NUI Galway
OÉ Gaillimh

Genetic analysis of *ATR*, a central regulator of genome stability

Marta Llorens Agost

Genome Stability Laboratory,
Centre for Chromosome Biology,
Biomedical Sciences, Dangan
National University of Ireland, Galway

A thesis submitted to the National University of Ireland, Galway
for the degree of Doctor of Philosophy

Supervisor: Prof. Noel F. Lowndes

May 2016

Table of contents

List of figures

List of tables

List of supplementary materials

Abbreviations

4NQO: 4-Nitroquinoline 1-Oxide
5-FOA: 5-Fluoroorotic Acid
53BP1: p53 Binding Protein 1
9-1-1: RAD9-HUS1-RAD1
9G8: Splicing factor also known as SRSF7 (serine/arginine-rich splicing factor 7)
AA: amino acid
AID: Auxin Inducible Degron
APS: Ammonium Persulphate
ATM: Ataxia Telangiectasia Mutated
ATR: Ataxia Telangiectasia and RAD3 related (Human)
Atr: Ataxia Telangiectasia and RAD3 related (Chicken)
ATRIP: ATR-interacting protein
AUX: Indole-3-acetic acid (Auxin)
BAX: Proapoptotic Bcl2-Associated X
BID: BH3 Interacting-domain Death agonist
BLM: Bloom Syndrome RecQ Helicase-Like
bp: Base Pair
BRCA1: Breast cancer associated gene 1
BRCA2: Breast cancer associated gene 2
BrdU: Bromodeoxyuridine
BSA: Bovine Serum Albumin
CBP: Calmodulin-binding protein
CDC25: Cell Division Cycle 25
CDK: Cyclin-Dependent kinase
cDNA: complementary DNA
CEP63/152: Centrosomal protein 63kDa or 152kDa
CHK1/2: Checkpoint kinase 1/2
CORE: COunterselectable REporter
CRISPR: Clustered regularly-interspaced short palindromic repeats
CS: Chicken Serum
DAPI: 4',6-diamidino-2-phenylindole
Dbp11: DNA Polymerase B (II)
Ddc2: DNA damage checkpoint 2
ddH2O: double distilled water
DDR: DNA Damage Response
DMEM: Dulbecco's Modified Eagles Medium
DMSO: Dimethyl Sulfoxide
DNA-PK: DNA-dependent protein kinase
DNA: Deoxyribonucleic Acid
dNTP: deoxynucleotide triphosphate
DSB: Double Strand Break
dsDNA: double stranded DNA
E: Elution
EDTA: Ethylenediaminetetraacetic Acid
EM: Electron microscopy

ERCC4: Excision Repair Cross-Complementation Group 4
ESE: Exonic Splicing Enhancer
ESS: Exonic Splicing Silencer
FACS: Fluorescence-Activated Cell Sorting
FAT-C: FAT-C terminal domain
FAT: FRAP-ATM-TRRAP
FBS: Foetal Bovine Serum
FITC: Fluorescein Isothiocyanate
FT: Flow Through
G1/2: Gap phase 1/2
gDNA: genomic DNA
gRNA: guide RNA
HA: Hemagglutinin
HCl: Hydrochloric acid
HEAT: Huntingtin, Elongation Factor 3 (EF3), Protein Phosphatase 2A (PP2A), and the yeast kinase TOR1
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HR: Homologous Recombination
HU: Hydroxyurea
IR: Ionizing radiation
ISE: Intronic Splicing Enhancer
ISS: Intronic Splicing Silencer
kb: kilobase
kDa: kilodaltons
MEC1: Meiotic Entry Checkpoint 1
MDC1: Mediator of DNA damage checkpoint 1
M-phase: Mitosis phase
MRN: Mre11-Rad50-Nbs1
mRNA: messenger RNA
mTOR: Mammalian Target of Rapamycin
NaCl: Sodium Chloride
NBS1: Nijmegen breakage syndrome 1
NEK1: Never in mitosis (NIMA) related kinase 1
NER: Nucleotide Excision Repair
NF1: Neurofibromin
NHEJ: Non-Homologous End Joining
NPC: Nuclear pore complex
Ova: Ovalbumin
PAGE: Polyacrylamide Gel Electrophoresis
PBS: Phosphate Buffered Saline
PCNA: Proliferating Cell Nuclear Antigen
PCNT: Pericentrin
PCR: Polymerase Chain Reaction
PI: Propidium Iodide
PI3K: Phosphatidylinositol-3 kinases
PIKK: Phosphatidylinositol 3-kinase-related kinases (PIKKs)
PIN1: Peptidylprolyl cis/trans Isomerase NIMA-interacting 1 protein
PLK1: Polo-like kinase 1
PP2A: Protein Phosphatase 2A
PRD: PIKK Regulator Domain
RAD: Radiosensitive
Rec: Recombination

RFC: Replication Factor C
RHINO: RAD9-HUS1-RAD1 Interacting Nuclear Orphan 1
RNA: Ribonucleic Acid
RNAse: Ribonuclease
RNP: Ribonucleoprotein
RNR: Ribonucleotide Reductase
RPA: Replication Protein A
RPMI: Roswell Park Memorial Institute Medium
SAC: Spindle Assembly Checkpoint
SCF: Skp1, Cullin and F-box protein
SD: Standard Deviation
SDS: Sodium Dodecyl Sulphate
siRNA: Small interfering RNA
SMC1: Structural maintenance of chromosomes 1A
SMG-1: Suppressor of morphogenesis in genitalia
SML1: Suppressor of Mec1 Lethality protein 1
SN: Supernatant
ssDNA: single stranded DNA
S-phase: Synthesis phase
SQ/TQ: Serine-Glutamine / Threonine-Glutamine motif
TCE: Total cell extract
Tel1: Telomere length regulation protein 1
TIM: Timeless protein
TIP60: 60kDa trans-acting regulatory protein of HIV type 1 (Tat) interacting protein
TIPIN: Timeless Interacting Protein
TIR1: Transport Inhibitor Response 1
TOPBP1: TopoisomeraseII binding protein I
Tra2Beta: Transformer 2 Beta Homolog
TRRAP: Tansformation/Transcription domain Associated Protein
UV: Ultraviolet
V(D)J recombination: Variable (Diversity) Joining recombination
WT: Wild type
XPF: Xeroderma Pigmentosum complementation group F
YB1: Y-box binding protein 1
γH2AX: phosphorylated H2AX
YNB: Yeast Nitrogen Base
YPD: Yeast Peptone and Dextrose medium

Acknowledgements

During the course of my Ph.D. and my stay in Galway, many people have helped and supported me. I would like to thank you all because without you I would not have come this far.

Firstly I would like to thank Noel Lowndes for giving me the opportunity to continue with this project in the Genome Stability Laboratory and your supervision throughout the years of my PhD.

I feel very lucky to have had the chance to work in such a wonderful place, full of amazing people. Thanks to all of you for balancing out the science with so many good memories. I will miss playing tag rugby with Scrummy Bears/George, the College Bar, our Christmas parties, morning tea breaks and ice cream breaks after long lab meetings.

Special thanks go to the 'oldies' Emma, Carla and Louise. Thanks for all your support and advice in and outside the lab. Special mention goes to Louise for being the first one reading this book! I would also like to express huge gratitude to John. Thank you for your advice and guidance, your infinite patience and for sharing your passion for science with all of us.

To my family, thank you for believing in me and supporting me even when that meant having me miles away. *Moltes gràcies mamà, papà i Anni per estar sempre al meu costat encara que estiguérem lluny. Gràcies a tots pel vostre suport i per estar orgullosos de mi.*

I cannot forget my Spanish girls. Thanks for all your support at the best and hardest times. You made me feel that Galway was my home. I won't forget our family meals, Saint Patrick's, coffee afternoons and trips together. *Gracias, Antonias. Por muchos más años juntas, aunque sea en la distancia.*

Special thanks go to Clara, my housemate for the last two years, you are like a sister to me. *Moltes gràcies Clareta per tot el que hem viscut juntes. Trobaré a faltar les nostres nits vegent 'Vikings', les nostres clases de ioga, i totes els riures i bons moments que hem passat en la nostra caseta.*

To my friends in Spain and further, Leire, Espe, Bárbara, Pilar and all the others, thank you for encouraging me from the start. I can't wait to share more sunny afternoons by the beach, in Castellón o wherever in the world you may be.

Finally, to Claas, thank you for being by my side on this journey. For all your unconditional support, love and understanding even at my most stressy/silly times. I'm excited for all the new adventures I will get to share with you, in Ireland or wherever life will take us.

My stay in Galway has marked me forever. I won't forget all those (many) rainy afternoons, the colourful houses and magic atmosphere. But I will especially remember my time here for all the wonderful friendships that I have made.

Thesis Declaration and Contributions

This PhD thesis was supported by a fellowship from the PhD programme "Molecular and Cellular Mechanisms Underlying Inflammatory Processes" which was funded by the Higher Education Authority of Ireland under Cycle 5 of the Irish Government's Programme for Research in Third Level Institutions (PRTLI; grant number RH5402). Research infrastructure was additionally supported by Science Foundation Ireland (SFI). Part of the research was also funded by the Beckman Fund Scholarship (grant number RNR921), which I was awarded with in 2013.

Ethical approval for the research was granted by the South Birmingham Ethics Committee (REC reference number 07/H1210/155). Informed consent was obtained and clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

I declare that I have not obtained any previous qualification from NUI Galway based on any results contained in this thesis. Part of the work done in Chapter 3 belongs to the final degree project that I carried out in Galway (as indicated below), as part of my degree in Biotechnology, awarded in 2011 in the Polytechnic University of Valencia.

I conducted all the experiments and wrote the thesis under the supervision of Professor Noel Lowndes. There are a few exceptions where experiments were performed by others but included in this thesis to complete the story. Some cell lines and plasmids were also obtained from other laboratories. For clarity these are indicated clearly in the figure legends or text when appropriate and included below.

Chapter 3

- The *AID-Atr* cell line, used as a control in this and following chapters, was generated by Dr. John Eykelenboom. More information can be found in the publication (Eykelenboom et al., 2013).

- The design of the window deletion mutants (*Atr*$^{\Delta HEATs}$) was performed by Dr. John Eykelenboom (Figure 23).
- Experiments done with the *Atr*$^{\Delta HEATs}$ cell lines (Figure 24-27) are part of my final degree project performed in the academic year 2010/2011.
- The design of the yeast replacement mutants was done in collaboration with Dr. Carla Abreu (Figure 28).
- The HFSC-Atr cell line was generated by Dr. John Eykelenboom, while the design of the HFSC tag was done by Dr. Fabio Pessina (Figure 32A) (Pessina and Lowndes, 2014).
- Electron microscopy data acquisition was carried out with the help of Pierce Lalor, senior technical officer at the Anatomy department in NUIG (Figure 33C-E).

Chapter 4

- The discovery of the *ATR*^{M1159I} and *ATR*^{K1665N} mutations in Seckel patients and initial human cell line characterization was carried out by Prof. Grant Stewart and colleagues, a collaborator based in Birmingham, UK (Figure 34A and 36).
- Patient-derived cell lines used in section 4.2.4 were a gift from Prof. Grant Stewart; while the *NF1* minigene vector used in section 4.2.5 was provided by Dr. Dawn O'Reilly, being originally generated in Prof. Diana Baralle's laboratory.

Chapter 5

- Image acquisition for karyotypic analysis (ie. metaphase spreads) was carried out by Dr. Muriel Voisin (Figure 48).

Abstract

To protect the DNA against the constant assaults of endogenous and environmental agents, cells have developed conserved signaling pathways known as the DNA damage response. Such responses are important to maintain genome stability and prevent human disease.

The ATR kinase plays a critical role during the DNA Damage Response, as well as during unperturbed DNA replication. Despite the vital importance of ATR, little is known about its structure and the precise regulation of its activity. In this study we investigate the structure-function relationship of ATR using DT40 chicken cells. In particular, we examine the role of tandem helical motifs, called HEAT repeats, that comprise the non-kinase portion of ATR. We found that *Atr* mutants, expressing an Atr protein version where single/multiple HEATs were removed, behaved similarly to the null cell line. In contrast, hybrid *Atr* mutants expressing equivalent regions of human ATR were fully functional. Our observations suggest that all ATR functions are tightly dependent on the integrity of HEAT repeats. Although we cannot exclude that these motifs mediate specific protein-protein interactions, it is likely that their main role is to contribute to an overall structural function, holding ATR configuration in place.

ATR is an essential gene and mutations in this gene have been implicated in different human diseases, such as cancer and Seckel syndrome. In this thesis, we have also modeled mutations implicated in such disorders using DT40 and human cells. In particular, we found that two novel missense *ATR* Seckel mutations (M1159I and K1665N) do not affect protein function, but instead impact ATR splicing, potentially by affecting binding of splice factors to sequence-specific regulators of splicing. Additionally, we have also investigated the effect of a non-Seckel mutation (Q2144R) linked to cancer predisposition in patients. Our results indicate that this mutation has a profound effect on ATR activity. Checkpoint signaling is abrogated in the presence of this substitution and, as a consequence, cells seem to accumulate chromosomal defects resulting in cell death. These findings indicate that Q2144 residue, which is part of a potential SQ site, could be key in the ATR– mediated response to DNA damage.

CHAPTER 1

General Introduction

ATR – An essential PIKK in the DNA Damage Response and DNA replication

Keywords: ATR, Cell cycle checkpoints, Cilia, DNA Damage Response, *Gallus gallus*, HEAT repeats, Mitochondria, Nuclear envelope, Oropharyngeal cancer, PIKK, *Saccharomyces cerevisiae***,** Seckel syndrome, Separation of function mutations, Splicing, SQ/TQ,

1.1 DNA Damage Response

The genome of all living organisms is continuously exposed to a variety of insults, which include both environmental DNA damaging agents, such as ionizing radiation (IR) and ultraviolet (UV) light, and also endogenous reactive metabolic by-products. Both exogenous and endogenous agents can lead to several types of DNA damage, including interstrand crosslinks (ICLs), double-strand breaks (DSBs), and single-strand breaks (SSBs), or cause replicative stress (Abraham, 2001; Ciccia and Elledge, 2010).

