 events regulated by ZRANB2 were also regulated in MCF7-Doxo-R *versus* MCF-7 cells (Fig. 4B), which is consistent

with ZRANB2 overexpression in MCF7-Doxo-R cells. The AS regulations by ZRANB2 were enriched in genes involved in cell death, microtubule dynamics, cell cycle and DNA repair (Fig. 4C), and we validated several of them by RT-PCR (Fig. 4D).

The core spliceosome component SYF2 controls AS of specific genes

SYF2 is a component of the spliceosome²⁶. To identify the AS events regulated by SYF2 in MCF7-Doxo-R cells, we analyzed the effects of its depletion using two independent siRNAs. This analysis identified 79 exons regulated by both siRNAs (Table S4). Remarkably, the main type of AS event regulated by SYF2 was A3SS (49% A3SS, 27% ASE, 12% A5SS; Fig. 5A), which is in sharp contrast with ZRANB2 (only 7% A3SS, 51% ASE and 27% A5SS; Fig. 4A) and with most splicing factors studied so far. 32 (41%) of the 79 events regulated by SYF2 were also found to be regulated in MCF7-Doxo-R *versus* MCF-7 cells (Fig. 5B). This proportion is lower than the 71% overlap that we observed between ZRANB2 and resistance datasets (Fig. 4B), and this is consistent with the milder overexpression of SYF2 compared to ZRANB2 in MCF7-Doxo-R cells as noted above (Fig. 3E). The AS regulations by SYF2 were enriched in genes involved in transcription, cell death, cell cycle and DNA repair (Fig. 5C), and we validated several A3SS regulations by RT-PCR (Fig. 5D).

ZRANB2 and SYF2 splicing programs converge on the ECT2-Ex5+ splice variant

We then investigated, whether the shared ability of ZRANB2 and SYF2 to promote Doxo resistance may be mediated by common splicing targets. Although they tended to regulate different types of AS events as noted above and as evidenced by the little overlap between their regulations (Suppl. Fig. S2A), we identified an AS event in the ECT2 gene that was robustly regulated in a similar manner by both factors (Fig. 6A-C). ECT2 exon 5 inclusion level was higher in MCF7-DoxoR than in MCF7 cells, and was reduced by depletion of either ZRANB2 or SYF2 (Fig. 6C).

To determine whether regulation of ECT2 Ex5 inclusion by ZRANB2 and SYF2 may be direct, we analyzed their association with ECT2 pre-mRNA by RIP-qPCR. We found a significant association of ZRANB2 with ECT2 pre-mRNA at the level of intron 5 (Fig. 6D), which is consistent with a public eCLIP dataset available in ENCODE (Suppl. Fig. S2B). Meanwhile,

SYF2 was found associated around the 3' splice sites of exons 5 and 6 (Fig. 6E), which is consistent with a factor that mainly regulates alternative 3' splice sites as noted above (Fig. 5A). These data suggest direct regulation of ECT2 Ex5 inclusion by both ZRANB2 and SYF2 acting through different mechanisms.

ZRANB2, SYF2 and the ECT2-Ex5+ isoform promote Doxo resistance through S phase arrest

To determine whether ECT2 Ex5 inclusion may play a role in Doxo resistance, we then transfected MCF7-Doxo-R cells with an siRNA targeting this exon (Suppl. Fig. S2C). Depletion of the ECT2 Ex5+ isoform led to a 25% decrease of MCF7-Doxo-R cell survival to Doxo 150 μ M, which represents between one half and one third of the effects of SYF2 and ZRANB2 depletion (Fig. 7A). Depletion of the ECT2 Ex5+ isoform had no impact on cell growth in the absence of Doxo (data not shown). The depletion of several other splice variants had little or no impact on MCF7-Doxo-R cell survival to Doxo 150 μ M (Fig. 7A).

To better understand the resistance phenotype controlled by ZRANB2, SYF2 and the ECT2 Ex5+ isoform, we analyzed cell cycle and cell death. In MCF-7 and other sensitive cells, Doxo typically induces G2-M and/or G1 arrest and cell death¹⁹. Treatment of MCF7-Doxo-R cells with Doxo at 150 μ M for 48 hours induced a striking decrease of the proportion of cells in G1 phase (from 50 to 14%) that could not be explained by the modest increases in G2-M and sub-G1 (cell death), but instead mainly resulted from a massive accumulation of cells in S phase (from 20 to 45%; Fig. 7B). These data indicate that the resistance phenotype of MCF7-Doxo-R cells is linked to their ability to arrest in S phase in response to high doses of Doxo (which may favor replication-coupled repair and prevent cell death; see discussion).

Depletion of either ZRANB2, SYF2 or the ECT2-Ex5+ isoform in MCF7-Doxo-R cells had little effect on cell cycle and cell death in the absence of Doxo (Suppl. Fig. S3), but reduced the accumulation of cells in S phase in response to Doxo (Fig. 7C). We also noted that in the absence of Doxo, cell death and polyploidy were strongly induced by depletion of total ECT2, but were not induced by selective depletion of the ECT2-Ex5+ isoform, or by depletion of either ZRANB2 or SYF2 (Suppl. Fig. S3). These data indicate that ZRANB2, SYF2 and the ECT2-Ex5+ isoform promote the Doxo-induced S-phase arrest phenotype of resistant cells.

The ECT2-Ex5+ splice variant is associated with resistance to chemotherapy in ER+ breast tumors

To determine whether AS of ECT2 may have relevance to chemotherapy resistance in breast cancer patients, we quantitated ECT2 isoforms by RT-qPCR in a large collection of clinically annotated breast tumors. Overall, when combining the different subtypes of breast cancer, higher levels of ECT2 Ex5 inclusion were associated with shorter survival in breast cancer patients that were treated with chemotherapy, but not in patients that were not treated with chemotherapy (Suppl. Fig. S4A). This association of higher ECT2 Ex5 inclusion levels with shorter survival specifically in cases of chemotherapeutic treatment, was even more pronounced in the ER+ subtype of breast cancer (Fig. 6D), but was not found in triple-negative breast cancer (Suppl. Fig. S4B). We also noted that total mRNA levels of ECT2 were associated with bad prognosis in ER+ breast cancer, whether or not patients were treated with chemotherapy (Suppl. Fig. S4C).

DISCUSSION

Recent studies have started to identify AS networks involved in oncogenesis, but little is known about AS regulatory pathways and genome-wide programs involved in resistance to anticancer treatments. In this study, we identify genome-wide AS programs associated with, as well as AS regulatory pathways controlling resistance of ER+ breast cancer cells to Doxo. These programs and pathways involve two poorly characterized splicing factors (ZRANB2 and SYF2), and converge on the ECT2-Ex5 AS event.

ECT2, also known as ARHGEF31, is a protooncogene with a well established role in the cytokinesis phase of the cell cycle²⁷. Interestingly, the exon 5 of ECT2 encodes a BRCT domain, which is often found in proteins involved in the DDR²⁸, and ECT2 was recently involved in cell death induction by DNA damage^{29,30}. Our data suggest that the ECT2-Ex5 isoform can promote resistance to Doxo through a role in S phase.

ZRANB2, also known as ZIS or ZNF265, is a ubiquitous RBP that was shown to interact with several splicing factors and a specific RNA motif, and to regulate AS of several minigenes, but little was known about its endogenous splicing targets and biological functions^{25,31,32}.

SYF2, also known as p29 or Ntc31, has a long-known dual function in both spliceosome and cell cycle, from yeast to human³³. In budding yeast, SYF2 mutation affects splicing of genes encoding tubulin, leading to spindle checkpoint activation³⁴. In human, SYF2 has been involved in DNA damage-induced cell cycle checkpoint activation, and was proposed to control the replication checkpoint through an association with MCM3 and PCNA during S phase^{35–37}. However, little was known about the splicing events controlled by human SYF2. Our findings strongly suggest that SYF2 promotes replication checkpoint through AS regulation of ECT2-Ex5. Thus, SYF2 appears to promote replication checkpoint through both splicing-dependent (this study) and independent mechanisms³⁵. Interestingly, within the spliceosome SYF2 is part of the yeast Prp19 complex and of the human Prp19-associated complex, two components of which also have a direct role in replication checkpoint activation or replication stress response^{26,38–40}. Thus, our findings may help understand the links between splicing regulation and cell-cycle checkpoint activation.

A more efficient replication checkpoint may favor replication-coupled repair and prevent cell death, as recently observed in cancer stem cells⁴¹. It remains to be determined, whether the AS regulatory pathways (involving ZRANB2, SYF2 and ECT2-Ex5) that we identify to promote replication checkpoint and genotoxic stress resistance in MCF7-Doxo-R cells, may be involved more generally in cancer stem cells.

Further identification of genome-wide AS programs associated with, as well as AS regulatory pathways controlling resistance to anticancer treatments, should help understand the complex gene regulatory networks involved in resistance. This approach should also help identify novel biomarkers of resistance, and potential therapeutic targets.

METHODS

Cell culture and treatment

Sensitive (MCF-7-Parental) and Doxo-resistant (MCF-7-Doxo-R) human breast carcinoma cells (Millour, Mol Cancer Ther 2011) were cultured in 4,5 g/L Glucose DMEM (Eurobio) supplemented with 10% (v/v) fetal bovine serum (Dutscher), and 2 mM L-Glutamine (Eurobio) in 5% CO₂ at 37°C. Doxo was obtained from Sigma (D-1515) and the resistant

phenotype was maintained by culturing the Doxo-R cells in 17 μM Doxo. For experiments, treatments in Doxo-R cells were done at 150 μM after one passage in the absence of Doxo.

siRNA transfections

siRNA pools used for the miniscreen were obtained by the cherry-pick library preparation service from Dharmacon. Individual siRNAs were from Eurogentec. Reverse transfection was performed at a final concentration of 100 nM with lipofectamine RNAiMax from ThermoFischer Scientific.

WST-1 cell survival assay

Cells were simultaneously reverse transfected with siRNA and seeded at an appropriate density in 96-well plates (3,000 cells/well for Parental and 5,000 cells/well for Doxo-R cells) on day 1, in 6 replicates per siRNA condition. 48 hours later on day 3, triplicates were treated with 150 μM Doxo. 72 hours later on day 6, cell survival was assayed using WST-1 (Sigma) according to the manufacturer's instructions.

RNA extraction and RT-(q)PCR

RNA from whole cells was extracted using TRIzol Reagent (ThermoFischer Scientific), and 1 μl of GlycoBlue (ThermoFisher Scientific) was added for RNA precipitation. Total RNA was treated with DNase I (TURBO DNasefree, ThermoFisher Scientific). Reverse transcription was performed using SuperScript™ Reverse Transcriptase (ThermoFisher Scientific) and random primers. PCR was performed using GoTaq Flexi DNA Polymerase (Promega), and PCR products were migrated on agarose gels. qPCR was performed using Power SYBR Green PCR Master Mix (ThermoFisher Scientific) on a CFX96 Real-Time PCR Detection System (BioRad).

RIP-qPCR

For RIP-qPCR, antibody was hybridized to Dynabeads Protein G (ThermoFisher Scientific) rotating at 4°C overnight before cell harvesting. RNA-protein complexes were crosslinked with a 1% formaldehyde incubation at RT for 10 minutes. The crosslink reaction was quenched using 125 mM glycine for 5 minutes at RT. Cells were washed twice with ice-cold PBS before harvesting and pelleted by centrifugation (4°C, 5 minutes, 800 g). Cells were resuspended in 1 ml RIPA (Sigma) with 1 μl RNase out and cOmplete™ Protease Inhibitor Cocktail (Sigma) at 1X concentration, and cell lysates were sonicated for 5 minutes. Supernatant was collected after centrifugation at 4°C for 10 minutes (10 000 g). One tenth of

the IP input was used for RNA input and was treated with proteinase K before RNA extraction with TRIzol Reagent (ThermoFisher Scientific). 400 µg of protein was used for each IP condition, and incubated with antibody-hybridized Dynabeads overnight rotating at 4°C. Supernatant was removed and the beads were washed twice with RIPA buffer before RNA-protein complex elution by incubation with elution buffer (Tris-HCl pH8 100 mM; Na₂-EDTA 10 mM; 1% SDS in H₂O) 3 minutes at 90°C. Protein was degraded with proteinase K treatment and RNA was extracted with TRIzol Reagent (ThermoFisher Scientific) for RT-qPCR analysis.

Western-blot

For immunoblotting, cells were lysed in RIPA buffer (Sigma) after two PBS wash and cell lysate was sonicated for 5 minutes. Supernatant was collected after centrifugation at 4°C for 10 minutes (10 000 g). Protein concentration was determined using the Pierce™ BCA protein assay kit from ThermoFisher Scientific and BSA standards from BioRad. Proteins were then separated in 4-12% NuPAGE® Bis-Tris precast gels (ThermoFisher Scientific) and transferred on nitrocellulose membrane using the iBlot2 dry transfer system from ThermoFisher Scientific. Membranes were incubated overnight at +4°C with primary antibodies and proteins were detected using horse-radish peroxidase-conjugated goat anti-mouse or anti-rabbit antibodies. After washing, the blots were revealed using Clarity™ Western ECL substrate (BioRad), and a ChemiDoc™ gel imaging system from BioRad.

Cell cycle analysis

For cell cycle analysis by FACS, culture media of each condition was removed and kept aside, and cells were washed twice with PBS. After trypsinization, cells were resuspended in their culture media and centrifuged at RT for 5 minutes (700 g). Supernatant was removed and cells were washed twice with PBS. For fixation, 1ml of cold 70% ethanol was added and cell pellets were incubated for at least 30 minutes at -20°C. Cells were then washed twice with ice cold PBS and resuspended in Vindelov solution (Tris-HCl 3,5 mM; NaCl 10 mM; IGEPAL o,1%) with RNase A (20 µg/ml) and propidium iodine (50 µg/ml) and incubated at RT in the dark for 30 minutes. Flow-cytometry analysis of 20 000 cells was performed on a FACSCanto™ flow cytometer from BD Biosciences. Data analysis was done on the FlowJo software.

Exon-junction array analyses

One μg of total RNA was processed and hybridized to Glue Grant Human Transcriptome arrays (Affymetrix) according to the manufacturer's instructions. The experiments, from cell treatment to array hybridization, were performed three times. Data were analyzed with Expression Console (Affymetrix) to perform quality assessment and normalized using quintile normalization. Background correction and probe selection were performed as previously described, using exon annotation from the Faster DB database¹⁹ (<http://fasterdb.lyon.unicancer.fr/>). Exonic regulation between two conditions was analyzed using the splicing index method in three independent ways, based either on probes in all annotated exons, on probes in annotated alternative exons and their neighbours, or on junction probes. Only the regulation events found by two methods (with fold change >1.8 and p value <0.05) were selected. Exonic regulations were classified into different types (ASE, etc.) based on annotation of known events. Data were visualized in gene context using the Elexir tool of Faster DB.

RNA-seq analyses

Libraries were made using one μg of total RNA and the TruSeq Stranded mRNA Library Preparation Kit (Illumina). Equimolar pool of libraries were sequenced on a Illumina HiSeq 2500 machine using paired-ends reads (PE, 2x101bp) and High Output run mode allowing to get 200 millions of raw reads per sample. Raw reads were mapped on the human reference genome hg19 using the STAR aligner (v.2.5.0a). PCR-duplicated reads and low mapping quality reads (MQ <20) were removed using Picard tools and SAMtools, respectively. We next used rMATS (v3.0.9), an event-based tool, to identify differentially spliced events using RNA-seq data⁴². Five distinct alternative splicing events were analyzed using rMATS: skipped exons (SE), alternative 3' splice sites (A3SS), alternative 5' splice sites (A5SS), mutually exclusive exons (MXE) and retained introns (RI). Briefly, rMATS uses a counts-based model, to calculate percent of spliced-in (PSI) value among replicates, using both spliced reads and reads that mapped the exon body. We used three different thresholds to identify differentially spliced events between two groups: splicing event has to be (i) supported by at least 15 unique reads, (ii) $|\Delta\text{PSI}| > 10\%$; (iii) FDR < 0.05 . Motif enrichment analysis was performed using the rMAPS software⁴³.

REFERENCES

1. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: An evolving paradigm. *Nat. Rev. Cancer* **13**, 714–726 (2013).
2. Wijdeven, R. H., Pang, B., Assaraf, Y. G. & Neefjes, J. Old drugs, novel ways out: Drug resistance toward cytotoxic chemotherapeutics. *Drug Resist. Updat.* (2016). doi:10.1016/j.drug.2016.07.001
3. Fu, X. D. & Ares, M. Context-dependent control of alternative splicing by RNA-binding proteins. *Nature Reviews Genetics* (2014). doi:10.1038/nrg3778
4. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* (2014). doi:10.1038/onc.2013.533
5. David, C. J. & Manley, J. L. Alternative pre-mRNA splicing regulation in cancer: Pathways and programs unhinged. *Genes Dev.* **24**, 2343–2364 (2010).
6. Dvinge, H., Kim, E., Abdel-Wahab, O. & Bradley, R. K. RNA splicing factors as oncoproteins and tumour suppressors. *Nat. Rev. Cancer* **16**, 413–430 (2016).
7. Kim, E. *et al.* SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell* (2015). doi:10.1016/j.ccell.2015.04.006
8. Shirai, C. L. *et al.* Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. *Cancer Cell* (2015). doi:10.1016/j.ccell.2015.04.008
9. Visconte, V. *et al.* Distinct iron architecture in SF3B1-mutant myelodysplastic syndrome patients is linked to an SLC25A37 splice variant with a retained intron. *Leukemia* (2015). doi:10.1038/leu.2014.170
10. Jin, S. *et al.* Splicing factor SF3B1K700E mutant dysregulates erythroid differentiation via aberrant alternative splicing of transcription factor TAL1. *PLoS One* (2017). doi:10.1371/journal.pone.0175523
11. Dehm, S. M. mRNA splicing variants: Exploiting modularity to outwit cancer therapy. *Cancer Res.* **73**, 5309–5314 (2013).
12. Pagliarini, V., Naro, C. & Sette, C. Splicing regulation: A molecular device to enhance cancer cell adaptation. *BioMed Research International* (2015). doi:10.1155/2015/543067
13. Siegfried, Z. & Karni, R. The role of alternative splicing in cancer drug resistance. *Curr. Opin. Genet. Dev.* **48**, 16–21 (2018).
14. Wang, B.-D. & Lee, N. Aberrant RNA Splicing in Cancer and Drug Resistance. *Cancers (Basel)*. **10**, 458 (2018).
15. Calabretta, S. *et al.* Modulation of PKM alternative splicing by PTBP1 promotes gemcitabine resistance in pancreatic cancer cells. *Oncogene* (2016). doi:10.1038/onc.2015.270
16. Liu, T. *et al.* TRA2A Promoted Paclitaxel Resistance and Tumor Progression in Triple-Negative Breast Cancers via Regulating Alternative Splicing. *Mol. Cancer Ther.* (2017). doi:10.1158/1535-7163.MCT-17-0026
17. Mohamed, A. M. *et al.* Oncogene- and drug resistance-associated alternative exon

- usage in acute myeloid leukemia (AML). *Oncotarget* (2016). doi:10.18632/oncotarget.3898
18. Sciarrillo, R. *et al.* Using RNA-sequencing to Detect Novel Splice Variants Related to Drug Resistance in *In Vitro* Cancer Models. *J. Vis. Exp.* 1–13 (2016). doi:10.3791/54714
 19. Dutertre, M. *et al.* A recently evolved class of alternative 3'-terminal exons involved in cell cycle regulation by topoisomerase inhibitors. *Nat. Commun.* **5**, 1–12 (2014).
 20. Millour, J. *et al.* ATM and p53 Regulate FOXM1 Expression via E2F in Breast Cancer Epirubicin Treatment and Resistance. *Mol. Cancer Ther.* (2011). doi:10.1158/1535-7163.MCT-11-0024
 21. Braeutigam, C. *et al.* The RNA-binding protein Rbfox2: An essential regulator of EMT-driven alternative splicing and a mediator of cellular invasion. *Oncogene* (2014). doi:10.1038/onc.2013.50
 22. Scheel, C. & Weinberg, R. A. Cancer stem cells and epithelial-mesenchymal transition: Concepts and molecular links. *Semin. Cancer Biol.* **22**, 396–403 (2012).
 23. Dean, M., Fojo, T. & Bates, S. Tumour stem cells and drug resistance. *Nature Reviews Cancer* (2005). doi:10.1038/nrc1590
 24. Bell, J. L. *et al.* Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): Post-transcriptional drivers of cancer progression? *Cellular and Molecular Life Sciences* (2013). doi:10.1007/s00018-012-1186-z
 25. Mangs, A. H. & Morris, B. J. ZRANB2: Structural and functional insights into a novel splicing protein. *International Journal of Biochemistry and Cell Biology* (2008). doi:10.1016/j.biocel.2007.08.007
 26. Chanarat, S. & Sträßer, K. Splicing and beyond: The many faces of the Prp19 complex. *Biochimica et Biophysica Acta - Molecular Cell Research* (2013). doi:10.1016/j.bbamcr.2013.05.023
 27. Basant, A. & Glotzer, M. Spatiotemporal Regulation of RhoA during Cytokinesis. *Current Biology* (2018). doi:10.1016/j.cub.2018.03.045
 28. Zou, Y. *et al.* Crystal structure of triple-BRCT-domain of ECT2 and insights into the binding characteristics to CYK-4. *FEBS Lett.* (2014). doi:10.1016/j.febslet.2014.07.019
 29. Srougi, M. C., Burridge, K. & Hotchin, N. The Nuclear Guanine Nucleotide Exchange Factors Ect2 and Net1 Regulate RhoB-Mediated Cell Death after DNA Damage. *PLoS One* (2011). doi:10.1371/journal.pone.0017108
 30. He, D., Xiang, J., Li, B. & Liu, H. The dynamic behavior of Ect2 in response to DNA damage. *Sci. Rep.* (2016). doi:10.1038/srep24504
 31. Loughlin, F. E. *et al.* The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5' splice site-like sequences. *Proc. Natl. Acad. Sci.* (2009). doi:10.1073/pnas.0802466106
 32. Yang, Y. H. J. *et al.* ZRANB2 localizes to supraspliceosomes and influences the alternative splicing of multiple genes in the transcriptome. *Mol. Biol. Rep.* (2013). doi:10.1007/s11033-013-2637-9

33. Ben-Yehuda, S. *et al.* Genetic and physical interactions between factors involved in both cell cycle progression and pre-mRNA splicing in *Saccharomyces cerevisiae*. *Genetics* (2000). doi:10.1515/cclm.2011.119
34. Dahan, O. & Kupiec, M. Mutations in genes of *Saccharomyces cerevisiae* encoding pre-mRNA splicing factors cause cell cycle arrest through activation of the spindle checkpoint. *Nucleic Acids Res.* (2002).
35. Chu, P. C. *et al.* Silencing of p29 affects DNA damage responses with UV irradiation. *Cancer Res.* (2006). doi:10.1158/0008-5472.CAN-05-3229
36. Chu, P. C. *et al.* Involvement of p29 in DNA damage responses and Fanconi anemia pathway. *Carcinogenesis* (2009). doi:10.1093/carcin/bgp204
37. Chen, C. H. *et al.* Disruption of murine mp29/Syf2/Ntc31 gene results in embryonic lethality with aberrant checkpoint response. *PLoS One* (2012). doi:10.1371/journal.pone.0033538
38. Maréchal, A. *et al.* PRP19 Transforms into a Sensor of RPA-ssDNA after DNA Damage and Drives ATR Activation via a Ubiquitin-Mediated Circuitry. *Mol. Cell* (2014). doi:10.1016/j.molcel.2013.11.002
39. Zhang, N., Kaur, R., Akhter, S. & Legerski, R. J. Cdc5L interacts with ATR and is required for the S-phase cell-cycle checkpoint. *EMBO Rep.* (2009). doi:10.1038/embor.2009.122
40. Wan, L. & Huang, J. The PSO4 protein complex associates with replication protein a (RPA) and modulates the activation of ataxia telangiectasia-mutated and RAD3-related (ATR). *J. Biol. Chem.* (2014). doi:10.1074/jbc.M113.543439
41. Manic, G. *et al.* Replication stress response in cancer stem cells as a target for chemotherapy. *Semin. Cancer Biol.* **53**, 31–41 (2018).
42. Shen, S. *et al.* rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci.* (2014). doi:10.1073/pnas.1419161111
43. Park, J. W., Jung, S., Rouchka, E. C., Tseng, Y. T. & Xing, Y. rMAPS: RNA map analysis and plotting server for alternative exon regulation. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw410

Table 1: List of splicing factors regulated in MCF7-Doxo-R versus MCF7 cells.

Symbol	Description	Regul ^o	Fold change
IGF2BP2	insulin-like growth factor 2 mRNA binding protein 2	up	17,8
IGF2BP3	insulin-like growth factor 2 mRNA binding protein 3	up	14,1
SNRPN	small nuclear ribonucleoprotein polypeptide N	up	13,7
IGF2BP1	insulin-like growth factor 2 mRNA binding protein 1	up	12,3
RBM7	RNA binding motif protein 7	up	4,9
CPEB1	cytoplasmic polyadenylation element binding protein 1	up	4,8
ELL2	elongation factor, RNA polymerase II, 2	up	3,5
KHDRBS3	KH domain containing, RNA binding, signal transduction associated 3	up	3,5
DHX35	DEAH (Asp-Glu-Ala-His) box polypeptide 35	up	3,1
ZRANB2	zinc finger, RAN-binding domain containing 2	up	3,1
CWC22	CWC22 spliceosome-associated protein homolog (<i>S. cerevisiae</i>)	up	3,0
QKI	quaking homolog, KH domain RNA binding (mouse)	up	2,8
RBM11	RNA binding motif protein 11	up	2,8
SYF2	SYF2 homolog, RNA splicing factor (<i>S. cerevisiae</i>)	up	2,7
ELAVL2	ELAV (embryonic lethal, abnormal vision, <i>Drosophila</i>)-like 2 (Hu antigen B)	up	2,5
RBM9	RNA binding motif protein 9	up	2,4
MAGOH	mago-nashi homolog, proliferation-associated (<i>Drosophila</i>)	up	2,4
CELF2	CUGBP, Elav-like family member 2	up	2,3
DEK	DEK oncogene	up	2,1
RBM47	RNA binding motif protein 47	down	11,0
ESRP2	epithelial splicing regulatory protein 2	down	10,4
RBPMS	RNA binding protein with multiple splicing	down	8,0
BCAS1	breast carcinoma amplified sequence 1	down	6,8
TTF2	transcription termination factor, RNA polymerase II	down	5,4
JUP	junction plakoglobin	down	5,2
RBM23	RNA binding motif protein 23	down	5,2
TRIM24	tripartite motif-containing 24	down	4,7
LEO1	Leo1, Paf1/RNA polymerase II complex component, homolog (<i>S. cerevisiae</i>)	down	4,4
RBM24	RNA binding motif protein 24	down	4,3
SRPK2	SRSF protein kinase 2	down	3,5
PCBP3	poly(rC) binding protein 3	down	3,2
POLR2J	polymerase (RNA) II (DNA directed) polypeptide J, 13.3kDa	down	3,2
ZFP36	zinc finger protein 36, C3H type, homolog (mouse)	down	3,0
ZFP36L2	zinc finger protein 36, C3H type-like 2	down	3,0
BCAS2	breast carcinoma amplified sequence 2	down	3,0
SRRT	serrate RNA effector molecule homolog (<i>Arabidopsis</i>)	down	2,9
SNRPA1	small nuclear ribonucleoprotein polypeptide A'	down	2,7
HNRNPAB	heterogeneous nuclear ribonucleoprotein A/B	down	2,5
ESRP1	epithelial splicing regulatory protein 1	down	2,4
HNRNPUL1	heterogeneous nuclear ribonucleoprotein U-like 1	down	2,2
PRPF4	PRP4 pre-mRNA processing factor 4 homolog (yeast)	down	2,0

FIGURE LEGENDS

Figure 1: Acquired resistance to Doxo is accompanied by widespread AS regulation. A-B, Types of alternative exons regulated in MCF7-Doxo-R *versus* MCF7 cells in exon-junction array (A) and RNA-seq data (B). **C,** Enriched functions (Ingenuity Pathway Analysis) in genes with exonic regulations in RNA-seq data. **D,** RT-PCR validations of AS events.

Figure 2: Acquired resistance to Doxo is accompanied by a switch of an EMT-related splicing network. A, The RBFOX binding motif is enriched around ASEs (cassette exons) regulated in MCF7-Doxo-R *versus* MCF7 cells. **B-C,** Expression regulation of EMT-related genes and splicing factors (B, qRT-PCR data) and splice variants (C, RT-PCR) between MCF7-Doxo-R and MCF7 cells. **D-E,** Depletion of RBFOX2 in MCF7-Doxo-R cells (D) or ESRPs in MCF7 cells (E) does not affect Doxo sensitivity.

Figure 3: An RNAi screen on splicing factors identifies ZRANB2 and SYF2 as mediators of Doxo resistance. A-B, RNAi screens in MCF7 (A) and MCF7-Doxo-R cells (B). Following transfection and two-day recovery, cells were grown for three days with or without Doxo, and cell survival was assessed (WST1). **C-D,** Validation of the effects of siRNAs on ZRANB2 (C) and SYF2 (D) on their protein levels (top panels) and Doxo sensitivity (bottom panels) in MCF7-Doxo-R cells. **E,** Analysis of ZRANB2 and SYF2 expression levels in MCF7-Doxo-R and MCF7 cells at the protein (top panel) and RNA levels (bottom panel).

Figure 4: Identification of exons regulated by ZRANB2. A, Types of AS events regulated by ZRANB2 depletion in MCF7-Doxo-R cells. **B,** Overlap with exons associated with Doxo resistance. **C,** Enriched functions (Ingenuity Pathway Analysis) in genes with exonic regulations by ZRANB2. **D,** RT-PCR validations of exonic regulations by ZRANB2.

Figure 5: The spliceosome component SYF2 controls AS of specific genes. A, SYF2 depletion in MCF7-Doxo-R cells mainly regulates alternative 3' splice sites. **B,** Overlap with exons associated with Doxo resistance. **C,** Enriched functions (Ingenuity Pathway Analysis) in genes with exonic regulations by SYF2. **D,** RT-PCR validations of A3SS regulations by SYF2.

Figure 6: ZRANB2 and SYF2 splicing networks converge on the ECT2-Ex5 splice variant. A, Overlap between Doxo-R associated splicing events regulated by ZRANB2 and SYF2. **B,**

Schematics of the ECT2 gene around exon 5. **C**, RT-PCR validation of ECT2-Ex5 AS regulation by ZRANB2 and SYF2 depletions, and in DoxoR vs MCF7 cells. Quantitations and representative gels are shown. **D-E**, RIP-qPCR analysis of ZRANB2 (D) and SYF2 (E) association with ECT2 pre-mRNA at the indicated locations.

Figure 7: The ECT2-Ex5 splice variant regulates Doxo sensitivity. **A**, Effect of siRNAs targeting ECT2-Ex5 and other alternative exons, on Doxo sensitivity in MCF7-Doxo-R cells. **B**, Effects of Doxo treatment at 150 μ M for 48 hours on cell cycle in MCF7-Doxo-R cells. **C**, Effects of ZRANB2, SYF2 and ECT2-Ex5 depletions on MCF7-Doxo-R cell cycle in the absence and presence of Doxo. **D**, High inclusion levels of ECT2-Ex5 (high V1/V2 ratio) in ER+ (HR+ ERBB2-) breast tumors correlate with bad prognosis in patients treated with chemotherapy. Kaplan-Meier curves. MFS, metastasis-free survival.

SUPPLEMENTARY FIGURES

Figure S1: Effects of splicing factor depletion on cell survival without Doxo. **A**, MCF7 cells. **B**, MCF7-Doxo-R cells.

Figure S2: **A**, Overlap between ZRANB2 and SYF2 splicing programs. **B**, ZRANB2 eCLIP data (from an ENCODE dataset) for the ECT2 and MPRIP genes. The regulated exons are circled. **C**, RT-PCR validation of siRNAs targeting specific exons in ECT2 and MPRIP.

Figure S3 : In the absence of Doxo, MCF7-Doxo-R cell cycle is affected by depletion of ECT2 total, but not of ECT2-Ex5, ZRANB2 or SYF2.

Figure S4 : Prognostic analysis of ECT2 RNA levels measured in breast tumors before treatment. **A**, Breast tumors of all types. V1/V2, ECT2-Ex5 inclusion levels. Patients with (left) or without (right) chemotherapy. **B**, Triple-negative breast tumors with chemotherapy treatment. ECT2-Ex5 inclusion levels (left) and total mRNA levels (right). **C**, ER+ breast tumors (HR+ ERBB2-). Total ECT2 mRNA levels. Patients with (left) or without (right) chemotherapy.

SUPPLEMENTARY TABLES

Table S1: List of exons regulated in MCF7-Doxo-R versus MCF7 cells (exon-junction arrays).

Table S2: List of exons regulated in MCF7-Doxo-R versus MCF7 cells (RNA-seq).

Table S3: List of exons regulated by ZRANB2 depletion in MCF7-Doxo-R (RNA-seq).

Table S4: List of exons regulated by SYF2 depletion in MCF7-Doxo-R (RNA-seq).

Figure 1

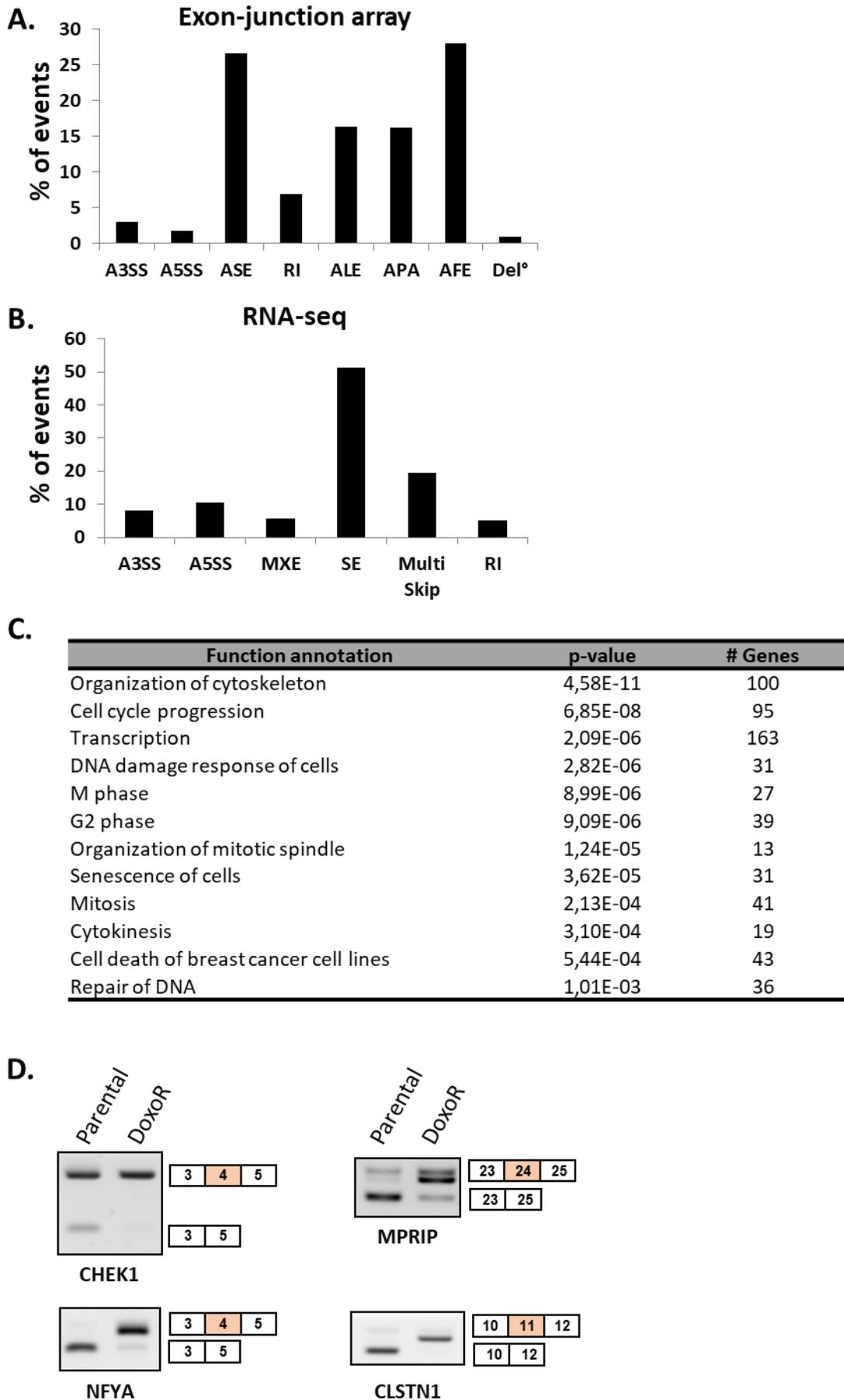


Figure 2

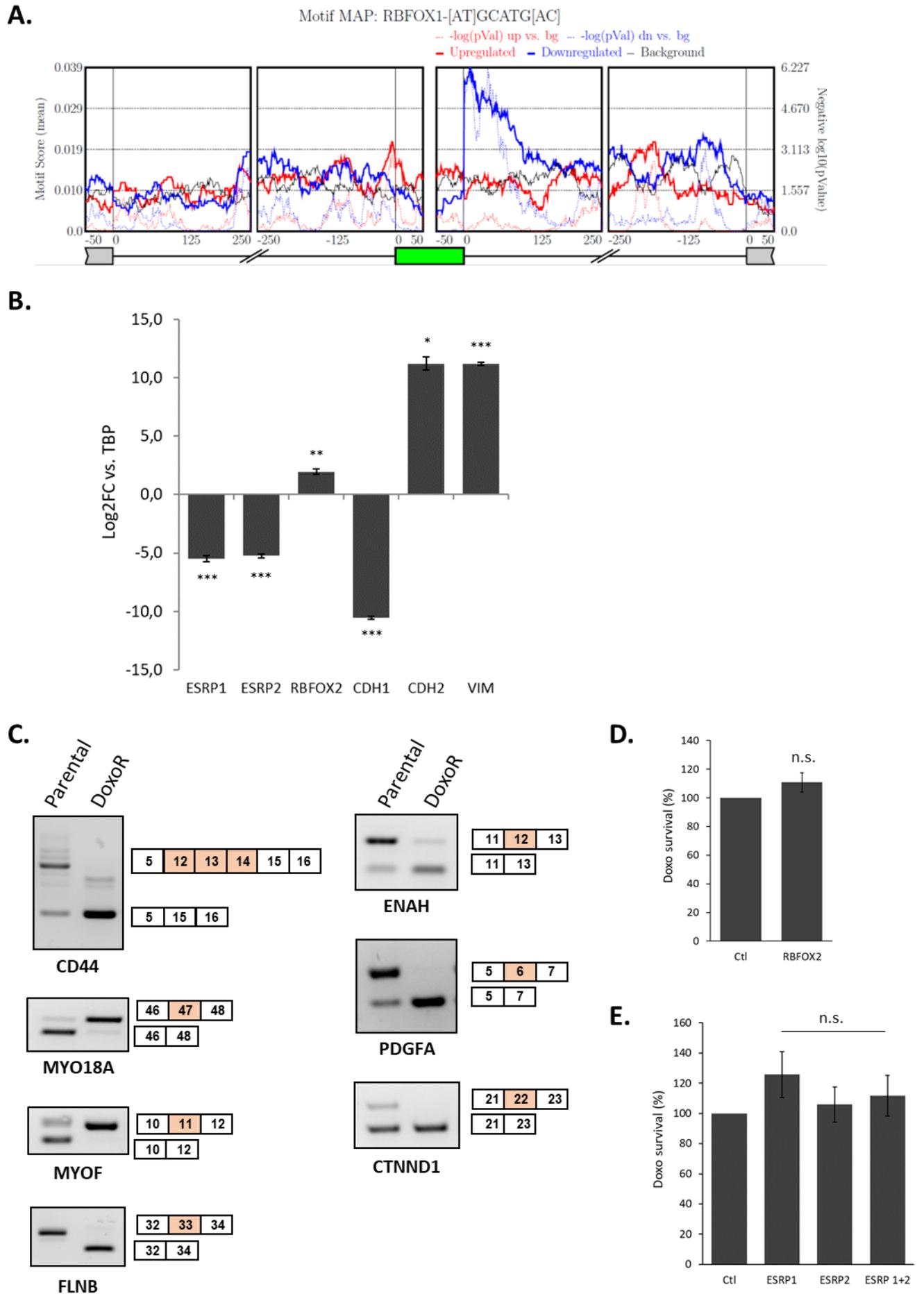


Figure 3

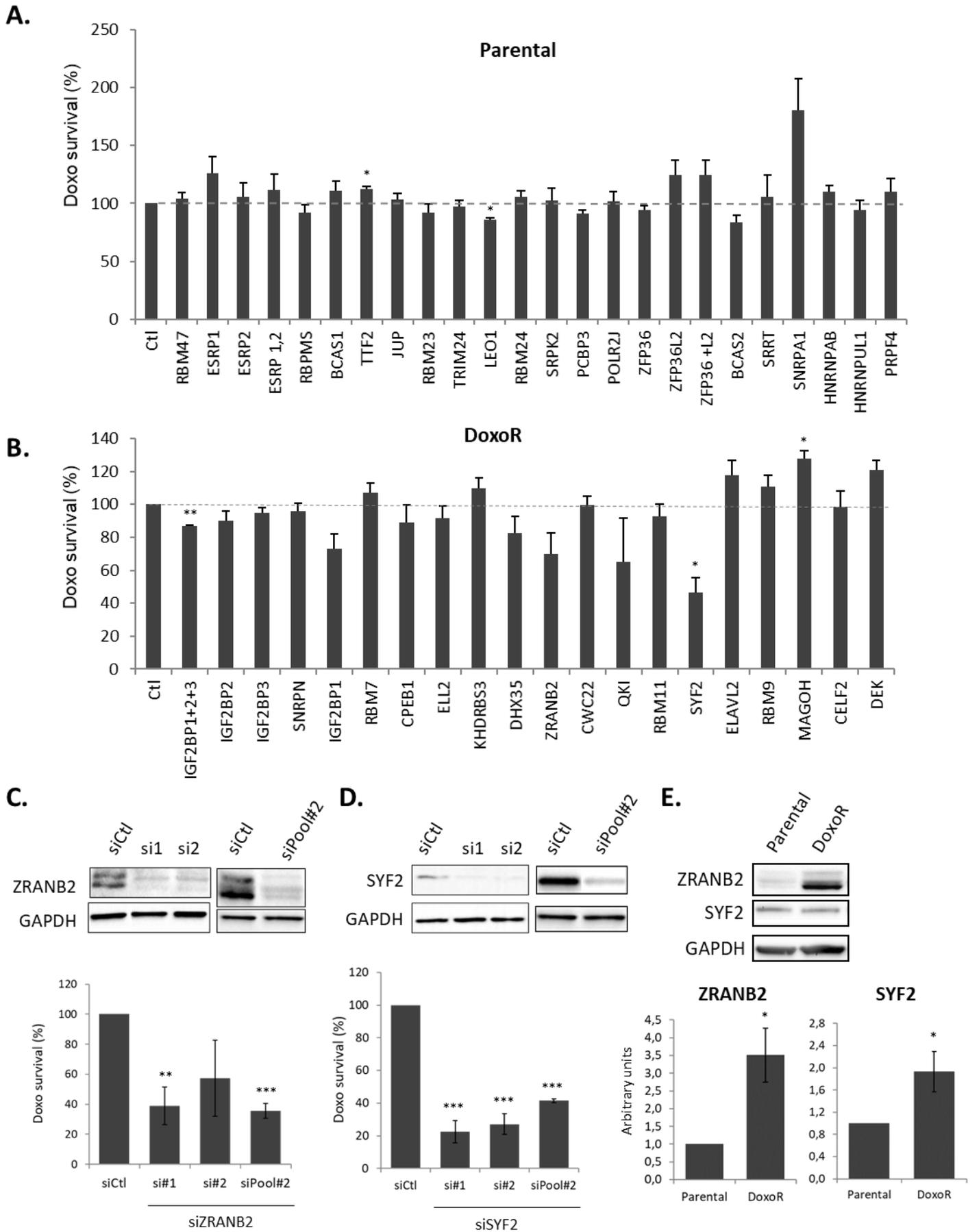
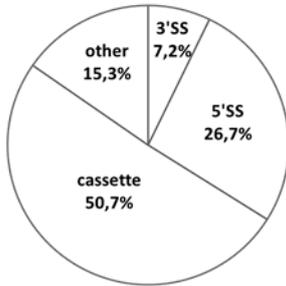


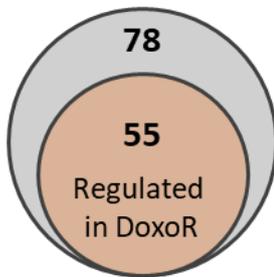
Figure 4

A.



B.

Regulated by ZRANB2



C.

Functions Annotation	p-value	# Genes
Apoptosis	4,95E-02	15
Organization of cytoskeleton	2,48E-02	8
Microtubule dynamics	1,73E-02	7
Invasion of carcinoma cell lines	4,21E-03	5
Formation of cellular protrusions	3,00E-02	5
G2 phase	4,05E-02	4
Double-stranded DNA break repair	1,92E-02	3

D.

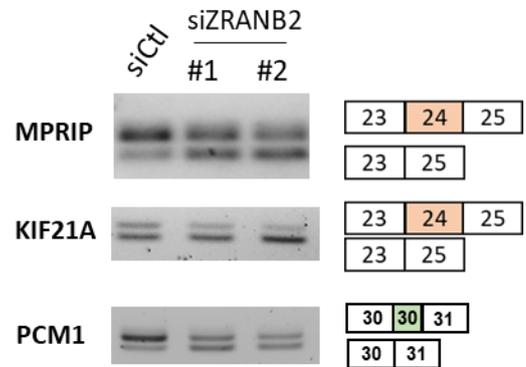
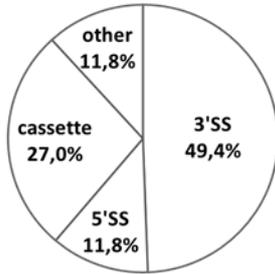


Figure 5

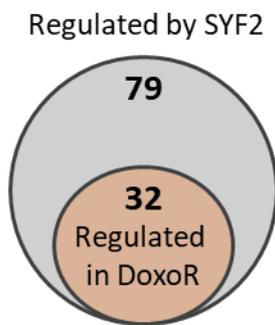
A.



C.

Functions Annotation	p-value	# Genes
Transcription	2,07E-03	17
Necrosis	4,51E-02	16
Cell death of tumor cell lines	3,56E-02	14
Activation of DNA endogenous promoter	2,85E-03	11
Organization of cytoplasm	2,20E-02	9
Cell cycle progression	3,56E-02	8
Repair of DNA	1,52E-02	5
Metabolism of DNA	2,98E-02	4

B.



D.

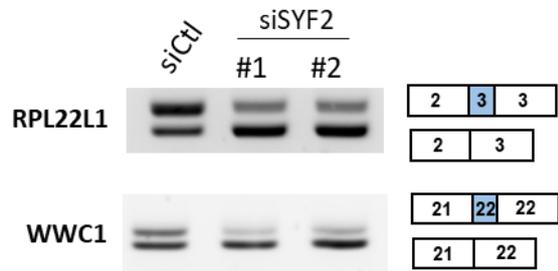


Figure 6

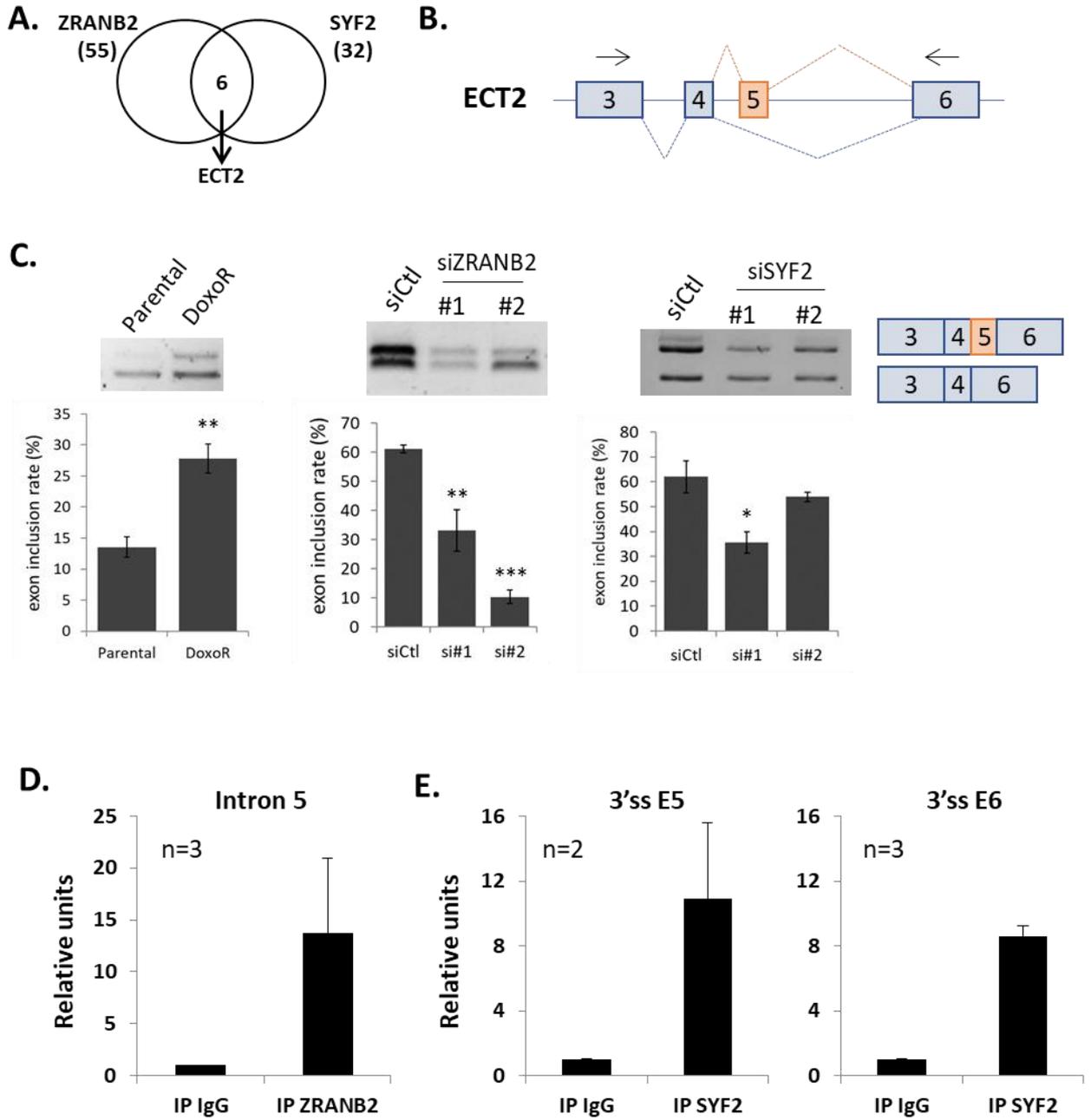


Figure 7

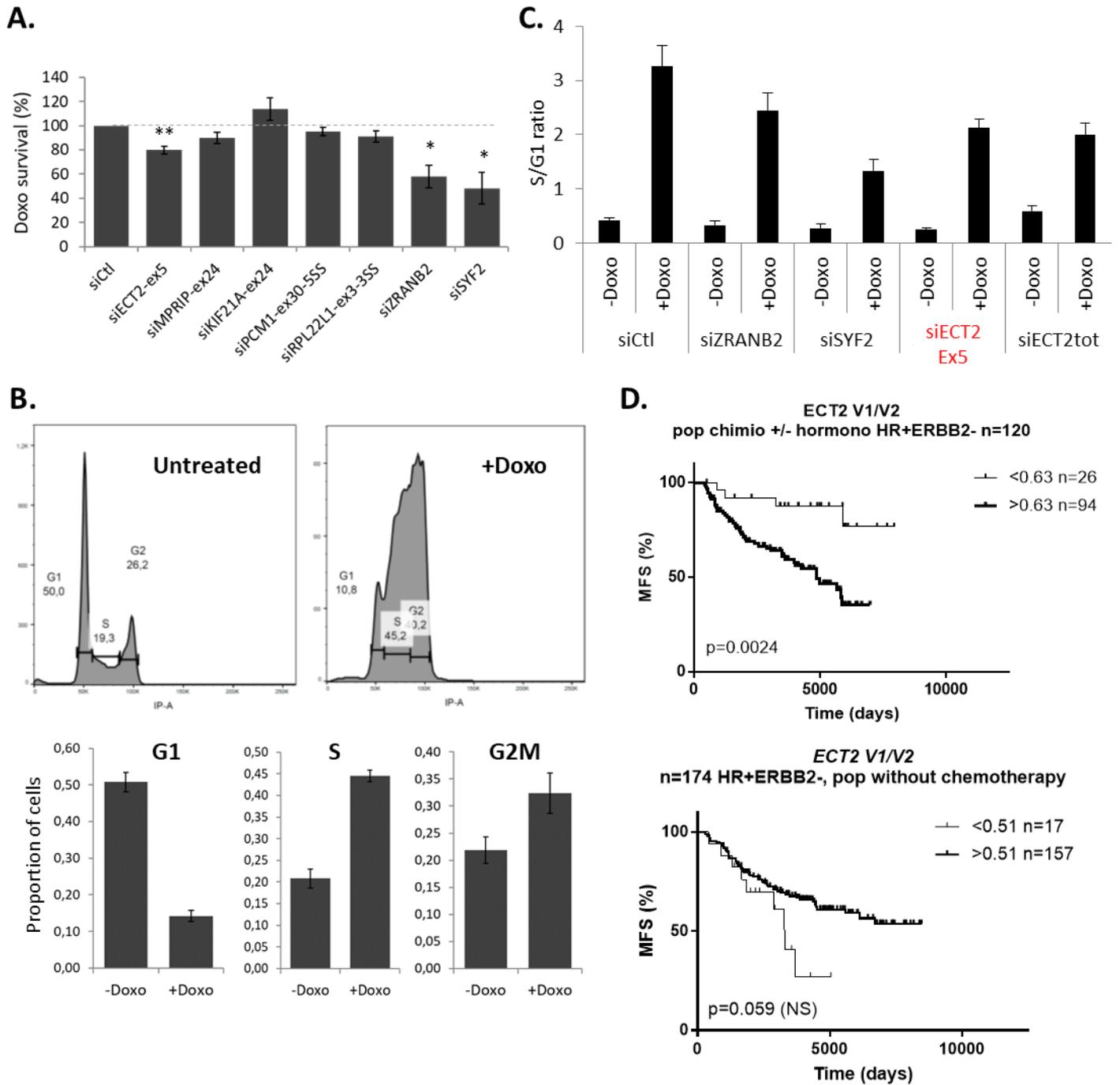
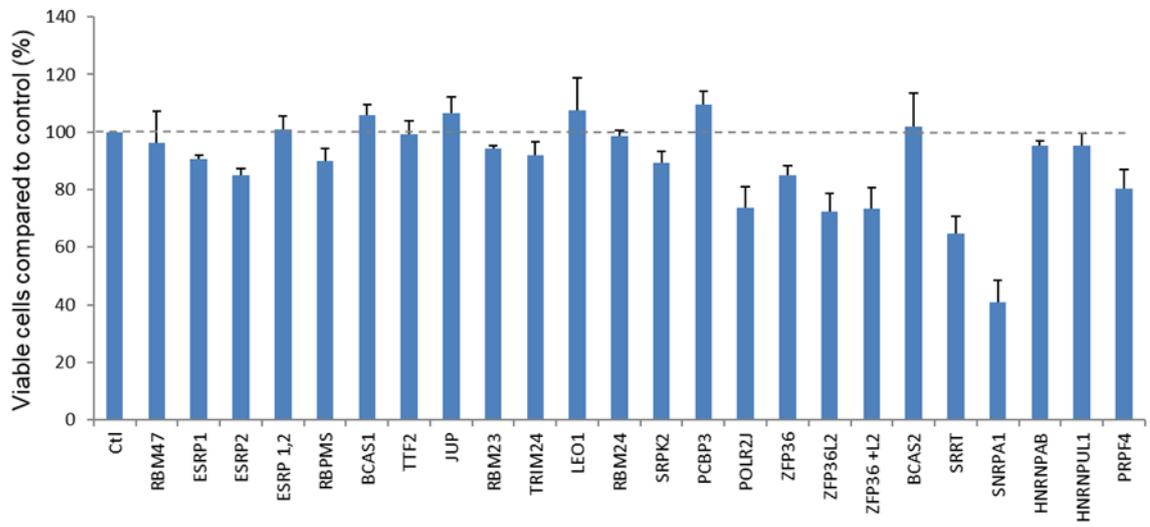


Figure S1

A.



B.

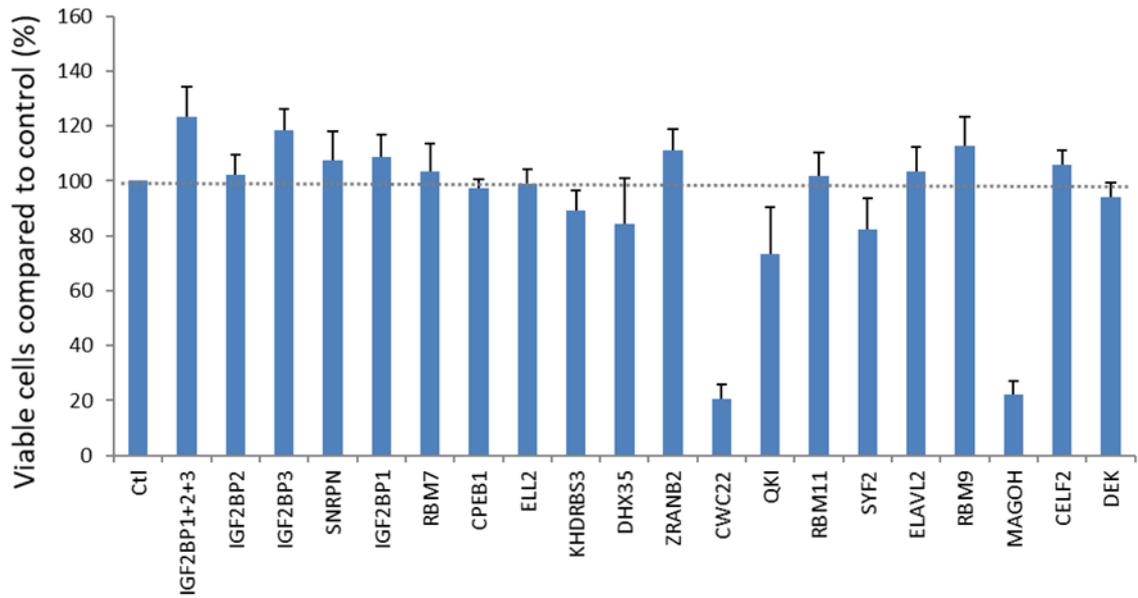
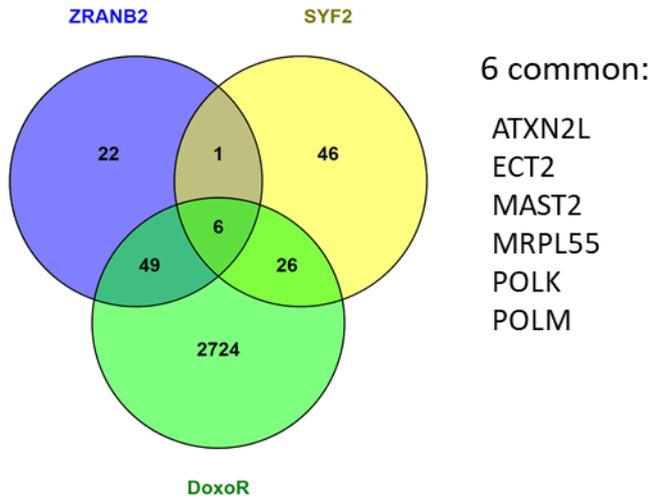
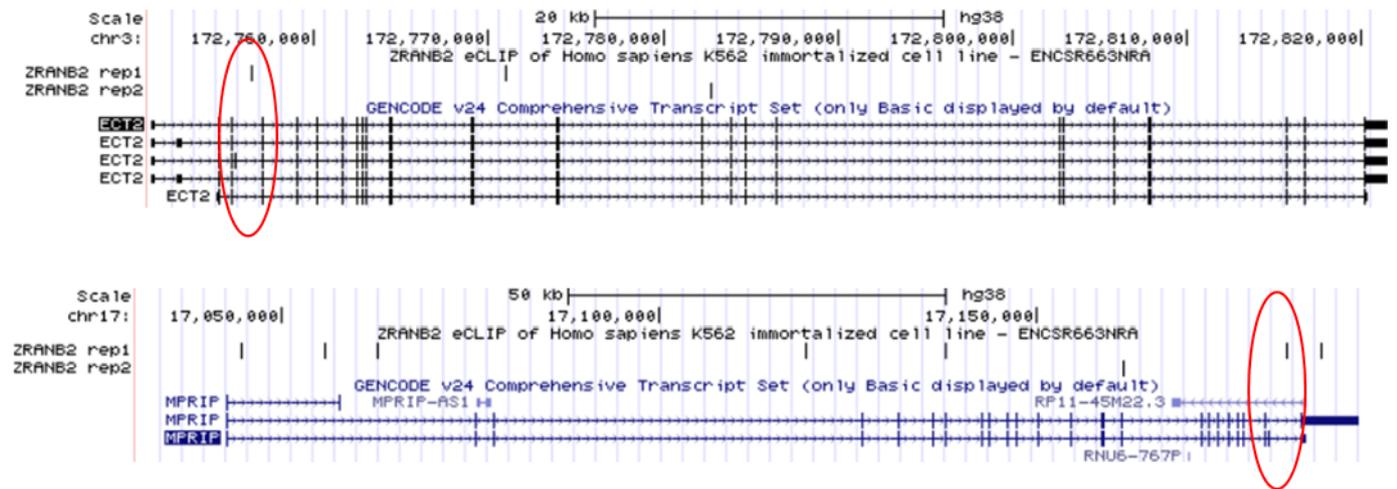


Figure S2

A.



B.



C.

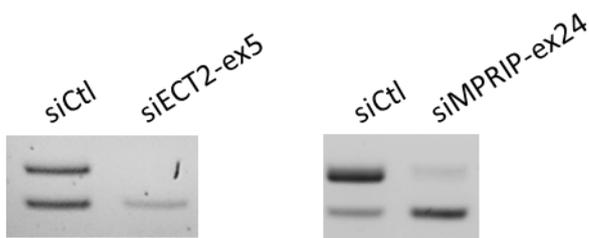


Figure S3

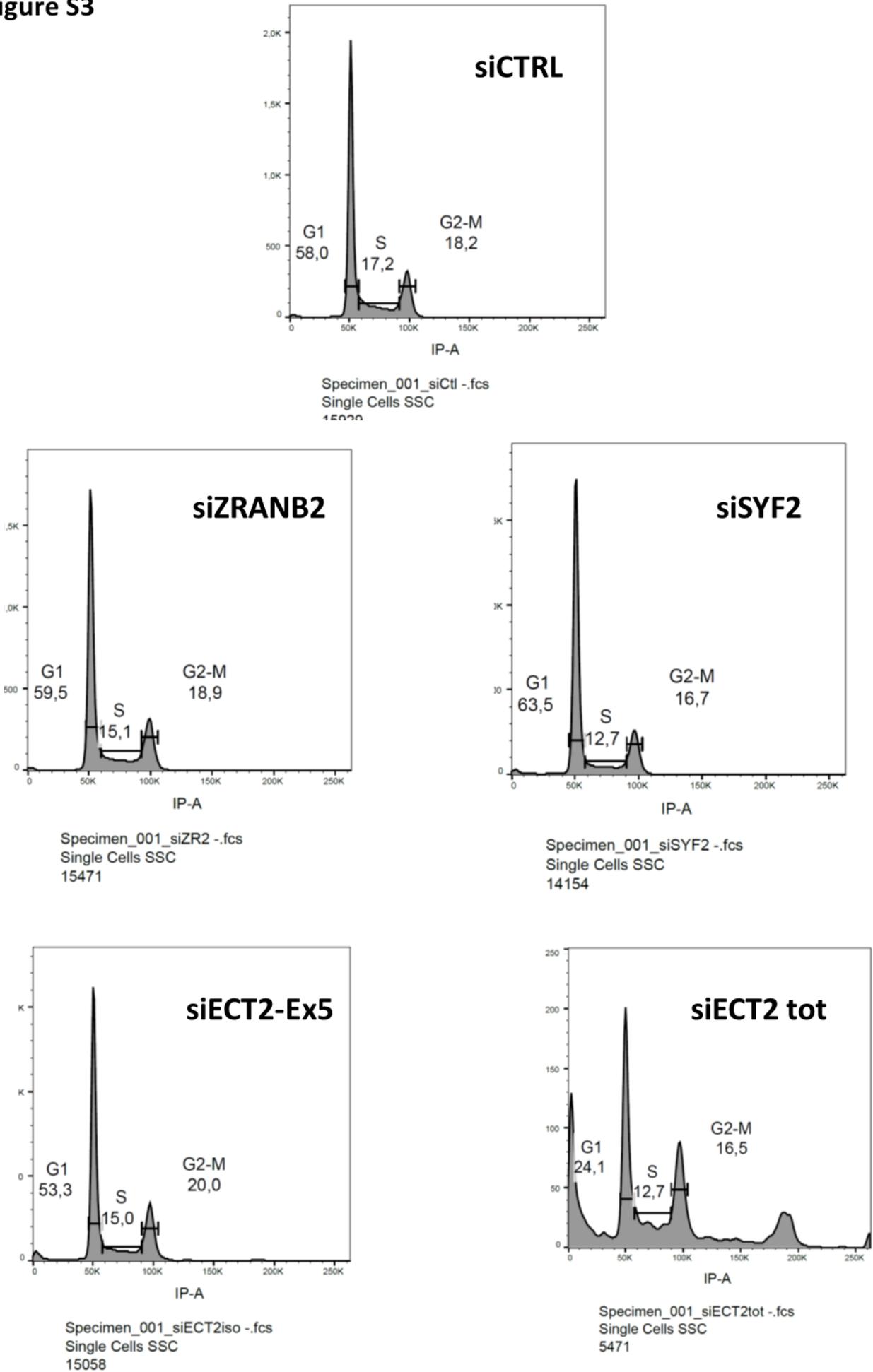
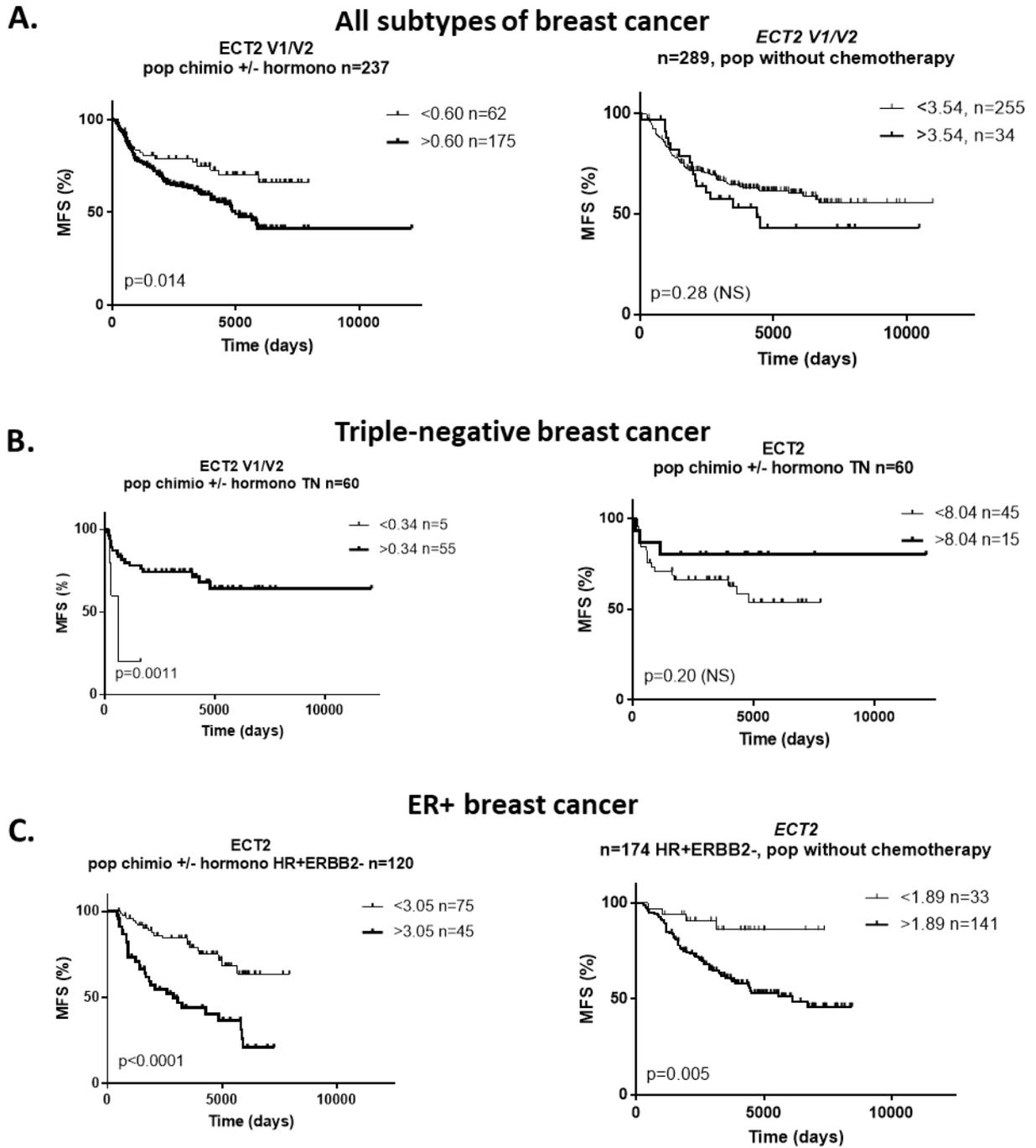


Figure S4



- II. Paper manuscript 2: “Intronic polyadenylation is linked to the antiproliferative effect of specific platinum compounds and generates differentially translated isoforms”

Intronic polyadenylation is linked to the antiproliferative effect of specific platinum compounds and generates differentially translated isoforms

Iris TANAKA¹⁻³, Mandy CADIX¹⁻⁴, Sophie MICHALLET⁵, Ludovic TESSIER¹⁻³, Hussein MORTADA⁶, Pierre GESTRAUD⁴, Virginie RAYNAL^{7,8}, Sylvain BAULANDE⁷, Didier AUBOEUF⁶, Nicolas SERVANT⁴, Stéphan VAGNER¹⁻³, Béatrice EYMIN⁵, Martin DUTERTRE^{1-3*}

¹Institut Curie, PSL Research University, CNRS UMR3348, F-91405, Orsay, France

²Université Paris Sud, Université Paris-Saclay, CNRS UMR3348, F-91405 Orsay, France

³Equipe Labellisée Ligue Contre le Cancer

⁴INSERM U900, Mines Paris Tech, Institut Curie, 75000 Paris, France

⁵INSERM U1209, CNRS UMR5309, Université Grenoble Alpes, Institute for Advanced Biosciences, 38000 Grenoble France

⁶CNRS UMR 5239, Ecole Normale Supérieure de Lyon, 69007 Lyon, France

⁷NGS platform, Institut Curie, 75000 Paris, France

⁸INSERM U830, Institut Curie, 75000 Paris, France

* To whom correspondence should be addressed. Tel: (+33) 1 69 86 31 04; Fax: (+33) 1 69 86 94 29; Email: martin.dutertre@curie.fr

ABSTRACT

Alternative last exons (ALEs) generated by intronic polyadenylation are repressed by doxorubicin in specific genes, but little is unknown about ALE regulation by other anticancer drugs, its link with drug antiproliferative effects, and its impact on the transcriptome. Here, using 3'-seq, we found that ALEs are a prevalent type of exons regulated by cisplatin in non-small cell lung cancer (NSCLC) cells. Unlike doxorubicin, cisplatin induced proximal-ALE isoforms when compared to full-length isoforms. Mechanistically, this regulation was mediated by an inhibition of transcription elongation, which was enriched in long genes. Functionally, cisplatin-induced ALE regulations were enriched in genes involved in cell cycle and cell death, and preceded these cellular responses. The lower antiproliferative effect of oxaliplatin compared to cisplatin in NSCLC cells correlated with dramatically less effects on ALEs despite similar numbers of up-regulated genes. Moreover, ALE regulations distinguished cisplatin-sensitive *versus* resistant cells. Finally, 3'-seq analyses on polysomes

identified an efficiently translated subset of cisplatin-induced proximal-ALE isoforms, and an inefficiently translated subset enriched in particularly short isoforms. Our findings identify proximal-ALE induction as a widespread mechanism of transcriptome regulation linked to cancer cell response to cisplatin -not oxaliplatin- and impacting the translome, and reveal a class of inefficiently translated ALE isoforms.