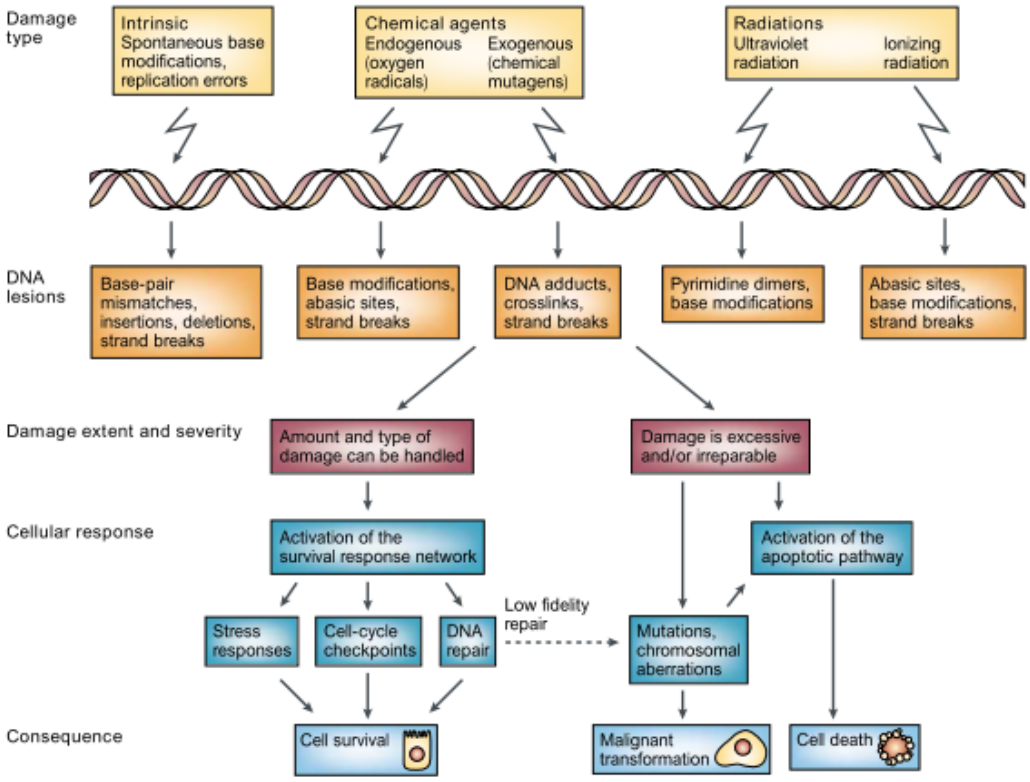


Figure 1. Cellular responses to DNA damage (Shiloh, 2003). Depending on the damaging agent, different types of lesions are produced and, consequently, the cell response to deal with the damage is also different. The survival, malignant transformation or death of the cells depends on the capacity of DDR to handle the damage produced.

To monitor genome integrity and respond to DNA damage, eukaryotes have developed a complex signal transduction pathway called the DNA damage response (DDR). Activation of this cascade affects multiple cellular processes and is essential to coordinate cell cycle transitions, DNA replication, DNA repair and in extreme cases, when damage is irreparable, cells undergo senescence or

programmed cell death termed apoptosis (Harper and Elledge, 2007). Failure to activate this specific response can lead to genome instability, resulting in cell death or propagation of mutations that can contribute to the development of cancer (Figure 1) (Cimprich and Cortez, 2008).

The DNA damage response involves proteins that are generally classified as sensors, signal transducers/mediators and effectors. Activation of DNA damage checkpoints occurs through the detection of the DNA damage by the sensor proteins, so this information is communicated from transducers to effectors via adaptor or mediator proteins, resulting in various physiological responses to handle the damage (Zhou and Elledge, 2000). This classification simplifies the checkpoints as unidirectional cascades, to facilitate the comprehension of the DDR. However, it has been shown that several checkpoint or DNA repair proteins can participate in different stages of this complex pathway (Putnam et al., 2010).

1.1.1 PIK kinase family: key regulators of the DDR

The master regulators of the DDR are serine/threonine kinases, which are part of the phosphatidylinositol 3-kinase-like protein kinase (PIKK) family. ATM (Ataxia Telangiectasia Mutated) and ATR (Ataxia Telangiectasia and Rad3-related) belong to this family of proteins, amongst others, and are essential for maintaining genome integrity (Lovejoy and Cortez, 2009).

After DNA damage is detected, PIK kinases are activated and relay the message by a phosphorylation cascade to stimulate downstream targets, including those that delay cell-cycle progression and induce the expression of DNA repair genes.

As seen in Figure 2, ATM is required for checkpoint activation primarily in response to DSBs (Riches et al., 2008; Suzuki et al., 1999), while ATR controls the response to ssDNA, which reflects a much broader spectrum of DNA damage and replication stress including breaks, crosslinks and base adducts (Cimprich and Cortez, 2008; Zou and Elledge, 2003). In addition, *ATM* is a non-essential gene, whilst *ATR* is essential for viability (Brown and Baltimore, 2000; Zhao et al., 1998).

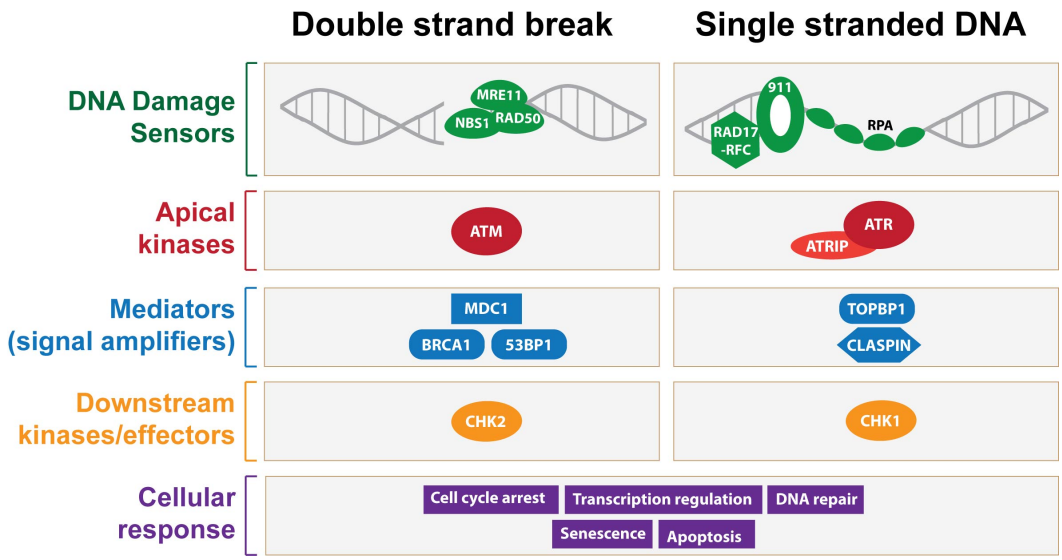


Figure 2. ATM versus ATR signaling. ATM orchestrates the response to DSBs, while ATR mediates the repair of exposed ssDNA. Key sensor (in green), mediator (in blue) and effector (in orange) proteins are compared in the figure between the two pathways.

The basic organization of DNA damage checkpoints has been conserved through evolution, and many of the checkpoint proteins are conserved from yeast to human (Table 1).

Table 1. Principal checkpoint proteins in the ATR and ATM pathways. Comparison of homologous proteins in *Saccharomyces cerevisiae*, *Saccharomyces pombe* and *Homo sapiens*. Adapted from (Navadgi-Patil and Burgers, 2009)

S. cerevisiae	*S. pombe*	Humans	Step	Proposed function(s)
Mre11/Rad50 / Xrs2	Rad32/Rad50/Nbs1	MRE11/RAD50 /NBS1	Sensor	Initial sensing and processing of DSBs. Involved in DSBs repair (NHEJ and HR), and telomere maintenance.
Rad24/RFC2-5 (Clamp-loader)	Rad17/RFC2-5	RAD17/ RFC2-5	Sensor	Loading of 9-1-1 complex onto DNA.
Rad17-Ddc1-Mec3 (9-1-1	Rad9-Rad1-Hus1	RAD9-RAD1-HUS1	Sensor	Complex structurally related to PCNA involved in signal transduction and possibly in

complex)				the recruitment of other checkpoint proteins.
Tel1	Tel1	ATM	Transducer	Central PIKK in DDR, involved in signaling DSBs. Also participates in telomere maintenance.
Mec1	Rad3	ATR	Transducer	Central PIKK in DDR, controls the response to not only DNA damage but also replication stress, including breaks, crosslinks and base adducts.
Ddc2	Rad26	ATRIP	Transducer	Mec1/Rad3/ATR regulatory subunit, which mediates its recruitment to the sites of damage.
Dbp11	Cut5	TOPBP1	Mediator	Replication initiation protein: checkpoint mediator recruited to sites of damage by the 9-1-1 complex, where it activates Mec1/Rad3/ATR.
Rad9	Crb2	53BP1/MDC1/ BRCA1	Adaptor/ Mediator	Required for signal transduction as an adaptor and involved also in DSB repair. Scaffold for Rad53 or Chk1 activation.
Rad53	Cds1	CHK2	Effector	Involved in signal transduction as an effector kinase.
Chk1	Chk1	CHK1	Effector	Involved in signal transduction as an effector kinase.

Cancer cells often show some degree of replication stress and therefore up-regulate ATR and CHK1 activity (Toledo et al., 2011). Besides, many cancer cells also have mutations in tumour suppressor genes, such as *ATM* (Jiang et al., 2009), which further increases the dependence on ATR signaling for survival, as these two proteins show partial functional redundancy and share overlapping substrate

specificity (Bensimon et al., 2011). For these reasons, the DDR regulators have become potential therapeutic targets in cancer therapy (Sanjiv et al., 2015; Weber and Ryan, 2014).

1.1.1.1 Response to DSBs: ATM signaling

When a DSB is generated, the MRE11-RAD50-NBS1 (MRN) complex is one of the first factors to be recruited to the site (Rupnik et al., 2008), acting as the initial damage sensor. MRN is responsible for ATM recruitment and activation at sites of damage. In particular, it is NBS1 that binds ATM through a particular structural motif (You et al., 2005), mediating its localization to the DSB. Before recruitment, ATM is kept inactive by interacting with PP2A (protein phosphatase 2A) (Goodarzi et al., 2004). To allow ATM's release from its inactive state, TIP60 (60kDa Trans-acting regulatory protein of HIV type 1 (Tat) – Interacting Protein) acetylates ATM on lysine 3016, favouring its dissociation from PP2A (Sun et al., 2005).

ATM auto-phosphorylation on serine 1981 is essential for its activation. This phosphorylation event also regulates the formation of active ATM monomers (Bakkenist and Kastan, 2003; Goodarzi et al., 2004). Once activated, ATM rapidly phosphorylates the histone variant H2AX on S139 (γH2AX). The formation of γH2AX spreads up to 2 megabases from the break, serving as a recruitment platform for MDC1 (Mediator of DNA damage Checkpoint Protein 1) (Paull et al., 2000; Rogakou et al., 1998, 1999; Stewart et al., 2003). Once recruited, MDC1 is phosphorylated by ATM and, together with the ubiquitin ligases RNF8 and RNF168 (Doil et al., 2009; Mailand et al., 2007), coordinates the recruitment to sites of damage of other mediator proteins, such as 53BP1 (p53 Binding Protein 1) and BRCA1 (BReast CAncer 1) (Kolas et al., 2007; Yan and Jetten, 2008).

The BRCA1 tumour suppressor is associated with hereditary breast and ovarian cancer and has an important role in the DDR. BRCA1 is capable of interacting with several partners, which allow its participation in distinct checkpoints and repair processes, such as the G2/M checkpoint (Wang et al., 2007; Yu and Chen, 2004), replication checkpoint (Xu et al., 2001) and DNA end

resection in homologous recombination (Sartori et al., 2007; Schlegel et al., 2006).

53BP1 is vital for ATM phosphorylation of a number of substrates including CHK2, SMC1, RPA2 and BRCA1. While 53BP1 is not directly required in repair by homologous recombination (HR), nor for 'classical' non-homologous end joining (cNHEJ); it regulates the balance between HR and NHEJ, by preventing double strand break end resection (Bunting et al., 2010). Mutual antagonism between 53BP1 and BRCA1 dictates whether the cell undertakes homologous recombination or non-homologous end joining repair (Bakr et al., 2015; Daley and Sung, 2014). Besides, long distance NHEJ relies on 53BP1, a process that occurs for example during V(D)J and class switch recombination (Panier and Boulton, 2014).

CHK2 (Checkpoint Kinase 2) is the main effector of the ATM signaling pathway. CHK2 activation occurs upon its phosphorylation by ATM on threonine 68 (T68) following treatment with IR, but not HU or UV (Matsuoka et al., 2000). CHK2 prevents cell cycle progression by inhibiting CDC25A, B and C (Falck et al., 2001). For further maintenance of cell cycle arrest, p53 is phosphorylated in a CHK2 dependent manner (Iliakis et al., 2003). If the damage is irreversible, CHK2 stimulates apoptosis by phosphorylating several targets such as p53, MDM4 and PML (Yang et al., 2002).

1.1.1.2 Response to ssDNA: ATR signaling

The ATR kinase orchestrates the response to ssDNA (Flynn and Zou, 2011). When ssDNA is detected, this is immediately coated by the Replication Protein A (RPA) complex, which allows the recruitment of ATR and its interacting partner, ATRIP (ATR Interacting Protein) (Zou and Elledge, 2003). Other sensor proteins, such as the RAD17 and the 9-1-1 (RAD9, RAD1 and HUS1) complexes, assist in ATR's localization to the sites of damage (Kobayashi et al., 2004; Lee et al., 2007). ATR is then activated by TOPBP1 (Topoisomerase II binding protein 1) (Kumagai et al., 2006; Mordes et al., 2008). This promotes the start of an ATR-dependent phosphorylation cascade, which culminates with the phosphorylation of CHK1. CHK1 (Checkpoint Kinase 1) is a central effector protein of the ATR pathway, and its activation results in cell cycle arrest and

DNA repair following damage or stalled replication (Liu et al., 2000; Zhao and Piwnica-Worms, 2001).

A detailed description of ATR response to ssDNA is included in section 1.3.1.

1.2 Cell cycle regulation and checkpoints

Cell division is a vital mechanism by which all living cells reproduce. This process must be tightly regulated in order to faithfully copy the entire genetic information in a cell and distribute it evenly into two daughter cells, so that each one of them receives a complete copy of the genome. The cell cycle is divided into four distinct phases, known as G1 (Gap 1), S (Synthesis), G2 (Gap 2) and M (Mitosis and cytokinesis) (Figure 3). G1, S and G2 phases can also be grouped and referred to as interphase, or the period between one M phase and the next one (Alberts et al., 2009).

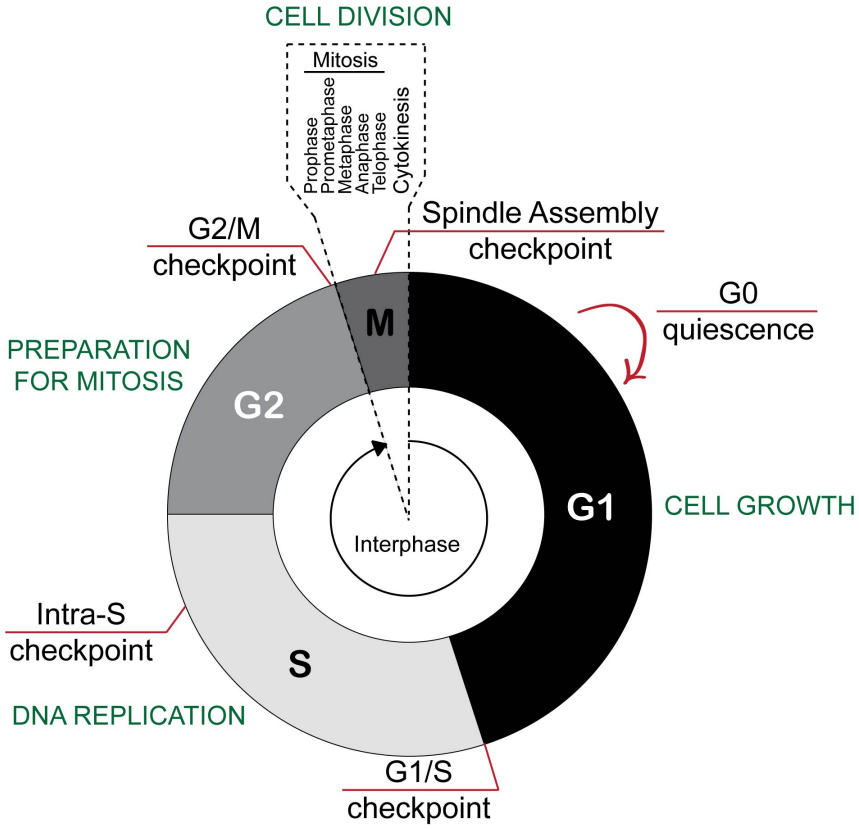


Figure 3. The cell cycle.