Keywords: Genotoxic anticancer drugs; platinum compounds; intronic polyadenylation; translome; transcription elongation.

INTRODUCTION

A major class of anticancer drugs used in chemotherapy is genotoxic drugs, which can induce various types and combinations of DNA lesions. For example, doxorubicin mainly acts by inhibiting DNA topoisomerase II, leading to double-strand DNA breaks, whereas cisplatin induces monoadducts, intrastrand and interstrand crosslinks. Lesion-specific signaling and DNA-damage response (DDR) involves the activation of specific DNA repair pathways, checkpoints that transiently arrest the cell cycle, and the transcriptional regulation of genes involved in DNA repair, cell cycle, apoptosis and other functions, thereby contributing to determine cell fate (*e.g.*, permanent cell-cycle arrest, cell death, etc.) (1,2). Genotoxic drugs also regulate gene expression at post-transcriptional levels, including pre-messenger RNA splicing, mRNA stability, and translation (2-4). In particular, genome-wide studies showed large-scale regulation of alternative splicing by several genotoxic anticancer drugs (5-10).

Among the various types of alternative splicing events and alternative exons, a less studied type is alternative last exons (ALEs), also often referred to as intronic polyadenylation (IPA). In contrast with alternative polyadenylation within the last exon of genes (3'UTR-APA), IPA corresponds to the use of a polyadenylation (polyA) site located upstream of the last exon (or distal ALE) of a gene, thus leading to a transcript ending in a proximal ALE (11-13). Proximal ALEs are found in thousands of human coding genes and come in two main flavors, skipped ALEs that are whole exons defined by a 3' splice site and a polyA site, and composite ALEs that have a polyA site downstream of a 5' splice site (11-13). Using exon-arrays, ALEs were found to be the main type of alternative exons regulated by doxorubicin, with about

100 cases identified (8). Strikingly, in 90% cases, doxorubicin repressed the proximal ALE relative to the last exon of genes. In many cases, the doxorubicin-induced distal ALE corresponded to a noncanonical isoform (lowly expressed and non-conserved). Thus, doxorubicin induces a specific pattern of ALE regulation. In parallel, camptothecin, a topoisomerase I inhibitor, induced either up- or down-regulation of proximal ALEs (depending on genes) in approximately equal proportions (8). The repression of proximal ALEs by doxorubicin and camptothecin was mediated at least in part by decreased binding of the RNA-binding protein HuR (8).

More generally, there have been several genome-wide studies of ALE regulation in cell proliferation and cancer, and different studies found different global patterns, with either preferential induction of proximal ALEs (14-16), preferential repression (8,17) or no preferential trend (18), suggesting that specific sets of ALEs may be differentially regulated in different contexts or by different factors. It is therefore important to identify specific ALE regulation patterns in different settings, and in particular to investigate the effects of additional anticancer drugs, such as platinum compounds. Cisplatin is a key agent used in the treatment of aggressive cancers, notably in non-small cell lung cancer (NSCLC) (19,20). Oxaliplatin is used in several cancers but not in NSCLC; it shows distinct profiles of antiproliferative effects across cancer cell lines when compared to cisplatin, and was recently shown to act primarily through a ribosome biogenesis stress rather than a genotoxic mechanism (21). Recent studies have reported widespread alternative splicing regulation by cisplatin, and oxaliplatin was shown to regulate alternative splicing of several genes, but the impact of these compounds on ALEs is poorly characterized (5,6,22).

Finally, little is known about the biological significance of ALE regulations. In particular, doxorubicin-regulated ALE isoforms of the *CENPN* gene were shown to be differentially active (8), but it is unknown, whether the genome-wide pattern of ALE regulation induced by the drug may be linked to its antiproliferative effect. Furthermore, while ALEs were shown to give rise to functional protein isoforms in several genes (11-14,16,23,24), little is known about the translation of ALE isoforms on a genome-wide scale. Indeed, genome-wide studies of the impact of APA isoforms on the translome or proteome have been mainly limited to the analysis of the impact of 3'UTR length (3'UTR-APA) on translation efficiency or output

(25-28). In fact, an ALE isoform of a coding gene was recently found to have a noncoding RNA function (29).

Thus, while doxorubicin represses a specific set of proximal ALEs, little is known on a genome-wide scale about the regulation of ALE isoforms by other anticancer drugs, and about their translation and biological significance. In this study, we found that cisplatin treatment of NSCLC cells widely induces proximal ALEs, mechanistically through an inhibition of transcription elongation, which was enriched in long genes. Cisplatin-induced ALE regulations were enriched in genes involved in cell cycle and cell death, and preceded these cellular responses. To further investigate the functional relevance of this regulation pattern, we tested genome-wide ALE regulation by more or less antiproliferative platinum compounds, and in more or less sensitive cells. These analyses revealed that proximal ALEs were poorly induced by oxaliplatin that had little antiproliferative effect in NSCLC cells, and distinguished cisplatin-sensitive from resistant cells. Finally, cisplatin-induced ALE regulations impacted the transcriptome; interestingly however, a subset of cisplatin-induced ALE isoforms were inefficiently translated. Thus, our findings indicate that proximal ALE induction is a widespread mechanism of transcriptome regulation that is linked to cancer cell response to cisplatin, discriminates different platinum compounds, and generates differentially translated isoforms.

MATERIALS AND METHODS

Cell culture and treatment

H358 and A549 sensitive and cisplatin-resistant(30) human lung carcinoma cells were cultured in RPMI-1640-Glutamax and DMEM medium, respectively (GibcoBRL, Life Technologies, Cergy Pontoise, France), supplemented with 10% (v/v) heat-inactivated fetal calf serum (GibcoBRL), in 5% CO₂ at 37°C. Cisplatin and oxaliplatin were obtained from Selleckchem (Euromedex, Souffelweyersheim, France). Unless otherwise stated, treatments with cisplatin and oxaliplatin were done at 100 μM. Treatment with vehicle (DMSO) was always done in parallel.

Western blot

For immunoblotting, cells washed three times in PBS were lysed in RIPA buffer (150mM NaCl, 50mM Tris HCl pH 8, 0.1% SDS, 1% Nonidet P40, 0.5% Na deoxycholate, 0.1mM PMSF, 2.5µg/ml pepstatin, 10µg/ml aprotinin, 5µg/ml leupeptin, 0.2mM Na₃VO₄) for 30 min on ice and pelleted. Protein concentration was determined using the Biorad Dc protein assay. Proteins (40µg) were then separated in 10-12% SDS-PAGE gels and electroblotted onto PVDF membranes. Membranes were incubated overnight at +4°C with primary antibodies and proteins were detected using horse-radish peroxidase-conjugated goat anti-mouse or anti-rabbit antibodies (Jackson ImmunoResearch Laboratories, West Grove, PA, USA). After washing, the blots were revealed using the ECL chemiluminescence method (Amersham, Les Ulis, France), according to the manufacturer's protocol. Primary antibodies were: anti-PRIM2 (PA5-48859; ThermoFisher Scientific) and anti-actin (Sigma-Aldrich, Lyon, France).

FACS analysis

For DNA content analysis, cells were fixed with 70% cold ethanol for 30 min on ice, treated with RNase A (20 µg/ml) for 20 min and stained with propidium iodide (10 µg/ml). Flow cytometric analysis of 10000 cells was performed on a FACScan flow cytometer (BD Biosciences) and data were recovered using the CellQuest software (BD Biosciences).

Cell fractionation, RNA extraction and RT-(q)PCR

Cytosol and polysome fractions were prepared as previously described(31). RNA from whole cells and cell fractions was extracted using TRIzol Reagent (ThermoFischer Scientific), and 1 µl of GlycoBlue (ThermoFisher Scientific) was added for RNA precipitation. Total and nuclear RNA were treated with DNase I (TURBO DNA-free, ThermoFisher Scientific). Reverse transcription was performed using SuperScript™ Reverse Transcriptase (ThermoFisher Scientific) and random primers. PCR was performed using GoTaq Flexi DNA Polymerase (Promega), and PCR products were migrated on agarose gels. qPCR was performed using Power SYBR Green PCR Master Mix (ThermoFisher Scientific) on a CFX96 Real-Time PCR Detection System (BioRad).

Exon-array experiments and bioinformatic analysis

Exon-array analyses were performed as previously described (8). Briefly, one µg of total RNA was processed according to the manufacturer's instructions and hybridized to GeneChip Human Exon 1.0 ST arrays (Affymetrix). The experiments, from cell treatment to array

hybridization, were performed three times. Exon-array data were analyzed with Expression Console (Affymetrix) to perform quality assessment and normalized using quintile normalization. Background correction and probe selection were performed as previously described, using exon annotation from the Faster DB database (8) (<http://fasterdb.lyon.unicancer.fr/>). The splicing index method corresponds to a comparison of gene-normalized exon intensity values between two analyzed experimental conditions. Exon-array data of each regulated exon were visualized in gene context using the exon-array visualization tool of Faster DB (Elexir), allowing to eliminate false positives and to classify exonic regulations into different types (ALE, etc.). For elongation analysis, each gene was divided into 3 bins, and genes with at least 3 intronic probes (median: 17 probes) in the 5' and 3' bins of the gene were used to compute a 3' to 5' ratio. Genes with at least 20% reduction of this ratio in cisplatin *versus* vehicle conditions ($p < 0.05$) were considered to be elongation inhibited.

3'-seq experiments and bioinformatic analysis

5 µg of total RNA were used for poly(A)+ RNA purification with Dynabeads mRNA DIRECT Micro kit (ThermoFisher Scientific). Poly(A)+ RNA was fragmented at 70°C for 5 min in RNA Fragmentation Reagents (ThermoFisher Scientific). After ethanol precipitation, RNA was reverse transcribed using anchored oligo-dT and SMARTScribe RT enzyme (Clontech) allowing to add adapter sequences to the 5' and 3' ends of RNA fragments. Libraries were amplified by 16 cycles of PCR with GoTaq Flexi DNA Polymerase (Promega), purified using Agencourt AMPure XP (Beckman Coulter) magnetic beads, and quantified with Quant-iT Picogreen dsDNA kit (ThermoFisher Scientific). A size selection step was done using SPRIselect (Beckman Coulter) magnetic beads to obtain fragments of 150–300 bp. Purified libraries were controlled by capillary electrophoresis (LabChip - Perkin) and quantitated by qPCR (KAPA Library Quantification Kits Illumina Platforms - Roche). Pooled libraries (12 pM) with 30% of phiX were subjected to single-end, 50 bp sequencing using the HiSeq 2500 machine (Illumina). Read 1 was read with primer HP6 (Illumina) with 3 dark cycles (first 3 bases of read 1 were not read). Index i7 (6 pb barcode) was read with primer HP8 (Illumina).

Raw reads were trimmed in their 5' and 3' ends to remove uninformative nucleotides due to primer sequences, nucleotides added by SMART Scribe RT enzyme and polyA tail of mRNAs. Trimmed reads of 25 bp or more were aligned on the human reference genome (hg19) using

Bowtie2 (version 2.2.5) (32). Only reads with a mapping quality score (MAPQ) of 20 or more were retained (Samtools version 1.1) for downstream analysis. Reads were then clustered along the genome using Bedtools (version 2.17.0) (33), allowing a maximum distance of 170 bp (max library fragment length (300bp) – min fragment length (60bp) – read length (50bp) – oligodT length (25bp)) and a minimum number of 5 reads per peak. Peaks with a stretch of 6 consecutive As (or 8 As out of 9 nucleotides) within 150 bp downstream were filtered out, as they are likely due to internal priming of oligo-dT. Overlapping peaks from all analyzed samples were merged to define a common set of genomic windows corresponding to polyA sites. To annotate peak location within genes, gene coordinates were obtained on the basis of overlapping Refseq transcripts with the same gene symbol. Peaks overlapping any intronic region of a gene were classified as intronic polyA (IPA) peaks. Peaks overlapping the last exon of a gene were classified as LE peaks. Differential analyses between two conditions were done using two independent biological replicates per condition. To compare the regulation of each IPA to the regulation of the gene's last exon (taken as the sum of the peaks in this exon), we used DESeq2 (version 1.4.5) (34) and the following statistical model:

$$Y_{ij} = \mu + L_i + C_j + (LC)_{ij} + E_{ij}$$

where Y_{ij} is the normalized counts of peak i in biological condition j , μ is the mean, L_i is the peak localization (IPA or LE), C_j is the biological condition, $(LC)_{ij}$ is the interaction between peak localization and biological condition, and E_{ij} is the residual. Adjusted P-values (Benjamini-Hochberg) were calculated. For H358 cells, data using a false discovery rate (FDR) of 10% are shown. For A549 cells, data with $p < 0.05$ are shown, to avoid under-estimations of cross-comparisons between lists (Fig. 5F and 6C). The complete bioinformatics pipeline (3'-SMART package) described above and in Fig. S1 can be freely downloaded at GitHub (<https://github.com/InstitutCurie/3-SMART>) and can be run through a configuration file and a simple command line. Annotated polyadenylation sites were retrieved from PolyA_DB_3 (35).

Other bioinformatic and statistical analyses

Functional gene annotation analyses were done using the DAVID software (36), using the human genome as a reference. For each experimental analysis, at least three independent experiments were performed. In all bar charts, error bars represent the standard error of the mean (SEM) that is the standard deviation divided by the square root of sample number. A

Student's paired t-test was used, except in Fig. 3B, 3D, and 4A, where a Mann-Whitney test was used. Tests were considered significant if $p < 0.05$.

RESULTS

Cisplatin regulates ALEs in many genes, with a strong bias toward proximal ALE induction

To analyze the impact of cisplatin on alternative exons, total RNA from the cisplatin-sensitive H358 lung adenocarcinoma cells treated for 7 or 24 hours with cisplatin at 100 μM or with vehicle were initially analyzed using Affymetrix exon-arrays, as previously done for doxorubicin and camptothecin (8). About 50 and 150 exon regulations were identified at 7 and 24 hours, respectively (Supplementary Table 1). The most prominent category of regulation events was ALEs, followed by alternative first exons (corresponding to alternative promoters) and cassette exons (alternatively spliced exons; Fig. 1A).

We then extended our analysis of ALE regulation using 3'-seq, an RNA-seq method that relies on the sequencing of regions preceding the 3'-terminal polyA tail of transcripts (37) (Supplementary Fig. S1A). Because 3'-seq analysis can be flawed by internal priming of the oligo(dT) at genomic stretches of A's, 3'-seq peaks ending near a genomic stretch of A's were discarded (Supplementary Fig. S1B, "peak filtering"). We also verified that most of the selected peaks contained a potential polyadenylation signal (AATAAA-like motif; Supplementary Fig. S1C). For each peak overlapping an annotated intron, we analyzed the effect of cisplatin on its expression relative to the last exon of the cognate gene (Supplementary Fig. S1B). With a false discovery rate (FDR) of 10%, 2963 intronic peaks in 1987 genes were regulated by cisplatin at 24 hours of treatment (Fig. 1B; Supplementary Table 2). Because internal priming artefacts can be difficult to filter out bioinformatically, we also crossed our list of regulated intronic peaks with a list of reliable polyA sites identified in normal tissues (annotated in PolyA_DB_3) (35). Among the 2963 regulated intronic peaks, 719 (24%) matched an annotated physiologic polyA site (Fig. 1B, Supplementary Fig. S1D). We found on average a similar percentage of annotated polyA sites in our subsequent 3'-seq analyses on cytosolic fractions (Supplementary Fig. S1D and see below), which are devoid of unspliced transcripts and thus are unlikely to be subject to internal priming in introns. These data suggest that many of the cisplatin-induced intronic peaks that we identified may

correspond to genuine polyA sites, that were not annotated in normal tissues. We therefore decided to keep all regulated peaks for subsequent analyses (which were also performed in parallel on annotated polyA sites for confirmation, see below). Finally, about half of the ALE regulations identified using exon-arrays were also found using 3'-seq, with lowly expressed genes yielding more regulations using exon-arrays (data not shown).

Strikingly, 90 to 95% of ALE regulations by cisplatin, identified by either exon-array or 3'-seq (annotated or not) corresponded to an up-regulation of the proximal ALE relative to the last exon of the gene (Fig. 1C). This pattern is in sharp contrast with the previously described effect of doxorubicin, which repressed proximal ALEs in 90% cases (8). For example, in intron 7 of the *PRIM2* gene, our 3'-seq data identified a proximal polyA site, which is annotated in PolyA_DB_3 (chr6:57299993:+) and was increased by cisplatin relative to the last exon of the gene (exon 14 ; Fig. 1D). RT-PCR analysis showed the existence of a spliced RNA containing exon 7 and an ALE within intron 7 (Fig. 1E-F), which is in agreement with an Ensembl transcript (PRIM2-205 ENST00000490313.1). RT-qPCR analysis validated the relative induction by cisplatin of this short *PRIM2* isoform compared to the full-length *PRIM2* mRNA, which was strongly decreased (Fig. 1G). Many other ALE regulations were validated by RT-PCR (Fig. 1H). Thus, ALE regulation is a prominent effect of cisplatin, with a strong bias toward up-regulation of proximal ALEs relative to full-length isoforms.

Cisplatin induction of proximal ALEs is mediated by an inhibition of transcription elongation in long genes

Our data indicate that proximal ALE up-regulation by cisplatin is mediated at least in part by an inhibition of transcription elongation. First, genes with proximal ALE up-regulation by cisplatin were on average much longer than non-regulated ALE-containing genes in the human genome (Fig. 2A, "LE"), and cisplatin-induced proximal ALEs were much closer to the gene start than the last exon of cognate genes (Fig. 2A, comparing "Prox. ALE" and "LE" of regulated genes). Second, cisplatin was reported to inhibit transcription elongation of several genes or gene constructs, but this has not been studied on a large scale (38-40). Using our exon-array data, we could compute a ratio of intronic RNA levels at the 3' versus 5' part of genes for 2545 genes, and we found that cisplatin inhibited transcription elongation in 753 (30%) of them by at least 20% at 7 and/or 24 hours of treatment ($p < 0.05$; Supplementary Fig. S2A and Supplementary Table 3). As in the case of ALE regulation,

elongation inhibition by cisplatin was enriched in long genes (Fig. 2A, right). Moreover, 48% of genes with proximal ALE up-regulation by cisplatin had their elongation inhibited by the drug (Fig. 2B, left). Elongation inhibition was twice more frequent for genes with proximal ALE up-regulation (Fig. 2B, middle), and *vice versa*, proximal ALE up-regulation was twice more frequent for genes with elongation inhibition (Fig. 2B, right). By RT-qPCR using primers in introns at the 3' *versus* 5' part of genes, we verified that cisplatin treatment for 7 hours inhibited transcription elongation in 8 out of 11 tested genes with proximal ALE up-regulation (Fig. 2C).

Third, for the genes tested, both the proximal ALE up-regulation and the elongation inhibition that were observed in response to cisplatin, were also observed in response to camptothecin (a topoisomerase I inhibitor), further suggesting a link between these two effects (Supplementary Fig. S2B-C). We also noted that the proximal ALE up-regulation could be observed in nuclear RNA (Supplementary Fig. S2D), which is in line with a mechanism relying on transcription rather than cytosolic stability of isoforms. Finally, a more detailed analysis on three genes indicated that in response to cisplatin, pre-mRNA levels were not affected in the first intron (suggesting no regulation of transcription initiation), but decreased in the intron preceding the up-regulated ALE, and were further decreased in the last intron (Fig. 2D). Reduced elongation between the proximal ALE and the last exon of the gene can directly explain the relative decrease in full-length mRNA levels. Altogether, our data suggest that cisplatin induction of proximal ALEs relative to the last exon of cognate genes is mediated at least in part by an inhibition of transcription elongation (see discussion).

Transcript truncation in response to cisplatin is enriched in cell-cycle and cell-death genes, and precedes these cellular responses

To investigate the biological relevance of ALE regulation by cisplatin, we first analyzed the functional annotation of genes with cisplatin up-regulation of proximal ALEs, taking either the full list (1875 genes) or only the subset corresponding to annotated polyA sites (577 genes; see Fig. 1B). In both cases, among the most enriched functions were functions related to DDR, cell cycle and cell death (Fig. 3A and Supplementary Fig. S3), which are key cellular effects of cisplatin. We then sought to determine, whether transcript truncation in response to cisplatin is a mere consequence of cell cycle arrest or cell death, or can precede these cellular responses. Our microarray analyses that were done at 7 and 24 hours of treatment

with cisplatin at 100 μ M in H358 cells, indicated that a significant subset of ALE regulation events and the majority of elongation repression events were detected at 7 hours of treatment (Fig. 1A and Supplementary Fig. S2A). In contrast, regarding cell death, the percentage of cells with a sub-G1 DNA content was not increased at 7 hours of treatment with cisplatin at 100 μ M, although it was increased at 24 hours of treatment (Fig. 3B). These data suggest that many ALE and elongation regulation events induced by cisplatin are not a consequence of cell death.

Regarding cell cycle, treatment with cisplatin at 100 μ M mainly induced an arrest in G2/M, which was detected at 7 hours (Fig. 3C). To go further, we compared the dose-response of cisplatin effects on cell cycle arrest and transcript truncation. For this, we took advantage of the cell cycle-related gene, *PRIM2*, in which we showed induction of an ALE within intron 7 in response to cisplatin treatment at 100 μ M for 24 hours (Fig. 1E-G). Following treatment with cisplatin at 25 μ M, cell cycle arrest was detected at 24 hours, but not at 7 hours (Fig. 3C-D). In contrast, at 7 hours of treatment with cisplatin at 25 μ M, we detected a decrease in *PRIM2* transcript elongation, and an increase in the ratio of short to full-length *PRIM2* mRNA isoforms (Fig. 3F-G). These data indicate that transcript truncation in response to cisplatin can precede cell cycle arrest.

We also showed that cisplatin treatment led to a decrease in full-length *PRIM2* protein levels within 7 hours of treatment, but we could not detect a short *PRIM2* protein isoform (Fig. 3E). Finally, a more detailed time-course analysis showed that cisplatin at 100 μ M inhibited *PRIM2* elongation at as early as one hour of treatment (Fig. 3H). Collectively, our data indicate that transcript truncation in response to cisplatin is an early event that can precede both cell cycle arrest and cell death, and is not a consequence thereof.

Oxaliplatin is much less efficient than cisplatin to regulate ALEs

To further investigate the potential link between proximal ALE induction by cisplatin and its antiproliferative effect, we performed analyses using different platinum compounds, and in more or less sensitive cells.

Oxaliplatin is another platinum derivative that shows different sensitivity profiles across cancer cell lines when compared to cisplatin, and was recently shown to act primarily through a non-genotoxic mechanism (21). The underlying differences between cisplatin and

oxaliplatin in terms of molecular mechanisms are not fully understood. In contrast with cisplatin, oxaliplatin at 100 μ M had no effect on H358 cell cycle at 7 hours of treatment, and had only little effects on cell cycle and cell death at 24 hours (Fig. 4A). While oxaliplatin up-regulated nearly as many genes as cisplatin, indicating that both drug treatments were efficient (Fig. 4B, left), it had much less effect than cisplatin on ALEs. Indeed, oxaliplatin significantly induced only 139 (including 52 annotated) proximal ALEs, as compared to 2817 (673 annotated) in the case of cisplatin (comparing Fig. 4C and Fig. 1B; 3'-seq experiments on both drugs were conducted in parallel). Consistently, oxaliplatin had much less effect than cisplatin on *PRIM2* transcript elongation, short/long isoforms ratio and full-length protein levels (Fig. 4D-F). We also noticed that oxaliplatin down-regulated less genes than cisplatin (Fig. 4B, right), which might be due to a lesser ability to inhibit transcription elongation. Thus, the much weaker antiproliferative effect of oxaliplatin when compared to cisplatin in H358 cells, was associated with a much lower ability to induce proximal ALEs.

ALE regulation distinguishes cisplatin-resistant from sensitive NSCLC cells

To further investigate the link between proximal ALE induction by cisplatin and its antiproliferative effect, we then examined ALE regulation in an NSCLC cell model of acquired resistance to cisplatin in the lung adenocarcinoma A549 background (30). While the growth of parental A549 cells was strongly inhibited by a 48-hour treatment with cisplatin at 25 μ M, the growth of the derived A549R cell line was not affected (Fig. 5A). 3'-seq analysis was performed in parallel on the parental and resistant cell lines, both in the presence and absence of cisplatin at 25 μ M for 16 hours (cytosolic RNA was analyzed as explained in the next section). In parental A549 cells, cisplatin mostly up-regulated proximal ALEs (Fig. 5B and Supplementary Table 4) as previously observed in H358 cells. In addition, a large subset (52%) of the up-regulations in A549 cells were also found in H358 cells (Fig. 5C).

We also identified sets of ALEs that were regulated in resistant *versus* parental cells (Fig. 5D-E and Supplementary Tables 5 and 6). In the absence of cisplatin treatment, there was no bias toward up- or down-regulation of proximal ALEs (Fig. 5D). In contrast, in resistant *versus* parental cells analyzed after cisplatin treatment, there was a bias toward down-regulation of proximal ALEs (Fig. 5E), which were enriched in proximal ALEs induced by cisplatin in parental cells as could be anticipated (Fig. 5F), and were also enriched in cell division-related genes (Fig. 5G). These data indicate that many ALEs are regulated in a cellular model of

cisplatin resistance, and that cisplatin-induced proximal ALEs can distinguish resistant from sensitive NSCLC cells.

Proximal ALE regulation impacts the translome, but some isoforms are inefficiently translated

To further investigate the biological significance of ALE regulation by cisplatin, we analyzed the translation of the ALE-generated mRNA isoforms in parental A549 cells. Indeed, little is known about ALE isoform translation on a genome-wide scale. For this, we carried out 3'-seq analyses on polysome fractions (which correspond to translated mRNAs) of treated and untreated A549 cells. Again, as in the case of cytosolic RNA analyzed in parallel (Fig. 5B), cisplatin mostly up-regulated proximal ALEs (Fig. 6A and Supplementary Table 7). These data suggested that at least a subset of cisplatin-induced proximal-ALE isoforms are translated.

To compare the translation efficiency of proximal-ALE isoforms and the corresponding full-length mRNA isoforms, we then analyzed the relative abundance of ALE isoforms in polysome *versus* cytosol fractions. Performing this analysis on all genes (not just cisplatin-regulated ALEs), we observed that a subset of proximal-ALE isoforms exhibited a differential translation efficiency (more often lower) relative to the corresponding full-length isoform (Fig. 6B, Supplementary Table 8 and Supplementary Fig. S4A).

Thanks to these analyses, we identified a subset of cisplatin-induced proximal-ALE isoforms that were efficiently translated: they were induced by cisplatin in both cytosol and polysomes, and did not have a translation efficiency (polysome/cytosol ratio) lower than the corresponding full-length isoform (N=91; Fig. 6C and Supplementary Table 9). We also identified a subset of cisplatin-induced proximal-ALE isoforms that were inefficiently translated: they were induced by cisplatin in the cytosol but not in polysomes, and had a lower translation efficiency than the corresponding full-length isoform (N=210; Fig. 6C and Supplementary Table 9). Compared to the efficiently translated proximal-ALE isoforms, the inefficiently translated ones were on average 3-times shorter, and were enriched in 5'IPA isoforms, which correspond to ALEs located in the 5' part of genes (Fig. 6D-E).

Interestingly, a recent study identified by RNA-seq, 84 genes with proximal-ALE isoforms induced by ultraviolet (UV-C) radiation, including a short isoform with a noncoding RNA function produced by the coding gene *ASCC3* (29). Interestingly, 34 (40%) of the 84 genes

with proximal ALE up-regulation by UV-C, including *ASCC3*, were also found in our 3'-seq dataset of cisplatin-induced proximal ALEs (Fig. 6F). Moreover, the *ASCC3*-short isoform, which was reported to encode a protein but had a noncoding RNA function (29), was part of the 210 inefficiently translated short isoforms induced by cisplatin (identified in Fig. 6C). RT-qPCR analysis showed that the *ASCC3* short/long RNA isoform ratio was up-regulated by cisplatin in the cytosol (Fig. 6G and Supplementary Fig. S4B), but was much lower in polysomes when compared to cytosol (Fig. 6H). These data indicate that the translation of the *ASCC3*-short isoform is inefficient compared to that of the full-length *ASCC3* mRNA. Altogether, these data suggest that the induction of proximal ALEs in long genes is a prominent effect of cisplatin, and that this regulation leads to two subsets of short transcripts: efficiently translated ones and inefficiently translated ones (Fig. 6I). The function of ALE regulation by cisplatin is further discussed below.

DISCUSSION

While doxorubicin represses a specific set of proximal ALEs, little is known on a genome-wide scale about the regulation of ALE isoforms by other anticancer drugs, and about their translation and biological significance. In this study, we show a widespread but specific pattern of proximal ALE induction in long genes, that is linked to the antiproliferative effect of cisplatin -but not oxaliplatin- in lung cancer cell lines, and that leads to two differentially translated subsets of proximal-ALE isoforms.

The global pattern of ALE regulation by cisplatin, that is the up-regulation of proximal ALEs relative to the last exon of genes in 95% cases, is very distinct from the effects previously described for other genotoxic anticancer drugs, with doxorubicin down-regulating proximal ALEs in 90% cases, and camptothecin inducing both patterns in approximately equal proportions (8). Our data strongly suggest a role for transcription elongation inhibition in proximal ALE induction by cisplatin. Previous analyses on gene constructs showed that cisplatin induces RNA polymerase stalling and inhibits transcription elongation; several mechanisms have been involved, including RNA polymerase modifications and reduced nucleosome mobility (38-40,42,43). Our analysis on about 2500 genes indicates that cisplatin specifically inhibits elongation of long genes. Reduced elongation between the proximal ALE

and the last exon of the gene can directly explain the relative decrease in full-length mRNA levels. In addition, because cleavage/ polyadenylation is extensively coupled to transcription, the detectable elongation inhibition upstream of the proximal polyA site (Fig. 2D) could favor its cotranscriptional recognition, which in turn should promote transcription termination downstream (11,24). A similar mechanism of proximal ALE induction (relative to the last exon of genes) due to elongation inhibition in long genes, as we describe here for cisplatin, was recently proposed for UV-C (29); consistently, many regulations induced by UV-C were found in our cisplatin dataset (Fig. 6F) (the limited dataset available for UV-C precludes the converse analysis). This scenario may also apply to camptothecin (in the cases of proximal ALE induction), which mimicked cisplatin effects on both elongation and proximal ALE induction in several tested genes (Supplementary Fig. S2B-C), and was shown to inhibit transcription elongation preferentially in long genes (44,45). Interestingly, depletion of CDK12 (an RNA polymerase II kinase) was also shown to inhibit elongation and induce proximal ALEs preferentially in long genes, but whether these effects are linked is unknown (16). In addition to elongation inhibition, other mechanisms might contribute to proximal ALE induction by cisplatin. Indeed, for UV-C, proximal ALE induction in several genes was shown to be mediated by repression of the U1 snRNP, which is a widespread repressor of intronic polyadenylation (41). Other potential mechanisms may include DNA-damage responsive cleavage/ polyadenylation factors and RNA-binding proteins that regulate the use of polyA sites (8,46-51). Clearly, the contributions of these various mechanisms to genome-wide ALE regulation by genotoxic agents remain to be established.

While ALEs affect both the coding region and the 3'UTR of mRNAs, the latter of which can control mRNA stability, export and translation, and while some ALE isoforms are known to encode functional protein isoforms, little is known about the fate of ALE isoforms genome-wide and their impact on the translome. In particular, while the impact of 3'UTR-APA on translation efficiency or output has been studied genome-wide by several groups, much less is known about the translation efficiency of ALE isoforms (25-28). Our 3'-seq analyses comparing polysomes and cytosol indicate that proximal-ALE isoforms tend to be less translated than full-length isoforms (Fig. 6B and Supplementary Fig. S4A), and identify two subsets of proximal-ALE isoforms induced by cisplatin: efficiently and inefficiently translated ones (Fig. 6C). Interestingly, we found that isoforms ending in the 5' part of genes (5'IPA

isoforms) were enriched in the inefficiently translated subset, when compared to the efficiently translated subset (Fig. 6E). While recent studies identified sets of 5'IPA isoforms that were degraded in the nucleus, thereby preventing their export and translation (52,53), our data identify a subset of 5'IPA isoforms that are significantly detected in the cytosol but are lowly associated with polysomes, suggesting that they are inefficiently translated. Consistent with this finding, a recent study identified a set of cell type-specific 5'IPA isoforms that lack an open-reading frame longer than 100 amino acids (17). The potential function of 5'IPA isoforms that are lowly associated with polysomes and/or have a low coding potential remains to be investigated on a large scale. Some of them may have a non-coding function, as recently shown for *ASCC3*-short (29), which is one of the cisplatin-induced 5'IPA isoforms that we found to be inefficiently translated (Fig. 6G-H). In addition, some 5'IPA isoforms may encode short polypeptides, as shown for the human *LEPR* (leptin receptor) gene, where two short RNA isoforms called *LEPROT* and *LEPROTL1* arise from ALEs located upstream of the start codon of the long mRNA, and encode short proteins that antagonize the function of the long protein despite sharing no similarity (54); our data indicate that these two short transcripts are induced by cisplatin (Fig. 1H). In fact, we cannot rule out that some 5'IPA isoforms may be lowly associated to polysomes despite being translated, if their open reading frames are too short to associate with several ribosomes (*e.g.*, micropeptides). Finally, our data also identify a set of cisplatin-induced ALE isoforms that are efficiently translated. While the function of most of these isoforms remains to be determined, it was shown for several genes that protein isoforms generated by ALEs can have distinct functional properties (11,14,23,24). Thus, ALE transcript isoforms that are induced by cisplatin may have different fates and translational outcomes.

Besides ALE regulations of the *ASCC3* and *CENPN* genes that were shown to play roles in cell responses to UV-C and doxorubicin, respectively (8,29), little is known about the role of ALE regulations in cell responses to DNA-damage and anticancer treatments. The enrichment of cisplatin-induced proximal ALEs in genes related to DDR, cell cycle and cell death (Fig. 3A), and the correlation of these ALE regulations with the antiproliferative effect of cisplatin (comparing more or less effective platinum compounds, and more or less sensitive cells), suggest their potential implication in cell response to cisplatin. For example, the strong decrease in full-length *PRIM2* mRNA synthesis and in the levels of the encoded protein (the

large subunit of DNA primase) is expected to prevent DNA replication, where this protein plays a key role (55). We detected the truncated *PRIM2* mRNA in polysomes, suggesting that it is translated (Supplementary Fig. S4C), but we did not detect a truncated PRIM2 protein, which might be unstable and should otherwise be defective for DNA primase activity, given the key role of the carboxy-terminal domain of PRIM2 (55). Another aspect of cell responses to DNA-damaging agents is the global regulation of gene expression and RNA metabolism. The *ASCC3*-short isoform that we found to be induced by cisplatin relative to full-length *ASCC3* (Fig. 6G) was previously shown to antagonize full-length *ASCC3* function and to promote transcription recovery and cell survival during the late response to UV-C irradiation (29). Our finding that cisplatin induces a 5'IPA isoform in the *ZFC3H1* gene (Fig. 1H), which was recently involved in the nuclear degradation of 5'IPA transcripts (52), raises the intriguing possibility of a positive feed-back mechanism, whereby *ZFC3H1* transcript truncation in response to cisplatin might enhance the cellular accumulation of 5'IPA isoforms. Likewise, we found that cisplatin induces a 5'IPA isoform of *INTS6* (Fig. 1H), a gene encoding a component of the integrator complex that controls UsnRNA 3' end maturation and transcription termination, as well as early elongation of specific coding genes (56,57). *INTS6* was also involved in the DDR (58). Thus, the large-scale regulation of ALE isoforms identified in this study in many genes related to DDR, cell cycle, cell death and gene regulation, likely constitutes a new component of the complex gene regulatory networks controlling tumor cell responses to cisplatin, and warrants further functional investigations in the future.

Our finding that oxaliplatin had much less ability than cisplatin to up-regulate proximal-ALE isoforms, may provide new molecular insights on the long-known but poorly understood discrepancies between these two major anticancer agents in terms of antiproliferative efficiency across cell lines (21). It was recently realized that oxaliplatin, unlike cisplatin, does not kill cells through a genotoxic mechanism, but through a ribosome biogenesis stress (21). It will be interesting to determine, whether and how the stronger effects of cisplatin on ALE regulation when compared to oxaliplatin, may be linked to its stronger genotoxicity in cancer cells. Furthermore, our finding that ALE isoform ratios are regulated in a cellular model of acquired resistance to cisplatin warrants further investigation to determine, whether ALE isoforms may be helpful to distinguish cisplatin-sensitive *versus* resistant tumors, given the dramatic issue of chemotherapy resistance (20). Further study of the poorly characterized

class of ALE isoforms may help understand the specific effects of different platinum compounds (and other genotoxic agents such as doxorubicin and camptothecin, as mentioned above) and identify drug-specific markers of sensitivity, with the aim of rationalizing their appropriate use in patients with different profiles.

In conclusion, in addition to the classical types of alternative splicing and alternative polyadenylation (cassette exons and 3'UTR-APA, respectively), there is increasing evidence that ALE (or IPA) isoforms are widely regulated in several biological processes (11-13). However, little is known about their translation status and biological significance. Our findings indicate that proximal ALE induction, typically in the 5' part of long genes, is a widespread mechanism of transcriptome regulation by cisplatin –not oxaliplatin– in lung cancer cell lines. This transcript truncation leads to two subsets of proximal-ALE isoforms, that are either efficiently or inefficiently translated (Fig. 6I). Our findings establish a link between IPA regulation by a drug and its antiproliferative effect, and may help understand the differential properties of different anticancer drugs. Our findings also identify links between different layers of gene expression regulation in response to a genotoxic agent (transcription elongation impacting IPA isoforms impacting the transcriptome). Finally, short 5'IPA isoforms with low translational output (this study) and/or low coding potential (17) may represent a new class of transcripts with noncoding functions generated by coding genes, as illustrated by the *ASCC3* gene (29).

AVAILABILITY

The complete bioinformatics pipeline for 3' seq analysis of ALEs (3'-SMART package) can be freely downloaded at GitHub (<https://github.com/InstitutCurie/3-SMART>).

ACCESSION NUMBERS

The datasets generated in this study have been deposited with the Gene Expression Omnibus repository (GEO) under accession number GSE119978 (private pre-publication link: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119978>; token: azazmciqfbatzyd) and are available on the UCSC genome browser (private pre-publication link:

https://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=mcadix&hgS_otherUserSessionName=Pt%2D3%2Dseq_draft).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

Supplementary Table 1 (.xls): Lists of cisplatin-regulated exonic events found by exon-array analysis in H358 cells.

Supplementary Table 2 (.xls): 3'-seq dataset of ALE regulation by cisplatin in H358 cells.

Supplementary Table 3 (.xls): Dataset of elongation analysis using exon-arrays.

Supplementary Table 4 (.xls): 3'-seq dataset of ALE regulation by cisplatin in A549 cytosol.

Supplementary Table 5 (.xls): 3'-seq dataset of ALE regulation in untreated A549R *versus* A549 cells.

Supplementary Table 6 (.xls): 3'-seq dataset of ALE regulation in cisplatin-treated A549R *versus* A549 cells.

Supplementary Table 7 (.xls): 3'-seq dataset of ALE regulation by cisplatin in A549 polysomes.

Supplementary Table 8 (.xls): 3'-seq dataset of ALE regulation in polysomes *versus* cytosol of cisplatin-treated A549 cells.

Supplementary Table 9 (.xls): Lists of cisplatin-induced proximal-ALE isoforms that are either efficiently or inefficiently translated.

ACKNOWLEDGEMENTS

We thank J.C. Soria for the cisplatin-resistant cell line, and L. Gratadou for help in exon-array experiments.

FUNDING

This work was supported by grants of the Fondation pour la Recherche Médicale (DBI20141231314) and Institut National du Cancer (2015-141) to MD, by a grant of Ligue Nationale Contre le Cancer (équipe labellisée) to SV, and by a fellowship from Fondation ARC to IT.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Experiments were performed by IT, SM and LT. Illumina sequencing was performed by VR and SB. Bioinformatics and biostatistics analyses were performed by MC, PG and NS for 3'-seq, and by HM and DA for exon-arrays. BE and MD supervised experiments. MD supervised 3'-seq analyses, and wrote the paper with help of BE. All authors read and approved the final manuscript.

REFERENCES

1. Ciccia, A. and Elledge, S.J. (2010) The DNA damage response: making it safe to play with knives. *Mol Cell*, **40**, 179-204.
2. Dutertre, M., Lambert, S., Carreira, A., Amor-Gueret, M. and Vagner, S. (2014) DNA damage: RNA-binding proteins protect from near and far. *Trends Biochem Sci*, **39**, 141-149.
3. Giono, L.E., Nieto Moreno, N., Cambindo Botto, A.E., Dujardin, G., Munoz, M.J. and Kornblihtt, A.R. (2016) The RNA Response to DNA Damage. *J Mol Biol*, **428**, 2636-2651.
4. Wickramasinghe, V.O. and Venkitaraman, A.R. (2016) RNA Processing and Genome Stability: Cause and Consequence. *Mol Cell*, **61**, 496-505.
5. Gabriel, M., Delforge, Y., Deward, A., Habraken, Y., Hennuy, B., Piette, J., Klinck, R., Chabot, B., Colige, A. and Lambert, C. (2015) Role of the splicing factor SRSF4 in cisplatin-induced modifications of pre-mRNA splicing and apoptosis. *BMC Cancer*, **15**, 227.
6. Anufrieva, K.S., Shender Vcapital O, C., Arapidi, G.P., Pavlyukov, M.S., Shakhparonov, M.I., Shnaider, P.V., Butenko, I.O., Lagarkova, M.A. and Govorun, V.M. (2018) Therapy-induced stress response is associated with downregulation of pre-mRNA splicing in cancer cells. *Genome medicine*, **10**, 49.

7. Dutertre, M., Sanchez, G., De Cian, M.C., Barbier, J., Dardenne, E., Gratadou, L., Dujardin, G., Le Jossic-Corcus, C., Corcos, L. and Auboeuf, D. (2010) Cotranscriptional exon skipping in the genotoxic stress response. *Nat Struct Mol Biol*, **17**, 1358-1366.
8. Dutertre, M., Chakrama, F.Z., Combe, E., Desmet, F.O., Mortada, H., Polay Espinoza, M., Gratadou, L. and Auboeuf, D. (2014) A recently evolved class of alternative 3'-terminal exons involved in cell cycle regulation by topoisomerase inhibitors. *Nature communications*, **5**, 3395.
9. Solier, S., Barb, J., Zeeberg, B.R., Varma, S., Ryan, M.C., Kohn, K.W., Weinstein, J.N., Munson, P.J. and Pommier, Y. (2010) Genome-wide analysis of novel splice variants induced by topoisomerase I poisoning shows preferential occurrence in genes encoding splicing factors. *Cancer Res*, **70**, 8055-8065.
10. Ip, J.Y., Schmidt, D., Pan, Q., Ramani, A.K., Fraser, A.G., Odom, D.T. and Blencowe, B.J. (2011) Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res*, **21**, 390-401.
11. Tian, B. and Manley, J.L. (2017) Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol*, **18**, 18-30.
12. Elkon, R., Ugalde, A.P. and Agami, R. (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet*, **14**, 496-506.
13. Ogorodnikov, A., Kargapolova, Y. and Danckwardt, S. (2016) Processing and transcriptome expansion at the mRNA 3' end in health and disease: finding the right end. *Pflugers Archiv : European journal of physiology*, **468**, 993-1012.
14. Lee, S.H., Singh, I., Tisdale, S., Abdel-Wahab, O., Leslie, C.S. and Mayr, C. (2018) Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature*, **561**, 127-131.
15. Elkon, R., Drost, J., van Haaften, G., Jenal, M., Schrier, M., Vrieling, J.A. and Agami, R. (2012) E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol*, **13**, R59.
16. Tien, J.F., Mazloomian, A., Cheng, S.G., Hughes, C.S., Chow, C.C.T., Canapi, L.T., Oloumi, A., Trigo-Gonzalez, G., Bashashati, A., Xu, J. *et al.* (2017) CDK12 regulates alternative last exon mRNA splicing and promotes breast cancer cell invasion. *Nucleic Acids Res*, **45**, 6698-6716.
17. Singh, I., Lee, S.H., Sperling, A.S., Samur, M.K., Tai, Y.T., Fulciniti, M., Munshi, N.C., Mayr, C. and Leslie, C.S. (2018) Widespread intronic polyadenylation diversifies immune cell transcriptomes. *Nature communications*, **9**, 1716.
18. Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A. and Burge, C.B. (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science*, **320**, 1643-1647.
19. Rossi, A. and Di Maio, M. (2016) Platinum-based chemotherapy in advanced non-small-cell lung cancer: optimal number of treatment cycles. *Expert review of anticancer therapy*, **16**, 653-660.
20. Fennell, D.A., Summers, Y., Cadranel, J., Benepal, T., Christoph, D.C., Lal, R., Das, M., Maxwell, F., Visseren-Grul, C. and Ferry, D. (2016) Cisplatin in the modern era: The backbone of first-line chemotherapy for non-small cell lung cancer. *Cancer treatment reviews*, **44**, 42-50.
21. Bruno, P.M., Liu, Y., Park, G.Y., Murai, J., Koch, C.E., Eisen, T.J., Pritchard, J.R., Pommier, Y., Lippard, S.J. and Hemann, M.T. (2017) A subset of platinum-containing

- chemotherapeutic agents kills cells by inducing ribosome biogenesis stress. *Nat Med*, **23**, 461-471.
22. Shkreta, L., Toutant, J., Durand, M., Manley, J.L. and Chabot, B. (2016) SRSF10 Connects DNA Damage to the Alternative Splicing of Transcripts Encoding Apoptosis, Cell-Cycle Control, and DNA Repair Factors. *Cell reports*, **17**, 1990-2003.
 23. Vorlova, S., Rocco, G., Lefave, C.V., Jodelka, F.M., Hess, K., Hastings, M.L., Henke, E. and Cartegni, L. (2011) Induction of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA activation. *Mol Cell*, **43**, 927-939.
 24. Sanchez, G., Bittencourt, D., Laud, K., Barbier, J., Delattre, O., Auboeuf, D. and Dutertre, M. (2008) Alteration of cyclin D1 transcript elongation by a mutated transcription factor up-regulates the oncogenic D1b splice isoform in cancer. *Proc Natl Acad Sci U S A*, **105**, 6004-6009.
 25. Spies, N., Burge, C.B. and Bartel, D.P. (2013) 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res*, **23**, 2078-2090.
 26. Sterne-Weiler, T., Martinez-Nunez, R.T., Howard, J.M., Cvitovik, I., Katzman, S., Tariq, M.A., Pourmand, N. and Sanford, J.R. (2013) Frac-seq reveals isoform-specific recruitment to polyribosomes. *Genome Res*, **23**, 1615-1623.
 27. Floor, S.N. and Doudna, J.A. (2016) Tunable protein synthesis by transcript isoforms in human cells. *eLife*, **5**.
 28. Gruber, A.R., Martin, G., Muller, P., Schmidt, A., Gruber, A.J., Gumienny, R., Mittal, N., Jayachandran, R., Pieters, J., Keller, W. *et al.* (2014) Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nature communications*, **5**, 5465.
 29. Williamson, L., Saponaro, M., Boeing, S., East, P., Mitter, R., Kantidakis, T., Kelly, G.P., Loble, A., Walker, J., Spencer-Dene, B. *et al.* (2017) UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. *Cell*, **168**, 843-855 e813.
 30. Michels, J., Vitale, I., Galluzzi, L., Adam, J., Olaussen, K.A., Kepp, O., Senovilla, L., Talhaoui, I., Guegan, J., Enot, D.P. *et al.* (2013) Cisplatin resistance associated with PARP hyperactivation. *Cancer Res*, **73**, 2271-2280.
 31. Boussemart, L., Malka-Mahieu, H., Girault, I., Allard, D., Hemmingsson, O., Tomasic, G., Thomas, M., Basmadjian, C., Ribeiro, N., Thuaud, F. *et al.* (2014) eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies. *Nature*, **513**, 105-109.
 32. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods*, **9**, 357-359.
 33. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841-842.
 34. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 550.
 35. Wang, R., Nambiar, R., Zheng, D. and Tian, B. (2018) PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res*, **46**, D315-D319.
 36. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, **4**, 44-57.

37. Lianoglou, S., Garg, V., Yang, J.L., Leslie, C.S. and Mayr, C. (2013) Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev*, **27**, 2380-2396.
38. Tremeau-Bravard, A., Riedl, T., Egly, J.M. and Dahmus, M.E. (2004) Fate of RNA polymerase II stalled at a cisplatin lesion. *J Biol Chem*, **279**, 7751-7759.
39. Jung, Y. and Lippard, S.J. (2003) Multiple states of stalled T7 RNA polymerase at DNA lesions generated by platinum anticancer agents. *J Biol Chem*, **278**, 52084-52092.
40. Zhu, G., Song, L. and Lippard, S.J. (2013) Visualizing inhibition of nucleosome mobility and transcription by cisplatin-DNA interstrand crosslinks in live mammalian cells. *Cancer Res*, **73**, 4451-4460.
41. Devany, E., Park, J.Y., Murphy, M.R., Zakusilo, G., Baquero, J., Zhang, X., Hoque, M., Tian, B. and Kleiman, F.E. (2016) Intronic cleavage and polyadenylation regulates gene expression during DNA damage response through U1 snRNA. *Cell discovery*, **2**, 16013.
42. Jung, Y. and Lippard, S.J. (2006) RNA polymerase II blockage by cisplatin-damaged DNA. Stability and polyubiquitylation of stalled polymerase. *J Biol Chem*, **281**, 1361-1370.
43. Todd, R.C. and Lippard, S.J. (2010) Consequences of cisplatin binding on nucleosome structure and dynamics. *Chemistry & biology*, **17**, 1334-1343.
44. Veloso, A., Biewen, B., Paulsen, M.T., Berg, N., Carmo de Andrade Lima, L., Prasad, J., Bedi, K., Magnuson, B., Wilson, T.E. and Ljungman, M. (2013) Genome-wide transcriptional effects of the anti-cancer agent camptothecin. *PloS one*, **8**, e78190.
45. Solier, S., Ryan, M.C., Martin, S.E., Varma, S., Kohn, K.W., Liu, H., Zeeberg, B.R. and Pommier, Y. (2013) Transcription poisoning by Topoisomerase I is controlled by gene length, splice sites, and miR-142-3p. *Cancer Res*, **73**, 4830-4839.
46. Kleiman, F.E. and Manley, J.L. (2001) The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression. *Cell*, **104**, 743-753.
47. Cevher, M.A., Zhang, X., Fernandez, S., Kim, S., Baquero, J., Nilsson, P., Lee, S., Virtanen, A. and Kleiman, F.E. (2010) Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage. *EMBO J*, **29**, 1674-1687.
48. Decorsiere, A., Cayrel, A., Vagner, S. and Millevoi, S. (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes Dev*, **25**, 220-225.
49. Li, W., Laishram, R.S., Ji, Z., Barlow, C.A., Tian, B. and Anderson, R.A. (2012) Star-PAP control of BIK expression and apoptosis is regulated by nuclear PIPKIalpha and PKCdelta signaling. *Mol Cell*, **45**, 25-37.
50. Newman, M., Sfaxi, R., Saha, A., Monchaud, D., Teulade-Fichou, M.P. and Vagner, S. (2017) The G-Quadruplex-Specific RNA Helicase DHX36 Regulates p53 Pre-mRNA 3'-End Processing Following UV-Induced DNA Damage. *J Mol Biol*, **429**, 3121-3131.
51. Fontana, G.A., Rigamonti, A., Lenzken, S.C., Filosa, G., Alvarez, R., Calogero, R., Bianchi, M.E. and Barabino, S.M. (2017) Oxidative stress controls the choice of alternative last exons via a Brahma-BRCA1-CstF pathway. *Nucleic Acids Res*, **45**, 902-914.
52. Ogami, K., Richard, P., Chen, Y., Hoque, M., Li, W., Moresco, J.J., Yates, J.R., 3rd, Tian, B. and Manley, J.L. (2017) An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev*, **31**, 1257-1271.

53. Chiu, A.C., Suzuki, H.I., Wu, X., Mahat, D.B., Kriz, A.J. and Sharp, P.A. (2018) Transcriptional Pause Sites Delineate Stable Nucleosome-Associated Premature Polyadenylation Suppressed by U1 snRNP. *Mol Cell*, **69**, 648-663 e647.
54. Touvier, T., Conte-Auriol, F., Briand, O., Cudejko, C., Paumelle, R., Caron, S., Bauge, E., Rouille, Y., Salles, J.P., Staels, B. *et al.* (2009) LEPROT and LEPROTL1 cooperatively decrease hepatic growth hormone action in mice. *J Clin Invest*, **119**, 3830-3838.
55. Pellegrini, L. (2012) The Pol alpha-primase complex. *Sub-cellular biochemistry*, **62**, 157-169.
56. Gardini, A., Baillat, D., Cesaroni, M., Hu, D., Marinis, J.M., Wagner, E.J., Lazar, M.A., Shilatifard, A. and Shiekhattar, R. (2014) Integrator regulates transcriptional initiation and pause release following activation. *Mol Cell*, **56**, 128-139.
57. Stadelmayer, B., Micas, G., Gamot, A., Martin, P., Malirat, N., Koval, S., Raffel, R., Sobhian, B., Severac, D., Rialle, S. *et al.* (2014) Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nature communications*, **5**, 5531.
58. Zhang, F., Ma, T. and Yu, X. (2013) A core hSSB1-INTS complex participates in the DNA damage response. *J Cell Sci*, **126**, 4850-4855.

FIGURE LEGENDS

Figure 1: Cisplatin regulates ALEs in many genes, with a strong bias toward proximal ALE induction. **A**, ALEs are the most prevalent type of exons regulated by cisplatin. Exon-array analysis on H358 cells treated with cisplatin at 100 μ M or vehicle for 7 or 24 hours. IR, intron retention; ASE, alternatively spliced exon; APA, 3'UTR-APA; AFE, alternative first exon; 3'down, down-regulation of exons at the gene 3'-end; Unclass., unclassified. **B**, Cisplatin regulates hundreds of ALEs. 3'-seq analysis on H358 cells treated with cisplatin at 100 μ M or vehicle for 24 hours (FDR 10%). FC, fold change. **C**, Cisplatin favors inclusion of proximal ALEs. **D**, Visualization of polyA sites and 3'-seq data for the *PRIM2* gene in the UCSC genome browser. **E**, Schematics of the *PRIM2* gene and PCR primers. **F**, RT-PCR detection of the short *PRIM2* isoform containing the proximal ALE. SM, size markers. **G**, RT-qPCR analysis of *PRIM2* isoforms in H358 cells treated with cisplatin (Cis-Pt) at 100 μ M or vehicle (Veh.) for 24 hours. **H**, RT-PCR validation of ALE regulation by cisplatin in various genes. Long isoforms and the position of proximal ALEs are indicated.

Figure 2: Cisplatin induction of proximal ALEs is mediated by transcription elongation inhibition in long genes. **A**, Proximal ALE induction and elongation inhibition by cisplatin are

enriched in long genes. Distance from transcription start site (TSS) to polyA site (showing median and SD). LE, last exon. **B**, Comparison of genes with proximal ALE up-regulation and elongation inhibition by cisplatin. **C-D**, RT-qPCR analysis of transcription elongation following cisplatin treatment for 7 hours, using the indicated primers spanning introns.

Figure 3: Transcript truncation precedes cell cycle arrest and cell death induction by cisplatin. **A**, ALE regulations by cisplatin are enriched in genes involved in DDR, cell cycle and cell death. Analysis with the DAVID software; the fold enrichment is indicated; only a subset of enriched functions are shown. **B-D**, FACS analysis of cell death (B) and cell cycle (C-D) using propidium iodide. **E**, Western blot analysis of PRIM2 protein levels. **F-G**, Dose-response analysis of cisplatin effects on *PRIM2* transcript elongation (F) and isoforms ratio (G) measured by RT-qPCR. Elongation was measured as in Fig. 2C. **H**, Time-course analysis of *PRIM2* elongation inhibition by cisplatin using RT-qPCR as in Fig. 2C.

Figure 4: Oxaliplatin is much less efficient than cisplatin to regulate ALEs. Comparison between the two drugs at 100 μ M in H358 cells. **A**, FACS analysis of cell cycle and cell death as in Fig. 3b-d. **B**, Number of genes either up- or down-regulated by at least 2-fold by cisplatin or oxaliplatin, based on 3'-seq reads in the last exon of genes. **C**, 3'-seq analysis of ALE regulation by oxaliplatin (100 μ M, 24 hours) in H358 cells (FDR 10%). **D-E**, RT-qPCR analysis of drug effects on *PRIM2* elongation and isoforms at 7 (D) and 24 hours of treatment (E). **F**, Western blot analysis of PRIM2 protein levels.

Figure 5: ALE regulation distinguishes cisplatin-sensitive and resistant cells. **A**, Cell growth analysis by WST1 assay. **B**, 3'-seq analysis of ALE regulation by cisplatin in A549 cells. **C**, Comparison of ALE regulation by cisplatin in H358 and A549 cells. **D-E**, 3'-seq analysis of ALE regulation in resistant *versus* parental cells, either without (D) or with cisplatin (E). **F**, Comparison of genes with proximal ALE regulation in response to cisplatin, and in resistant *versus* sensitive cells. **G**, ALE regulations in resistant *versus* sensitive cells are enriched in DDR, cell-cycle and cell-death genes. Analysis with the DAVID software; the fold enrichment is indicated; only a subset of enriched functions are shown.

Figure 6: Proximal-ALE regulation impacts the transcriptome, but some isoforms are inefficiently translated. **A**, 3'-seq analysis of ALE regulation by cisplatin in polysomal RNA from A549 cells. **B**, 3'-seq analysis of ALE regulation in polysomes *versus* cytosol from cisplatin-treated A549 cells. **C**, Subsets of cisplatin-induced ALE isoforms that are efficiently or inefficiently translated. Comparison between the indicated lists of proximal ALEs (overlap between 3'-seq peaks, allowing a gap of 300 bp). **D-E**, Comparison between the efficiently and inefficiently translated subsets of cisplatin-induced proximal-ALE isoforms. **D**, Distance from transcription start site (TSS) to intronic polyA site (showing median and SD). **E**, Percent of cisplatin-induced ALEs that are located in the 5' part of the gene (5'IPA isoforms). **F**, Comparison of genes with ALE regulations by cisplatin or UV-C radiations. **G-H**, ASCC3-short isoform is induced by cisplatin in A549 cytosol (G) but is inefficiently translated (H, cisplatin-treated cells). RT-qPCR analysis of ASCC3 isoforms. Cyto, cytosol. Poly, polysomes. **I**, Model (see main text).