Cells cycle through G1 (Gap1 phase), S (Synthesis phase), G2 (Gap2 phase) and M (cell division), which can be subdivided into distinct sub-phases including prophase,

prometaphase, metaphase, anaphase, telophase and cytokinesis. Cells can escape from the cell cycle and stop proliferating, by entering a quiescence or prolonged non-dividing state called G0 (Martin and Stein, 1976).

Genome integrity is surveyed at three major DNA damage checkpoints, G1/S, Intra-S and G2/M. If damage is detected in cells, these checkpoints are activated, pausing cell cycle progression until the damage is resolved. The spindle assembly checkpoint (SAC) monitors the transition through mitosis, ensuring proper segregation of the DNA.

Progression through the cell cycle is determined by cyclin dependent kinases (CDKs) and their partners the cyclins (Johnson and Walker, 1999; Sherr and Roberts, 2004). CDKs are constitutively expressed in cells, whereas cyclins are synthesised at specific stages of the cell cycle, in response to various stimuli. Cyclin-CDK complexes result in irreversible transitions at three main stages in the cell cycle: the G1/S transition, the G2/M transition and the metaphase to anaphase transition (Morgan, 1997). A summary of these CDK-cyclin complexes and the cell cycle phase that they play a role in is shown in Figure 4.

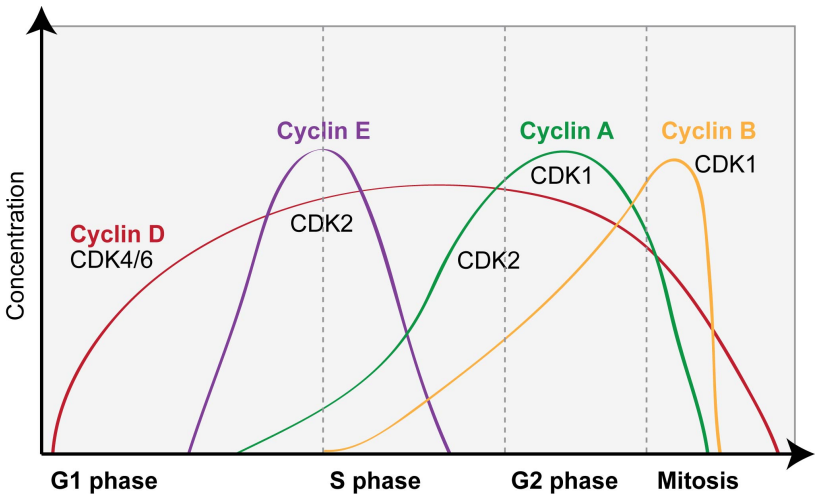


Figure 4. Profile of cyclin and CDK complexes through the cell cycle. Expression levels of the four principal cyclins (cyclin D, cyclin E, cyclin A and cyclin B) oscillate throughout the cell cycle, as represented by the coloured lines. Progression through the cell cycle depends on the interactions between cyclins and specific CDKs.

To ensure that cells duplicate their genome and organelles, and divide in an orderly manner, eukaryotic cells have developed a network of regulatory pathways known as checkpoints. These monitor the cell cycle and only allow progression if all quality control criteria are met. This prevents improper

chromosome segregation and the propagation of DNA damage on to the next generation of daughter cells (Abraham, 2001).

DNA lesions cause checkpoint activation, halting the cell cycle at specific points: the G1/S transition, mid-S phase or the G2/M transition. This allows time to complete DNA repair prior to re-entering the cell cycle or for the cell to decide to undergo apoptosis in case of excessive damage (Bartek and Lukas, 2007; Flynn and Zou, 2011; Lowndes and Murguia, 2000). ATM and ATR are essential governors of DNA damage checkpoints (Abraham, 2001; Ciccia and Elledge, 2010).

1.2.1 The G1/S checkpoint

The G1/S checkpoint is activated when cells are exposed to DNA damage in G1 phase. In an unperturbed cell cycle, entry into S-phase is regulated by Cyclin E/CDK2 association, which phosphorylates p27, a CDK inhibitor that is then degraded by the proteasome (Rizzardi and Cook, 2012). CDC25A also participates in this transition by removing inhibitory phosphate groups in CDK2, allowing CDK2 to promote S-phase progression (Satyanarayana and Kaldis, 2009). Abundance and activity of the Cyclin E/CDK2 complexes are also regulated by Myc and Rb pathways, which regulate genes critical for the G1/S transition (Bartek and Lukas, 2001).

In the presence of DNA damage, ATR and ATM activate their downstream targets CHK1 and CHK2, respectively. Once activated, these effector kinases phosphorylate CDC25A leading to its degradation (Mailand et al., 2000). In the absence of CDC25, CDK2 is unable to promote S-phase entry (Falck et al., 2001). This response is rapidly activated, however, it is only transient and cells re-enter the cell cycle after only a few hours. If repair remains incomplete, p21 plays a vital role in maintaining the G_1/S DNA damage checkpoint. p21, which becomes transcriptionally activated in a p53 dependent manner, binds and inhibits CDK2, impairing the formation of the Cyclin E/CDK2 complex and thereby delaying cell cycle progression (Agarwal et al., 1995; Bartek and Lukas, 2001).

1.2.2 The S-phase checkpoint

In response to replication fork stalling, cells activate an S-phase checkpoint pathway, also known as the replication checkpoint (Paulsen and Cimprich, 2007). Replication forks can be blocked by a variety of different factors, such as the uncoupling of the MCM helicase from the DNA polymerase machinery (Byun et al., 2005) or the presence of genotoxic agents like hydroxyurea (Hakansson et al., 2006). During the S-phase checkpoint, DNA synthesis is slowed down and origin firing is inhibited, to allow fork stabilization and restart (Santocanale and Diffley, 1998; Shirahige et al., 1998; Zegerman and Diffley, 2010). The replication checkpoint also plays an important role in preventing fork reversal (Hu et al., 2012).

The S-phase checkpoint can occur via two distinct mechanisms. The first effector branch relies on ATM/ATR-dependent proteolysis of CDC25A. CHK2- and CHK1-mediated phosphorylation events cooperate to target CDC25A for degradation, which ultimately leads to the disruption of Cyclin A–CDK2 complexes. This impedes the loading of the CDC45 initiation factor onto chromatin, preventing firing of the remaining pool of replication origins (Falck et al., 2002; Takisawa et al., 2000). The second effector branch of the replication checkpoint is not very well understood yet. NBS1 is the central player of this pathway, which can act as a platform/mediator protein for the ATM and ATR pathways. NBS1 phosphorylation in an ATM/ATR-dependent manner is required for the S-phase checkpoint (Falck et al., 2002; Pichierri and Rosselli, 2004). This is followed by further phosphorylation of SMC1 and FANCD2, events that lead to replication inhibition by different mechanisms. SMC1 phosphorylation regulates the cohesion between the template and the sister chromatids, hence affecting the progression of the replication fork (Yazdi et al., 2002). While upon FANCD2 phosphorylation, its monoubiquitination is slightly enchanced (Ho et al., 2006), stimulating its downstream functions in DNA repair. Additionally, Cyclin E degradation is mediated by Artemis, allowing free CDK2 to now bind Cyclin A, which permits re-entry into the cell cycle (Wang et al., 2009).

Bartek and colleagues have also described two other checkpoints that occur in S-phase: the S/M checkpoint and the intra-S checkpoint, which are commonly mistaken with the described replication checkpoint. The S/M checkpoint is in charge of halting cell division until cells have faithfully

duplicated their entire genome, thus its main target is the cyclin-B–CDK1 complex. The intra-S checkpoint is activated by DSBs that are generated in the genomic regions that fall outside active replicons and is, therefore, independent of replication. Despite having distinct characteristic features, all the above checkpoint pathways are closely coordinated and share some components (Bartek et al., 2004).

1.2.3 The G2/M checkpoint

Upon DNA damage, the G2/M checkpoint targets the Cyclin B/CDK1 complex, which regulates mitotic entry in unperturbed cells. As in the other DNA damage checkpoints, ATM and ATR are responsible for CHK2 and CHK1 stimulation. Activation of the G2/M checkpoint machinery promotes CDK1 phosphorylation at residues T14/Y15 by the WEE1 and MYT1 kinases, inhibiting its function (Lindqvist et al., 2009). Moreover, CDC25C, which normally counteracts WEE1 inhibitory function on CDK1, is downregulated or degraded (Abraham, 2001; Thanasoula et al., 2012). Both the G1/S and G2/M checkpoints, unlike the S-phase checkpoint, are dependent on p53. As previously explained, p53 activates p21 transcription and promotes CDK2 inhibition (Agarwal et al., 1995; Taylor and Stark, 2001). Cyclin B and CDK1 interaction is also directly disrupted by GADD45. Further action from the 14-3-3 regulatory proteins stimulates CDK1 degradation by the proteasome (Peng et al., 1997).
Checkpoint recovery, or progression through mitosis, occurs when PLK1 phosphorylates and inhibits two CHK1 activators, WEE1 and CLASPIN, promoting CDC25C accumulation (Mamely et al., 2006; Van Vugt et al., 2004).

1.3 The ATR kinase

The present project will focus on the vertebrate ATR (Ataxia Telangiectasia Mutated and RAD3-related) protein, the functional homolog of RAD3 and MEC1 checkpoint proteins from *S. pombe* and *S. cerevisiae*, respectively (Cortez et al., 2001). *ATR* is an essential gene in mammalian cells (Cortez et al., 2001; O'Driscoll, 2009) and mutations in this gene can be linked to cancer predisposition in humans (O'Driscoll and Jeggo, 2003; Tanaka et al.,

2012). In addition, hypomorphic mutations in *ATR* have been linked to rare cases of Seckel syndrome, which is characterized by microcephaly and developmental defects (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012).

1.3.1 ATR signaling

Some DNA damaging agents, such as ultraviolet (UV) light, replication inhibitors such as hydroxyurea, or endogenous replication stress can cause the formation of single-stranded DNA (Abraham, 2001). When such damage is detected, the ATR kinase is activated and orchestrates a response that will lead to suppression of cell cycle progression to allow enough time to stabilize and repair damaged replication forks (Cimprich and Cortez, 2008) (Figure 5).

1.3.1.1 Upstream events in the ATR signaling pathway: ATR recruitment and activation

ATR is one of the apical kinases of the eukaryotic DDR (Shiloh, 2003) and, although it is activated in response to many different types of DNA damage, including double-strand breaks (DSB), base adducts, crosslinks and replication stress, the common feature required for ATR activation is the presence of single stranded regions of DNA (ssDNA) (Cimprich and Cortez, 2008).

One potential trigger of ATR activation can occur during DNA replication, when uncoupling of the MCM helicase from the DNA polymerase machinery can occur if there are DNA lesions that block the polymerase activity and lead to the production of regions of ssDNA (Byun et al., 2005). The ssDNA is recognised and becomes bound by RPA, forming an RPA polymer and generating a platform that recruits ATR through direct interaction with its interacting partner ATRIP (Zou and Elledge, 2003). RPA consists of three subunits, RPA1, RPA2 and RPA3 (of 70, 34 and 14 kDa in size respectively), and it localizes to all ssDNA, preventing its degradation (Binz et al., 2004). ATR and RPA interaction is dependent on ATRIP and occurs via the larger RPA subunit (RPA1) (Xu et al., 2008).

Although RPA-coated ssDNA allows the ATR-ATRIP localizing function to the damaged site, it is not sufficient for ATR activation. The start of the ATR

signaling cascade depends on the co-localization of the ATR-ATRIP complex and the 9-1-1 complex. The 9-1-1 complex is loaded onto the DNA by the DNA damage clamp loader replication factor C (RAD17-RFC) (Bermudez et al., 2003), which consists of the RAD17 protein and the four small RFC2-5 subunits of RFC. This allows the 9-1-1 complex to bind to a DNA end adjacent to the RPA-coated ssDNA, and facilitates the interaction of ATR with TOPBP1, a critical activator (Lee et al., 2007). The recruitment of TOPBP1 is thought to be coordinated by the MRN complex (Duursma et al., 2013).

TOPBP1 function is conserved in both fission and budding yeast, where TOPBP1 orthologues (Rad4/Cut5 and Dpb11, respectively) are also essential for efficient CHK1 activation (Garcia et al., 2005). Some evidence suggests that TOPBP1 binds primarily to a region within ATRIP (301-338 amino acids), but also to ATR. The TOPBP1-interacting region in ATR is present in its C-terminus (2483-2597 amino acids) and has been termed as the PIKK regulatory domain (PRD) (Burrows and Elledge, 2008; Mordes et al., 2008). Although TOPBP1 plays a crucial role on ATR activation, which is required for further downstream ATR-dependent phosphorylation events, it does not seem to regulate ATR recruitment to ssDNA (Liu et al., 2006).

1.3.1.2 Later events in the ATR signaling pathway: Start of the phosphorylation cascade and regulation of cell cycle progression

TOPBP1 stimulates ATR kinase activity, promoting the initiation of the transduction cascade signal. The best-studied effector of the ATR-dependent signaling pathway is the serine-threonine kinase CHK1 (Cimprich and Cortez, 2008; Lupardus et al., 2002; Zou and Elledge, 2003), which is activated by phosphorylation on S317, S345 and S366 (Walker et al., 2009), within its C-terminal regulatory domain. These phosphorylation events are crucial for CHK1 activation (Zhao and Piwnica-Worms, 2001), but may also be important for other functions, such as stimulating CHK1 dissociation from chromatin (Smits et al., 2006) or limiting its export from the nucleus (Jiang et al., 2003). Phospho-specific antibodies recognizing phospho-serine 345 reveal that CHK1 is dramatically phosphorylated on S345 in response to HU and UV light (Zhou and Elledge, 2000)

These phosphorylation events are mediated by CLASPIN, a protein that is recruited to ATR-phosphorylated RAD17 subunit of the RAD17-RFC complex, and this interaction is important for sustaining CHK1 phosphorylation (Wang et al., 2006). CLASPIN was initially identified as an adapter protein involved in the ATR signaling pathway in *Xenopus laevis* (Kumagai and Dunphy, 2000), which has been later shown to mediate CHK1 phosphorylation (Liu et al., 2006). CLASPIN binds directly to CHK1 kinase domain, mediating its recruitment to ATR (Jeong et al., 2003), only to disengage once CHK1 is phosphorylated and activated. However, more recent evidence confirms that CHK1 activation is not only dependent on CLASPIN (Petermann et al., 2008), but it could also rely on other mediator/adaptor proteins such as MDC1 or BRCA1 (Stewart et al., 2003; Yarden et al., 2002) or other new DDR players such as TIM/TIPIN (Smith et al., 2009). Deficiency in these proteins impacts in CHK1 phosphorylation, although the mechanisms behind this process are still unclear. Regulation of the CLASPIN-dependent pathway relies on CDC7 (Kim et al., 2008; Rainey et al., 2013), a kinase vital for in initiation of DNA replication.