Fig. 1

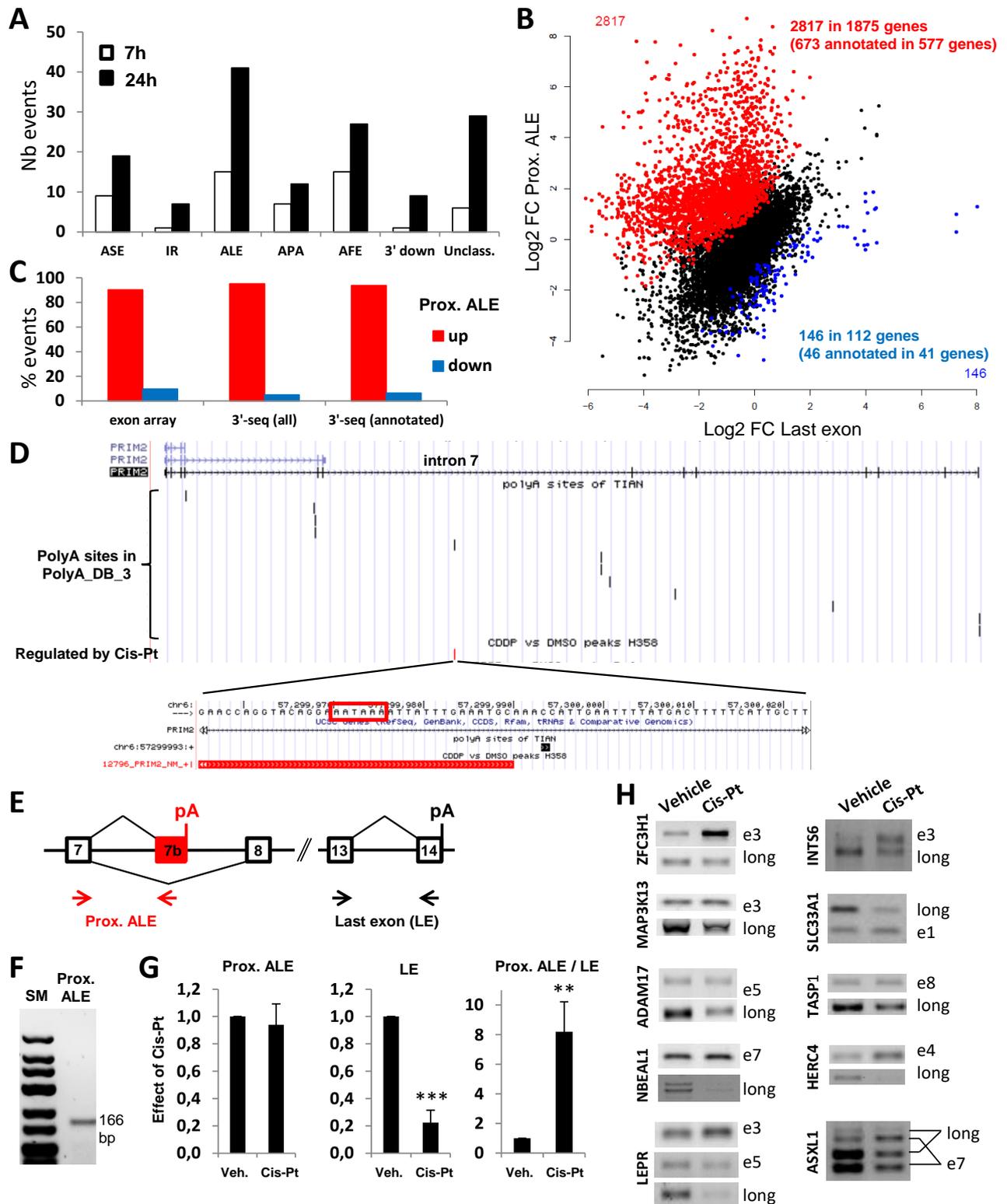


Fig. 2

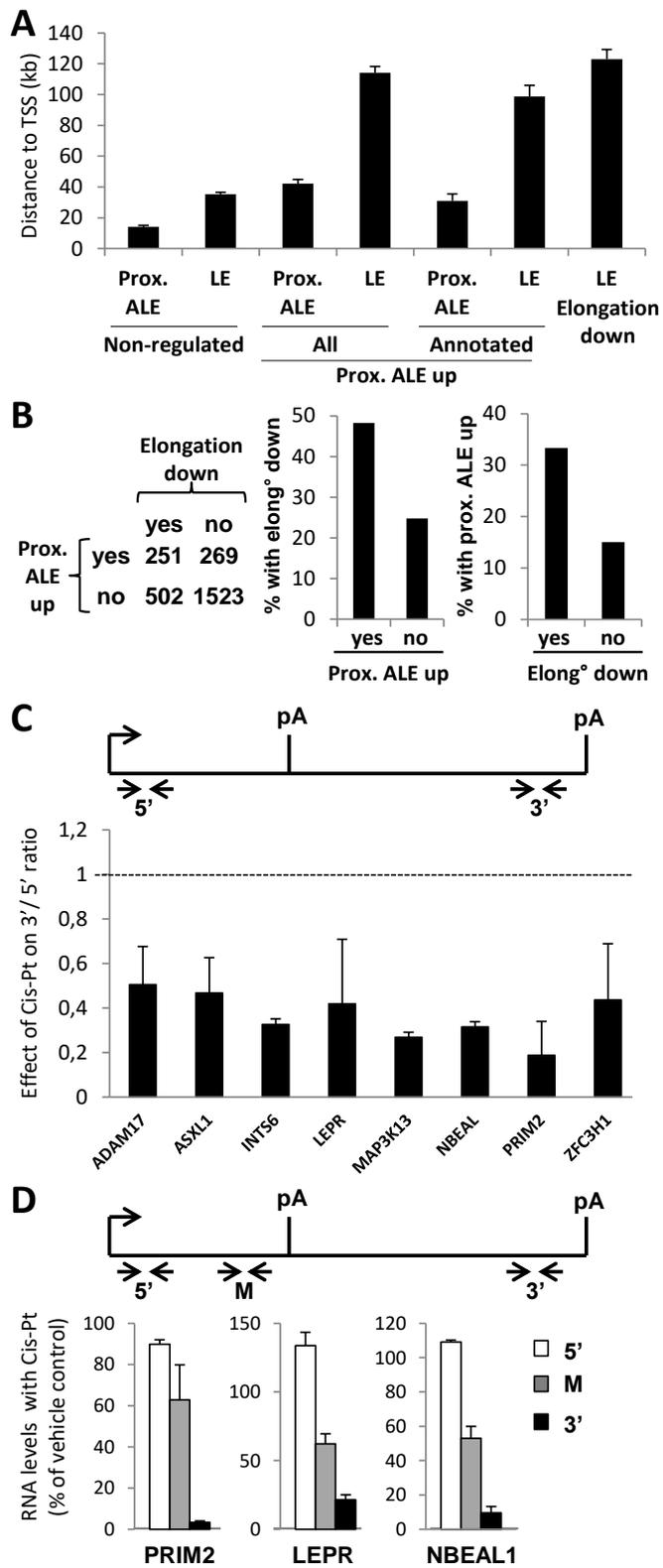


Fig. 3

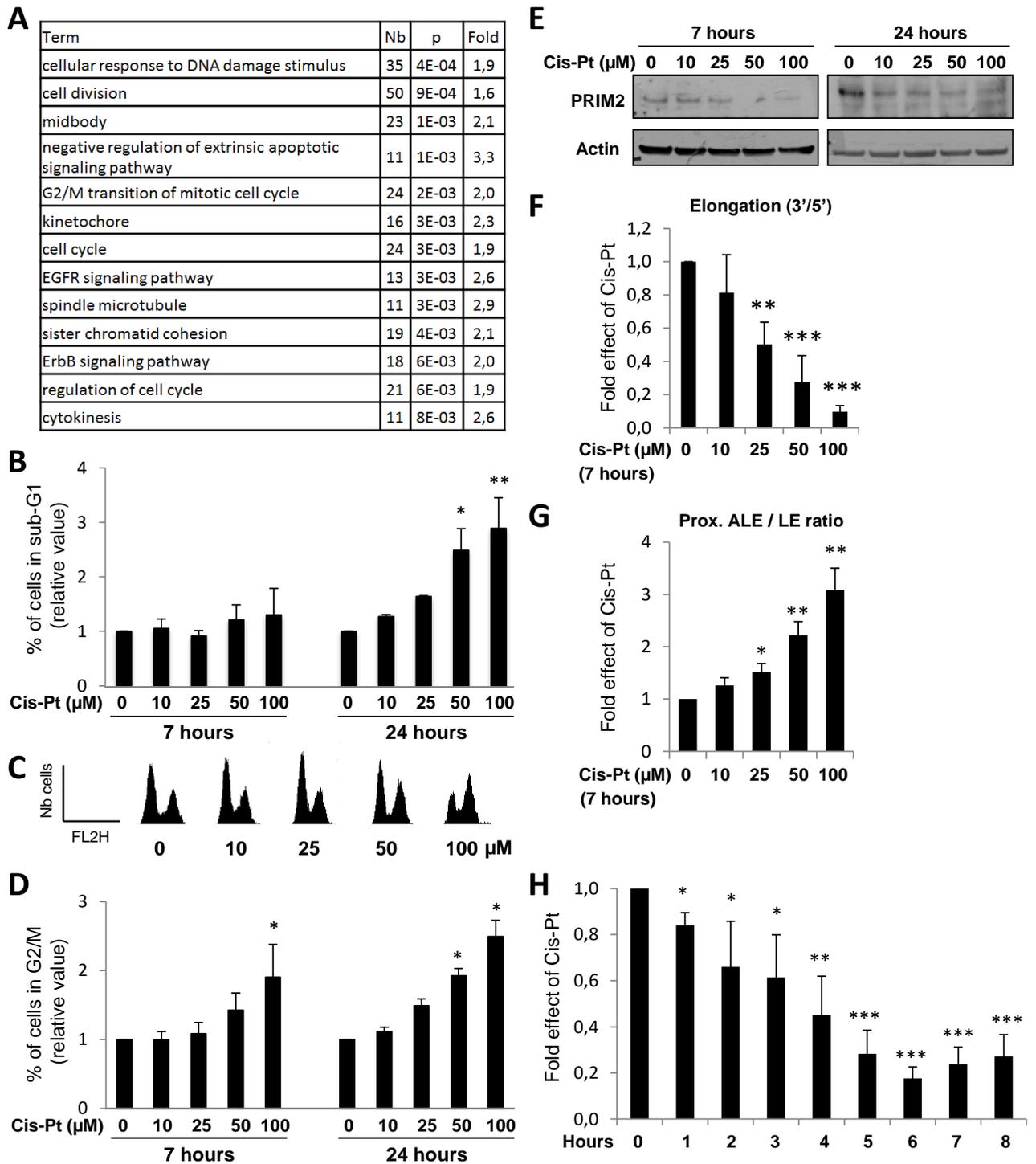


Fig. 4

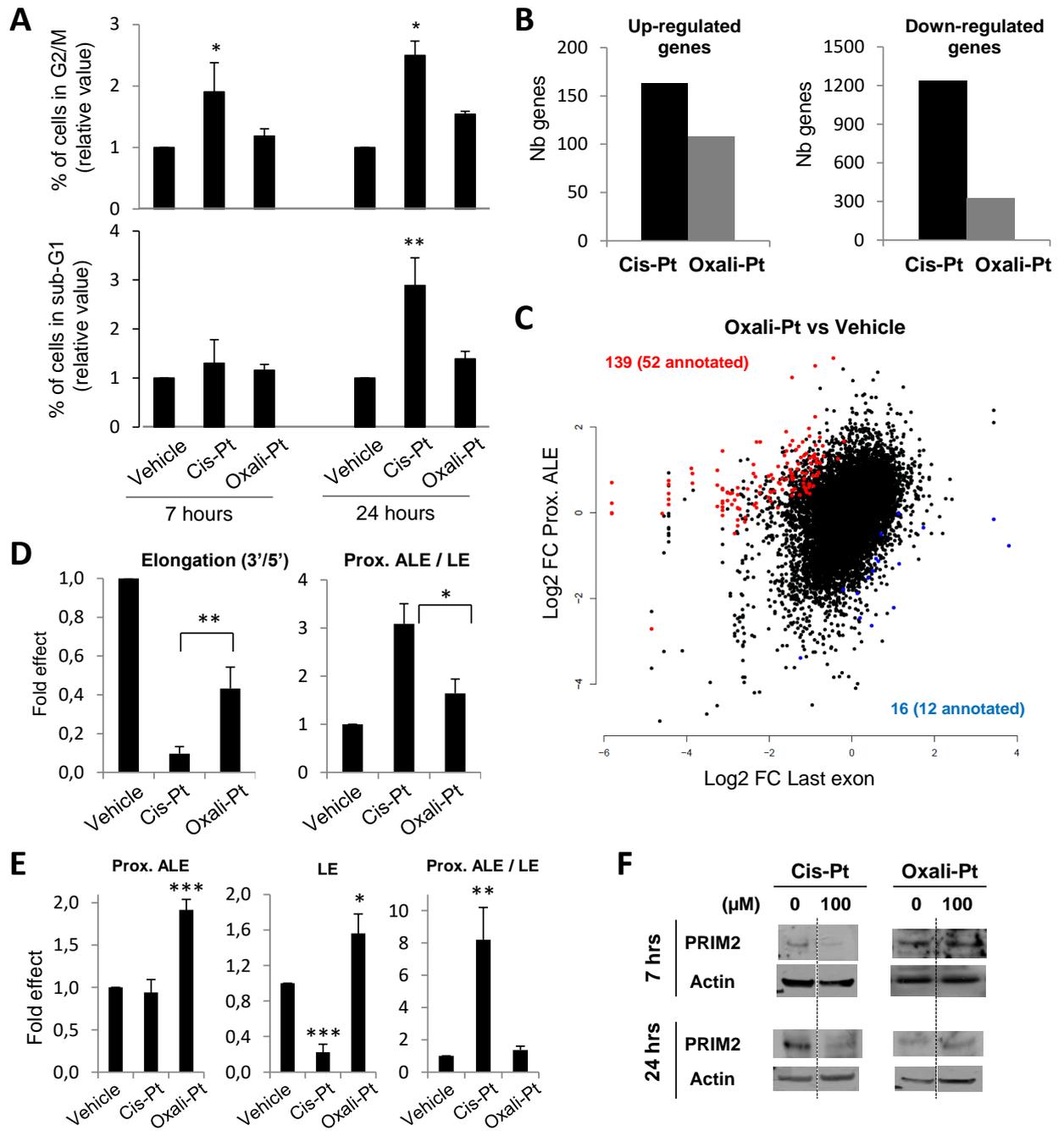
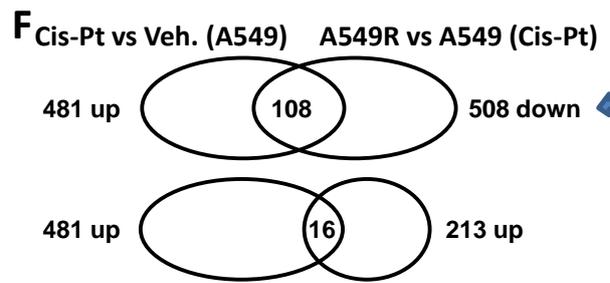
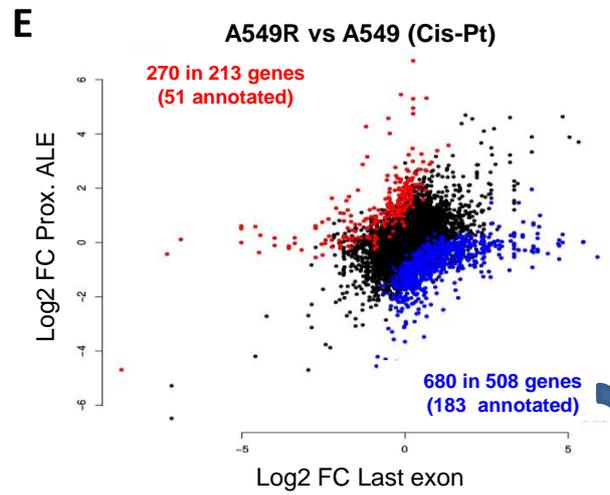
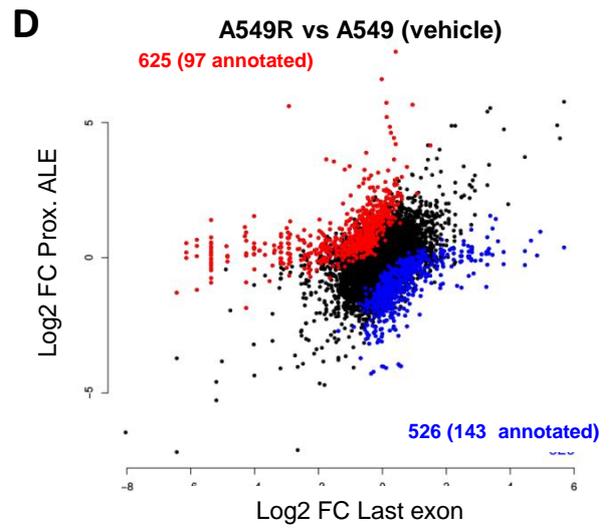
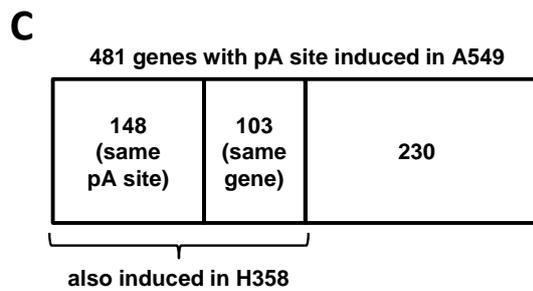
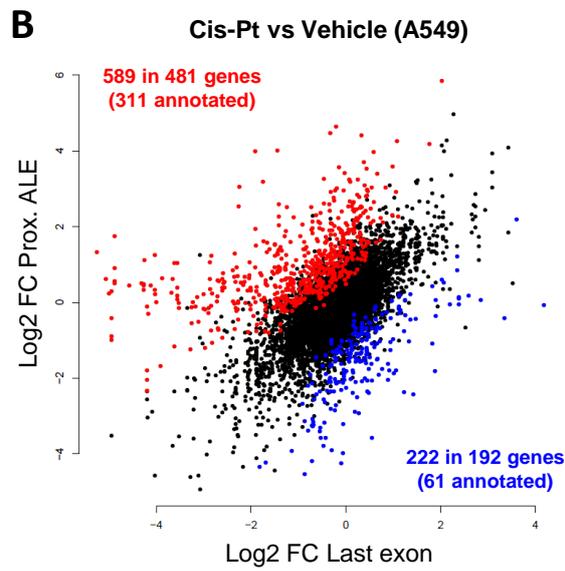
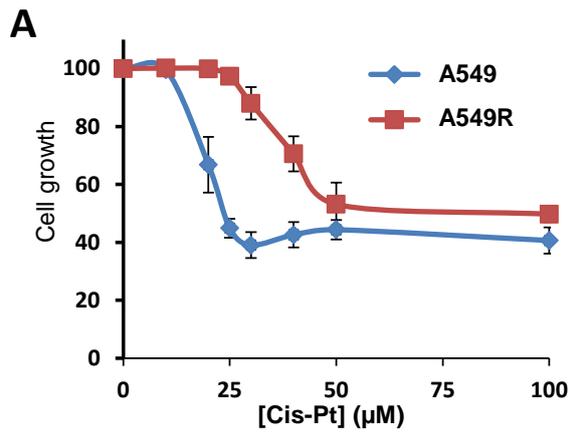
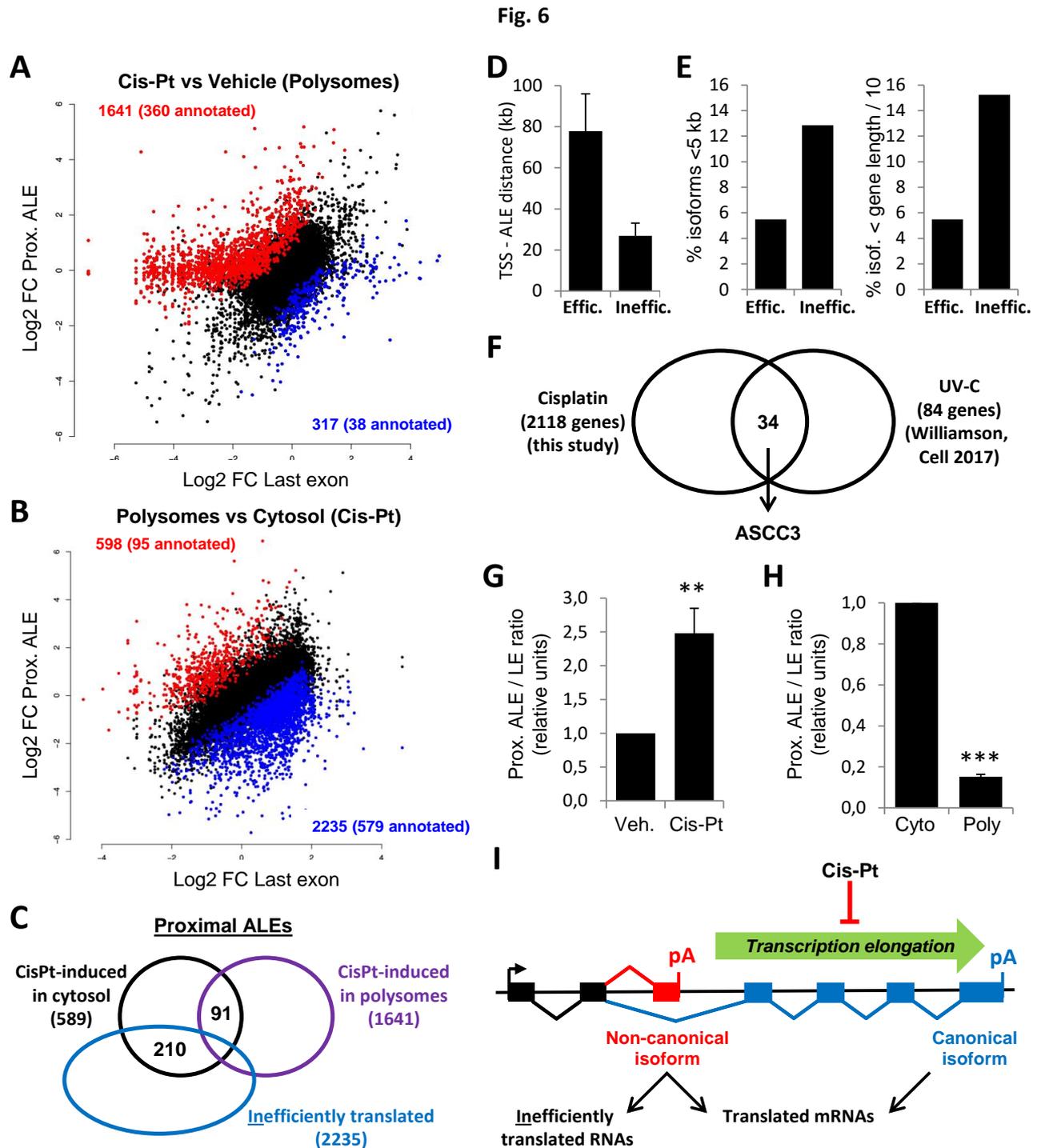


Fig. 5



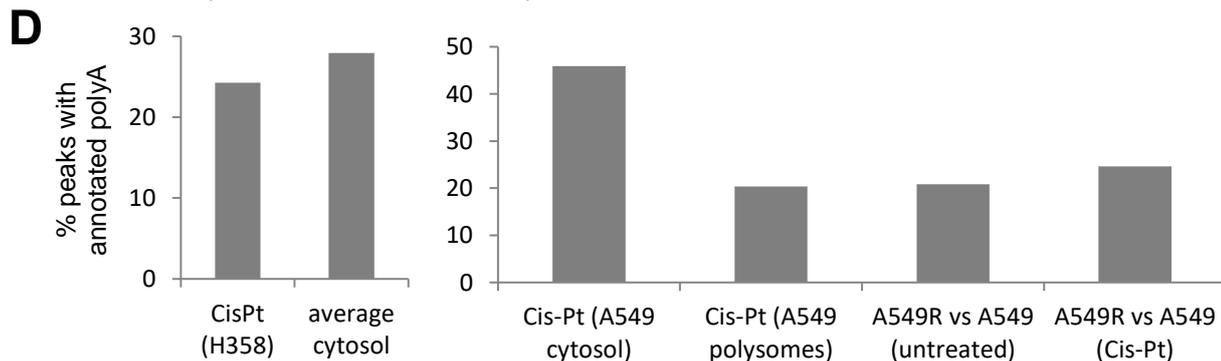
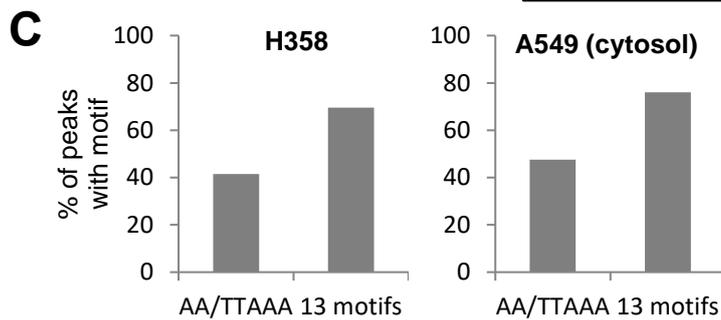
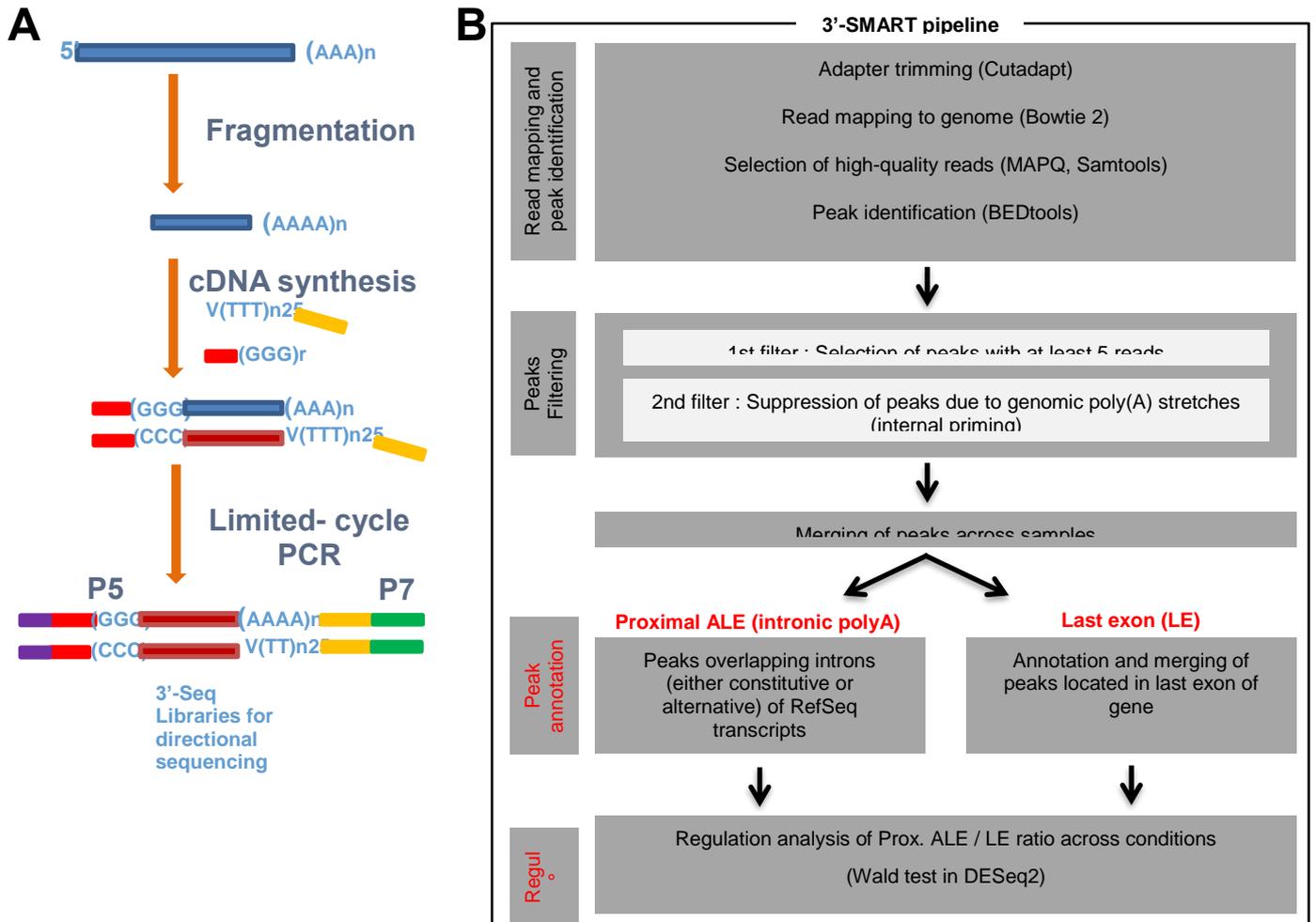
G

Term	Nb	p val	Fold
Midbody	12	5.2 E-4	3.6
Cell division	18	1.7E-2	1.9



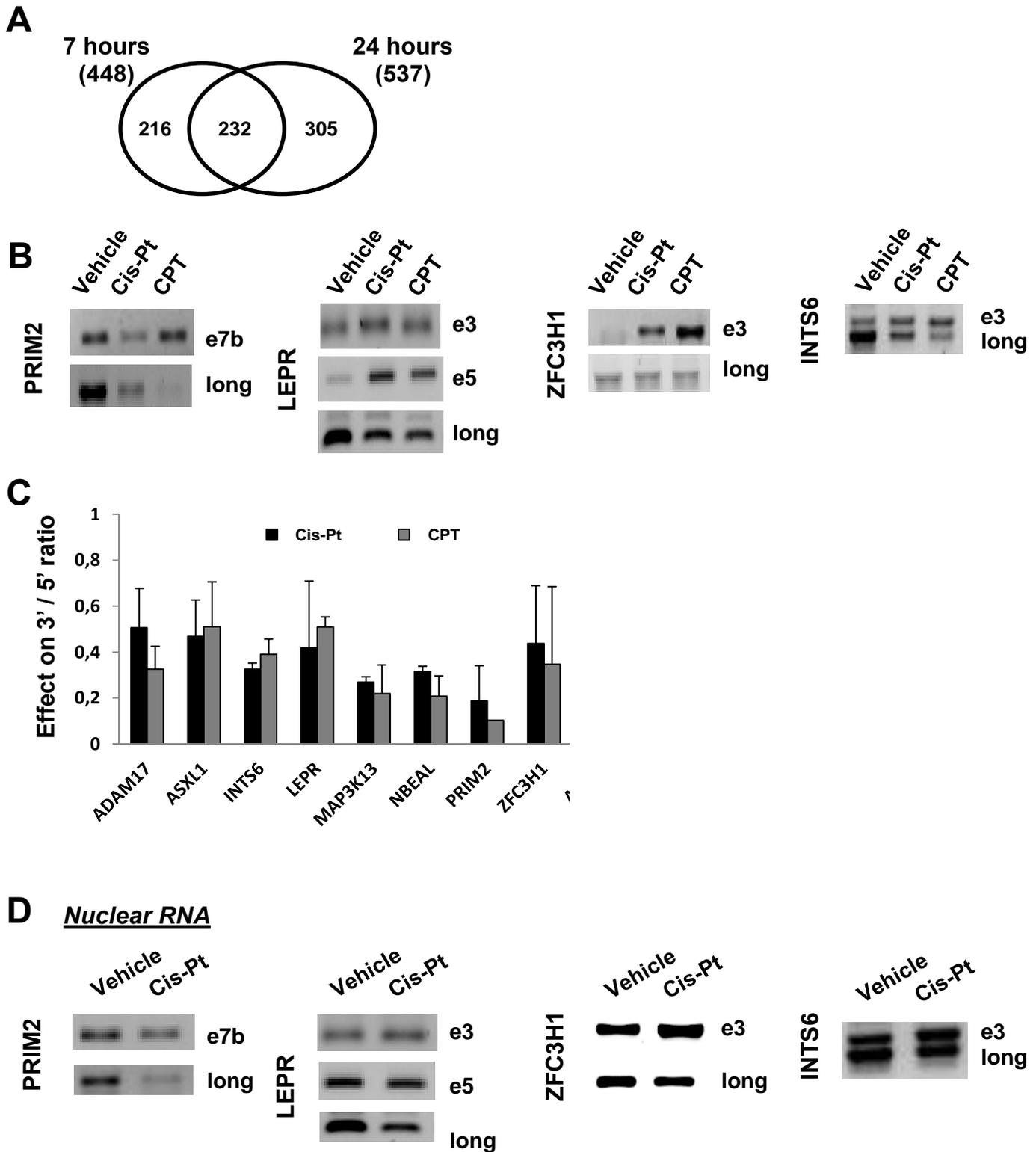
Suppl. Fig. S1: 3'-seq analysis.

A, 3'-seq method. **B**, Bioinformatic pipeline for ALE regulation analysis. **C**, Proportion of peaks with a potential polyA signal, either focusing on the main two motifs (AATAAA and ATTTAA) or including other previously identified variants. **D**, Percentage of regulated peaks that correspond to a polyA site in PolyA_DB_3. "Average cytosol" is an average of the data presented in the right panel.



Suppl. Fig. S2:

Proximal ALE up-regulation and elongation inhibition by cisplatin and camptothecin. **A**, Comparison of genes repressed by cisplatin at the elongation level at 7 and 24 hours. **B**, ALE regulation by cisplatin and camptothecin (CPT) in total RNA. **C**, Elongation inhibition by cisplatin and camptothecin. **D**, ALE regulation by cisplatin in nuclear RNA.



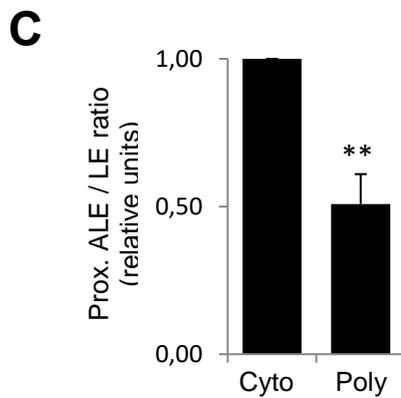
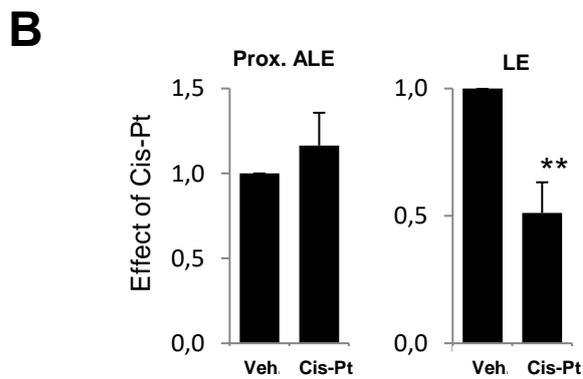
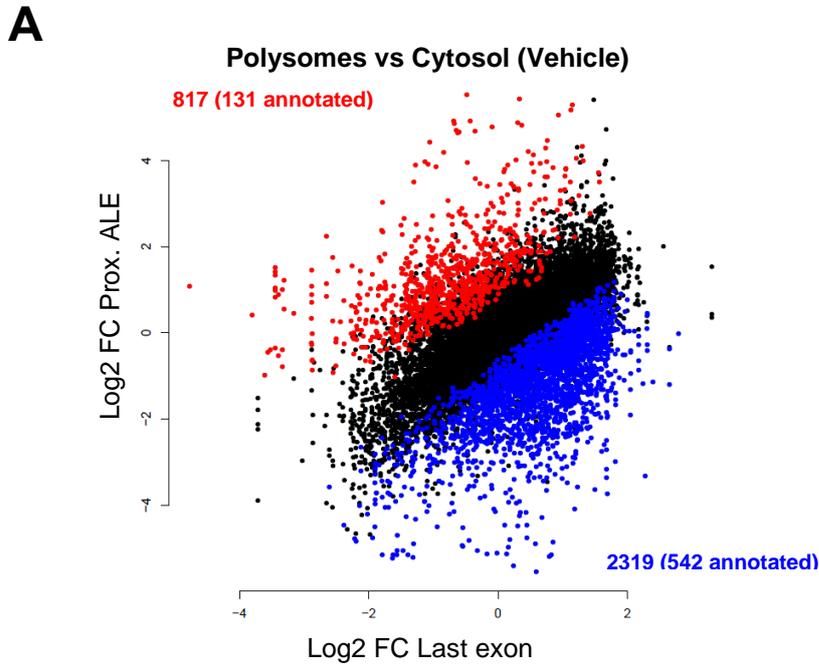
Suppl. Fig. S3 :

Enriched functions of genes where cisplatin up-regulates a proximal ALE that is annotated in PolyA_DB_3. Analysis with the DAVID software; the fold enrichment is indicated; only a subset of enriched functions are shown.

Term	Nb	p	Fold
negative regulation of EGFR signaling pathway	7	7E-04	6,4
negative regulation of extrinsic apoptotic signaling pathway	7	9E-04	6,1
mitotic spindle assembly	6	4E-03	5,5
epidermal growth factor receptor signaling pathway	7	7E-03	4,1
spindle	9	2E-02	2,6
negative regulation of cell growth	9	3E-02	2,5
MAPK cascade	15	3E-02	1,9
ErbB signaling pathway	8	4E-02	2,5
spindle microtubule	5	4E-02	3,9
response to stress	6	4E-02	3,2
double-strand break repair	6	5E-02	3,0

Suppl. Fig. S4:

A, 3'-seq analysis of ALE regulation in polysomes *versus* cytosol of untreated A549 cells (FDR 10%). **B**, RT-qPCR analysis of ASCC3 isoforms regulation by cisplatin in the cytosol of A549 cells. **C**, RT-qPCR analysis of PRIM2 isoforms ratio in the cytosol and polysomes of cisplatin-treated A549 cells.



DISCUSSION

Little is currently known about the role of AS and APA regulation in anticancer therapy sensitivity, despite the fact that post-transcriptional regulations widely modulate cancer cells transcriptomes and that therapy resistance is a major clinical issue. My thesis project aimed at increasing our understanding of AS and APA regulations in genotoxic anticancer drug sensitivity, specifically AS regulations in breast cancer cell resistance to doxorubicin and APA regulation in NSCLC cell response to cisplatin.

AS networks of resistance to doxorubicin in breast cancer cells

In my first project, I identified for the first time an AS pathway involved in breast cancer cell resistance to doxorubicin. In doxorubicin resistant vs. sensitive cells, thousands of AS events were regulated with enrichment in cell cycle and DNA repair genes. Poorly studied splicing factors ZRANB2 and SYF2, which regulate about 80 alternative exons each in our MCF-7 doxorubicin resistance model, partially re-sensitized the resistant cells to doxorubicin. Interestingly, among the very few target genes they had in common, inclusion of exon 5 of the oncogene ECT2 (a gene with double functions in cell cycle and DNA repair) was decreased when ZRANB2 and SYF2 were depleted with siRNAs. Further, depletion of ECT2-ex5 isoform also partially re-sensitized resistant cells to doxorubicin, and reduced doxorubicin-induced S phase accumulation. Finally, high expression of the ECT2-ex5 isoform detected by qPCR in ER+ breast cancer patients was associated with poor prognosis when they were treated with chemotherapy including anthracyclines, but not hormonotherapy.

Over the course of this project, I wondered if ZRANB2 or SYF2 depletion had the same effect on doxorubicin survival in other breast cancer cell lines (ER positive or negative). However, in the small panel of cell lines tested, ZRANB2 and/or SYF2 depletion was already toxic on its own without doxorubicin treatment (data not shown). I could not conclude on doxorubicin survival *per se* because of this unexpected toxicity effect, but this observation was still interesting: do some cells rely on AS pathways involving ZRANB2 and SYF2 for cell proliferation? Particularly high expression of ZRANB2 and/or SYF2 expression in DoxoR cells, might explain the lack of toxicity of their depletion in these cells. Indeed, for ZRANB2, mRNA and protein levels detected by qPCR and western-blot, respectively, were higher in DoxoR cells compared to the others (data not shown). Also, even at high siRNA concentration, I only managed to have a 60% depletion at the mRNA level for ZRANB2, suggesting that residual amounts of ZRANB2 could still be present in siRNA-depleted DoxoR cells. For instance, complete suppression (by CRISPR-Cas9 approach for example) could have been toxic in DoxoR cells even in unstressed conditions. To complement findings in this project, it would be interesting to overexpress ZRANB2, SYF2, and ECT2-ex5 isoform in MCF7-Parental cells to assess reverse effects, and to perform rescue experiments (doxorubicin survival, cell cycle arrest) by ECT2 \pm exon5 transfection in DoxoR cells after ZRANB2 or SYF2 depletion, in the presence and absence of doxorubicin.

The survival analysis was done on a cohort available at the Institut Curie (I. Bièche). It would be interesting to do *in silico* analysis on publicly available breast cancer patient RNA-seq data from the TCGA to confirm these findings. Survival curves are important indicators of the clinical relevance of the expression of a gene in a given context. However, they do not reflect causation. One of my ongoing experiments, with the help of the animal facility at Orsay research centre of the Institut Curie, aims to show this causal link of the ZRANB2/SYF2 → ECT2-ex5 AS pathway. We are analysing MCF7-DoxoR cells xenograft tumour progression in nude mice, in the presence or absence of doxorubicin treatment, after depletion of ZRANB2, SYF2, or ECT2-ex5 isoform by antisense *vivo*-morpholinos (VMOs, Genetools). For this experiment, I hypothesized that depletion of either ZRANB2, SYF2, or ECT2-ex5 would result in decreased tumour growth in response to doxorubicin treatment. If my hypothesis is confirmed, this would be the first demonstration of the role of converging AS pathways that involve ZRANB2, SYF2, and ECT2 in tumour progression. Indeed, ZRANB2 and SYF2 functions are poorly characterized. Further investigations on the other target exons specific to ZRANB2 and SYF2 and the cellular process they control would increase our understanding of these splicing networks.

An interesting result not fully pursued in the publication manuscript is the enrichment of RBFOX2 binding motif surrounding the AS events regulated in DoxoR vs. Parental cells. When we did a motif enrichment analysis, RBFOX2 motif had the strongest enrichment. RBFOX2 is overexpressed in the resistant cells, which is consistent with their mesenchymal-like state. RBFOX2 depletion did not however, re-sensitize the resistant cells to doxorubicin, probably because RBFOX2-mediated AS regulations do not participate directly in the doxorubicin resistance mechanisms. Remarkably however, I found that RBFOX2 regulates the AS of ZRANB2 (Figure 23). RBFOX2 upregulation could therefore be an upstream or early event controlling ZRANB2 expression levels, suggesting the existence of an AS network involving RBFOX2, ZRANB2, and SYF2 in breast cancer cell resistance to doxorubicin.

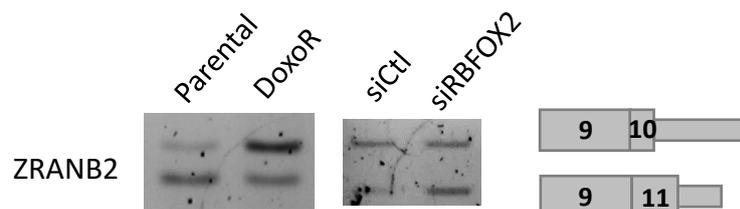


Figure 23: ZRANB2 AS regulation by RBFOX2.

A widespread IPA response to cisplatin in NSCLC cells

In my second project, I identified by 3'-seq widespread IPA upregulation in response to cisplatin in two NSCLC cell lines. These regulations were enriched in upregulation (with respect to the full length isoform) of particularly short isoforms in long genes, especially cell cycle and cell death related genes. Notably, IPA were poorly regulated by oxaliplatin, a

cisplatin analogue whose main cytotoxic mechanism was recently revealed not to be by genotoxicity. Further, 3'-seq analysis on polysomes showed that a large subset of these short IPA were not efficiently translated, suggesting (at least for some of them) non-coding functions.

Several observations suggest that induction of ALE by cisplatin is mediated by a defect in Pol II elongation. A more direct way to demonstrate this effect would be to carry a Pol II ChIP-seq and a global-run-on RNA sequencing (GRO-seq). The Pol II ChIP-seq and GRO-seq will reflect global Pol II occupancy on the chromatin and transcriptionally engaged Pol II, respectively. Doing so in cisplatin treated and untreated conditions, I should be able to confirm higher occupancy and activity of Pol II upstream of the polyadenylation sites of the ALE isoforms induced by cisplatin, compared to downstream. Additionally, it would be interesting to rigorously demonstrate transcription termination mechanisms at ALE sites. This could be done by investigating on which residues of Pol II CTD are phosphorylated (Ser2, Ser5, Tyr1 ?), and by which factors (CDK9, CDK12 ?) in cisplatin treated cells. Because cisplatin and Top1 inhibitors have been shown to have synergistic effects (Van Waardenburg *et al.*, 2004), I am currently looking into the role of Top1 in cisplatin-induced IPA regulation.

Ongoing experiments include functional characterization of PRIM2 and ASCC3 short ALE isoforms. Are these short isoforms necessary for cell survival after cisplatin-induced DNA damage? To test this, I am performing cell survival experiments similar to those in the AS and doxorubicin resistance project: isoform-specific depletion followed by drug treatment and cell survival assay. Hopefully, this will extend Williamson *et al.* findings on ASCC3 short isoform in response to UV in another context, and will describe for the first time a biological role for a truncated PRIM2 isoform. In addition, IPA regulation will also be analysed by 3'-seq in cisplatin responsive and non-responsive NSCLC patient tumours, in collaboration with Curie hospital. This will give further insights into the clinical relevance of IPA regulation in cisplatin (intrinsic) resistance.

More on the role of APA in cisplatin sensitivity

For the cisplatin project, we chose to focus on a poorly studied type of APA, IPA. Nonetheless, our 3'-seq data concern all 3' ends, and therefore detect 3' UTR-APA regulation as well. Genome-wide 3'UTR-APA regulation in response to cisplatin is not currently described in the literature, and will be an interesting complementary analysis to the one described in this manuscript. Given that UV irradiation for example have been shown to regulate 3'UTR-APA (Devany *et al.*, 2016), it would be expected that 3'UTR-APA is also modulated in response to cisplatin-induced genotoxic stress. In fact, by doing a quick visualisation of our raw data on UCSC Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>), by a candidate-gene approach inspired by 3' UTR-APA events described in the literature, some 3' UTR peaks seem regulated between the DMSO and Cisplatin conditions. This is the case for example for Cyclin D2 gene CCND2, as shown in Figure 24. In response to cisplatin, a distal polyadenylation site seem to be upregulated with respect to the couple of proximal peaks (3' UTR elongation) in H358 cells, suggesting a more important

microRNA repression after cisplatin-induced DNA damage. Proper 3' UTR-APA analysis in CDDP vs. DMSO cells, combined with microRNA and RBP motif search in the 3' UTR regions lost in the cisplatin response would be an interesting starting point to identify 3' UTR-APA regulation networks in the NSCLC cell response to cisplatin.

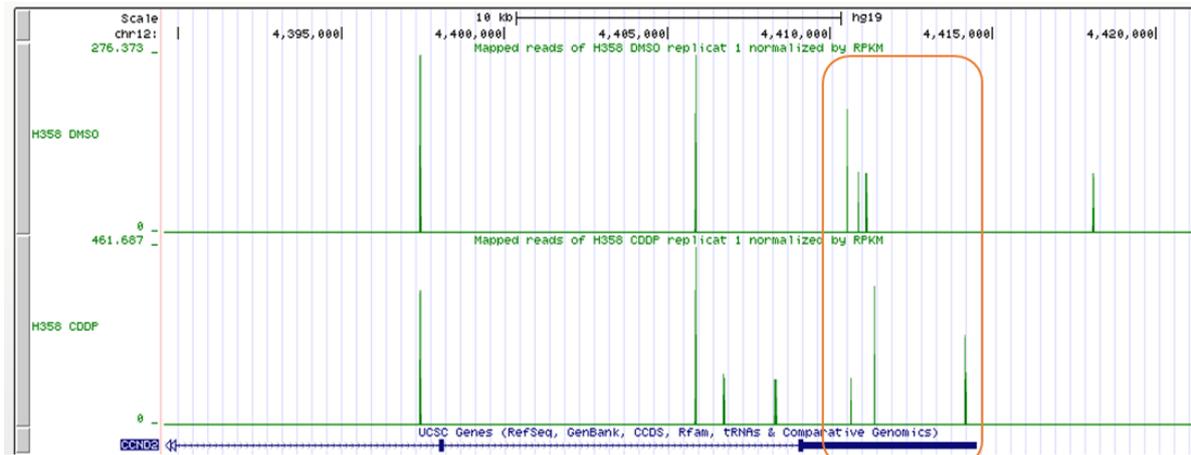


Figure 24: Raw 3'-seq peaks of the Cyclin D2 gene 3'UTR.

Other unexploited data are IPA regulations in untreated A549 cisplatin resistant vs. sensitive cells, shown in Figure 25. There are similar numbers of upregulated and downregulated IPA isoforms in resistant cells vs. sensitive cells. This suggests that IPA isoform relative amounts to the full-length isoform are regulated in the acquisition of cisplatin resistance in A549 cells. It would be interesting to investigate which switch of isoforms are involved in the resistance phenotype. IPA isoform often have a different, or antagonist function to the full-length isoform. Therefore, in A549 cisplatin resistant cells, upregulated IPA would suggest that the function of the full-length protein is deleterious for the resistant phenotype, and on the contrary, that downregulated ones have an advantageous function in the full-length isoform. To identify IPA switches responsible for the resistant phenotype, it would be interesting to perform a miniscreen similar to the one in the doxorubicin resistance project, but with IPA isoforms. After shortlisting a dozen of candidates from their functions as well as their IPA/ALE regulation levels (strong events), this miniscreen would test cisplatin survival after short and long isoform specific siRNA depletion, and hopefully identify isoform switches associated with cisplatin resistance.

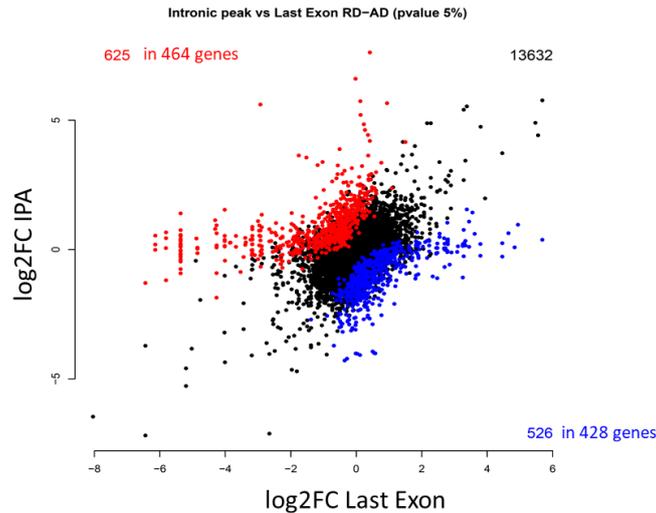


Figure 25: ALE regulation in A549 cisplatin-resistance vs. sensitive cells. (red: ALE/LE up, blue: ALE/LE down)

IPA as a way to express lncRNAs ?

A novel concept that emerges from ours and Williamson *et al.*'s studies, is the expression of DDR-involved lncRNAs by IPA in response to genotoxic stress. For this manuscript, I only tested and validated ASCC3 short isoform upregulation by cisplatin and its very low association with polysomes. However, there are potentially many other lncRNA induced by cisplatin in a similar manner (Figure 6C of the publication manuscript). It would be interesting to further investigate on these inefficiently translated, cisplatin-induced short transcripts. lncRNAs are particularly difficult to analyse because they are a class of ncRNAs mainly defined by exclusion, namely a lack of translated ORF, and definitive inclusion criteria have not been found yet. In fact, lncRNA and mRNA biogenesis and molecular characteristics are very similar. Assays to prove that a transcript is non-coding also relies on excluding criteria. Lack of protein signal at the expect molecular weight by western-blot analysis can show that this transcript do not encode a protein, provided the antibody is efficient and specific. One can also transfect the presumed lncRNA cDNA with mutation in the start or stop codon, and show that it is not its protein-coding ability that govern its function.

lncRNA candidates in the context of NSCLC cells response to cisplatin could be selected *in silico* and further individually investigated according to the flowchart below (Figure 26).

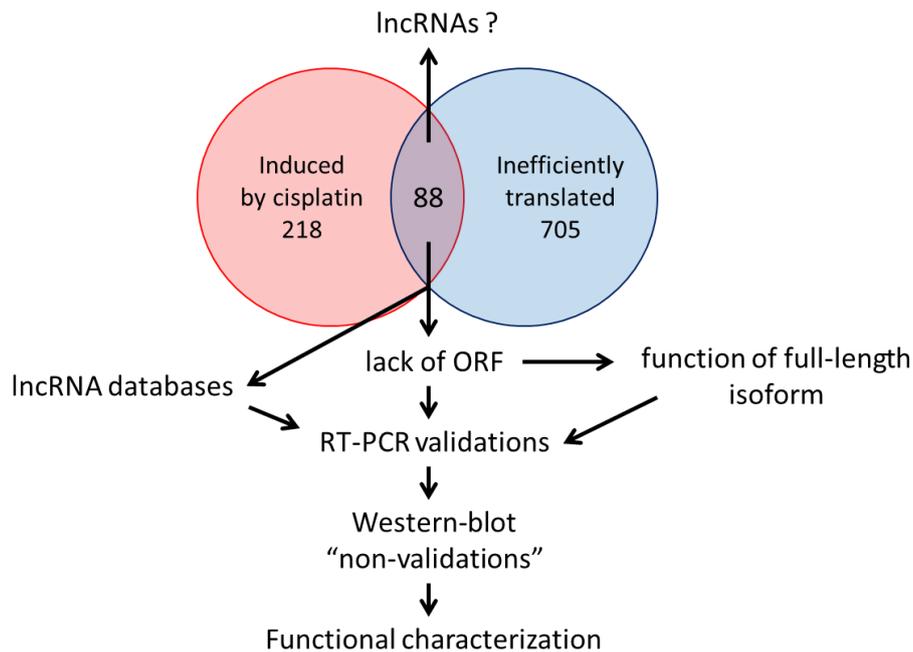


Figure 26: Workflow of cisplatin-induced, IPA-generated IncRNA investigation in A549 cells.

Here, I chose to put a $\log_2FC > 2$ cut-off on the IPA vs. Last Exon ratio to select for the strongest effects both in the cisplatin-induced and the inefficiently translated conditions. Additionally, I would not consider transcript length although IncRNA are often described as > 200 nucleotides. Indeed, this is more of an empirical and arbitrary cut-off. A 100 nucleotides transcript for example, would be too long to be a short ncRNA (e.g. microRNA) but would be excluded from this analysis. It would also be useful to verify if some of these candidates are known IncRNAs, by inquiring ncRNA databases such as Incrnadb (<http://www.lncrnadb.org>), Incipedia (<https://hg19.lncipedia.org>), noncode (<http://www.noncode.org>), and RNA central (<https://rnacentral.org>). After shortlisting a dozen of candidates this way and including some with interesting functions of the host gene, I would perform RT-PCR validations to verify the IPA/IncRNA transcript expression. Of those which would be validated by RT-PCR, verification of the absence of a peptide will be done before functional characterization starting with the IPA/IncRNA effect on cisplatin survival.

At this stage, one or two top candidate(s) would be selected for mechanistic investigations. As a IncRNA can act in various ways (Figure 27a), it would be tricky to identify its precise role in cellular processes. Identification of its direct RNA or protein interactors would be the key to understanding their functions. To do so, RNA-centric approaches of RNA-protein and RNA-RNA interactions analysis are needed. In this regard, RNA affinity purification (RAP) developed by the M. Guttman lab in Caltech is particularly interesting. Direct and indirect RNA and protein interactors of a IncRNA can be co-purified by pull down using biotinylated probes which form stable RNA-DNA hybrids with the IncRNA of interest, then the purified interacting RNAs and proteins are submitted to RNA-seq and mass-spectrometry, respectively (Figure 27b). With this method, depending on the protein

digestion and RNA fragmentation conditions, direct or indirect interactions can be separately assessed. Combining RAP-RNA-seq and RAP-MS would allow a comprehensive lncRNA interactome description, which can be used as a starting point for selected lncRNA in depth mechanistic investigations.

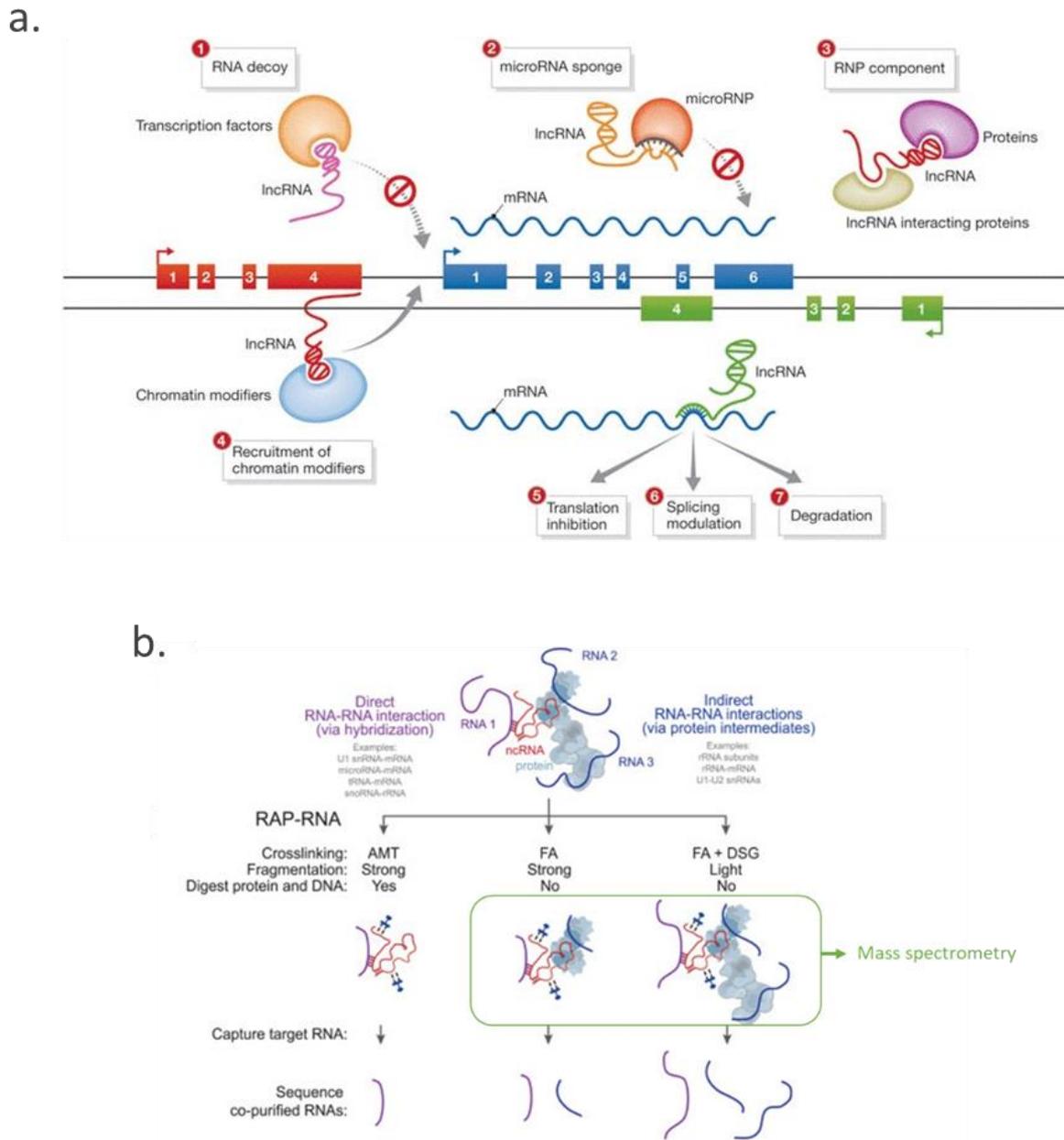


Figure 27: a) Modes of action of lncRNAs. (adapted from W. Hu et al., EMBO reports, 2012)

b) RAP-seq and RAP-MS methods. (adapted from Engreitz et al., Cell 2014)

Clinical relevance of AS and APA networks: from bedside to the bench

Molecular characterization of RNA processing networks is fundamental for our understanding of the cellular machineries complexity in normal and pathological conditions, in stressed and unstressed conditions. To link these mechanisms to the clinical context,

hypothesis testing in patient samples is necessary. Alternatively, RNA processing networks identifications in clinical samples would give us “real-life” demonstration of their relevance.

For example, a complementary approach of RBPome and transcriptome analysis from RNA extracted from a selected cohort of cancer patients would give novel insights on clinically relevant AS and APA regulations. In a hypothetic novel study, I would want to identify AS and APA regulations and their mechanisms that are important for triple-negative breast cancer (TNBC) patients’ response to chemotherapy. The rationale behind this choice of patients is because TNBC patients cannot benefit from hormone or HER2-targeted therapies, and their prognosis is generally quite poor. Identifying mechanisms that predict therapy response would help manage treatment regimens and eventually lead to the development of new targeted therapies. The general workflow for this project is as shown below in Figure 28a.

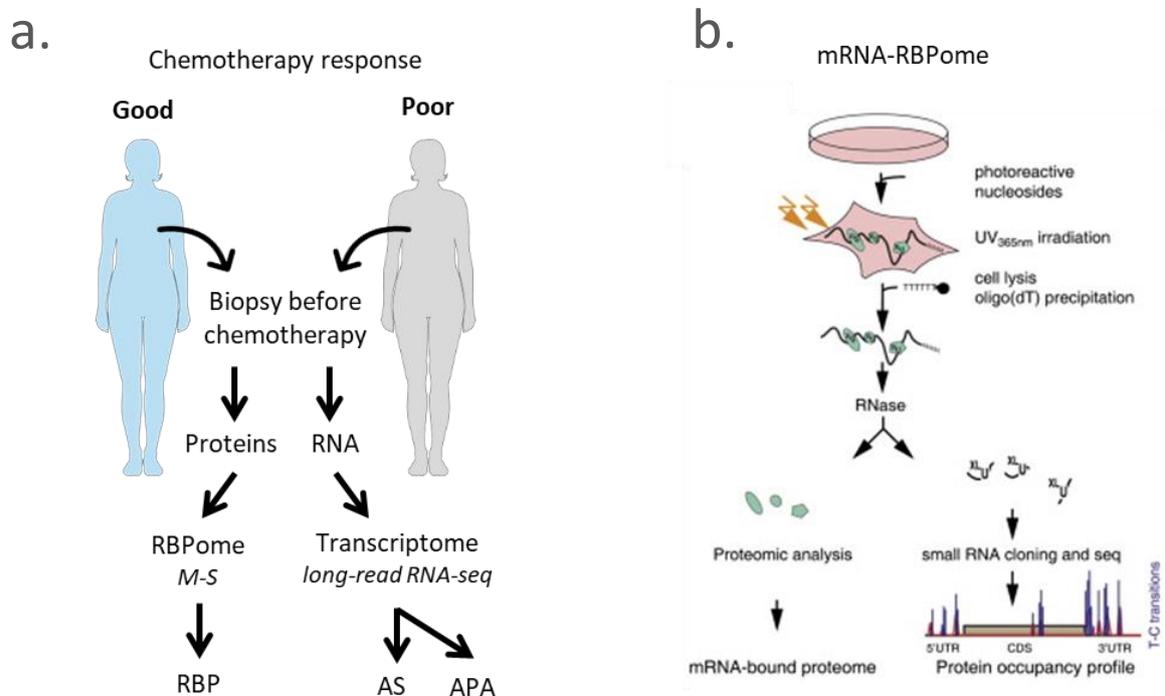


Figure 28: Investigation on clinically relevant AS and APA networks. a) Workflow of RBP, AS and APA networks identification in patient samples. b) Method for mRNA RBPome identification. (adapted from Baltz et al., Mol Cell 2012)

For this project, tumour samples of TNBC patients before neoadjuvant chemotherapy and the information on their response to chemotherapy would be needed. RNA and protein would be extracted from these tumours, separated in groups of “good responders” and “poor responders”. Tumour protein and RNA would be used to perform RBPome and transcriptome analysis, respectively. RBPome analysis would be done by mass-spectrometry (M-S) on purified proteins from tumours. A more interesting, but technically difficult approach would be to identify RBPs and their binding sites in parallel, by an mRNA-bound RBPome analysis developed by M. Landthaler lab (Figure 28b). To perform this on tumour biopsies, it would be necessary to put them in primary culture first in order to be able to do

the UV crosslink of RBP-RNA complexes. AS and APA analysis could be done by RNA-seq and 3'-seq, but a more innovative approach would be to use long-read RNA-seq (with PacBio technology for example) to identify transcript isoforms in full, and thus all combinations of AS and APA variants. This sequencing technique would reflect better the relative amounts of specific isoforms in the samples. RBP and their binding sites, as well as AS and APA transcript isoforms being identified, verification of their interactions and regulations could be done in cell line models. For validated networks, further mechanistical and functional characterization could also be pursued in cell lines.

The major limitation of this project is the initial input of RNA and proteins. The techniques presented here require large amount of high-quality RNA and proteins, which is not easily obtained from tumour samples. Better tumour protein and RNA extraction methods as well as optimization of CLIP and long-read sequencing techniques to reduce the RNA input in the future could allow the realization of this hypothetical project.

Hopes for the future: RNA therapeutics

Given the now recognized AS involvement in oncogenesis, strategies to target splicing, whether it is core spliceosome components, splicing factors, or AS variants, in cancer therapy have already started to emerge. There are various strategies to therapeutically target AS, as shown in Figure 29. Antisense oligonucleotides (ASOs) for example hybridize RNA in a sequence-specific manner and can be designed to modulate splicing by shielding splicing regulatory elements (splice-switching oligonucleotides, SSO). While SSOs have already shown clinical efficacy for treatment of neuromuscular diseases, oncogenic AS manipulation by SSO in cancer seems to be more challenging. This can be due to general AS alteration in cancer cells, loose specificity of splicing regulatory elements, and compensatory mechanisms of splicing factors. One very recent and encouraging study however demonstrated the efficacy of a SSO promoting the expression of the tumour-suppressive variant of MKNK2 (Mnk2a) involved in the activation of the p38-MAPK apoptosis pathway, in impairing glioblastoma development *in vivo* (Mogilevsky *et al.*, 2018).

Further research on AS networks and on the development of molecules that modulate AS are needed to increase our understanding of oncogenic AS mechanisms and their therapeutical targetability. In the doxorubicin project, I am currently testing the effect of ECT2-exon5 targeting SSO on xenograft tumour growth in response to doxorubicin. It would also be interesting to target IPA isoforms associated with cisplatin response or resistance identified in my second project. ASOs are particularly interesting for their almost infinite versatility. As they are custom-made, they can theoretically target any RNA sequence and any type of RNA, including splice sites of ALEs, 3' end processing cis-regulatory sequences, 5' regions of mRNAs important for translation initiation, and ncRNAs such as microRNAs and lncRNAs. They could also potentially be administered in a cocktail to target multiple levels of an oncogenic, or resistance-associated RNA processing network.

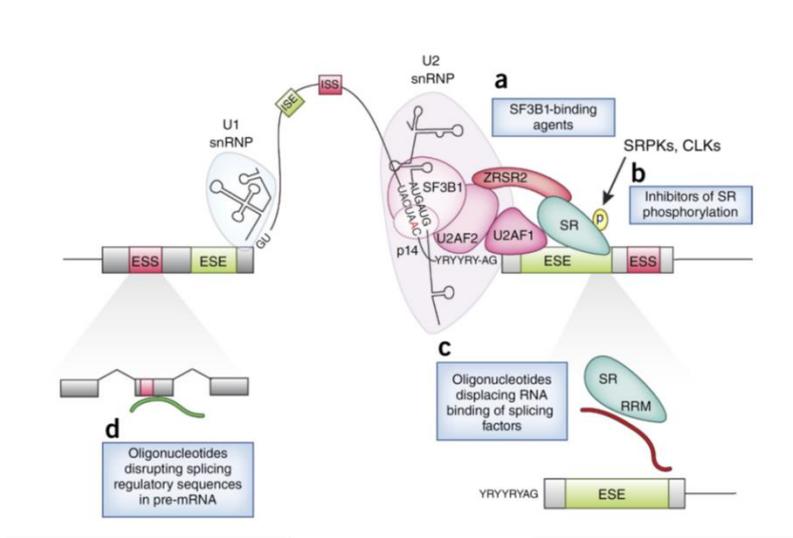


Figure 29: Therapeutic strategies targeting AS. (adapted from Lee et al., Nat Med Rev 2016)

SYNTHÈSE EN FRANÇAIS

Dans la prise en charge des cancers, malgré l'émergence des thérapies ciblées, les chimiothérapies conventionnelles génotoxiques sont encore largement utilisées en clinique. Pour beaucoup de cancers elles restent la stratégie thérapeutique principale. Malheureusement, émergent souvent chez les patients des résistances aux traitements anti-cancéreux, qu'ils soient génotoxiques ou non. Il est donc important et nécessaire de mieux comprendre les mécanismes moléculaires de chimiorésistance afin de surmonter ce problème clinique.

Outre d'importantes régulations transcriptionnelles, on sait aujourd'hui que les régulations post-transcriptionnelles, et notamment l'épissage alternatif (AS, alternative splicing) et la polyadénylation alternative (APA, alternative polyadenylation), jouent un rôle clé dans l'oncogenèse. On sait aussi que les agents génotoxiques utilisés en chimiothérapie, tels que la doxorubicine et le cisplatine, régulent l'AS et l'APA. En revanche, l'AS et l'APA ont été peu étudiés dans la résistance aux traitements anti-cancéreux. On connaît donc très peu l'étendue, les mécanismes et les fonctions de ces régulations dans la résistance aux agents anticancéreux génotoxiques. L'AS, qui concerne plus de 90% des gènes, permet de générer des isoformes d'ARN-messager codant des isoformes protéiques avec des fonctions différentes. L'implication de l'AS dans le cancer est maintenant admise. L'APA, moins étudiée, peut concerner jusqu'à 60% des gènes, et permet de générer entre autres des isoformes tronquées utilisant des exons terminaux alternatifs (ALE, alternative last exons) par polyadénylation intronique. L'APA a été démontrée récemment comme étant régulés à grande échelle dans les cancers, mais très peu d'études prennent en compte les ALE et la polyadénylation intronique.

Mon hypothèse est que la sensibilité aux chimiothérapies est en partie médiée par une modulation de la maturation des ARNs pré-messagers. Cette modulation serait orchestrée par des facteurs contrôlant un ensemble d'exons alternatifs dans des gènes impliqués dans des processus liés à la chimio-sensibilité. Afin de mieux comprendre le rôle de ces modulations dans la sensibilité aux agents anti-cancéreux génotoxiques, mon projet de thèse a 2 objectifs : 1) Déterminer l'étendue, les mécanismes, et les fonctions des régulations d'AS dans la résistance acquise à la doxorubicine, et 2) Déterminer la régulation des ALE dans la réponse et la résistance au cisplatine. La doxorubicine et le cisplatine sont deux agents anti-cancéreux génotoxiques parmi les plus anciens et les plus couramment utilisés, sur les cancers du sein et du poumon respectivement. Cependant, le rôle des réseaux d'AS et d'APA dans la réponse et la résistance à ces agents est très peu connu.

Dans l'introduction de ce manuscrit de thèse, j'ai d'abord exposé les connaissances actuelles sur les mécanismes des régulations de l'AS et l'APA dans des conditions normales, pathologiques (cancer), et de stress génotoxique. Dans une seconde partie, j'ai décrit quelques généralités sur les agents anti-cancéreux génotoxiques et la sensibilité cellulaire à

ces agents, plus particulièrement la doxorubicine et le cisplatine. Enfin, j'ai synthétisé dans une troisième partie les dernières avancées sur les relations entre l'AS, l'APA, et la réponse aux dommages de l'ADN ainsi que la chimiorésistance.

La partie résultats de ce manuscrit de thèse consiste en deux manuscrits d'articles, prêts à soumission : 1) Identification of splicing programs and pathways involved in breast cancer cell resistance to doxorubicin, et 2) Intronic polyadenylation is linked to the antiproliferative effect of specific platinum compounds and generates differentially translated isoforms.

Dans le premier projet, j'ai identifié plus de 2000 évènements d'AS régulés dans des cellules de cancer du sein MCF7 résistantes à la doxorubicine par rapport aux cellules parentales sensibles. Un mini-screen siARN sur 40 facteurs d'épissage régulés dans ce modèle de résistance m'a permis d'identifier 2 facteurs d'épissage peu connus, ZRANB2 et SYF2, dont la déplétion resensibilise partiellement les cellules résistantes à la doxorubicine. Par RNA-seq sur les cellules résistantes après transfection de siARNs ciblant ZRANB2 et SYF2, j'ai identifié environ 80 évènements d'AS régulés par chaque facteur, principalement des exons cassettes et des sites 3' d'épissage, respectivement. De façon intéressante, ces régulations sont enrichies dans des fonctions liées à la résistance à la doxorubicine (prolifération, cycle, réparation de l'ADN). L'exon 5 de l'oncogène ECT2 codant pour un domaine BRCT (BRCA1 C-Terminal, important dans la réparation de l'ADN) est le seul à être régulé à la fois par ZRANB2 et SYF2. La déplétion spécifique de l'isoforme incluant l'exon 5 resensibilise aussi partiellement les cellules à la doxorubicine, et limite l'accumulation en phase S du cycle cellulaire induit par la doxorubicine dans ces cellules. De plus, l'inclusion de l'exon 5 de ECT2 est associé à un mauvais pronostic de survie sans métastases chez un sous-groupe de patientes de cancer du sein traitées à la chimiothérapie comprenant de la doxorubicine.

Dans le deuxième projet, j'ai pu identifier par 3'-seq que le principal type d'exon induit par le cisplatine dans les cellules de poumon non à petites cellules sont les ALEs. Dans 95% des cas, le cisplatine induit des ALE promoteur-proximaux, potentiellement codant des isoformes protéiques tronquées. Ces régulations sont enrichies dans de longs gènes impliqués dans la réponse aux dommages de l'ADN, du cycle et de la mort cellulaire, et sont médiées par une inhibition de l'élongation de la transcription. Ces régulations d'ALEs ont également permis de distinguer les cellules sensibles et résistantes au cisplatine. Un autre platine, l'oxaliplatine, qui a moins d'effets que le cisplatine sur l'inhibition de la prolifération cellulaire, n'a pas été capable d'induire de telles régulations, bien qu'il induise une régulation importante au niveau de l'expression des ARNs messagers. De plus, des analyses 3'-seq sur des fractions de polysomes (ARN messagers associés à des ribosomes) ont démontré l'impact du cisplatine sur le traductome : un groupe d'isoformes ALE particulièrement courtes sont très peu traduites en réponse au traitement, et notamment celle d'un transcrit dont la fonction non-codante a été révélée récemment.

En conclusion, ces travaux mettent en évidence de nouveaux concepts de réseaux de régulation post-transcriptionnelles et leur complexité dans le contexte de la chimiorésistance. Sur le modèle de cancer du sein, j'ai pu mettre en évidence pour la première fois un programme d'AS impliqué dans la résistance à la doxorubicine. Sur le modèle de cancer du poumon, j'ai pu identifier une régulation à grande échelle des ALE par le cisplatine et leur impact sur le traductome. Au plus long terme, les facteurs d'épissage, les isoformes d'AS ou d'ALE identifiés pourront potentiellement servir de nouveaux marqueurs moléculaires de chimiorésistance, voire de cible thérapeutique.

Dans la partie discussion de ce manuscrit de thèse, j'aborde plusieurs perspectives possibles à ces travaux, à court et moyen-long terme, notamment l'étude de l'expression de longs ARNs non-codant (lncRNAs) par polyadénylation intronique et l'étude à grande échelle de réseaux d'AS et d'APA à partir de tumeurs de patients. En effet, un concept intéressant émergeant du projet sur les ALEs en réponse au cisplatine, est un mécanisme favorisant l'expression de lncRNAs par polyadénylation intronique en réponse à un stress génotoxique. Il serait intéressant d'étudier plus en profondeur ce type de transcrits, en confirmant leur fonctionnalité non-codante, et en étudiant leur effet sur la réponse au cisplatine ainsi que les protéines liant ces lncRNAs par des approches de précipitation d'ARN couplé à une spectrométrie de masse, dans les cellules de cancer non à petite cellules. L'autre principale perspective discutée vise à identifier les transcrit issus d'AS et d'APA et leur régulateurs (des protéines liant l'ARN) directement dans des tumeurs de patients. Je propose dans cette partie d'étudier les cancers du sein triple-négatifs (n'exprimant pas les récepteurs hormonaux et ERBB2) pour leur caractère particulièrement agressif et difficile à traiter. Dans des groupes de tumeurs issues de patients répondant et non répondant à la chimiothérapie, les transcrits d'AS et d'APA seraient identifiés par une approche innovante de séquençage « long-read » et les protéines par spectrométrie de masse. Ceci permettrait de mettre en évidence les réseaux d'AS et d'APA directement avec une relevance clinique, impliqués dans la résistance intrinsèque des cancers du sein triple-négatifs à la chimiothérapie.

REFERENCES

- Aabo, K. *et al.* (1998) 'Chemotherapy in advanced ovarian cancer: Four systematic meta-analyses of individual patient data from 37 randomized trials', *British Journal of Cancer*. doi: 10.1038/bjc.1998.710.
- Adamson, B. *et al.* (2012) 'A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response', *Nature Cell Biology*. doi: 10.1038/ncb2426.
- Adesso, L. *et al.* (2013) 'Gemcitabine triggers a pro-survival response in pancreatic cancer cells through activation of the MNK2/eIF4E pathway', *Oncogene*. doi: 10.1038/onc.2012.306.
- Aebi, S. *et al.* (1996) 'Loss of DNA mismatch repair in acquired resistance to cisplatin', *Cancer Research*.
- Albert, H. *et al.* (2012) 'Genotoxic stress modulates CDC25C phosphatase alternative splicing in human breast cancer cell lines', *Molecular Oncology*. doi: 10.1016/j.molonc.2012.06.003.
- Alekseyenko, A. V., Kim, N. and Lee, C. J. (2007) 'Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes', *RNA*. doi: 10.1261/rna.325107.
- Alsafadi, S. *et al.* (2016) 'Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage', *Nature Communications*. Nature Publishing Group, 7, pp. 1–12. doi: 10.1038/ncomms10615.
- Alt, F. W. *et al.* (1980) 'Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends', *Cell*. doi: 10.1016/0092-8674(80)90615-7.
- Amable, L. (2016) 'Cisplatin resistance and opportunities for precision medicine', *Pharmacological Research*. doi: 10.1016/j.phrs.2016.01.001.
- Anczuków, O. *et al.* (2015) 'SRSF1-Regulated Alternative Splicing in Breast Cancer', *Molecular Cell*, 60(1), pp. 105–117. doi: 10.1016/j.molcel.2015.09.005.
- Anvar, S. Y. *et al.* (2018) 'Full-length mRNA sequencing uncovers a widespread coupling between transcription initiation and mRNA processing', *Genome Biology*. *Genome Biology*, 19(1), pp. 1–18. doi: 10.1186/s13059-018-1418-0.
- Arai, N. *et al.* (1986) 'Immunologically distinct p53 molecules generated by alternative splicing.', *Molecular and cellular biology*. doi: 10.1128/MCB.6.9.3232.
- Ashour, M. E., Atteya, R. and El-Khamisy, S. F. (2015) 'Topoisomerase-mediated chromosomal break repair: An emerging player in many games', *Nature Reviews Cancer*. doi: 10.1038/nrc3892.
- Bates, D. O. *et al.* (2002) 'VEGF165b, an inhibitory splice variant of vascular endothelial growth factor, is down-regulated in renal cell carcinoma', *Cancer Research*. doi: 10.1038/35025220.
- Bava, F. A. *et al.* (2013) 'CPEB1 coordinates alternative 3'-UTR formation with translational regulation', *Nature*, 495(7439), pp. 121–125. doi: 10.1038/nature11901.
- Beaudoing, E. *et al.* (2000) 'Patterns of variant polyadenylation signal usage in human genes', *Genome Research*. doi: 10.1101/gr.10.7.1001.
- Bebee, T. W. *et al.* (2015) 'The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development', *eLife*. doi: 10.7554/eLife.08954.
- Beisang, D., Reilly, C. and Bohjanen, P. R. (2014) 'Alternative polyadenylation regulates CELF1/CUGBP1 target transcripts following T cell activation', *Gene*. Elsevier B.V., 550(1), pp. 93–100.

doi: 10.1016/j.gene.2014.08.021.

Bentley, D. L. (2014) 'Coupling mRNA processing with transcription in time and space', *Nature Reviews Genetics*. Nature Publishing Group, 15(3), pp. 163–175. doi: 10.1038/nrg3662.

Berg, M. G. *et al.* (2012) 'U1 snRNP determines mRNA length and regulates isoform expression', *Cell*. doi: 10.1016/j.cell.2012.05.029.

Berget, S. M., Moore, C. and Sharp, P. A. (1977) 'Spliced segments at the 5' terminus of adenovirus 2 late mRNA', *Proceedings of the National Academy of Sciences*, 74(8), pp. 3171–3175. doi: 10.1073/pnas.74.8.3171.

Betticher DC1, Thatcher N, Altermatt HJ, Hoban P, Ryder WD, H. J. (1995) 'Alternate splicing produces a novel cyclin D1 transcript', *Oncogene*.

Beyer, A. L. and Osheim, Y. N. (1988) 'Splice site selection, rate of splicing, and alternative splicing on nascent transcripts.', *Genes & development*, 2(6), pp. 754–765. doi: 10.1101/gad.2.6.754.

Biamonti, G. *et al.* (2014) 'The alternative splicing side of cancer', *Seminars in Cell and Developmental Biology*. doi: 10.1016/j.semcdb.2014.03.016.

Blackford, A. N. and Jackson, S. P. (2017) 'ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response', *Molecular Cell*. doi: 10.1016/j.molcel.2017.05.015.

Blencowe, B. J. (2017) 'The Relationship between Alternative Splicing and Proteomic Complexity', *Trends in Biochemical Sciences*. Elsevier Ltd, 42(6), pp. 407–408. doi: 10.1016/j.tibs.2017.04.001.

Boelens, W. C. *et al.* (1993) 'The human U1 snRNP-Specific U1A protein inhibits polyadenylation of its own pre-mRNA', *Cell*. doi: 10.1016/0092-8674(93)90577-D.

Boise, L. H. *et al.* (1993) 'Bcl-X, a Bcl-2-Related Gene That Functions As a Dominant Regulator of Apoptotic Cell Death', *Cell*, 74(4), pp. 597–608. doi: 10.1016/0092-8674(93)90508-N.

Bonnet, A. *et al.* (2017) 'Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability', *Molecular Cell*. doi: 10.1016/j.molcel.2017.07.002.

Bosetti, C. *et al.* (2013) 'Cancer mortality in Europe, 2005-2009, and an overview of trends since 1980', *Annals of Oncology*. doi: 10.1093/annonc/mdt301.

Boutz, P. L. *et al.* (2007) 'A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons', *Genes and Development*. doi: 10.1101/gad.1558107.

Bouwman, P. and Jonkers, J. (2012) 'The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance', *Nature Reviews Cancer*. doi: 10.1038/nrc3342.

Bowden, N. A. (2014) 'Nucleotide excision repair: Why is it not used to predict response to platinum-based chemotherapy?', *Cancer Letters*. doi: 10.1016/j.canlet.2014.01.005.

Braunschweig, U. *et al.* (2014) 'Widespread intron retention in mammals functionally tunes transcriptomes', *Genome Research*. doi: 10.1101/gr.177790.114.

Breglio, A. M. *et al.* (2017) 'Cisplatin is retained in the cochlea indefinitely following chemotherapy', *Nature Communications*. doi: 10.1038/s41467-017-01837-1.

Brody, E. and Abelson, J. (1985) 'The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction.', *Science (New York, N.Y.)*, 228(4702), pp. 963–967. doi: 10.1126/science.3890181.

Brumbaugh, J. *et al.* (2018) 'Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling', *Cell*. doi: 10.1016/j.cell.2017.11.023.