Once activated, CHK1 is released from chromatin and phosphorylates many substrates. As a result, cell cycle progression is delayed and stalled replication forks are stabilized and consequently re-started in order to complete chromosome replication (Ciccia and Elledge, 2010). The mechanism by which CHK1 prevents inappropriate cell cycle progression is well understood. CHK1 phosphorylates the CDC25 family, promoting their ubiquitination and degradation by the proteasome (Mailand et al., 2000). This stops CDC25A from dephosphorylating and activating its Cyclin-CDK substrates (Bartek et al., 2004).

Another CHK1 function is to control replication origin firing (Feijoo et al., 2001; Syljuåsen et al., 2005). It is known that the ATR-CHK1 mediated response to ssDNA feeds back to RPA to prevent its exhaustion. In the absence of ATR, ssDNA accumulates until sequestering all available RPA, resulting in the generation of unprotected ssDNA and its rapid conversion to DSBs. Therefore, propagation of CHK1, which diffuses away to suppress dormant origins, is essential to ensure all existent ssDNA is shielded by RPA (Toledo et al., 2013).

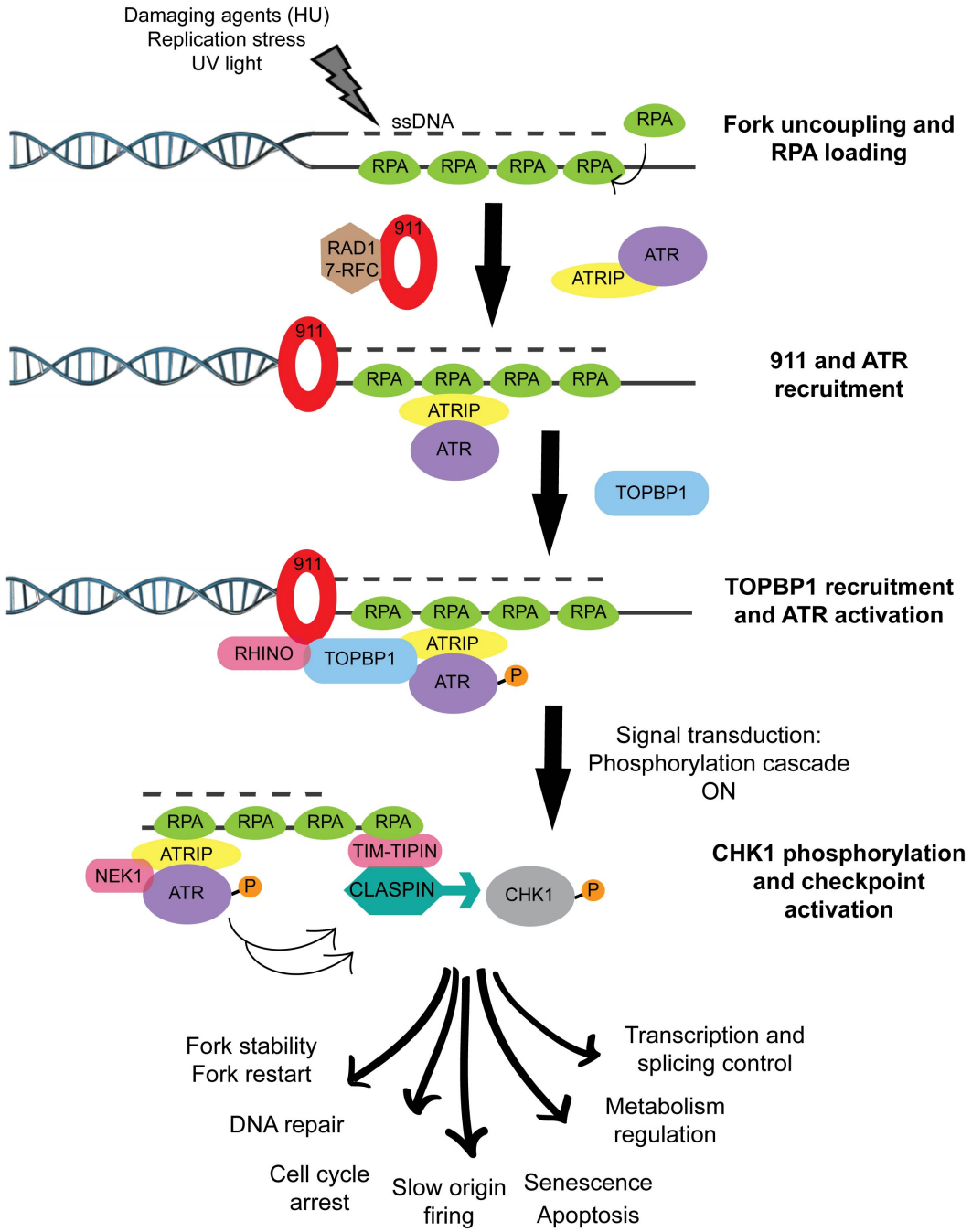


Figure 5. ATR recruitment and functionality in the DDR, adapted from (Cimprich and Cortez, 2008). ATR activation leads to the effector kinase CHK1 phosphorylation that promotes DNA repair, controls cell cycle progression, regulates transcription and determines cell fate (recovery, senescence or cell death).

The ATR-CHK1 pathway is essential to control the firing of replication origins, the repair of damaged replication forks and to prevent the premature onset of mitosis (Brown et al., 2014; Cimprich and Cortez, 2008; Eykelenboom et al., 2013). CHK1 and CHK2 (another effector protein that is activated in response to

DSBs), have also been associated with repair by homologous recombination. They phosphorylate RAD51, and promote its dissociation from BRCA2 and DNA (Bahassi et al., 2008; Sørensen et al., 2005), allowing the initiation of this repair pathway.

1.3.1.3 Emerging players in the ATR-CHK1 signaling pathway

Although the ATR-mediated response to DNA damage has been studied in detail and it is a well-understood process, there are many new players that are added to the list every year. These new and less well-known players come from very different backgrounds, linking different cell functions, such as proteins involved in the cell cycle control or the circadian clock and DNA helicases, to the DNA damage response. Some examples of these new DDR proteins are described below.

- **RecQ helicases**

The RecQ family of helicases, so-called 'genome caretakers', are known to prevent genomic instability. In human cells there are up to five RecQ helicases, named RECQ1, BLM (Bloom Syndrome RecQ Helicase-Like; also RECQ2), WRN (Werner Syndrome RecQ Helicase-Like; also RECQ3), RECQ4 and RECQ5. Mutations in at least three of the five members (*BLM*, *WRN* and *RECQ4* genes) give rise to autosomal recessive disorders associated with cancer predisposition and premature ageing (Chu and Hickson, 2009).

The high incidence of cancer and chromosomal abnormalities found in patients carrying mutations in *BLM* and *WRN* genes could be due to their role in replication regulation and the ATR-CHK1 pathway. The Hickson lab has reported a role for BLM in fork restart and suppression of origin firing upon replication stress. They have shown that phosphorylation of BLM by ATR at T99 is essential to support these functions (Davies et al., 2007). The WRN helicase has also been associated with ATR. There is growing evidence that WRN works together with ATR to suppress fragile site instability (Pirzio et al., 2008) and regulates the ATR–CHK1-induced S-phase checkpoint (Patro et al., 2011). This occurs through

its interaction with the 9-1-1 complex, in particular with the RAD1 subunit (Pichierri et al., 2011).

- **RHINO**

Recently, RHINO was identified as a new player of the ATR-CHK1 signaling pathway. The *RHINO* gene seems to be only present in vertebrate genomes, suggesting it may be a unique regulatory element in higher eukaryotes.
It has been shown that RHINO binds to both 9-1-1 and TOPBP1 and that its knockdown partially reduces CHK1 activation (Cotta-Ramusino et al., 2011), which occurs both when cells are exposed to IR, but also in unperturbed conditions. This suggests that RHINO may bridge the interaction between 9-1-1 and TOPBP1, allowing the activation of ATR. In particular, RHINO forms a stoichiometric complex with the 9-1-1 clamp, now a heterotetrameric complex (Lindsey-Boltz et al., 2015). RHINO is predicted to interact with RAD9, facilitating the loading of TOPBP1 and ATR activation.

- **TIM/TIPIN**

The Timeless protein (TIM), which is well known for its functions in the circadian clock (Barnes et al., 2003), also plays a role in the ATR-dependent checkpoint response (Ünsal-Kaçmaz et al., 2005). TIM forms a tight complex with a smaller protein called TIPIN (Timeless Interacting Protein) and, together, they promote ATR-mediated CHK1 phosphorylation (Kemp et al., 2010; Smith et al., 2009; Ünsal-Kaçmaz et al., 2007). In particular, the TIPIN subunit interacts with RPA2, stabilizing both the TIM-TIPIN complex but also CLASPIN on RPA-coated ssDNA (Kemp et al., 2010). This shows how RPA, TIM-TIPIN and CLASPIN work in the same pathway to facilitate CHK1 phosphorylation at sites of damage. This brings more light into how ATR interacts with its substrates, in this case CHK1. This interaction might be transient to allow a quick release of CHK1 and promote its nuclear functions.

Apart from linking TIM-TIPIN to the ATR-CHK1 DNA damage responses, this complex is also part of the replisome and is required for continued DNA synthesis and stabilization of replication forks in unperturbed cells (Smith et

al., 2009). It is possible that TIPIN is responsible for slowing down fork progression, while TIM plays an inhibitory effect on its partner to allow fork restart, in the absence of replication problems (Ünsal-Kaçmaz et al., 2007).

- **NEK1**

In fungi, the Never-In-Mitosis A (NIMA) kinase functions in cell cycle control and checkpoint regulation (Osmani et al., 1988). In mammals, there are 11 NIMA homolog proteins known as NEKs (NIMA-related kinases), the function of which is not very well understood. In 2008, the function of NEK1 was briefly described in relation to the DNA damage response and checkpoint control (Chen et al., 2008). NEK1 has been found to form foci after IR, co-localizing with γH2AX and MDC1. Furthermore, NEK1 defective cells are unable to activate CHK1 and CHK2. As a result of the defective checkpoint response, cells fail to repair DNA breaks and ultimately develop chromosome instability (Chen et al., 2011). In a more recent publication, Zou's group has narrowed down NEK1's interplay in the DNA damage response. They have demonstrated that NEK1 is necessary to maintain ATRIP levels and stabilize the ATR-ATRIP complex before damage, to allow a robust ATR-dependent response to DNA damage. NEK1 associates directly with ATR-ATRIP and also regulates the phosphorylation events in the ATR cascade, including ATR autophosphorylation on residue T1989 (Liu et al., 2013).

1.3.2 Common structural features within the PIKK family

In human cells, the PIKK family includes not only ATR and ATM, but also DNA-PKcs (DNA-dependent Protein Kinase catalytic subunit), mTOR (mammalian Target Of Rapamycin), SMG-1 (Suppressor of Morphogenesis in Genitalia) and TRRAP (Transformation/Transcription domain Associated Protein). PIKKs are unusually large proteins (~300-500kDa) that share common structural features (Figure 6). These proteins contain a conserved kinase domain at their carboxyl-terminus, which shows significant amino acid sequence homology to the catalytic domains of the phosphoinositide 3-kinases (PI3-K) (Abraham, 2004). The catalytic domain of the PIKKs accounts for only 5-10% of total

sequence (Perry and Kleckner, 2003), but it is essential for their function. Some studies have also shown that overexpression of kinase dead *ATR* mutants causes sensitivity to DNA-damaging agents as well as checkpoint defects (Cliby et al., 1998; Paciotti et al., 2001). This probably occurs when the kinase inactive version of ATR, which might still be able to interact with ATR partners, competes with WT ATR for substrates and impedes proper DDR signaling (dominant negative effect).

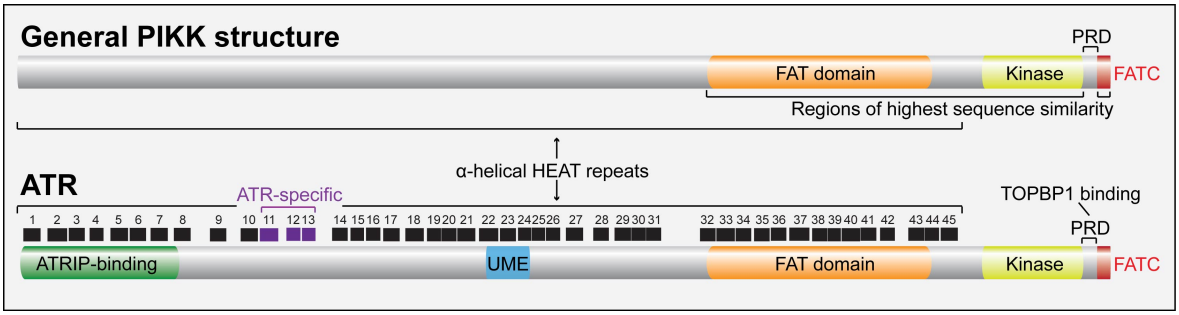


Figure 6. Functional domains of PIKKs, adapted from (Cimprich and Cortez, 2008). The architecture of PIKKs is conserved within the different members of the family. Distinct domains can be distinguished in their configuration. The catalytic or kinase domain (in yellow) is located at the C-terminus and is flanked by two regions of high sequence similarity named the FAT (in orange) and FATC (in red) domains. Furthermore, the PIK regulatory domain (PRD) between the kinase and FATC domains is poorly conserved between family members but highly conserved within orthologues in different organisms. This region is not essential for basal kinase activity but is a regulatory domain in ATR, mediating its interaction with TOPBP1. The N-terminus is composed of helical-rich regions or HEAT repeats (shown as 45 black boxes for ATR), which mediate interactions with the protein cofactors, such as ATRIP for ATR (in green). Characteristic features of the ATR subfamily members, such as the ATR-specific HEAT repeats (in purple) or the UME domain (in blue), are also indicated.

Two conserved domains termed FAT and FATC flank the kinase domain. The FAT domain was identified by Bosotti and colleagues as a conserved domain spanning ~500 amino acids N-terminal to the kinase domain of three main protein subfamilies, whose representatives give the name to the domain: FRAP (alternative name for mTOR), ATM, and TRRAP (Bosotti et al., 2000). At the other end of the kinase domain of PIK kinases, lies the FAT-C domain, which is a smaller motif (~35 amino acids) that is also found in these subfamilies. The FAT and FATC domains only appear in combination, suggesting they could interact with each other. It is possible that together they conform a structure that allows for proper function of the kinase domain. Since the FATC might be too small to

fold independently but it is more conserved than the FAT domain (34% versus 16% average identity) (Bosotti et al., 2000), it may have a major role for catalytic activity compared to the FAT domain. Although the function of the FAT domain is still not well understood, it is thought that it may have some structural role, either as a scaffold or mediating protein-protein interactions. In a recent crystallography study, it has been shown that the FAT domain of mTOR encircles its kinase domain and interacts directly with it through conserved residues, potentially regulating its function (Yang et al., 2013). Between the kinase domain and the FATC domain there is a short region that is involved in regulation of the kinase activity (PRD: PIK regulatory domain). This region mediates the interaction between ATR and its crucial activator TOPBP1 (Mordes et al., 2008).

At their N-terminal region, PIKKs have no apparent homology and differ significantly in size and sequence. However, deeper bioinformatic analyses have shown that the non-kinase portions of PIKK proteins are formed almost entirely of HEAT repeats (Baretić and Williams, 2014; Perry and Kleckner, 2003), named after the four proteins in where these repeats were initially found: Huntingtin, Elongation factor 3, A subunit of protein phosphatase 2A, and TOR1 (Andrade and Bork, 1995). About 85% of ATR's structure is made up of 45 HEAT repeats, seven of which comprise the ATRIP interacting region located at the amino-terminus of the protein (Ball et al., 2007). However, the remaining 38 HEAT repeats are uncharacterised. Interestingly, these include three repeats that are specific to ATR (HEAT repeats 11-13) (Perry and Kleckner, 2003) and the region that constitutes the UME (UVSB PI3K, MEI-41, ESR1) domain (1141-1247 amino acids) found in all reported ATR homologues and which is predicted to be required for protein-protein interactions (Rounds and Larsen, 2008). The relevance of these HEAT repeats is further discussed in section 3.1.1.1.

1.3.2.1 HEAT repeats

A single HEAT repeat unit is a pair of anti-parallel helices linked by a flexible 'loop' (Figure 7A). HEAT repeats often appear in proteins in repetitive series (Figure 7B), with adjacent HEAT units connected by flexible inter-unit loops (Andrade et al., 2001a). The helical regions are usually 10-20 residues long, while intra-unit loops tend to be 5-8 residues in size and connected to a

neighbouring unit by a ≥ 1 residue corresponding to the inter-unit loop (Perry and Kleckner, 2003).

HEAT repeats can be quite diverse in length and amino acid composition (Andrade et al., 2001a, 2001b), which often hinders the identification of these structural blocks. However, although the sequence similarities among individual repeats are low, a specific pattern of hydrophobic and hydrophilic amino acids can be detected, so a common helical architecture can be predicted (Kobe et al., 1999). Most of the HEAT repeat units present in PIKKs are shared amongst the different subfamilies; however, there are a few key units that are subfamily-specific. Interestingly, as studied by Perry and Kleckner, some of the HEAT repeats only exist in particular members of a subfamily (ATRs, ATMs, or TORs) but are absent in the members of the other subfamilies (Perry and Kleckner, 2003). These differences could possibly reflect the disparities in PIKK functions.

In HEAT-rich proteins such as PIKKs, these motifs occur in series with adjacent units piling together to form elongated helical structures also referred to as ‘solenoids’ or ’rings’. When visualised by crystallography, HEAT-rich proteins form super-helical scaffold matrices often when engaged with other macromolecules (Chook and Blobel, 1999; Cingolani et al., 1999; Groves and Barford, 1999; Groves et al., 1999).

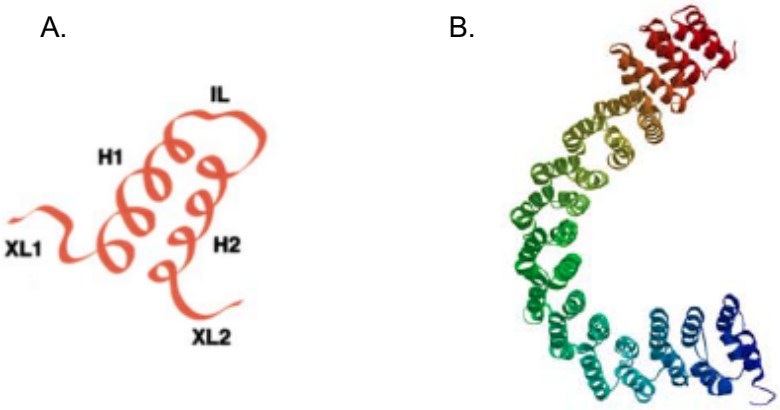


Figure 7. Structural features of HEAT repeats and HEAT repeat-containing proteins.

A. General structure of a HEAT repeat unit: XL, external or interunit loop; IL, intraunit loop; H, helix (Perry and Kleckner, 2003).

B. Example of a HEAT repeat containing protein: Structure of Phosphatase 2 A-subunit, composed of 15 HEAT repeats (Groves et al., 1999; Kobe et al., 1999).

1.3.2.2 Three-dimensional structure of PIKKs

The three-dimensional (3D) structure of several PIKK members has been solved by electron microscopy and/or crystallography analyses (Arias-Palomo et al., 2011; Chiu et al., 1998; Llorca et al., 2003; Sibanda et al., 2010; Yang et al., 2013). As these proteins share analogous structural features to ATR, this structural information may be helpful to consider for this study.

Structural analyses have revealed DNA-PKcs, ATM, SMG-1 and mTOR 3D configurations are similar, reflecting the probable structural homology within the PIKK family members (Arias-Palomo et al., 2011; Boskovic et al., 2003; Chiu et al., 1998; Leuther et al., 1999; Llorca et al., 2003; Sibanda et al., 2010; Spagnolo et al., 2006; Yang et al., 2013). Schematic models based on those findings are shown in Figure 8A to highlight the structural homology between PIKKs. Early structural studies showed that each of these proteins are comprised of two main regions: (1) a 'head-like' domain where the kinase domain is located and from which extensions corresponding to the FAT and FATC domains project; (2) a curved and flattened tubular 'arm-like' domain, associated to the less conserved HEAT rich N-terminus of the proteins (Llorca et al., 2003; Rivera-Calzada et al., 2005; Spagnolo et al., 2006). In later studies, the crystal structure of DNA-PKcs with a resolution of 6.6Å was solved, demonstrating for the first time that the 'arm-like' domain is shaped as a ring-like structure (Sibanda et al., 2010) (Figure 8B). Recently, part of the mTOR structure was also crystallized with a higher resolution (3.2 Å), showing that the FAT domain wraps around the kinase domain. This interaction occurs through conserved residues among PIKK family members (Yang et al., 2013).

The major differences amongst PIKKs are expected to be present in the HEAT-rich arm, due to the lack of homology between the different proteins at this region. As indicated above, the lack of sequence homology between HEAT repeat containing regions, does not necessarily mean lack of HEAT-unit homology. Therefore, the 'arm-like' domain of PIKKs would be the region to be considered as the least conserved in amino acidic sequence, however, this region could possibly mediate the interaction of PIKKs with different proteins (Baretić and Williams, 2014) leading to the sequence diversity. As some PIKKs have been reported to form dimers to regulate their activation (Bakkenist and Kastan, 2003; Boskovic et al., 2003; Spagnolo et al., 2006), it could be speculated that this event

could also be mediated by the 'arm-like' domain of PIKKs (Rivera-calzada et al., 2015).

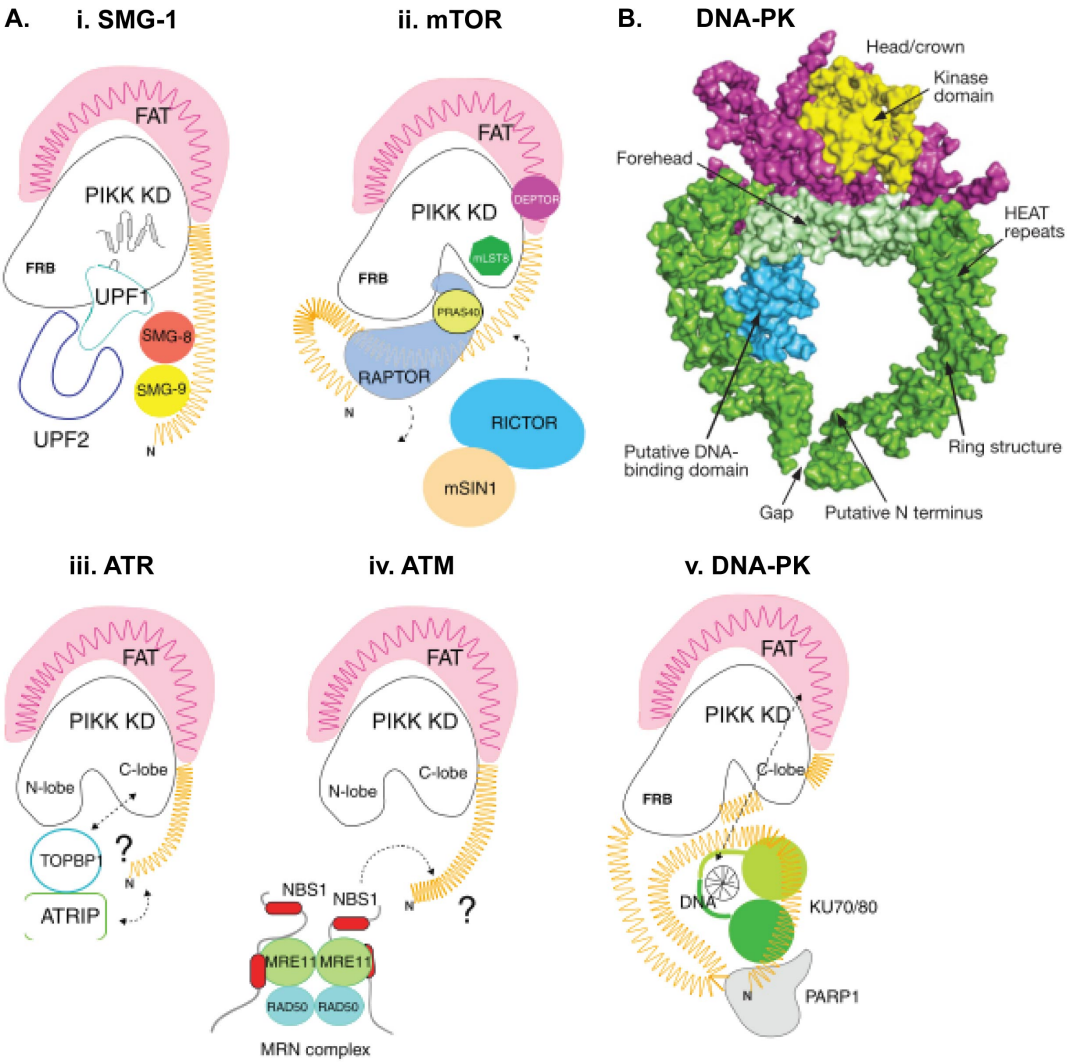


Figure 8. Common structural features within the PIKK family.

A. Schematic models of PIKK structure (Baretić and Williams, 2014). The FAT domain (in pink) and 'arm-like' domain (in yellow), which are made up of HEAT repeats, are represented as spirals; while the kinase domain is represented as a globular silhouette. The N-terminus of each protein is indicated (N), and interactions with associated regulatory partners and known interactors are suggested (dashed lines). The question marks indicate unknown geometries of the solenoid with respect to the kinase domains of ATR and ATM.

B. Crystal structure of DNA-PK, resolved by Sibanda and colleagues (Sibanda et al., 2010). The different parts of the molecules are colour-coded, as indicated.

These studies assist us in the understanding of some aspects of the general organization of the PIKK family. In summary, it is now known that PIKK enzymes have a conserved core, consisting of a kinase domain encircled by the FAT domain, plus a more divergent helical arm that may mediate protein-protein interactions. It is expected that the ATR kinase share similar three-dimensional features and also similar DNA recruitment kinetics. It could be possible that some of the ATR HEAT repeats might be essential for recognising and binding to the damaged DNA, while other HEAT repeats could be required for regulating the kinase activity of ATR or the interaction with other key proteins of the DDR pathway.

1.3.3 Phosphorylation sites in ATR

As phosphorylation events are crucial for PIKK activation and cellular signaling in response to DNA damage (Bensimon et al., 2011), it is important to identify putative phosphorylation sites in ATR. Most characterized PIKK phosphorylation sites are serine or threonine followed by glutamine (SQ/TQ) (Kim et al., 1999; Lees-Miller et al., 1992). ATR has 19 putative SQ/TQ sites; but only 16 of them are conserved in mouse and several are also conserved in *Xenopus laevis* and *S. cerevisae* (Nam et al., 2011a). These residues could be targets for ATR autophosphorylation or other PIKKs phosphorylation. ATR also contains other non-SQ/TQ sites that have been shown to be important for its functions (Figure 9).

Many proteomic studies have looked at the phosphorylation of different proteins across the cell cycle (Daub et al., 2008; Dephoure et al., 2008), but also at particular interactors and phosphorylation events of DDR factors (Matsuoka et al., 2007). Some *in silico* tools have also been developed to identify novel ATM and ATR targets by the presence of SQ/TQ motifs (Cara et al., 2014). Abnormal activation of kinase signaling pathways is commonly associated with numerous syndromes, thus the study of the cell phosphoproteome will be key to better understanding cancer and other diseases (Harsha and Pandey, 2010).

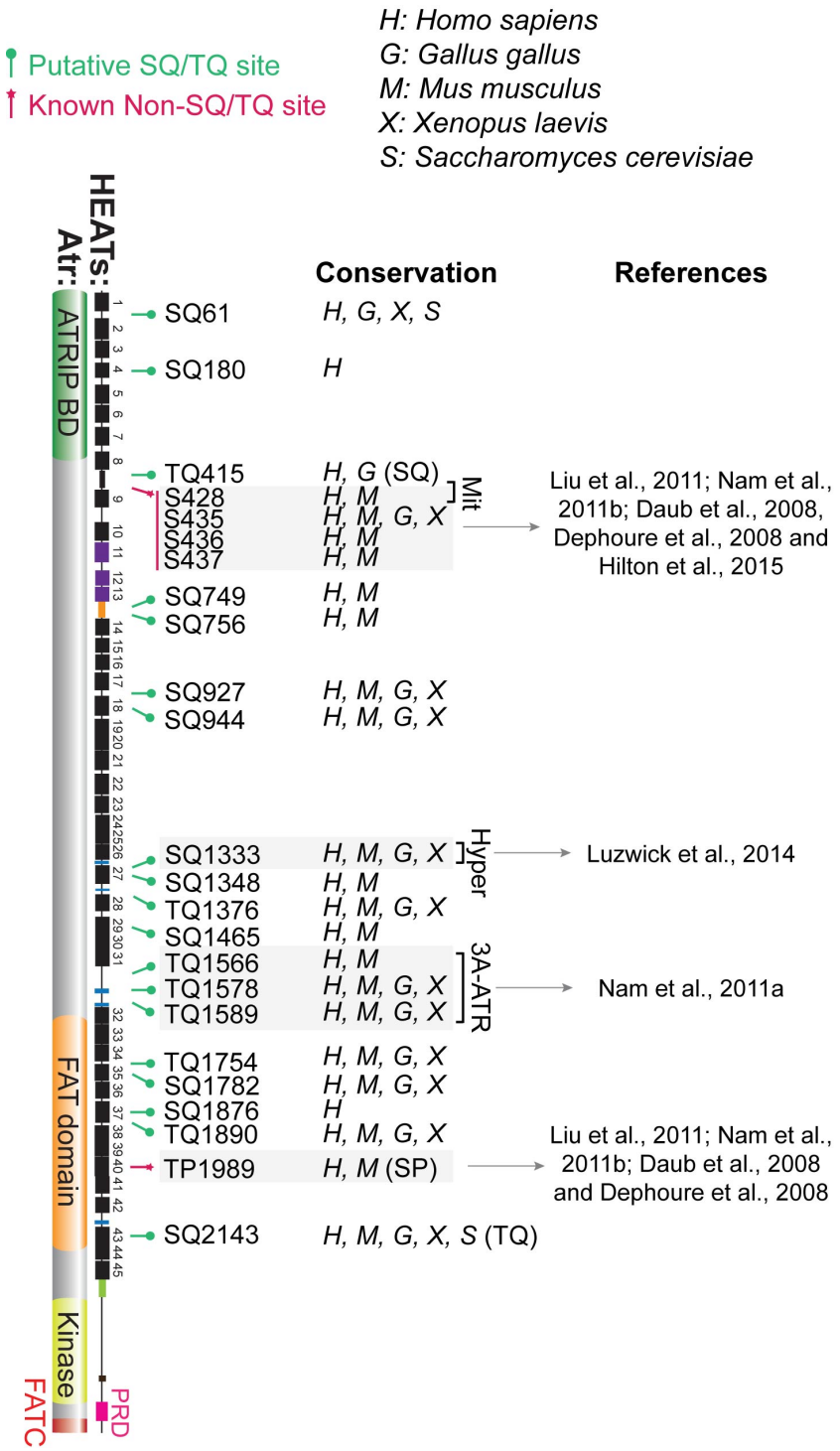


Figure 9. Putative SQ/TQ and known phosphorylation sites in ATR, adapted from (Nam et al., 2011a). Schematic shows the 19 putative SQ/TQ residues (in green) and other known phosphorylation sites (in pink) in human ATR (H). Their level of conservation amongst other species, including mouse (M), chicken (G), frog (X) and budding yeast (S), is also indicated. References for known phosphorylated residues are included, as well as for S1333, mutation of which results in hyperactivation of the ATR kinase (Hyper); the T1566/T1578/T1589 cluster, which mutation into A results in the separation of ATR functions (3A-ATR); and S428, which phosphorylation is crucial for a recently described function of ATR in the mitochondria (Mit). All of which are described in more detail in the sections below.