Bruno, P. M. *et al.* (2017) 'A subset of platinum-containing chemotherapeutic agents kills cells by

- inducing ribosome biogenesis stress', *Nature Medicine*. doi: 10.1038/nm.4291.
- Calabretta, S. *et al.* (2016) 'Modulation of PKM alternative splicing by PTBP1 promotes gemcitabine resistance in pancreatic cancer cells', *Oncogene*. doi: 10.1038/onc.2015.270.
- Carrillo Oesterreich, F. *et al.* (2016) 'Splicing of Nascent RNA Coincides with Intron Exit from RNA Polymerase II', *Cell*. doi: 10.1016/j.cell.2016.02.045.
- Castellani, P. *et al.* (1994) 'The fibronectin isoform containing the ed-b oncofetal domain: A marker of angiogenesis', *International Journal of Cancer*. doi: 10.1002/ijc.2910590507.
- Castelo-Branco, P. *et al.* (2004) 'Polypyrimidine tract binding protein modulates efficiency of polyadenylation.', *Molecular and cellular biology*. doi: 10.1128/MCB.24.10.4174.
- Ceccaldi, R., Sarangi, P. and D'Andrea, A. D. (2016) 'The Fanconi anaemia pathway: New players and new functions', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm.2016.48.
- Cepeda, V. *et al.* (2007) 'Biochemical Mechanisms of Cisplatin Cytotoxicity', *Anti-Cancer Agents in Medicinal Chemistry*. doi: 10.2174/187152007779314044.
- Cevher, M. A. *et al.* (2010) 'Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage', *EMBO Journal*. doi: 10.1038/emboj.2010.59.
- Cevher, M. A. and Kleiman, F. E. (2010) 'Connections between 3'-end processing and DNA damage response', *Wiley Interdisciplinary Reviews: RNA*. doi: 10.1002/wrna.20.
- Chabner, B. A. and Roberts, T. G. (2005) 'Chemotherapy and the war on cancer', *Nature Reviews Cancer*. doi: 10.1038/nrc1529.
- Chandler, D. S. *et al.* (2006) 'Genotoxic stress induces coordinately regulated alternative splicing of the p53 modulators MDM2 and MDM4', *Cancer Research*. doi: 10.1158/0008-5472.CAN-05-4271.
- Chang, J. W. *et al.* (2015) 'MRNA 3'-UTR shortening is a molecular signature of mTORC1 activation', *Nature Communications*. doi: 10.1038/ncomms8218.
- Chen, F. C. (2015) 'Alternative RNA structure-coupled gene regulations in tumorigenesis', *International Journal of Molecular Sciences*, 16(1), pp. 452–475. doi: 10.3390/ijms16010452.
- Chen, H. H. W. *et al.* (2008) 'Elevated Glutathione Levels Confer Cellular Sensitization to Cisplatin Toxicity by Up-Regulation of Copper Transporter hCtr1', *Molecular Pharmacology*. doi: 10.1124/mol.108.047969.
- Chen, H. H. W. *et al.* (2012) 'Predictive and prognostic value of human copper transporter 1 (hCtr1) in patients with stage III non-small-cell lung cancer receiving first-line platinum-based doublet chemotherapy', *Lung Cancer*. doi: 10.1016/j.lungcan.2011.06.011.
- Chen, M. and Manley, J. L. (2009) 'Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm2777.
- Chen, S. *et al.* (2018) 'Interaction of WBP2 with ER α increases doxorubicin resistance of breast cancer cells by modulating MDR1 transcription', *British Journal of Cancer*. doi: 10.1038/s41416-018-0119-5.
- Cheng, J. *et al.* (1994) 'Protection from Fas-mediated apoptosis by a soluble form of the Fas molecule', *Science*. doi: 10.1126/science.7510905.
- Cheung-Ong, K., Giaever, G. and Nislow, C. (2013) 'DNA-damaging agents in cancer chemotherapy: Serendipity and chemical biology', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2013.04.007.
- Chien, A. J. and Moasser, M. M. (2008) 'Cellular Mechanisms of Resistance to Anthracyclines and Taxanes in Cancer: Intrinsic and Acquired', *Seminars in Oncology*. doi: 10.1053/j.seminoncol.2008.02.010.
- Chow, L. T. *et al.* (1977) 'An amazing sequence arrangement at the 5' ends of adenovirus 2

messenger RNA', *Cell*, 12(1), pp. 1–8. doi: 10.1016/0092-8674(77)90180-5.

Christofk, H. R. *et al.* (2008) 'The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth', *Nature*. doi: 10.1038/nature06734.

Cieply, B. and Carstens, R. P. (2015) 'Functional roles of alternative splicing factors in human disease', *Wiley Interdisciplinary Reviews: RNA*, 6(3), pp. 311–326. doi: 10.1002/wrna.1276.

Clauson, C., Schäfer, O. D. and Niedernhofer, L. (2013) 'Advances in understanding the complex mechanisms of DNA interstrand cross-link repair.', *Cold Spring Harbor perspectives in biology*. doi: 10.1101/cshperspect.a012732.

Climente-González, H. *et al.* (2017) 'The Functional Impact of Alternative Splicing in Cancer', *Cell Reports*, 20(9), pp. 2215–2226. doi: 10.1016/j.celrep.2017.08.012.

Cobbold, L. C. *et al.* (2010) 'Upregulated c-myc expression in multiple myeloma by internal ribosome entry results from increased interactions with and expression of PTB-1 and YB-1', *Oncogene*. doi: 10.1038/onc.2010.31.

Colgan, D. F. *et al.* (1996) 'Cell-cycle related regulation of poly(A) polymerase by phosphorylation', *Nature*. doi: 10.1038/384282a0.

Collesi, C. *et al.* (1996) 'A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype.', *Molecular and Cellular Biology*. doi: 10.1128/MCB.16.10.5518.

Comiskey, D. F. *et al.* (2015) 'Splicing factor SRSF1 negatively regulates alternative splicing of MDM2 under damage', *Nucleic Acids Research*. doi: 10.1093/nar/gkv223.

Corsini, L. *et al.* (2007) 'U2AF-homology motif interactions are required for alternative splicing regulation by SPF45', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb1260.

Daguenet, E., Dujardin, G. and Valcarcel, J. (2015) 'The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches', *EMBO reports*, 16(12), pp. 1640–1655. doi: 10.15252/embr.201541116.

Dai, W., Zhang, G. and Makeyev, E. V. (2012) 'RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage', *Nucleic Acids Research*. doi: 10.1093/nar/gkr783.

Danan-Gotthold, M. *et al.* (2015) 'Identification of recurrent regulated alternative splicing events across human solid tumors', *Nucleic Acids Research*. doi: 10.1093/nar/gkv210.

Das, S. *et al.* (2012) 'Oncogenic Splicing Factor SRSF1 Is a Critical Transcriptional Target of MYC', *Cell Reports*. doi: 10.1016/j.celrep.2011.12.001.

Das, S. and Krainer, A. R. (2014) 'Emerging Functions of SRSF1, Splicing Factor and Oncoprotein, in RNA Metabolism and Cancer', *Molecular Cancer Research*. doi: 10.1158/1541-7786.MCR-14-0131.

David, C. J. *et al.* (2010) 'HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer', *Nature*. doi: 10.1038/nature08697.

David, C. J. and Manley, J. L. (2010) 'Alternative pre-mRNA splicing regulation in cancer: Pathways and programs unhinged', *Genes and Development*, 24(21), pp. 2343–2364. doi: 10.1101/gad.1973010.

Decorsière, A. *et al.* (2011) 'Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage', *Genes and Development*. doi: 10.1101/gad.607011.

Delgado, J. L. *et al.* (2018) 'Topoisomerases as anticancer targets', *Biochemical Journal*. doi: 10.1042/BCJ20160583.

Derti, A. *et al.* (2012) 'A quantitative atlas of polyadenylation in five mammals', *Genome Research*.

doi: 10.1101/gr.132563.111.

Devany, E. *et al.* (2016) 'Intronic cleavage and polyadenylation regulates gene expression during DNA damage response through U1 snRNA', *Cell Discovery*. doi: 10.1038/celldisc.2016.13.

DeVita, V. T. and Chu, E. (2008) 'A history of cancer chemotherapy', *Cancer Research*. doi: 10.1158/0008-5472.CAN-07-6611.

Dilruba, S. and Kalayda, G. V. (2016) 'Platinum-based drugs: past, present and future', *Cancer Chemotherapy and Pharmacology*. doi: 10.1007/s00280-016-2976-z.

Dolatshad, H. *et al.* (2015) 'Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells', *Leukemia*. doi: 10.1038/leu.2014.331.

Dujardin, G. *et al.* (2014) 'How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping', *Molecular Cell*, 54(4), pp. 683–690. doi: 10.1016/j.molcel.2014.03.044.

Dutertre, M. *et al.* (2010) 'Cotranscriptional exon skipping in the genotoxic stress response', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.1912.

Dutertre, M., Chakrama, F. Z., *et al.* (2014) 'A recently evolved class of alternative 3'-terminal exons involved in cell cycle regulation by topoisomerase inhibitors', *Nature Communications*, 5, pp. 1–12. doi: 10.1038/ncomms4395.

Dutertre, M., Lambert, S., *et al.* (2014) 'DNA damage: RNA-binding proteins protect from near and far', *Trends in Biochemical Sciences*. doi: 10.1016/j.tibs.2014.01.003.

Dutertre, M. and Vagner, S. (2017) 'DNA-Damage Response RNA-Binding Proteins (DDRBP): Perspectives from a New Class of Proteins and Their RNA Targets', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2016.09.019.

Dvinge, H. *et al.* (2016) 'RNA splicing factors as oncoproteins and tumour suppressors', *Nature Reviews Cancer*. Nature Publishing Group, 16(7), pp. 413–430. doi: 10.1038/nrc.2016.51.

Dvinge, H. (2018) 'Regulation of alternative mRNA splicing: old players and new perspectives', *FEBS Letters*, 592, pp. 2987–3006. doi: 10.1002/1873-3468.13119.

Dvinge, H. and Bradley, R. K. (2015) 'Widespread intron retention diversifies most cancer transcriptomes', *Genome Medicine*. doi: 10.1186/s13073-015-0168-9.

Edmond, V. *et al.* (2011) 'Acetylation and phosphorylation of SRSF2 control cell fate decision in response to cisplatin', *EMBO Journal*. doi: 10.1038/emboj.2010.333.

Elkon, R. *et al.* (2012) 'E2F mediates enhanced alternative polyadenylation in proliferation', *Genome Biology*, 13(7). doi: 10.1186/gb-2012-13-7-r59.

Ellis, M. J. *et al.* (2012) 'Whole-genome analysis informs breast cancer response to aromatase inhibition', *Nature*. doi: 10.1038/nature11143.

Enoiu, M., Jiricny, J. and Schärer, O. D. (2012) 'Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis', *Nucleic Acids Research*, 40(18), pp. 8953–8964. doi: 10.1093/nar/gks670.

Eom, Y. W. *et al.* (2005) 'Two distinct modes of cell death induced by doxorubicin: Apoptosis and cell death through mitotic catastrophe accompanied by senescence-like phenotype', *Oncogene*. doi: 10.1038/sj.onc.1208627.

Ezkurdia, I. *et al.* (2014) 'Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes', *Human Molecular Genetics*. doi: 10.1093/hmg/ddu309.

Farber, S. *et al.* (1948) 'Temporary Remissions in Acute Leukemia in Children Produced by Folic Acid

Antagonist, 4-Aminopteroyl-Glutamic Acid (Aminopterin)', *New England Journal of Medicine*. doi: 10.1056/NEJM194806032382301.

Fischer, K. R. *et al.* (2015) 'Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance', *Nature*. doi: 10.1038/nature15748.

Floor, S. N. and Doudna, J. A. (2016) 'Tunable protein synthesis by transcript isoforms in human cells', *eLife*. doi: 10.7554/eLife.10921.

Florea, A. M. and Büsselberg, D. (2011) 'Cisplatin as an anti-tumor drug: Cellular mechanisms of activity, drug resistance and induced side effects', *Cancers*. doi: 10.3390/cancers3011351.

Fong, N. *et al.* (2014) 'Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate', *Genes and Development*. doi: 10.1101/gad.252106.114.

Fonseca, D. *et al.* (2017) 'The mRNA processing factor CstF-50 and ubiquitin escort factor p97 are BRCA1/BARD1 cofactors involved in chromatin remodeling during DNA damage response.', *Molecular and Cellular Biology*. doi: 10.1128/MCB.00364-17.

Frendewey, D. and Keller, W. (1985) 'Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences', *Cell*, 42(1), pp. 355–367. doi: 10.1016/S0092-8674(85)80131-8.

Friboulet, L. *et al.* (2013) 'ERCC1 Isoform Expression and DNA Repair in Non-Small-Cell Lung Cancer', *New England Journal of Medicine*. doi: 10.1056/NEJMoa1214271.

Fu, X. D. and Ares, M. (2014) 'Context-dependent control of alternative splicing by RNA-binding proteins', *Nature Reviews Genetics*. Nature Publishing Group, 15(10), pp. 689–701. doi: 10.1038/nrg3778.

Fu, Y. *et al.* (2011) 'Differential genome-wide profiling of tandem 3'UTRs among human breast cancer and normal cells by high-throughput sequencing', *Genome Research*. doi: 10.1101/gr.115295.110.21.

Gabriel, M. *et al.* (2015) 'Role of the splicing factor SRSF4 in cisplatin-induced modifications of pre-mRNA splicing and apoptosis', *BMC Cancer*. doi: 10.1186/s12885-015-1259-0.

Galej, W. P. (2018) 'Structural studies of the spliceosome : past , present and future perspectives', *Biochemical Society Transactions*, (August), pp. 1–16.

Galluzzi, L. *et al.* (2012) 'Molecular mechanisms of cisplatin resistance', *Oncogene*. doi: 10.1038/onc.2011.384.

Gao, R. *et al.* (2014) 'Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2-DNA and Top2-RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2)', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M114.565374.

Garneau, D. *et al.* (2005) 'Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the apoptotic mediator Bcl-x', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M501070200.

van Gelder, C. W. *et al.* (1993) 'A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A protein is required for regulation of polyadenylation.', *The EMBO journal*.

Genovese, I. *et al.* (2017) 'Not only P-glycoprotein: Amplification of the ABCB1-containing chromosome region 7q21 confers multidrug resistance upon cancer cells by coordinated overexpression of an assortment of resistance-related proteins', *Drug Resistance Updates*. doi: 10.1016/j.drug.2017.10.003.

Germann, S. *et al.* (2012) 'Splicing programs and cancer', *Journal of Nucleic Acids*. doi: 10.1155/2012/269570.

Giglia-Mari, G., Zotter, A. and Vermeulen, W. (2011) 'DNA damage response', *Cold Spring Harbor*

Perspectives in Biology. doi: 10.1101/cshperspect.a000745.

Gilman, A., Philips, F. S. and Hedgpeth, J. W. (1946) 'The biological actions and therapeutic applications of the B-chloroethyl amines and sulfides', *Science*. doi: 10.1126/science.103.2675.409.

Giono, L. E. *et al.* (2016) 'The RNA response to DNA damage.', *Journal of molecular biology*. Elsevier Ltd, 428(12), pp. 2636–51. doi: 10.1016/j.jmb.2016.03.004.

Goldstein, M. and Kastan, M. B. (2015) 'The DNA Damage Response: Implications for Tumor Responses to Radiation and Chemotherapy', *Annual Review of Medicine*. doi: 10.1146/annurev-med-081313-121208.

Gómez-Herrerros, F. *et al.* (2013) 'TDP2-Dependent Non-Homologous End-Joining Protects against Topoisomerase II-Induced DNA Breaks and Genome Instability in Cells and In Vivo', *PLoS Genetics*. doi: 10.1371/journal.pgen.1003226.

Gonzalez, V. M. *et al.* (2001) 'Is Cisplatin-Induced Cell Death Always Produced by Apoptosis?', *Molecular pharmacology*. doi: 10.1124/mol.59.4.657.

Graber, J. H. *et al.* (2013) 'DNA damage induces targeted, genome-wide variation of poly(A) sites in budding yeast', *Genome Research*. doi: 10.1101/gr.144964.112.

Grabowski, P. J., Seiler, S. R. and Sharp, P. A. (1985) 'A multicomponent complex is involved in the splicing of messenger RNA precursors', *Cell*, 42(1), pp. 345–353. doi: 10.1016/S0092-8674(85)80130-6.

Grifone, R. *et al.* (2014) 'The RNA-binding protein Rbm24 is transiently expressed in myoblasts and is required for myogenic differentiation during vertebrate development', *Mechanisms of Development*. doi: 10.1016/j.mod.2014.08.003.

Gruber, A. J. *et al.* (2016) 'A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation', *Genome Research*, 26(8), pp. 1145–1159. doi: 10.1101/gr.202432.115.

Gruber, A. R., Martin, G., Müller, P., *et al.* (2014) 'Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells', *Nature Communications*, 5. doi: 10.1038/ncomms6465.

Gruber, A. R., Martin, G., Keller, W., *et al.* (2014) 'Means to an end: Mechanisms of alternative polyadenylation of messenger RNA precursors', *Wiley Interdisciplinary Reviews: RNA*, 5(2), pp. 183–196. doi: 10.1002/wrna.1206.

Gunderson, S. I. *et al.* (1997) 'Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation', *Genes Dev*. doi: 10.1101/gad.11.6.761.

Günthert, U. *et al.* (1991) 'A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells', *Cell*. doi: 10.1016/0092-8674(91)90403-L.

Hahn, C. N. *et al.* (2015) 'Splice factor mutations and alternative splicing as drivers of hematopoietic malignancy', *Immunological Reviews*. doi: 10.1111/imr.12241.

Han, H. *et al.* (2017) 'Multilayered Control of Alternative Splicing Regulatory Networks by Transcription Factors', *Molecular Cell*. doi: 10.1016/j.molcel.2017.01.011.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: The next generation', *Cell*. doi: 10.1016/j.cell.2011.02.013.

Hande, K. R. (1998) 'Clinical applications of anticancer drugs targeted to topoisomerase II', *Biochimica et Biophysica Acta - Gene Structure and Expression*. doi: 10.1016/S0167-4781(98)00134-1.

Harbour, J. W. *et al.* (2013) 'Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma', *Nature Genetics*. doi: 10.1038/ng.2523.

- Harrington, E. D. *et al.* (2004) 'Estimating rates of alternative splicing in mammals and invertebrates', *Nature Genetics*. doi: 10.1038/ng0904-916.
- Havens, M. A. *et al.* (2012) 'Biogenesis of mammalian microRNAs by a non-canonical processing pathway', *Nucleic Acids Research*. doi: 10.1093/nar/gks026.
- Haynes, B. *et al.* (2015) 'Crosstalk between translesion synthesis, Fanconi anemia network, and homologous recombination repair pathways in interstrand DNA crosslink repair and development of chemoresistance', *Mutation Research - Reviews in Mutation Research*. doi: 10.1016/j.mrrev.2014.11.005.
- Henninger, C. and Fritz, G. (2017) 'Statins in anthracycline-induced cardiotoxicity: Rac and Rho, and the heartbreakers', *Cell Death and Disease*. doi: 10.1038/cddis.2016.418.
- Herzel, L. *et al.* (2017) 'Splicing and transcription touch base: Co-transcriptional spliceosome assembly and function', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(10), pp. 637–650. doi: 10.1038/nrm.2017.63.
- Himeji, D. *et al.* (2002) 'Characterization of caspase-8L: A novel isoform of caspase-8 that behaves as an inhibitor of the caspase cascade', *Blood*. doi: 10.1182/blood.V99.11.4070.
- Le Hir, H., Nott, A. and Moore, M. J. (2003) 'How introns influence and enhance eukaryotic gene expression', *Trends in Biochemical Sciences*. doi: 10.1016/S0968-0004(03)00052-5.
- Hoeijmakers, J. H. J. (2001) 'Genome maintenance mechanisms for preventing cancer', *Nature*. doi: 10.1038/35077232.
- Hoffman, Y. *et al.* (2016) '3'UTR Shortening Potentiates MicroRNA-Based Repression of Pro-differentiation Genes in Proliferating Human Cells', *PLoS Genetics*, 12(2), pp. 1–15. doi: 10.1371/journal.pgen.1005879.
- Hollerer, I. *et al.* (2014) 'mRNA 3'end processing: A tale of the tail reaches the clinic', *EMBO Molecular Medicine*, 6(1), pp. 16–26. doi: 10.1002/emmm.201303300.
- Holmila, R. *et al.* (2003) 'Splice mutations in the p53 gene: case report and review of the literature.', *Human mutation*. doi: 10.1002/humu.9104.
- Holohan, C. *et al.* (2013) 'Cancer drug resistance: An evolving paradigm', *Nature Reviews Cancer*. Nature Publishing Group, 13(10), pp. 714–726. doi: 10.1038/nrc3599.
- Hoque, M. *et al.* (2013) 'Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing', *Nature Methods*. doi: 10.1038/nmeth.2288.
- Hsin, J. P. and Manley, J. L. (2012) 'The RNA polymerase II CTD coordinates transcription and RNA processing', *Genes and Development*. doi: 10.1101/gad.200303.112.
- Hu, Z. *et al.* (2015) 'Revealing Missing Human Protein Isoforms Based on Ab Initio Prediction, RNA-seq and Proteomics', *Scientific Reports*. doi: 10.1038/srep10940.
- Hubé, F. and Francastel, C. (2015) 'Mammalian introns: When the junk generates molecular diversity', *International Journal of Molecular Sciences*. doi: 10.3390/ijms16034429.
- Huelga, S. C. *et al.* (2012) 'Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins', *Cell Reports*. doi: 10.1016/j.celrep.2012.02.001.
- Hull, J. *et al.* (2007) 'Identification of common genetic variation that modulates alternative splicing', *PLoS Genet*. doi: 10.1371/journal.pgen.0030099.
- Ip, J. Y. *et al.* (2011) 'Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation', *Genome Research*. doi: 10.1101/gr.111070.110.
- Ishida, S. *et al.* (2002) 'Uptake of the anticancer drug cisplatin mediated by the copper transporter

- Ctr1 in yeast and mammals', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.162491399.
- Jackson, S. P. and Bartek, J. (2010) 'The DNA-damage response in human biology and disease', *Nature*. doi: 10.1038/nature08467.The.
- Jean-Philippe, J., Paz, S. and Caputi, M. (2013) 'hnRNP A1: The Swiss Army Knife of gene expression', *International Journal of Molecular Sciences*. doi: 10.3390/ijms140918999.
- Jenal, M. *et al.* (2012) 'The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites', *Cell*. doi: 10.1016/j.cell.2012.03.022.
- Ji, Z. and Tian, B. (2009) 'Reprogramming of 3' Untranslated Regions of mRNAs by Alternative Polyadenylation in Generation of Pluripotent Stem Cells from Different Cell Types', *PLoS ONE*, 4(12), p. e8419. doi: 10.1371/journal.pone.0008419.
- Jiang, L. *et al.* (2016) 'Genomic Landscape Survey Identifies SRSF1 as a Key Oncodriver in Small Cell Lung Cancer', *PLoS Genetics*. doi: 10.1371/journal.pgen.1005895.
- Jiricny, J. (2006) 'The multifaceted mismatch-repair system', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm1907.
- Johnstone, T. C., Suntharalingam, K. and Lippard, S. J. (2016) 'The Next Generation of Platinum Drugs: Targeted Pt(II) Agents, Nanoparticle Delivery, and Pt(IV) Prodrugs', *Chemical Reviews*. doi: 10.1021/acs.chemrev.5b00597.
- Jones, R. (2016) 'Cytotoxic chemotherapy: Clinical aspects', *Medicine (United Kingdom)*. doi: 10.1016/j.mpmed.2015.10.007.
- Jordan, P. *et al.* (1999) 'Cloning of a novel human Rac1b splice variant with increased expression in colorectal tumors', *Oncogene*. doi: 10.1038/sj.onc.1203233.
- Jung, H. *et al.* (2015) 'Intron retention is a widespread mechanism of tumor-suppressor inactivation', *Nature Genetics*. doi: 10.1038/ng.3414.
- Jung, Y. and Lippard, S. J. (2007) 'Direct cellular responses to platinum-induced DNA damage', *Chemical Reviews*. doi: 10.1021/cr068207j.
- Kachalaki, S. *et al.* (2016) 'Cancer chemoresistance; Biochemical and molecular aspects: A brief overview', *European Journal of Pharmaceutical Sciences*. doi: 10.1016/j.ejps.2016.03.025.
- Kaida, D. *et al.* (2010) 'U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation', *Nature*. doi: 10.1038/nature09479.
- Kaida, D. (2016) 'The reciprocal regulation between splicing and 3'-end processing', *Wiley Interdisciplinary Reviews: RNA*, 7(4), pp. 499–511. doi: 10.1002/wrna.1348.
- Kanagasabai, R. *et al.* (2016) 'Alternative RNA Processing of Topoisomerase II in Etoposide-Resistant Human Leukemia K562 Cells: Intron Retention Results in a Novel C-Terminal Truncated 90-kDa Isoform', *Journal of Pharmacology and Experimental Therapeutics*. doi: 10.1124/jpet.116.237107.
- Karasawa, T. and Steyger, P. S. (2015) 'An integrated view of cisplatin-induced nephrotoxicity and ototoxicity', *Toxicology Letters*. doi: 10.1016/j.toxlet.2015.06.012.
- Karni, R. *et al.* (2007) 'The gene encoding the splicing factor SF2/ASF is a proto-oncogene', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb1209.
- Kato, N. *et al.* (2017) 'Sensing and Processing of DNA Interstrand Crosslinks by the Mismatch Repair Pathway', *Cell Reports*. doi: 10.1016/j.celrep.2017.10.032.
- Kelland, L. (2007) 'The resurgence of platinum-based cancer chemotherapy', *Nature Reviews Cancer*. doi: 10.1038/nrc2167.

- Keren, H., Lev-Maor, G. and Ast, G. (2010) 'Alternative splicing and evolution: Diversification, exon definition and function', *Nature Reviews Genetics*. Nature Publishing Group, 11(5), pp. 345–355. doi: 10.1038/nrg2776.
- Kim, E. *et al.* (2015) 'SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition', *Cancer Cell*. doi: 10.1016/j.ccell.2015.04.006.
- Kim, E., Magen, A. and Ast, G. (2007) 'Different levels of alternative splicing among eukaryotes', *Nucleic Acids Research*, 35(1), pp. 125–131. doi: 10.1093/nar/gkl924.
- Kleiman, F. E. and Manley, J. L. (2001) 'The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression', *Cell*. doi: 10.1016/S0092-8674(01)00270-7.
- De Klerk, E. *et al.* (2012) 'Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation', *Nucleic Acids Research*. doi: 10.1093/nar/gks655.
- Knowles, D. A. *et al.* (2018) 'Determining the genetic basis of anthracycline-cardiotoxicity by molecular response QTL mapping in induced cardiomyocytes', *eLife*. doi: 10.7554/eLife.33480.001.
- Knox, R. J. *et al.* (1986) 'Mechanism of cytotoxicity of anticancer platinum drugs: Evidence that cis-diamminedichloroplatinum(ii) and cis-diammine-(1,1-cyclobutanedicarboxylato)platinum(ii) differ only in the kinetics of their interaction with dna', *Cancer Research*.
- Ko, B. and Gunderson, S. I. (2002) 'Identification of new poly(A) polymerase-inhibitory proteins capable of regulating pre-mRNA polyadenylation', *Journal of Molecular Biology*. doi: 10.1016/S0022-2836(02)00240-1.
- Koh, C. M. *et al.* (2015) 'MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis', *Nature*. doi: 10.1038/nature14351.
- Kornblihtt, A. R. *et al.* (2013) 'Alternative splicing: A pivotal step between eukaryotic transcription and translation', *Nature Reviews Molecular Cell Biology*, 14(3), pp. 153–165. doi: 10.1038/nrm3525.
- Kroemer, G. *et al.* (2013) 'Immunogenic cell death in cancer therapy.', *Annual review of immunology*. doi: 10.1146/annurev-immunol-032712-100008.
- Krysko, D. V. *et al.* (2012) 'Immunogenic cell death and DAMPs in cancer therapy', *Nature Reviews Cancer*. doi: 10.1038/nrc3380.
- Kuehner, J. N., Kaufman, J. W. and Moore, C. (2017) '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*. doi: 10.1016/j.dnarep.2017.07.006.
- Kuehner, J. N., Pearson, E. L. and Moore, C. (2011) 'Unravelling the means to an end: RNA polymerase II transcription termination', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm3098.
- Kumar, A. (2009) 'An overview of nested genes in eukaryotic genomes', *Eukaryotic Cell*. doi: 10.1128/EC.00143-09.
- Kyburz, A. *et al.* (2006) 'Direct Interactions between Subunits of CPSF and the U2 snRNP Contribute to the Coupling of Pre-mRNA 3' End Processing and Splicing', *Molecular Cell*. doi: 10.1016/j.molcel.2006.05.037.
- Lackford, B. *et al.* (2014) 'Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal', *EMBO Journal*. doi: 10.1002/embj.201386537.
- Ladomery, M. (2013) 'Aberrant alternative splicing is another hallmark of cancer', *International Journal of Cell Biology*. doi: 10.1155/2013/463786.
- Ledesma, F. C. *et al.* (2009) 'A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage', *Nature*. doi: 10.1038/nature08444.

- Lee, S. H. *et al.* (2018) 'Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia', *Nature*. doi: 10.1038/s41586-018-0465-8.
- Lents, N. H. *et al.* (2008) 'Identification and characterization of a novel Mdm2 splice variant acutely induced by the chemotherapeutic agents adriamycin and actinomycin D', *Cell Cycle*. doi: 10.4161/cc.7.11.5985.
- Leroy, B., Anderson, M. and Soussi, T. (2014) 'TP53 mutations in human cancer: Database reassessment and prospects for the next decade', *Human Mutation*. doi: 10.1002/humu.22552.
- Lev Maor, G., Yearim, A. and Ast, G. (2015) 'The alternative role of DNA methylation in splicing regulation', *Trends in Genetics*. doi: 10.1016/j.tig.2015.03.002.
- Li, Q., Lee, J. A. and Black, D. L. (2007) 'Neuronal regulation of alternative pre-mRNA splicing', *Nature Reviews Neuroscience*. doi: 10.1038/nrn2237.
- Li, W. *et al.* (2012) 'Star-PAP control of BIK expression and apoptosis is regulated by nuclear PIPK1 α and PKC δ signaling', *Molecular Cell*. doi: 10.1016/j.molcel.2011.11.017.
- Li, W. *et al.* (2015) 'Systematic Profiling of Poly(A)⁺ Transcripts Modulated by Core 3' End Processing and Splicing Factors Reveals Regulatory Rules of Alternative Cleavage and Polyadenylation', *PLoS Genetics*, 11(4), pp. 1–28. doi: 10.1371/journal.pgen.1005166.
- Li, W. *et al.* (2016) 'Overcoming ABC transporter-mediated multidrug resistance: Molecular mechanisms and novel therapeutic drug strategies', *Drug Resistance Updates*. doi: 10.1016/j.drug.2016.05.001.
- Li, W. *et al.* (2017) 'Distinct regulation of alternative polyadenylation and gene expression by nuclear poly(A) polymerases', *Nucleic Acids Research*. doi: 10.1093/nar/gkx560.
- Li, Y. *et al.* (2017) 'Revealing the Determinants of Widespread Alternative Splicing Perturbation in Cancer', *Cell Reports*. Elsevier Company., 21(3), pp. 798–812. doi: 10.1016/j.celrep.2017.09.071.
- Liang, Z. D. *et al.* (2012) 'Mechanistic Basis for Overcoming Platinum Resistance Using Copper Chelating Agents', *Molecular Cancer Therapeutics*. doi: 10.1158/1535-7163.MCT-12-0580.
- Licatalosi, D. D. *et al.* (2008) 'HITS-CLIP yields genome-wide insights into brain alternative RNA processing', *Nature*. doi: 10.1038/nature07488.
- Lin, Y. *et al.* (2018) 'Global profiling of Rbm24 bound RNAs uncovers a multi-tasking RNA binding protein', *International Journal of Biochemistry and Cell Biology*. doi: 10.1016/j.biocel.2017.11.002.
- Liu, Y. *et al.* (2017) 'Impact of Alternative Splicing on the Human Proteome', *Cell Reports*, 20(5), pp. 1229–1241. doi: 10.1016/j.celrep.2017.07.025.
- Loi, S. *et al.* (2013) 'CD73 promotes anthracycline resistance and poor prognosis in triple negative breast cancer', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1222251110.
- Long, D. T. *et al.* (2014) 'BRCA1 promotes unloading of the CMG Helicase from a stalled DNA replication fork', *Molecular Cell*. doi: 10.1016/j.molcel.2014.08.012.
- López-Saavedra, A. *et al.* (2016) 'MAD2 γ , a novel MAD2 isoform, reduces mitotic arrest and is associated with resistance in testicular germ cell tumors', *Cell Cycle*. doi: 10.1080/15384101.2016.1198863.
- Lord, C. J. and Ashworth, A. (2012) 'The DNA damage response and cancer therapy', *Nature*. doi: 10.1038/nature10760.
- Luco, R. F. *et al.* (2011) 'Epigenetics in alternative pre-mRNA splicing', *Cell*. Elsevier Inc., 144(1), pp. 16–26. doi: 10.1016/j.cell.2010.11.056.
- Luo, W., Johnson, A. W. and Bentley, D. L. (2006) 'The role of Rat1 in coupling mRNA 3'-end

- processing to transcription termination: Implications for a unified allosteric-torpedo model', *Genes and Development*. doi: 10.1101/gad.1409106.
- Mallinjoed, P. *et al.* (2014) 'Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin', *Genome Research*. doi: 10.1101/gr.162933.113.
- Maniatis, T. and Reed, R. (2002) 'An extensive network of coupling among gene expression machines', *Nature*. doi: 10.1038/416499a.
- Mansfield, K. D. and Keene, J. D. (2012) 'Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation', *Nucleic Acids Research*. doi: 10.1093/nar/gkr1114.
- Maréchal, A. *et al.* (2014) 'PRP19 Transforms into a Sensor of RPA-ssDNA after DNA Damage and Drives ATR Activation via a Ubiquitin-Mediated Circuitry', *Molecular Cell*. doi: 10.1016/j.molcel.2013.11.002.
- Markey, M. and Berberich, S. J. (2008) 'Full-length hdmX transcripts decrease following genotoxic stress', *Oncogene*. doi: 10.1038/onc.2008.266.
- Martin, G. *et al.* (2012) '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*. doi: 10.1016/j.celrep.2012.05.003.
- Martin, M. *et al.* (2013) 'Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3', *Nature Genetics*. doi: 10.1038/ng.2674.
- Martino, E. *et al.* (2017) 'The long story of camptothecin: From traditional medicine to drugs', *Bioorganic and Medicinal Chemistry Letters*. doi: 10.1016/j.bmcl.2016.12.085.
- Marullo, R. *et al.* (2013) 'Cisplatin induces a mitochondrial-ros response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions', *PLoS ONE*. doi: 10.1371/journal.pone.0081162.
- Masamha, C. P. *et al.* (2014) 'CFIm25 links alternative polyadenylation to glioblastoma tumour suppression', *Nature*. doi: 10.1038/nature13261.
- Mastrocola, A. S. *et al.* (2013) 'The RNA-binding protein fused in sarcoma (FUS) functions downstream of poly(ADP-ribose) polymerase (PARP) in response to DNA damage', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M113.497974.
- Matsuoka, S. *et al.* (2007) 'ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage', *Science*. doi: 10.1126/science.1140321.
- Mayr, C. (2016) 'Evolution and Biological Roles of Alternative 3'UTRs', *Trends in Cell Biology*. doi: 10.1016/j.tcb.2015.10.012.
- Mayr, C. and Bartel, D. P. (2009) 'Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells', *Cell*. Elsevier Ltd, 138(4), pp. 673–684. doi: 10.1016/j.cell.2009.06.016.
- McManus, C. J. and Graveley, B. R. (2011) 'RNA structure and the mechanisms of alternative splicing', *Current Opinion in Genetics and Development*. doi: 10.1016/j.gde.2011.04.001.
- Melnikov, S. V. *et al.* (2016) 'Insights into RNA binding by the anticancer drug cisplatin from the crystal structure of cisplatin-modified ribosome', *Nucleic Acids Research*. doi: 10.1093/nar/gkw246.
- Merdzhanova, G. *et al.* (2008) 'E2F1 controls alternative splicing pattern of genes involved in apoptosis through upregulation of the splicing factor SC35', *Cell Death and Differentiation*. doi: 10.1038/cdd.2008.135.
- Millevoi, S. *et al.* (2002) 'A novel function for the U2AF 65 splicing factor in promoting pre-mRNA 3'-

- end processing', *EMBO Reports*. doi: 10.1093/embo-reports/kvf173.
- Millevoi, S. *et al.* (2006) 'An interaction between U2AF 65 and CF Ilinks the splicing and 3' end processing machineries', *EMBO Journal*. doi: 10.1038/sj.emboj.7601331.
- Millevoi, S. and Vagner, S. (2009) 'Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation', *Nucleic Acids Research*, 38(9), pp. 2757–2774. doi: 10.1093/nar/gkp1176.
- Minotti, G. (2004) 'Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity', *Pharmacological Reviews*. doi: 10.1124/pr.56.2.6.
- Mitra, M. *et al.* (2018) 'Alternative polyadenylation factors link cell cycle to migration', *Genome Biology*. *Genome Biology*, 19(1), p. 176. doi: 10.1186/s13059-018-1551-9.
- Mogilevsky, M. *et al.* (2018) 'Modulation of MKNK2 alternative splicing by splice-switching oligonucleotides as a novel approach for glioblastoma treatment', *Nucleic Acids Research*. doi: 10.1093/nar/gky921.
- Mohamed, A. M. *et al.* (2016) 'Oncogene- and drug resistance-associated alternative exon usage in acute myeloid leukemia (AML)', *Oncotarget*. doi: 10.18632/oncotarget.3898.
- Mollet, I. G. *et al.* (2010) 'Unconstrained mining of transcript data reveals increased alternative splicing complexity in the human transcriptome', *Nucleic Acids Research*. doi: 10.1093/nar/gkq197.
- Morris, A. R. *et al.* (2012) 'Alternative cleavage and polyadenylation during colorectal cancer development', *Clinical Cancer Research*, 18(19), pp. 5256–5266. doi: 10.1158/1078-0432.CCR-12-0543.
- Mueller, A. A. *et al.* (2016) 'Intronic polyadenylation of PDGFR α in resident stem cells attenuates muscle fibrosis', *Nature*. doi: 10.1038/nature20160.
- Muñoz, M. J. *et al.* (2009) 'DNA Damage Regulates Alternative Splicing through Inhibition of RNA Polymerase II Elongation', *Cell*, 137(4), pp. 708–720. doi: 10.1016/j.cell.2009.03.010.
- Naro, C. *et al.* (2015) 'The interplay between DNA damage response and RNA processing: The unexpected role of splicing factors as gatekeepers of genome stability', *Frontiers in Genetics*, 6(MAR), pp. 1–10. doi: 10.3389/fgene.2015.00142.
- Nazeer, F. I. *et al.* (2011) 'P53 inhibits mRNA 3' processing through its interaction with the CstF/BARD1 complex', *Oncogene*. doi: 10.1038/onc.2011.29.
- Newman, M. *et al.* (2017) 'The G-Quadruplex-Specific RNA Helicase DHX36 Regulates p53 Pre-mRNA 3'-End Processing Following UV-Induced DNA Damage', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2016.11.033.
- Ni, T. K. and Kuperwasser, C. (2016) 'Premature polyadenylation of MAGI3 produces a dominantly-acting oncogene in human breast cancer', *eLife*, 5(MAY2016), pp. 1–21. doi: 10.7554/eLife.14730.
- Nilsen, T. W. and Graveley, B. R. (2010) 'Expansion of the eukaryotic proteome by alternative splicing', *Nature*, 463(7280), pp. 457–463. doi: 10.1038/nature08909.
- Nitiss, J. L. (2009) 'Targeting DNA topoisomerase II in cancer chemotherapy', *Nature Reviews Cancer*, 9(5), pp. 338–350. doi: 10.1038/nrc2607.
- Niwa, M. and Berget, S. M. (1991) 'Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns', *Genes and Development*. doi: 10.1101/gad.5.11.2086.
- Niwa, M., Rose, S. D. and Berget, S. M. (1990) 'In vitro polyadenylation is stimulated by the presence of an upstream intron', *Genes and Development*. doi: 10.1101/gad.4.9.1552.
- Nojima, T. *et al.* (2015) 'Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing', *Cell*, 161(3), pp. 526–540. doi: 10.1016/j.cell.2015.03.027.