1.3.3.1 T1989 phosphorylation

Phosphorylation of residue T1989 has previously been identified as a crucial event in the activation of ATR simultaneously by two independent groups (Liu et al., 2011; Nam et al., 2011b). Phosphoproteomic studies, such as the Elledge screen, have also identified this threonine as a phosphorylated residue (Daub et al., 2008; Dephoure et al., 2008). However, T1989 does not fit this consensus (TP) and it is also not a conserved residue. T1989 is only conserved in primates, changed to a serine in mouse, and not conserved in frogs, flies (Nam et al., 2011b) or chicken, which is one of the model systems that we are using in this study.

Phosphorylation of this residue is thought to rely on ATR kinase activity (autophosphorylation). A kinase dead ATR mutant does not exhibit phosphorylation of T1989, while specific inhibition of ATM and DNA-PK do not affect this event (Liu et al., 2011; Nam et al., 2011b). Besides, phosphorylation of T1989 is dependent on RPA, ATRIP and ATR activation (Liu et al., 2011; Nam et al., 2011b). However, it is not very clear in the field how this event occurs. Nam and colleagues propose that phosphorylation of T1989 is dependent on ATR interaction with TOPBP1. They have shown that a mutant that is unable to bind TOPBP1 (K2589E), and therefore is unable to complete ATR activation, is not phosphorylated on residue T1989 in response to HU (Nam et al., 2011b). In contrast, Zou's research group suggest that T1989 phosphorylation is previous to TOPBP1 recruitment and that phosphorylation of this residue is important for the interaction between ATR and TOPBP1 after DNA damage (Liu et al., 2011).

Following Liu and colleagues' work, it is possible that ATR autophosphorylation on residue T1989 occurs after ATR-ATRIP complex binds to ssDNA coated with RPA. This would allow TOPBP1 to be recruited and mediate full activation of ATR and subsequent phosphorylation events. This hypothesis supports the notion that ATR might interact with TOPBP1 through both T1989 and the PRD. Alternatively, based on research by Nam *et al.*, it is possible that the initial interaction between TOPBP1 and ATR occurs through the PRD, allowing ATR activation and autophosphorylation in residue T1989. This phosphorylation event would be the start of the signaling cascade and would serve as a platform for TOPBP1 to bind and stabilize its union with ATR. Further experiments need

to be carried out to clarify whether T1989 phosphorylation occurs before or after the recruitment of TOPBP1.

1.3.3.2 T1566/T1578/T1589 regulatory cluster

Cortez's group has briefly looked at the importance of all 16 conserved SQ/TQ sites in ATR. Mutation of all 16 residues to an alanine in a single cDNA, resulted in separation of ATR functions. The 16A-ATR mutant was able to maintain kinase and G2 checkpoint activities, but it failed to support cell viability and to promote replication recovery from a transient exposure to replication stress (S phase defects) (Nam et al., 2011a).

Nam and colleagues showed that mutations responsible for these defects were T1566A/T1578A/T1589A (3A-ATR mutant); however, these residues do not seem to be phosphorylated. They could possibly have a regulatory function. They are located in a region that is conserved among PIKKs (ATM and DNA-PK) and could potentially regulate phosphorylation of other residues, such as T1989 (Nam et al., 2011a).

1.3.3.3 S1333

During the analysis of putative ATR phosphorylation sites, the Cortez group found another SQ site (S1333) that seemed to be important for modulating ATR kinase activity (Luzwick et al., 2014). Although there is no evidence for phosphorylation of S1333 (Luzwick et al., 2014; Nam et al., 2011a), mutation of this residue results in a hyper active kinase. Cells expressing ATR^{S1333A} showed higher basal phosphorylation levels of ATR substrates but no difference was seen with regards to checkpoint or replication functions. These results suggest that residue S1333 could be important for regulating the kinase domain. Since S1333 is located in HEAT repeat 27, regulation could be direct or perhaps due to structural HEAT-repeat modifications, which could affect ATR folding.

1.3.3.4 Other phosphorylated residues

Apart from T1989, ATR can be phosphorylated at residues S428, S435 and possibly S436 and S437 (Liu et al., 2011). A phospho-specific antibody to ATR S428 is available commercially and, although it has been suggested that it may not be a useful marker for ATR activity (Nam et al., 2011b), phospho-S428 has been recently linked to a previously unknown function of ATR at mitochondria (further information in section 1.3.7.2) (Hilton et al., 2015). As for T1989, phosphorylation of S435 has been also reported in large-scale phospho-proteomic studies (Daub et al., 2008; Dephoure et al., 2008). Contrary to the T1989A mutant, neither S428A nor S435/436/437A failed to activate CHK1 (Liu et al., 2011), suggesting they are not crucial residues for ATR-CHK1 signaling.

Further studies regarding these and other potential ATR phosphorylation sites are necessary to better understand and dissect how ATR activation and interactions occur.

1.3.4 ATR functions: essential vs. checkpoint

PIKKs play essential roles in the repair of DNA damage by regulating the activity of proteins that participate in the repair process. However, loss of ATR functions also leads to defects during the unchallenged life of the cell, indicating this protein is required even in the absence of any exogenous damage (Kato and Ogawa, 1994).

It is well known that Mec1, the ATR homologue in budding yeast, is a key checkpoint regulator, which plays a critical role in the maintenance of genomic stability (Carballo and Cha, 2007). In the presence of DNA damage or replication blocks, Mec1 is required for arrest or slow cell cycle progression. At the same time, during the normal cell cycle, Mec1 has an essential role in removing the inhibitory effect of Sml1 on the ribonucleotide reductase (RNR), which controls the dNTP pools, to facilitate DNA replication (Zhao et al., 1998).

According to the Mec1 essential role, it has been suggested that ATR also has an essential role in coordinating DNA replication (Zhou and Elledge, 2000). For instance, it has been shown in our laboratory by Eykelenboom *et al.* that ATR is involved in modulating replication during unperturbed S phase. Cells lacking ATR have a defective replication, showing slower fork progression and higher

origin firing to compensate for the decreased replication rate. Furthermore, in the absence of ATR, cells accumulate several defects, such as chromosome breaks and gaps, anaphase bridges or micronuclei, which culminate in eventual cell death. This occurs as a consequence of their inability to activate the G2/M checkpoint and therefore undergo mitosis in the presence of un-replicated DNA (Eykelenboom et al., 2013). Other researchers have also observed consistent roles at replication forks for both ATR and Mec1 during unperturbed cellular proliferation (Brown et al., 2014; Cha and Kleckner, 2002; Couch et al., 2013; Labib and De Piccoli, 2011).

It was not until recently that it was confirmed that ATR has also a conserved function in RNR regulation. Two separate studies have come to the conclusion that RRM2 function, a regulatory subunit of the RNR, is regulated by ATR (Buisson et al., 2015; Lopez-Contreras et al., 2015). This brings light into the previous findings, explaining the mechanism behind ATR's role in replication control. Fernandez-Capetillo's group found out that RRM2 upregulation results in the reduction of chromosome breakage in response to ATR inhibition. Besides, increased levels of RRM2 also prolong the life span of Seckel mice, which show reduced levels of ATR expression (Lopez-Contreras et al., 2015). In parallel, Zou's group also pointed out that ATR has a crucial role in S phase by coordinating RRM2 accumulation and origin firing (Buisson et al., 2015).

Together, these findings suggest that one of the essential roles of ATR is to regulate nucleotide pools and sustain a well-timed cell cycle to prevent premature entry into mitosis, ensuring the maintenance of genomic stability. Separation of function studies, described in detail in next section, have shown dissociation of G2 and S phase ATR functions, which confirms ATR's role in replication control.

1.3.5 Reported *ATR* mutations

For the development of this project, it is important to understand the different *ATR* mutations described in literature that are related to human syndromes. But it is also critical to be aware of the few studies that have shown the separation of ATR functions. Below is a brief description of the reports that we considered more relevant for our purposes.

1.3.5.1 Separation of function mutations in *ATR/Mec1*

In 2001, the Longhese lab was showed that Mec1 S and G2 phase functions could be dissociated (Paciotti et al., 2001). They used random mutagenesis in budding yeast to generate the *mec1-100* and *mec1-101* alleles. These mutants have an intact G2 /M checkpoint in response to UV light, whereas they are unable to activate the G1 /S and intra-S DNA damage checkpoints. Although neither of these mutants showed hypersensitivity to MMS and UV radiation, only the *mec1-100* mutant lost viability upon HU treatment. This behavior can be explained by comparing the Mec1, Rad3 and ATR sequences. The point mutations present in the *mec1-101* allele are not conserved amongst ATR homologues, while amino acid changes in the *mec1-100* allele affect residues that are identical in Mec1 and Rad3 and are located in a regions that is also well conserved in human ATR (Paciotti et al., 2001). The equivalent mutations in ATR would affect residues within HEAT repeat 28 (L1405) and the linker between 38 and 39 (A1934). These results highlight key residues in Mec1/ATR, which could be important for its S phase functions and the response to replication fork stalling.

As explained in section 1.3.3, Nam and colleagues were also able to identify separation of function mutations in *ATR* using higher cells, which confirm the yeast results (Nam et al., 2011a). They focused their studies on the 16 conserved SQ/TQ residues present in human ATR and found an ATR regulatory cluster, when they mutated all these residues to alanine (A) in a single cDNA. The 16A-ATR mutant was proficient in kinase and G2 checkpoint activities, but it was unable to maintain cell viability or stimulate replication recovery after a transient replication block. This mutant also had reduced ability to phosphorylate CHK1 in response to HU.

The sites responsible for these defects were narrowed down to three: T1566A/T1578A/ T1589A (3A-ATR mutant). These three threonines are located in the linker between HEAT repeats 31 and 32 of ATR, which aligns with autophosphorylation regulatory domains of ATM and DNA-PK, also present in long inter-HEAT repeat loops. Although ATR does not seem to be autophosphorylated in this region *in vitro*, these similarities among PIKKs suggest that these uncommonly long loops might constitute an important regulatory domain involved in monitoring PIKK function (Nam et al., 2011a).

Both yeast and human cell studies show a separation of G2 and S phase roles, confirming the hypothesis that the essential function of ATR might be to regulate replication.

Other studies in fruit fly have also reported separation of function alleles in *Mei-41*, the *Drosophila* ATR/ATM homologue. These analyses show that two *Mei-41* null alleles, which contain early truncations in the protein, are able to support growth, although show multiple severe defects. Mutation of *Mei-41* leads to a number of phenotypes, including sensitivity to MMS, reduction in recombination, increased chromosome loss and checkpoint defects (Laurençon et al., 2003). This study also supports the idea of separate ATR-dependent checkpoint and viability functions. However, it has to be considered that other ATM/ATR homologues in *Drosophila* might compensate for Mei-41 roles; such as CG6535, identified in the Genome Project as a second ATM/ATR homologue more closely related to ATM.

1.3.5.2 *ATR* mutations related to human diseases

Mutations in *ATR* are associated with Seckel syndrome, a human hereditary disorder characterised by premature aging, severe microcephaly and dwarfism. A synonymous mutation (2101A>G) located in exon 9 of *ATR* was identified by O'Driscoll and colleagues to compromise *ATR* splicing. In particular, this silent change is thought to affect a splice regulator site, which results in exon 9 skipping. As a consequence, this mutation, from now on referred to as the classic Seckel mutation, causes a severe reduction of ATR expression, which promotes the development of the syndrome (O'Driscoll et al., 2003). Since 2012, other *ATR* mutations have been reported to cause Seckel syndrome (Mokrani-Benhelli et al., 2013; Ogi et al., 2012), all of which affect splicing resulting in ATR hypomorphism. Detailed explanation of these mutations is enclosed in chapter 4, together with the description of two novel *ATR* missense Seckel mutations. Our results show that, although these mutations cause amino acid changes, they surprisingly also affect *ATR* splicing.

Alterations in the *ATR* gene have also been observed in some cancers (Poehlmann and Roessner, 2010); however, in 2012 a specific mutation in this gene was pointed out as the direct cause of cancer predisposition. Mutation of

residue Q2144 was reported in an American family who suffered from oropharyngeal cancer predisposition (Tanaka et al., 2012). This study shows that the ATR^{Q2144R} mutation compromises ATR activity and leads to the development of this disorder. This glutamine is part of a potential SQ phosphorylation site; however, there is no evidence so far in the literature for phosphorylation of this residue. Detailed analysis of its effect on ATR function is given in chapter 5.

1.3.6 Overview of splicing regulation

The removal of introns during mRNA processing, or splicing, is a very complex and precise process. Thus, it is not surprising that alteration of such a critical aspect of gene expression may lead to human hereditary syndromes or cancer (Faustino and Cooper, 2003; Supek et al., 2014).

Splicing is catalysed by the spliceosome, a large ribonucleoprotein (RNP) complex that is made up of five small nuclear RNAs (snRNAs U1, U2, U4, U5, and U6) and more than 100 core proteins. Additional regulatory proteins contribute to the splicing of particular pre-mRNAs (Zhou et al., 2002).

The splicing reaction takes place in two steps (or trans-esterification reactions), which occur at the exon/intron junctions. These exon/intron borders are defined by rather short and degenerate sequences known as core splicing signals (the 5' splice site, the branch site and the 3' splice site). These signals become bound by the spliceosome components, with the U1 snRNP binding the 5' splice site, and the U2 interacting with the branch point. The processive component of the spliceosome is U6 snRNP, which assembles onto the spliceosome as a complex together with U4 and U5 (Hastings and Krainer, 2001; Madhani and Guthrie, 1994).

As the core splicing signals are very short and diverse, these consensus elements are not sufficient for the splicing machinery to discern between real and pseudo splice sites. Therefore, to increase overall fidelity of the splicing process, additional regulatory elements are present in exons and introns, known as splicing enhancers or silencers (ESE/ISEs; exonic or intronic splicing enhancers and ESS/ISSs; exonic or intronic splicing silencers) (Wang and Burge, 2008).

The best-characterised splicing silencers are recognised by the hnRNP (heterogeneous nuclear ribonucleoprotein) family, which participate in many

aspects of RNA biogenesis, including mRNA splicing. In particular, hnRNP A1 downregulates splicing by indirectly blocking recruitment of the U2 snRNP to the splice sites (Tange et al., 2001). On the other hand, the best-known splicing enhancers are purine-rich sequences that interact with members of the SR (serine-arginine) family. This family of proteins stimulate splicing by recruiting components of the core splicing machinery through their SR-rich domains (Wu and Maniatis, 1993). SR proteins and hnRNPs often have opposing roles within the same splicing regulatory unit. They might compete for common binding sites or limit their access to enhancers and silencers motifs, respectively. The balance between hnRNPs and SR proteins determines the ratio of inclusion/exclusion of exons (Jean-Philippe et al., 2013; Long and Caceres, 2009).

A general overview of the splicing reaction is given on Figure 10, highlighting the role of core signals, spliceosome and additional regulators.

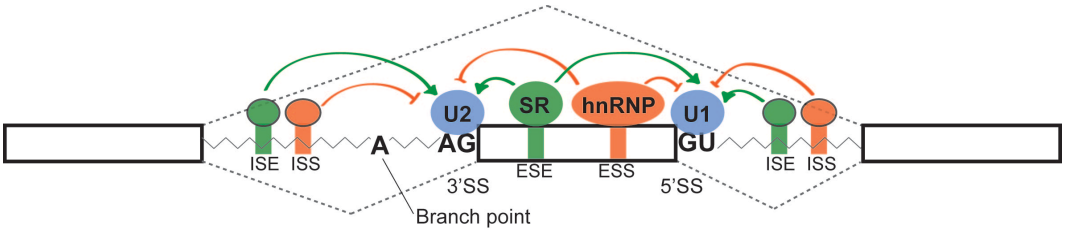


Figure 10. Diagram of splicing reaction, adapted from (Wang and Burge, 2008). Exons are represented as white boxes, while introns are drawn as spiky lines. Consensus sequences for the core splicing signals are shown in the figure. The start of every intron is defined by a GU (5' splice site), while the end is marked by an AG (3' splice site). In the middle of the 5' and 3' splice sites, the branch point (A) marks the location for the first trans-esterification reaction. Spliceosome components (U1 and U2), which bind the core splicing signals, are shown in blue. Enhancer signals (ESE, ISE) and regulators (SR) are marked in green, while silencers (ESS, ISS and hnRNP) are marked in orange. The dotted lines indicate the different splicing possibilities: exon inclusion (bottom) or exclusion (top).

Alteration of any of these splicing modulators can lead to different splicing defects. The outcomes for these defects comprise exon skipping, intron retention, or use of alternative 3' or 5' splice sites (Figure 11) (Cartegni et al., 2002).

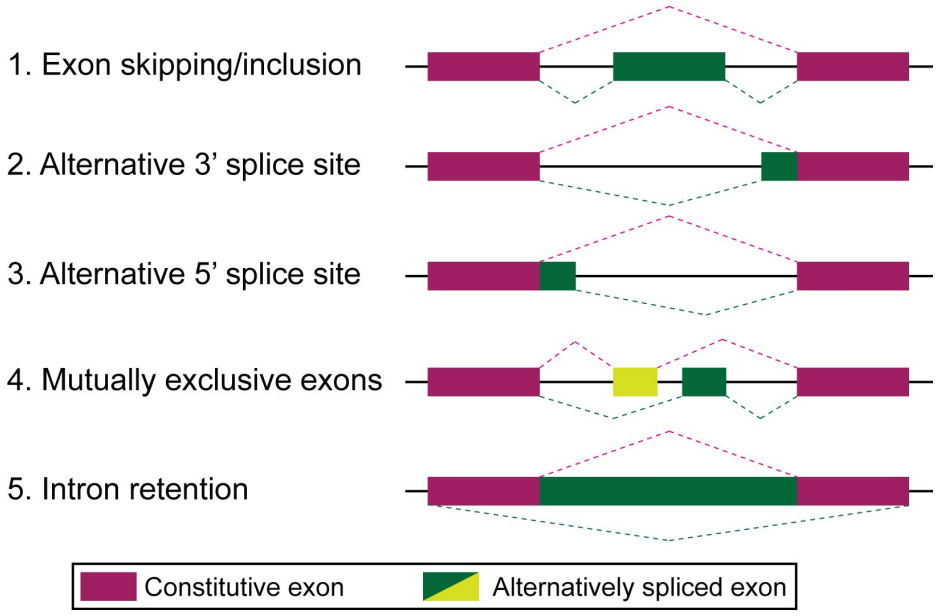


Figure 11. Landscape of splicing possibilities, adapted from (Cartegni et al, 2002). Several genes have different isoforms thanks to alternative splicing. Variations in the splicing reactions can also be the result of splicing mutations. These splicing possibilities/defects are shown in the figure and include exon skipping/inclusion, use of alternative 5'/3' sites, mutually exclusive exons and intron retention. Introns are indicated as black lines, while exons are drawn as purple or green boxes, depending on whether they are expressed constitutively or due to alternative splicing.

1.3.6.1 Splicing mutations and human disease

Disease-causing mutations are often thought to directly affect protein function. However, there is growing evidence that many silent or missense mutations actually affect the splicing pattern of genes instead of protein activity (Baralle and Baralle, 2005; Faustino and Cooper, 2003; Li et al., 2016; Lopez-Bigas et al., 2005). Genetic analyses performed only at the level of exonic DNA sequences are not adequate to detect and characterise these mutations, but instead, further mRNA splicing studies are required for proper diagnosis.

Some disease-causing mutations have been reported to alter splicing by creating novel splice sites or by disrupting regulatory elements such as splicing enhancers or silencers. Splicing mutations may play a more significant role than previously thought in human hereditary disorders (Baralle et al., 2009).

Apart from the mutation that causes classic Seckel syndrome, there is a wide group of mutations described in the literature that also result in defective splicing. Other progeria syndromes, such as Ataxia-Telangiectasia (A-T), are also

caused by defects in splicing. For instance, there is a 4bp deletion in intron 20 of *ATM* that was found in A-T patients. This mutation affects binding of the U1 ribonucleoprotein to the 5' splice site, resulting in incorrect processing of intron 20. As a consequence, part of intron 20 is inserted between exon 20 and 21 in the transcript (phenomenon known as intron retention) (Pagani et al., 2002).

Several splicing mutations have also been reported in the neurofibromin (*NF1*) gene, which encodes for a protein required to form the myelin sheaths or fatty coverings that insulate neurons. Splicing mutations of the *NF1* gene have been identified in different exons, such as exon 3 and 29, which are skipped due to possible alteration of U1 binding sites or other splicing regulators (Baralle et al., 2003; Raponi et al., 2009). These mutations give rise to a genetic disorder called NeuroFibromatosis type 1.

A long list of other syndromes, such as Myotonic Dystrophy, Cystic Fibrosis and Retinitis Pigmentosa, are also known to be caused by defective splicing (Orengo et al., 2011; Pagani et al., 2000; Tanackovic et al., 2011). These occur through disruption of canonical or alternative core splice sites, or alteration of particular components of the splicing machinery (Faustino and Cooper, 2003).

1.3.7 ATR functions outside the DDR

ATR plays a vital role in the regulation of the DNA damage checkpoints. Its nuclear checkpoint functions are well established, but little is known about its functions outside this cell compartment. A recent paper looking into SQ/TQ rich domains has identified potential ATM/ATR targets in cellular pathways, such as vesicle trafficking, actin cytoskeleton and neural development, where a role for ATM/ATR remains unknown or is poorly understood (Cara et al., 2016). This suggests that ATR spectrum of functions could be wider that its known nuclear roles. In recent years, some reports have shown that ATR is also active at the nuclear envelope, mitochondria and cilia.

1.3.7.1 ATR role at the nuclear envelope

In eukaryotes, transcription is coupled with mRNA splicing and export to the cytoplasm (Aguilera, 2005; Köhler and Hurt, 2007). These processes are

coordinated by protein complexes that tether the newly synthesized mRNA to the nuclear pore complex (NPC), which is known as gene gating. Topological barriers might appear at sites where transcribed genes bind to fixed nuclear envelope structures. The approach of replication forks to these regions might contribute to a higher state of DNA supercoiling, ultimately leading to the collision of transcription and replication machinery and impeding fork progression (McGlynn et al., 2012).

Marco Foiani group's findings reveal that the replication checkpoint might have a role in coordinating the transcription and replication processes to avoid fork collapse and reversal. In particular, they suggest that Mec1/ATR might be able to sense fork pausing at transcribed regions and allow fork progression by stimulating the release of transcribed genes from nuclear pores (Figure 12). They have shown that Rad53, the CHK2 homologue and Mec1 substrate, phosphorylates key nucleoporins and counteracts gene gating, counterbalancing the topological stress generated at nuclear pore gated genes (Bermejo et al., 2011).

How ATR/Mec1 are able to sense fork pausing has also been reported in a more recent publication. These proteins are able to sense vibrations derived from chromosome dynamics and torsional stress on the nuclear membrane. In fact, upon mechanical stress, a fraction of ATR, ATRIP and pCHK1 colocalize with Nup153 at the nuclear envelope (Kumar et al., 2014). Interestingly, ATR recruitment to the nuclear membrane is not dependent on other classic DDR players, such as RPA, RAD17 and TOPBP1. Once localized at the nuclear envelope, ATR/Mec1 modulate membrane plasticity and its association with chromatin, enabling cells to cope with this type of stress. HEAT repeats might be the key as to how these big kinases sense mechanical stress, since they can behave as elastic connectors (Grinthal et al., 2010).

These findings support that, apart from the DNA damage response, ATR may play a crucial role in cell communication, by rapidly connecting external signals and epigenetic modifications. This would allow the synchronization of membrane transport and signaling with gene expression.

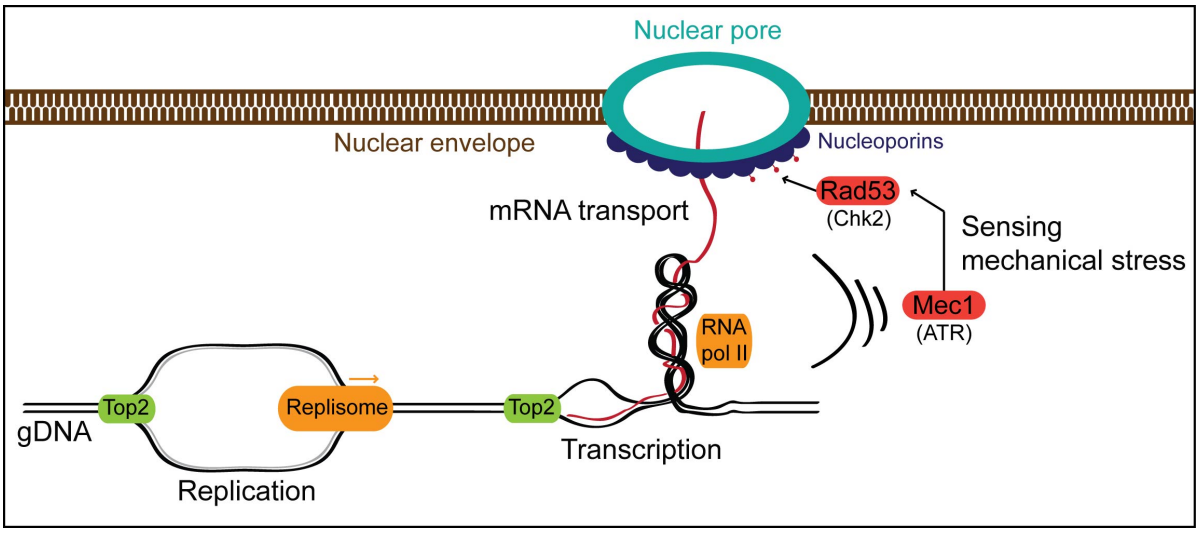


Figure 12. Schematic representation of the proposed topological features of forks meeting nuclear pore-gated transcribed genes, adapted from (Bermejo et al., 2011). In the figure, mRNA is represented in red, while DNA is shown in black. It is speculated that, when the replication forks and gated genes clash, Mec1/ATR is able to sense the resulting vibrations thanks to the elasticity of its HEAT repeats. As a result, Rad53/CHK2 becomes activated in a Mec1/ATR dependent manner, which stimulates the phosphorylation of several nucleoporins that form the nuclear pores (events shown as small red circles) to allow release of transcribed genes.

1.3.7.2 ATR role at the mitochondria

It is well documented that ATR is involved in the response to UV damage, protecting our skin cells from accumulating UV-induced mutations (Jarrett et al., 2014). Last year, an American group reported a new anti-apoptotic function for ATR in response to UV (Figure 13), which takes place in the mitochondria and is distinct from its nuclear checkpoint function (Hilton et al., 2015). Using SDS-PAGE, they detected the accumulation of a 'heavier', or slower migrating, form of ATR in the mitochondria (mtATR or ATR-H, for heavy) in response to UV. This electrophoretic mobility shift reflected a conformational change in the ATR protein mediated by the PIN1 (Peptidylprolyl cis/trans Isomerase NIMA-interacting 1) protein. PIN1 is a crucial regulator of several biological processes, which recognizes phosphorylated S/T followed by P as catalytic sites for isomerization, regulating protein conformation and, as a consequence, protein function (Liou et al., 2011; Lu and Hunter, 2014).

In unstressed cells, PIN1 isomerizes ATR at phosphorylated S428 residue, converting the heavier ATR (ATR-H) into the lighter or regular ATR (ATR-L) that exists ubiquitously in the absence of damage. However, upon UV treatment,

PIN1 is phosphorylated and thus inactivated, preventing the cytoplasmic conversion of ATR-H into ATR-L. This conformational change potentially exposes a BH3-like domain (Chittenden, 2002) (amino acids 462–474), which might be buried in nuclear ATR and allows it to function at mitochondria. This new ATR domain overlaps with the first helix within HEAT repeat 9 and mediates the interaction with BID (BH3 Interacting-domain Death agonist), blocking BID-dependent recruitment of BAX (Proapoptotic Bcl2-Associated X) and thus suppressing initiation of apoptosis (Hilton et al., 2015).