- O'Malley, F. P. *et al.* (2009) 'Topoisomerase II alpha and responsiveness of breast cancer to adjuvant chemotherapy', *Journal of the National Cancer Institute*. doi: 10.1093/jnci/djp067.
- Oltean, S. and Bates, D. O. (2014) 'Hallmarks of alternative splicing in cancer', *Oncogene*. Nature Publishing Group, 33(46), pp. 5311–5318. doi: 10.1038/onc.2013.533.
- Pagliarini, V., Naro, C. and Sette, C. (2015) 'Splicing regulation: A molecular device to enhance cancer cell adaptation', *BioMed Research International*. doi: 10.1155/2015/543067.
- Pan, Q. *et al.* (2008) 'Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing', *Nature Genetics*, 40(12), pp. 1413–1415. doi: 10.1038/ng.259.
- Pandit, S. *et al.* (2013) 'Genome-wide Analysis Reveals SR Protein Cooperation and Competition in Regulated Splicing', *Molecular Cell*. Elsevier Inc., 50(2), pp. 223–235. doi: 10.1016/j.molcel.2013.03.001.
- Pang, B. *et al.* (2013) 'Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin', *Nature Communications*. doi: 10.1038/ncomms2921.
- Papasaikas, P. *et al.* (2015) 'Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery', *Molecular Cell*. doi: 10.1016/j.molcel.2014.10.030.
- Park, E. *et al.* (2018) 'The Expanding Landscape of Alternative Splicing Variation in Human Populations', *American Journal of Human Genetics*. The Authors, 102(1), pp. 11–26. doi: 10.1016/j.ajhg.2017.11.002.
- Paronetto, M. P. *et al.* (2007) 'The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x', *Journal of Cell Biology*. doi: 10.1083/jcb.200701005.
- Paronetto, M. P., Miñana, B. and Valcárcel, J. (2011) 'The Ewing Sarcoma Protein Regulates DNA Damage-Induced Alternative Splicing', *Molecular Cell*. doi: 10.1016/j.molcel.2011.05.035.
- Passacantilli, I. *et al.* (2017) 'Alternative polyadenylation of ZEB1 promotes its translation during genotoxic stress in pancreatic cancer cells', *Cell Death and Disease*. Nature Publishing Group, 8(11), pp. e3168-8. doi: 10.1038/cddis.2017.562.
- Paulsen, R. D. *et al.* (2009) 'A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability', *Molecular Cell*. doi: 10.1016/j.molcel.2009.06.021.
- Perry, W. L. 3rd *et al.* (2005) 'Human Splicing Factor SPF45 (RBM17) Confers Broad Multidrug Resistance to Anticancer Drugs When Overexpressed--A Phenotype Partially Reversed By Selective Estrogen Receptor Modulators.', *CANCER RESEARCH*. doi: 10.1158/0008-5472.CAN-03-3675.
- Pinto, P. A. B. *et al.* (2011) 'RNA polymerase II kinetics in polo polyadenylation signal selection', *EMBO Journal*. doi: 10.1038/emboj.2011.156.
- Plowman, S. J. *et al.* (2006) 'The K-Ras 4A isoform promotes apoptosis but does not affect either lifespan or spontaneous tumor incidence in aging mice', *Experimental Cell Research*. doi: 10.1016/j.yexcr.2005.10.004.
- Pommier, Y. (2009) 'DNA topoisomerase I Inhibitors: Chemistry, biology, and interfacial inhibition', *Chemical Reviews*. doi: 10.1021/cr900097c.
- Pommier, Y. *et al.* (2010) 'DNA topoisomerases and their poisoning by anticancer and antibacterial drugs', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2010.04.012.
- Pommier, Y. (2013) 'Drugging topoisomerases: Lessons and Challenges', *ACS Chemical Biology*. doi: 10.1021/cb300648v.
- Ponta, H., Sherman, L. and Herrlich, P. A. (2003) 'CD44: From adhesion molecules to signalling regulators', *Nature Reviews Molecular Cell Biology*, 4(1), pp. 33–45. doi: 10.1038/nrm1004.

- Porrua, O. and Libri, D. (2015) 'Transcription termination and the control of the transcriptome: Why, where and how to stop', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm3943.
- Poulikakos, P. I. *et al.* (2011) 'RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E)', *Nature*. doi: 10.1038/nature10662.
- Prinos, P. *et al.* (2011) 'Alternative splicing of SYK regulates mitosis and cell survival', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.2040.
- Proudfoot, N. J., Furger, a and Dye, M. J. (2002) 'Integrating rRNA processing with transcription', *Cell*, 108, p. 501–512 ST–Integrating rRNA processing with tr. doi: 10.1016/S0092-8674(02)00617-7.
- Raponi, M. *et al.* (2014) 'BRCA1 exon 11 a model of long exon splicing regulation', *RNA Biology*. doi: 10.4161/rna.28458.
- Rauch, J. *et al.* (2011) 'c-Myc regulates RNA splicing of the A-Raf kinase and its activation of the ERK pathway', *Cancer Research*. doi: 10.1158/0008-5472.CAN-10-4447.
- Raymond, E. *et al.* (2002) 'Cellular and molecular pharmacology of oxaliplatin', *Molecular Cancer ...* doi: 10.1023/a:1008213732429.
- Rebucci, M. and Michiels, C. (2013) 'Molecular aspects of cancer cell resistance to chemotherapy', *Biochemical Pharmacology*. doi: 10.1016/j.bcp.2013.02.017.
- Revil, T. *et al.* (2009) 'Heterogeneous nuclear ribonucleoprotein K represses the production of pro-apoptotic Bcl-xS splice isoform', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M109.019711.
- Ridinger, J. *et al.* (2018) 'Dual role of HDAC10 in lysosomal exocytosis and DNA repair promotes neuroblastoma chemoresistance', *Scientific Reports*. doi: 10.1038/s41598-018-28265-5.
- Rogozin, I. B. *et al.* (2012) 'Origin and evolution of spliceosomal introns', *Biology Direct*. doi: 10.1186/1745-6150-7-11.
- Roos, W. P. and Kaina, B. (2013) 'DNA damage-induced cell death: From specific DNA lesions to the DNA damage response and apoptosis', *Cancer Letters*. doi: 10.1016/j.canlet.2012.01.007.
- Rosenberg, B. *et al.* (1969) 'Platinum compounds: a new class of potent antitumour agents', *Nature*. doi: 10.1038/222385a0.
- Rosenberg, B., Van Camp, L. and Krigas, T. (1965) 'Inhibition of cell division in Escherichia coli by electrolysis products from a platinum electrode [17]', *Nature*. doi: 10.1038/205698a0.
- Ruby, J. G., Jan, C. H. and Bartel, D. P. (2007) 'Intronic microRNA precursors that bypass Drosha processing', *Nature*. doi: 10.1038/nature05983.
- Saldi, T. *et al.* (2016) 'Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing', *Journal of Molecular Biology*. Elsevier Ltd, 428(12), pp. 2623–2635. doi: 10.1016/j.jmb.2016.04.017.
- Sampath, J. *et al.* (2003) 'Human SPF45, a Splicing Factor, Has Limited Expression in Normal Tissues, Is Overexpressed in Many Tumors, and Can Confer a Multidrug-Resistant Phenotype to Cells', *American Journal of Pathology*. doi: 10.1016/S0002-9440(10)63538-9.
- Sandberg, R. *et al.* (2008) 'Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites.', *Science (New York, N.Y.)*. doi: 10.1126/science.1155390.
- Savage, K. I. *et al.* (2014) 'Identification of a BRCA1-mRNA Splicing Complex Required for Efficient DNA Repair and Maintenance of Genomic Stability', *Molecular Cell*. doi: 10.1016/j.molcel.2014.03.021.
- Sebestyén, E., Zawisza, M. and Eyras, E. (2015) 'Detection of recurrent alternative splicing switches in tumor samples reveals novel signatures of cancer', *Nucleic Acids Research*. doi: 10.1093/nar/gku1392.

- Seiler, M. *et al.* (2018) 'Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types', *Cell Reports*, 23(1), p. 282–296.e4. doi: 10.1016/j.celrep.2018.01.088.
- Senturk, J. C., Bohlman, S. and Manfredi, J. J. (2017) 'Mdm2 selectively suppresses DNA damage arising from inhibition of topoisomerase II independent of p53', *Oncogene*. doi: 10.1038/onc.2017.229.
- Shalpour, S. *et al.* (2015) 'Immunosuppressive plasma cells impede T-cell-dependent immunogenic chemotherapy', *Nature*. doi: 10.1038/nature14395.
- Shi, Y. (2012) 'Alternative polyadenylation: new insights from global analyses.', *RNA (New York, N.Y.)*. doi: 10.1261/rna.035899.112.
- Shi, Y. (2017) 'Mechanistic insights into precursor messenger RNA splicing by the spliceosome', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(11), pp. 655–670. doi: 10.1038/nrm.2017.86.
- Shkreta, L. *et al.* (2008) 'Anticancer drugs affect the alternative splicing of Bcl-x and other human apoptotic genes', *Molecular Cancer Therapeutics*. doi: 10.1158/1535-7163.MCT-08-0192.
- Shkreta, L. *et al.* (2016) 'SRSF10 Connects DNA Damage to the Alternative Splicing of Transcripts Encoding Apoptosis, Cell-Cycle Control, and DNA Repair Factors', *Cell Reports*. doi: 10.1016/j.celrep.2016.10.071.
- Shkreta, L. and Chabot, B. (2015) 'The RNA splicing response to DNA damage', *Biomolecules*. doi: 10.3390/biom5042935.
- Siddik, Z. H. (2003) 'Cisplatin: Mode of cytotoxic action and molecular basis of resistance', *Oncogene*, 22(47 REV. ISS. 6), pp. 7265–7279. doi: 10.1038/sj.onc.1206933.
- Sigalas, I. *et al.* (1996) 'Alternatively spliced mdm2 transcripts with loss of p53 binding domain sequences: Transforming ability and frequent detection in human cancer', *Nature Medicine*. doi: 10.1038/nm0896-912.
- Singh, I. *et al.* (2018) 'Widespread intronic polyadenylation diversifies immune cell transcriptomes', *Nature Communications*. Springer US, 9(1). doi: 10.1038/s41467-018-04112-z.
- Singh, R. and Valcárcel, J. (2005) 'Building specificity with nonspecific RNA-binding proteins', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb961.
- Sistigu, A. *et al.* (2014) 'Cancer cell-autonomous contribution of type I interferon signaling to the efficacy of chemotherapy', *Nature medicine*. doi: 10.1038/nm.3708.
- Solier, S. *et al.* (2004) 'Topoisomerase I and II inhibitors control caspase-2 pre-messenger RNA splicing in human cells', *Molecular cancer research : MCR*.
- Solier, S. *et al.* (2010) 'Genome-wide analysis of novel splice variants induced by topoisomerase i poisoning shows preferential occurrence in genes encoding splicing factors', *Cancer Research*. doi: 10.1158/0008-5472.CAN-10-2491.
- Song, X. *et al.* (2017) 'Alternative splicing in cancers: From aberrant regulation to new therapeutics', *Seminars in Cell and Developmental Biology*. Elsevier Ltd, 75, pp. 13–22. doi: 10.1016/j.semcd.2017.09.018.
- Spies, N., Burge, C. B. and Bartel, D. P. (2013) '3' UTR-Isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts', *Genome Research*, 23(12), pp. 2078–2090. doi: 10.1101/gr.156919.113.
- Stebbing, J. *et al.* (2018) 'LMTK3 confers chemo-resistance in breast cancer', *Oncogene*. doi: 10.1038/s41388-018-0197-0.

- Sterne-Weiler, T. *et al.* (2013) 'Frac-seq reveals isoform-specific recruitment to polyribosomes', *Genome Research*. doi: 10.1101/gr.148585.112.
- Straussman, R. *et al.* (2012) 'Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion', *Nature*. doi: 10.1038/nature11183.
- Sulli, G., Di Micco, R. and Di Fagagna, F. D. A. (2012) 'Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer', *Nature Reviews Cancer*. doi: 10.1038/nrc3344.
- Sun, Y. *et al.* (2012) 'Treatment-induced damage to the tumor microenvironment promotes prostate cancer therapy resistance through WNT16B', *Nature Medicine*. doi: 10.1038/nm.2890.
- Supek, F. *et al.* (2014) 'Synonymous mutations frequently act as driver mutations in human cancers', *Cell*. doi: 10.1016/j.cell.2014.01.051.
- Sveen, A. *et al.* (2016) 'Aberrant RNA splicing in cancer; Expression changes and driver mutations of splicing factor genes', *Oncogene*. Nature Publishing Group, 35(19), pp. 2413–2427. doi: 10.1038/onc.2015.318.
- Szakács, G. *et al.* (2006) 'Targeting multidrug resistance in cancer', *Nature Reviews Drug Discovery*. doi: 10.1038/nrd1984.
- Tacar, O., Srimornsak, P. and Dass, C. R. (2013) 'Doxorubicin: An update on anticancer molecular action, toxicity and novel drug delivery systems', *Journal of Pharmacy and Pharmacology*. doi: 10.1111/j.2042-7158.2012.01567.x.
- Takagaki, Y. *et al.* (1996) 'The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation', *Cell*. doi: 10.1016/S0092-8674(00)82000-0.
- Tamaro, C. *et al.* (2012) 'BRCA1 exon 11 alternative splicing, multiple functions and the association with cancer', *Biochemical Society Transactions*. doi: 10.1042/BST20120140.
- Tejedor, J. R., Papasaikas, P. and Valcárcel, J. (2015) 'Genome-Wide Identification of Fas/CD95 Alternative Splicing Regulators Reveals Links with Iron Homeostasis', *Molecular Cell*. doi: 10.1016/j.molcel.2014.10.029.
- Tian, B. *et al.* (2005) 'A large-scale analysis of mRNA polyadenylation of human and mouse genes', *Nucleic Acids Research*. doi: 10.1093/nar/gki158.
- Tian, B. and Manley, J. L. (2016) 'Alternative polyadenylation of mRNA precursors', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(1), pp. 18–30. doi: 10.1038/nrm.2016.116.
- Tresini, M. *et al.* (2015) 'The core spliceosome as target and effector of non-canonical ATM signalling', *Nature*. doi: 10.1038/nature14512.
- Tress, M. L., Abascal, F. and Valencia, A. (2017) 'Alternative Splicing May Not Be the Key to Proteome Complexity', *Trends in Biochemical Sciences*. Elsevier Ltd, 42(2), pp. 98–110. doi: 10.1016/j.tibs.2016.08.008.
- Tsai, Y. S. *et al.* (2015) 'Transcriptome-wide identification and study of cancer-specific splicing events across multiple tumors', *Oncotarget*. doi: 10.18632/oncotarget.3145.
- Turner, N., Biganzoli, L. and Di Leo, A. (2015) 'Continued value of adjuvant anthracyclines as treatment for early breast cancer', *The Lancet Oncology*. doi: 10.1016/S1470-2045(15)00079-0.
- Ule, J. *et al.* (2005) 'Nova regulates brain-specific splicing to shape the synapse', *Nature Genetics*. doi: 10.1038/ng1610.
- Vagner, S. *et al.* (2000) 'Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP', *RNA*. doi: 10.1017/S1355838200991854.
- Vagner, S., Vagner, C. and Mattaj, I. W. (2000) 'The carboxyl terminus of vertebrate poly(A)

- polymerase interacts with U2AF 65 to couple 3'-end processing and splicing', *Genes and Development*. doi: 10.1101/gad.14.4.403.
- Vasudevan, S., Peltz, S. W. and Wilusz, C. J. (2002) 'Non-stop decay - A new mRNA surveillance pathway', *BioEssays*. doi: 10.1002/bies.10153.
- Vejpongsa, P. and Yeh, E. T. H. (2014) 'Prevention of anthracycline-induced cardiotoxicity: Challenges and opportunities', *Journal of the American College of Cardiology*. doi: 10.1016/j.jacc.2014.06.1167.
- Venables, J. P. *et al.* (2009) 'Cancer-associated regulation of alternative splicing', *Nature Structural and Molecular Biology*, 16(6), pp. 670–676. doi: 10.1038/nsmb.1608.
- Vethantham, V., Rao, N. and Manley, J. L. (2007) 'Sumoylation Modulates the Assembly and Activity of the Pre-mRNA 3' Processing Complex', *Molecular and Cellular Biology*. doi: 10.1128/MCB.01186-07.
- Vidal, S. J. *et al.* (2014) 'Targeting cancer stem cells to suppress acquired chemotherapy resistance', *Oncogene*. doi: 10.1038/onc.2013.411.
- Villamizar, O. *et al.* (2016) 'Long noncoding RNA Saf and splicing factor 45 increase soluble Fas and resistance to apoptosis', *Oncotarget*. doi: 10.18632/oncotarget.7329.
- Vivarelli, S. *et al.* (2013) 'Paraquat Modulates Alternative Pre-mRNA Splicing by Modifying the Intracellular Distribution of SRPK2', *PLoS ONE*. doi: 10.1371/journal.pone.0061980.
- Vorlová, S. *et al.* (2011) 'Induction of Antagonistic Soluble Decoy Receptor Tyrosine Kinases by Intronic PolyA Activation', *Molecular Cell*, 43(6), pp. 927–939. doi: 10.1016/j.molcel.2011.08.009.
- Vos, S. M. *et al.* (2011) 'All tangled up: How cells direct, manage and exploit topoisomerase function', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm3228.
- Vuong, C. K., Black, D. L. and Zheng, S. (2016) 'The neurogenetics of alternative splicing', *Nature Reviews Neuroscience*. doi: 10.1038/nrn.2016.27.
- Van Waardenburg, R. C. A. M. *et al.* (2004) 'Platinated DNA adducts enhance poisoning of DNA topoisomerase I by camptothecin', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M410103200.
- Wan, Y. *et al.* (2014) 'Landscape and variation of RNA secondary structure across the human transcriptome', *Nature*. doi: 10.1038/nature12946.
- Wang, A. T. *et al.* (2015) 'A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination', *Molecular Cell*. doi: 10.1016/j.molcel.2015.07.009.
- Wang, D. and Lippard, S. J. (2005) 'Cellular processing of platinum anticancer drugs', *Nature Reviews Drug Discovery*. doi: 10.1038/nrd1691.
- Wang, E. T. *et al.* (2008) 'Alternative isoform regulation in human tissue transcriptomes', *Nature*. doi: 10.1038/nature07509.
- Wang, H. *et al.* (2011) 'Identification of an Exon 4-Deletion Variant of Epidermal Growth Factor Receptor with Increased Metastasis-Promoting Capacity', *Neoplasia*. doi: 10.1593/neo.101744.
- Wang, L. *et al.* (1994) 'Ic-1, an Ice/ced-3-related gene, encodes both positive and negative regulators of programmed cell death', *Cell*. doi: 10.1016/S0092-8674(94)90422-7.
- Wang, L. *et al.* (2016) 'Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia', *Cancer Cell*. Elsevier Inc., 30(5), pp. 750–763. doi: 10.1016/j.ccell.2016.10.005.
- Wang, Y. *et al.* (2014) 'The Splicing Factor RBM4 Controls Apoptosis, Proliferation, and Migration to Suppress Tumor Progression', *Cancer Cell*. Elsevier Inc., 26(3), pp. 374–389. doi:

10.1016/j.ccr.2014.07.010.

Wang, Y. *et al.* (2016) 'The BRCA1- Δ 11q alternative splice isoform bypasses germline mutations and promotes therapeutic resistance to PARP inhibition and cisplatin', *Cancer Research*. doi: 10.1158/0008-5472.CAN-16-0186.

Warzecha, C. C. *et al.* (2009) 'ESRP1 and ESRP2 Are Epithelial Cell-Type-Specific Regulators of FGFR2 Splicing', *Molecular Cell*. doi: 10.1016/j.molcel.2009.01.025.

Wassarman, K. M. and Steitz, J. A. (1993) 'Association with terminal exons in pre-mRNAs: A new role for the U1 snRNP?', *Genes and Development*. doi: 10.1101/gad.7.4.647.

Weatheritt, R. J., Sterne-Weiler, T. and Blencowe, B. J. (2016) 'The ribosome-engaged landscape of alternative splicing', *Nature Structural and Molecular Biology*, 23(12). doi: 10.1038/nsmb.3317.

Wheate, N. J. *et al.* (2010) 'The status of platinum anticancer drugs in the clinic and in clinical trials', *Dalton Transactions*. doi: 10.1039/c0dt00292e.

Wijdeven, R. H. *et al.* (2015) 'Genome-wide identification and characterization of novel factors conferring resistance to topoisomerase II poisons in cancer', *Cancer Research*. doi: 10.1158/0008-5472.CAN-15-0380.

Wijdeven, R. H. *et al.* (2016) 'Old drugs, novel ways out: Drug resistance toward cytotoxic chemotherapeutics', *Drug Resistance Updates*. doi: 10.1016/j.drug.2016.07.001.

Williams, H. L., Gottesman, M. E. and Gautier, J. (2012) 'Replication-Independent Repair of DNA Interstrand Crosslinks', *Molecular Cell*. doi: 10.1016/j.molcel.2012.05.001.

Williamson, L. *et al.* (2017) 'UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene', *Cell*. Elsevier, 168(5), p. 843–855.e13. doi: 10.1016/j.cell.2017.01.019.

Wilson, M. D., Harreman, M. and Svejstrup, J. Q. (2013) 'Ubiquitylation and degradation of elongating RNA polymerase II: The last resort', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. doi: 10.1016/j.bbarm.2012.08.002.

Wolf, C. R. *et al.* (1987) 'Cellular heterogeneity and drug resistance in two ovarian adenocarcinoma cell lines derived from a single patient', *International Journal of Cancer*. doi: 10.1002/ijc.2910390607.

Xia, Z. *et al.* (2014) 'Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'2-UTR landscape across seven tumour types', *Nature Communications*. Nature Publishing Group, 5, pp. 1–13. doi: 10.1038/ncomms6274.

Xue, Y. *et al.* (2009) 'Genome-wide Analysis of PTB-RNA Interactions Reveals a Strategy Used by the General Splicing Repressor to Modulate Exon Inclusion or Skipping', *Molecular Cell*. doi: 10.1016/j.molcel.2009.12.003.

Xue, Z. *et al.* (2018) 'Recurrent tumor-specific regulation of alternative polyadenylation of cancer-related genes', *BMC Genomics*. BMC Genomics, 19(1), pp. 1–12. doi: 10.1186/s12864-018-4903-7.

Yamagishi, T. *et al.* (2013) 'P-glycoprotein mediates drug resistance via a novel mechanism involving lysosomal sequestration', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M113.514091.

Yan, G. C. *et al.* (1993) 'Exon Switching and Activation of Stromal and Embryonic Fibroblast Growth Factor (FGF)-FGF Receptor Genes in Prostate Epithelial Cells Accompany Stromal Independence and Malignancy', *Mol Cell Biol*. doi: 10.1128/MCB.13.8.4513.Updated.

Yang, J. *et al.* (2014) 'RBM24 Is a major regulator of muscle-specific alternative splicing', *Developmental Cell*. doi: 10.1016/j.devcel.2014.08.025.

Yang, X. *et al.* (2016) 'Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing', *Cell*. Elsevier Inc., 164(4), pp. 805–817. doi: 10.1016/j.cell.2016.01.029.

- Yeo, G. *et al.* (2004) 'Variation in alternative splicing across human tissues.', *Genome biology*. doi: 10.1186/gb-2004-5-10-r74.
- Yoon, Y. *et al.* (2013) 'Anxiety-associated alternative polyadenylation of the serotonin transporter mRNA confers translational regulation by hnRNP', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1301485110.
- Yoshida, K. *et al.* (2011) 'Frequent pathway mutations of splicing machinery in myelodysplasia', *Nature*. doi: 10.1038/nature10496.
- Yu, Y. and Reed, R. (2015) 'FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1506282112.
- Zhang, H., Rigo, F. and Martinson, H. G. (2015) 'Poly(A) Signal-Dependent Transcription Termination Occurs through a Conformational Change Mechanism that Does Not Require Cleavage at the Poly(A) Site', *Molecular Cell*. Elsevier Inc., 59(3), pp. 437–448. doi: 10.1016/j.molcel.2015.06.008.
- Zhang, S. *et al.* (2012) 'Identification of the molecular basis of doxorubicin-induced cardiotoxicity', *Nature Medicine*. doi: 10.1038/nm.2919.
- Zhang, Y. *et al.* (2013) 'Circular Intronic Long Noncoding RNAs', *Molecular Cell*. doi: 10.1016/j.molcel.2013.08.017.
- Zhang, Z., Fu, J. and Gilmour, D. S. (2005) 'CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11', *Genes and Development*. doi: 10.1101/gad.1296305.
- Zhao, W. and Manley, J. L. (1998) 'Deregulation of poly(A) polymerase interferes with cell growth', *Molecular and Cellular Biology*. doi: 10.1128/MCB.18.9.5010.
- Zheng, S., Kim, H. and Verhaak, R. G. W. (2014) 'Silent mutations make some noise', *Cell*. doi: 10.1016/j.cell.2014.02.037.
- Zheng, X. *et al.* (2015) 'Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer', *Nature*. doi: 10.1038/nature16064.
- Zhitomirsky, B. and Assaraf, Y. G. (2015) 'Lysosomal sequestration of hydrophobic weak base chemotherapeutics triggers lysosomal biogenesis and lysosome-dependent cancer multidrug resistance', *Oncotarget*. doi: 10.18632/oncotarget.2732.
- Zhu, N. *et al.* (2004) 'An alternatively spliced survivin variant is positively regulated by p53 and sensitizes leukemia cells to chemotherapy', *Oncogene*. doi: 10.1038/sj.onc.1208038.

Titre : Régulation de l'épissage et de la polyadénylation alternatifs par les agents anticancéreux génotoxiques

Mots clés : *épissage alternatif, polyadénylation alternative, chimiothérapie, résistance*

Résumé : La plupart des gènes humains codants génèrent des transcrits alternatifs (isoformes) par épissage alternatif (alternative splicing, AS) et polyadénylation alternative (APA) en général dans la région codante et la région 3' non traduite (3'UTR), respectivement. Le rôle de l'AS et la 3'UTR-APA est de plus en plus reconnu dans l'oncogénèse. En particulier, des réseaux d'AS connectant des facteurs d'épissage et des variants d'épissage ont récemment été identifiés. L'AS est aussi largement régulé par les agents anticancéreux génotoxiques, tel que la doxorubicine et le cisplatine (induisant des différents types de lésions sur l'ADN), qui sont régulièrement utilisés dans les traitements du cancer du sein et du poumon non-à-petites-cellules (non-small-cell lung cancer, NSCLC), respectivement. Étant donné l'apparition fréquente de résistances aux chimiothérapies, comprendre les mécanismes sous-jacents est crucial pour surmonter ce problème clinique. Il existe des exemples d'évènements d'AS associés à la résistance aux agents anticancéreux, mais l'implication des facteurs d'épissage et des réseaux d'AS est très peu connue. De plus, une étude précédente a démontré que la doxorubicine réprime un grand groupe d'exon terminaux alternatifs (alternative last exons, ALE), qui correspondent à l'utilisation de sites de polyadénylation introniques (intronic polyadenylation, IPA). Les ALEs ont un rôle émergent dans le cancer, mais on ne sait encore que très peu sur leur régulation par d'autres agents anticancéreux, tel que le cisplatine. Afin de mieux comprendre le rôle des régulations d'AS et d'APA dans la réponse et la résistance cellulaire à la chimiothérapie, mon projet de thèse avait deux objectifs principaux : 1) déterminer l'étendue, les réseaux régulateurs, et les fonctions des régulations d'AS dans la résistance à la doxorubicine des cellules de cancer du sein, et 2) déterminer l'étendue, les mécanismes, et l'impact des régulations d'ALE en réponse au cisplatine dans des cellules de NSCLC. Dans la première partie, j'ai identifié par RNA-seq des milliers d'évènements d'AS et des dizaines de facteurs d'épissage régulés dans un modèle cellulaire de cancer du sein ER+ résistant à la doxorubicine. Par un miniscreen siARN, j'ai identifié deux facteurs, ZRANB2 et SYF2, impliqués dans la résistance à la doxorubicine. D'autres analyses RNA-seq ont révélé les évènements d'AS régulés par ces deux facteurs peu étudiés, ainsi que leur convergence vers l'exon 5 alternatif de l'oncogène ECT2. La déplétion de ZRANB2, SYF2, et du variant ECT2-ex5 réduit l'arrêt en phase S induit par la doxorubicine et la résistance des cellules. De plus, un niveau élevé d'inclusion de l'exon 5 d'ECT2 corrèle avec une mauvaise survie spécifiquement de patientes ER+ traitées par chimiothérapie. Dans la deuxième partie, j'ai identifié par 3'-seq que le traitement cisplatine (mais pas oxaliplatine) induit des ALEs/IPAs dans des milliers de gènes enrichis en gènes de cycle et de mort cellulaire. Cet effet est lié à une inhibition de la processivité de l'élongation dans les longs gènes. Une analyse 3'-seq sur polysomes m'a permis de montrer que ces régulations d'ALEs impactent le traductome, et a révélé un groupe d'isoformes particulièrement courtes peu efficacement traduites, dont un transcrit connu avec une fonction non-codante. En conclusion, j'ai pu identifier durant ma thèse un nouveau réseau d'AS impliqué dans la résistance à la doxorubicine des cancers du sein ER+, et une importante régulation d'ALEs impactant le traductome en réponse au cisplatine dans des cellules NSCLC. Ces travaux améliorent notre compréhension du rôle de l'AS et des ALE/IPA dans la réponse et la résistance cellulaire à la chimiothérapie anticancéreuse. Au plus long terme, les transcrits alternatifs et les régulateurs identifiés constituent des biomarqueurs candidats de chimiorésistance.

Title : Regulation of alternative splicing and polyadenylation by genotoxic anticancer agents

Keywords : *alternative splicing, alternative polyadenylation, chemotherapy, resistance*

Abstract : Most human coding genes generate alternative transcripts (isoforms) through alternative splicing (AS) and alternative polyadenylation (APA), most often within the coding region and the 3' untranslated region (3'UTR), respectively. Both AS and 3'UTR-APA regulations have been increasingly involved in oncogenesis. In particular, AS networks connecting oncogenic splicing factors and oncogenic splicing variants have been recently identified. AS is also widely regulated by genotoxic anticancer drugs, like doxorubicin and cisplatin that induce different types of DNA lesions and are widely used in breast cancer and non-small-cell lung cancer (NSCLC) therapy, respectively. Given the frequent occurrence of resistance to chemotherapy, understanding the underlying mechanisms is crucial to overcome this major issue. There are examples of AS events associated with anticancer drug resistance, but very little is known about the splicing factors and therefore the AS networks involved. In addition, a previous study showed that doxorubicin represses a large set of alternative last exons (ALE) corresponding to the use of intronic polyadenylation (IPA) sites. ALEs have an emerging role in cancer, but little is known about its regulation by other anticancer drugs, like cisplatin. In order to better understand the role of AS and APA regulation in cell response and resistance to chemotherapy, my PhD project had two main aims: 1) determine the extent, regulatory networks and function of AS regulation in breast cancer cell resistance to doxorubicin, and 2) determine the extent, mechanism and impact of ALE regulation in response to cisplatin in NSCLC cells. In the first part, I identified by RNA-seq thousands of AS events and dozens of splicing factors regulated in a cell model of acquired resistance to doxorubicin in ER+ breast cancer. Through an siRNA miniscreen, I found two splicing factors, ZRANB2 and SYF2, involved in doxorubicin resistance. Further RNA-seq analyses revealed the AS events regulated by depletion of these poorly characterized splicing factors, and their convergence on the alternative exon 5 of the oncogene ECT2. Depletion of ZRANB2, SYF2 and the ECT2-Ex5 variant reduces doxorubicin-induced S phase arrest and doxorubicin resistance. In addition, high inclusion levels of ECT2-Ex5 correlate with poor survival specifically in ER+ breast cancer treated with chemotherapy. In the second part, I found by 3'-seq that in NSCLC cell treatment with cisplatin (but not oxaliplatin) induces ALE/IPA in thousands of genes enriched in cell cycle and cell death. This effect is linked to an inhibition of transcription elongation processivity in long genes. 3'-seq analysis on polysomes showed that this ALE regulation impacts the transcriptome, and revealed a set of particularly short isoforms that were inefficiently translated, including a transcript with a non-coding function. In conclusion, during my thesis, I could identify a novel AS network involved in doxorubicin resistance in ER+ breast cancer, and widespread ALE regulation impacting the transcriptome in lung cancer cisplatin response. This work increases our understanding of AS and IPA role in cell response and resistance to anti-cancer chemotherapy. In the longer term, the identified alternative transcripts and regulators constitute candidate biomarkers of chemoresistance.