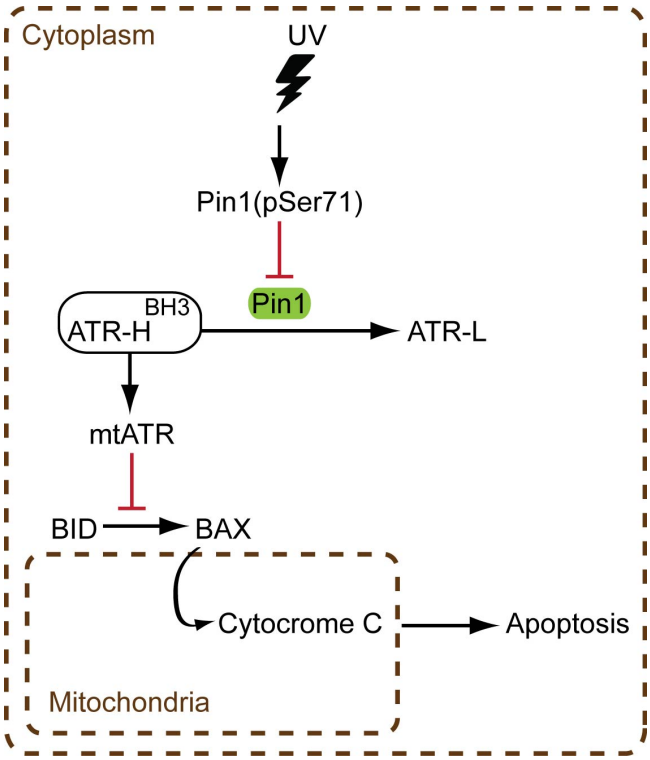


Figure 13. ATR functions in the mitochondria in response to UV damage, adapted from (Hilton et al., 2015)**.** In response to UV, PIN1 is inactivated by phosphorylation at S71, which allows the formation of ATR-H in the cytoplasm. ATR-H can then localize to the mitochondria, where it interacts with BID via its now accessible BH3-like domain, and/or associate to cytosolic BID before entering mitochondria. As a result, ATR-H acts as an anti-apoptotic protein by blocking further recruitment of BAX to mitochondria and thus impeding the release of cytochrome-c and subsequent apoptosis.

Surprisingly, ATR-H did not interact with ATRIP and its mitochondrial functions are independent from its kinase activity. It is possible that ATRIP binding in the nucleus favours the formation of ATR-L in this cell compartment. How this occurs still remains to be elucidated.

This paper is the first evidence that, although phosphorylation of S428 might not be directly involved in the ATR-CHK1 signaling (as explained in section 1.3.3), it is a crucial residue for the production of ATR-L, which is the only ATR form in the nucleus. Moreover, the description of a new domain within ATR contributes to a better understanding of its structure and the role of HEAT repeats.

1.3.7.3 ATR role in cilia signaling

This year, a new role for ATR in ciliogenesis and cilia signaling has also been described (Stiff et al., 2016). Cilia are elongated structures that originate from centrosomes, which function is to support cell motility as well as to transmit and receive extracellular signals. They also participate in signaling pathways that are important during development (Satir and Christensen, 2007). It has been shown that ATR depletion or inhibition has a consistent but small impact on cilia length, and a striking impact on cilia signaling and function during development. These results were confirmed using both human cells and the zebrafish model system, which is widely used in the cilia field (Fraenkel et al., 2005).

This new role of ATR in cilia function can be dissociated from its replication functions. Accordingly, RPA and CHK1 were dispensable for cilia signaling; whilst both ATR kinase activity and TOPBP1, were required for this function (Stiff et al., 2016). The mechanism by which ATR is activated to regulate cilia function remains unknown. A possibility is that ATR is able to sense mechanical forces imposed by the formation of cilia, in a similar way that it functions at the nuclear envelope (Kumar et al., 2014).

This discovery creates a provocative link between cilia regulation and microcephaly. Mutations in the *ATR* gene have been previously identified in a subset of Seckel patients (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012), who suffer from microcephaly and developmental defects. In addition, this syndrome, along with other disorders conferring microcephalic primordial dwarfism, are also caused by genetic defects in other genes that encode for centrosome or spindle-associated proteins (Griffith et al., 2008; Kaindl et al., 2010; Kalay et al., 2011; Willems et al., 2010). The clinical manifestations of Seckel patients caused by *ATR* mutations have been usually attributed to the canonical function of ATR in regulating the DDR (Murga et al., 2009; O'Driscoll

et al., 2003). Stiff and colleagues suggest that the defect in cilia signaling, conferred by ATR deficiency, could be an independent or additional trigger to the developmental and neuronal defects seen in Seckel sufferers. In fact, depletion of ATR in zebrafish led to a microcephaly-like phenotype (Stiff et al., 2016).

Further analyses are needed to understand the precise details of ATR contribution to cilia signaling.

1.4 *Gallus gallus* and *Saccharomyces cerevisiae* as model systems

In this project, DT40 chicken cells were used as a model system to study the functional relevance of HEAT repeats in ATR, as well as to investigate the effect that different *ATR* mutations found in patients have on protein function. Budding yeast, a model system that has been widely used in the past to study pathways that are conserved from yeast to humans (Duina et al., 2014), was also used in a small part of chapter 3.

Within this report and from here onwards, we will refer to the human, chicken and yeast ATR homologues using different nomenclature. The human protein/gene will be referred to as ATR (in capital letters), the chicken orthologue will be denoted as Atr (only the first letter in capitals), while the yeast orthologue is known as Mec1.

There is high sequence similarity between ATR and Atr, while Mec1 is not as well conserved (see protein alignment in appendixes, section 8.5). Detailed HEAT repeat and domain comparison between ATR and Atr can be found in chapter 3; whereas a comprehensive study of the Mec1 protein and its structural comparison to human ATR can be found in the thesis of a former PhD student from our laboratory (De Castro Abreu, 2012).

The main advantage for using both the *S. cerevisiae* and the DT40 systems is the high frequency of homologous recombination that takes place in both these cells. Gene constructs transfected into yeast cells integrate preferentially by homologous recombination into the locus of their cellular counterpart. In contrast, transfection of higher cells usually leads to integration of the exogenous DNA predominantly at random chromosomal positions and only infrequently into the homologous gene locus (Adair et al., 1989). However, as shown by Buerstedde

and Takeda, the chicken DT40 cell line shows high levels of homologous recombination allowing a high rate of gene targeting compared to random integration (Buerstedde and Takeda, 1991).

We combine our investigations in chicken and yeast model systems to infer relevance in human cells, with an overall aim of elucidating the mechanism of ATR function and its role in disease progression.

1.4.1 DT40 cell line.

The DT40 chicken cell line is derived from a bursal lymphoma that was transformed by an avian leukosis virus (ALV) (Baba et al., 1985). The interest of this cell line is based on two main properties: First, as mentioned above, homologous recombination between transfected and genomic DNA in DT40 is as efficient as in budding yeast (Buerstedde and Takeda, 1991). Therefore, DT40 cells are ideal for gene targeting experiments and, as shown in Table 2, they have been used for disrupting many genes as a way of studying their function (Brown et al., 2003; Winding and Berchtold, 2001). Second, DT40 cells have a cell cycle lasting about 8 hours, which is about three times shorter than most human cell lines studied (Cooper, 2000). In addition, these cells have a stable, nearly euploid karyotype, are grown in suspension, easy to maintain and have a genome with lower proportion of interspersed repetitive DNA than mouse and human (Brown et al., 2003).

The DT40 cell line has proven to be a useful model system for studying the function of numerous genes, including several implicated in the DDR (such as MRE11, DNA-PK or ATM, as indicated in the table below).

Table 2. Examples of genes that have been disrupted in DT40 chicken cells (Brown et al., 2003; Buerstedde and Takeda, 2006; Fant et al., 2010; de Groote et al., 2011; Inanç et al., 2013; Oestergaard et al., 2012; Prosser and Morrison, 2015; Samejima et al., 2015; Tuul et al., 2013; Zachos et al., 2003).

Pathway or function	Disrupted genes
DNA recombination, DNA repair and DNA replication	*RAD50, RAD51, RAD52, RAD54, RAD54B, RAD51B, RAD51C, RAD51D, XRCC2, XRCC3, MRE11, Ligase IV, KU70, DNA-PK, RAG2, AID, FEN1, RAD18, XPA, POL*

	κ, NBS1, REV1, FANCG, FANCC, FANCD2
RecQ DNA helicase	*WRN, BLM*
Damage checkpoint	*ATM, ATR, RAD9, RAD17, RNF8, RNF168, CHK1*
Cell cycle	*Cyclin D1, CDH1, SUMO, Incenp1, Survivin1*
Transcription and RNA metabolism	*CSTF, ASF/SF2, TATA Binding Protein, cTAF(II), cMYB, GAS41, Poly(A) Polymerase*
Apoptosis	*MEKK1 kinase, CAD/CPAN/DFF40, Caspase6, Annexin5*
Chromatin	*Histone H1, H2A, H2B, H2B-V, H3-IV/H3-V, H4, HDAC-1, -2 and -3, HAT-1, HMG-17, HMG-14a*
Chromosome structure and function	*CENP-C, CENP-H, CENP-I, CENP-K, CENP-M, CENP-O, CENP-N, CENP-R, CENP-S, SCC1/RAD21/MDC1, Centrin2, CEP-135*
Heat shock factors	*HSF1, HSF3*
Signal transduction	*LYN, SYK, BTK, CD45 phosphates, PLC, ABL, γCBL-B, Inositol 1,4,5-Trisphosphate Receptor (types 1,2 and 3), BLNK, BAM32, SHP1, SHP2, BCAP, GRB2, GRAP, SHIP, SHC, TRP1, VAV3, PI3Kp-α*
Metabolism and scavenger proteins	*HPRT, Thioredoxin*
Protein secretion and cellular homeostasis	*CHC, TRPC1, TRPC7, TRPM7, ZnT5, ZnT6, ZnT7*

1.4.2 *Saccharomyces cerevisiae*: Similarities and differences between ATR and Mec1 signaling pathways

Although *S. cerevisiae* is one of the simplest eukaryotic organisms, several of the essential cellular processes are conserved from yeast and humans. For this reason, yeast is an important model organism, which has been used for gene identification and initial functional characterization, but ultimately to understand basic molecular processes in humans (Duina et al., 2014). Besides, yeast cells are easy to handle and divide as fast as once every 90 min.

In relation to our study, the DDR pathways and its regulators are conserved through evolution, as described in Table 1. In particular, the budding

yeast Mec1 protein participates in the DDR in an equivalent way as its homologue ATR. However, there are some differences in Mec1 activation and signalling.

Mec1 is often considered as the main PIKK in *S. cerevisiae* since *mec1* mutants are very sensitive to several DNA damaging agents (Harrison and Haber, 2006) while Tel1 (ATM yeast orthologue) function is mostly apparent when numerous DSBs are generated (Mantiero et al., 2007). These two proteins have key roles mostly similar to their vertebrate homologues. ATM/Tel1 is required for checkpoint activation primarily in response to DSBs (Suzuki et al., 1999), while ATR/Mec1 controls the response to a much broader spectrum of DNA damage and replication stress including breaks, crosslinks and base adducts (Cimprich and Cortez, 2008). In addition, *Tel1/ATM* is a non-essential gene, whereas *Mec1/ATR* is essential for sustaining viability (Brown and Baltimore, 2000; Desany et al., 1998; Zhao et al., 1998).

While the activation of ATR in higher eukaryotes appears to require a single pathway, in which both the 9-1-1 complex and the TOPBP1 function independently of the cell cycle stage; it has been shown that Mec1 activation differs at the various phases of the cell cycle (Navadgi-Patil and Burgers, 2011) (Figure 14). In *S. cerevisiae*, the proper loading of the Ddc1-Rad17-Mec3 complex (yeast orthologue of the 9-1-1 complex) onto the ssDNA/dsDNA junctions is required for the activation of both G1 and G2 checkpoints, but the mechanism differs between these two phases of the cell cycle (Majka et al., 2006; Navadgi-Patil and Burgers, 2009; Puddu et al., 2011). On the other hand, Dbp11 (TOPBP1 homologue protein) is only required for Mec1 activation in G2-phase (Navadgi-Patil and Burgers, 2009). Finally, the mechanism of Mec1 activation in S-phase is not yet clear, but it seems that the Ddc1-Rad17-Mec3 complex and Dpb11 are both required to activate Mec1 independently along with other components of the replication machinery (Puddu et al., 2011).

Apart from differences the recruitment of Mec1 to ssDNA, there also variations in the Mec1 dependent signaling. Once activated, Mec1 promotes the Rad9-dependent phosphorylation of both Chk1 and Rad53 (CHK1 and CHK2 homologues) (Gilbert et al., 2001; Sanchez et al., 1999; Sweeney et al., 2005). The role of Rad9 in Rad53 phosphorylation has been deeply studied in budding yeast (Sweeney et al., 2005). In fact, Rad53 phosphorylation is a key step in the

signal transduction cascade used as a marker to monitor full checkpoint activation (Pellicioli et al., 1999). It is important to mention that Rad53 is the homologue of CHK2, which is not an ATR target in higher organisms but an ATM target (Shiloh, 2003). Besides, it has been shown that in *S. cerevisiae*, Mrc1, the CLASPIN orthologue, seems to predominantly control Rad53 activation during the DNA replication checkpoint (Tanaka, 2010) and a role in Chk1 activation remains to be reported. There is some evidence that in both budding and fission yeasts the prototypical mediator, Rad9/Crb2 (homologues of 53BP1/MDC1/BRCA1 mediators), has a role in Chk1 activation after DNA damage (Blankley and Lydall, 2004).

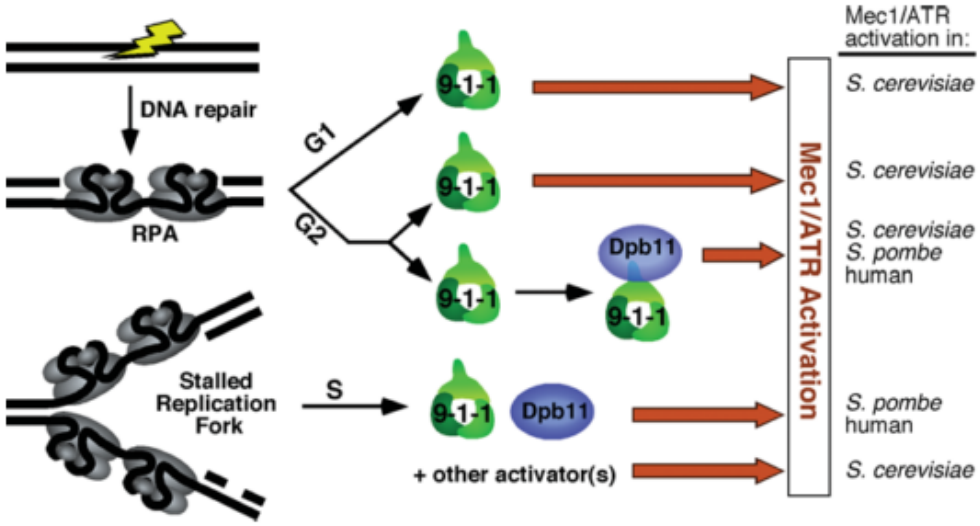


Figure 14. Differential cell cycle-dependent activation pathways in *S. cerevisiae* versus *S. pombe* and human (Navadgi-Patil and Burgers, 2011). The 9-1-1 complex and Dbp11 (TOPBP1 homologue) proteins play different roles in the activation of Mec1, dependent on the cell cycle stage. These features are not shared in other species, such as fission yeast (*S. pombe*) or human.

1.5 Conditional-null approaches to study essential genes

The directed introduction of null or knockout (KO) mutations into defined genes is a powerful tool for elucidating gene function in a variety of experimental organisms (Le and Sauer, 2000). However, the disruption of a gene that is essential for development often leads to lethal phenotypes (Banaszynski and Wandless, 2006). As this is the case for the *ATR* gene, it was not possible to work with a parental cell line in which the gene was knocked out. Nevertheless, a

variety of methods to conditionally regulate gene function have been developed for use in vertebrates, in which gene disruption, protein depletion or protein inactivation can be induced in a specific controlled manner. Some of the conditional-null strategies used in the study of essential genes are shown in the table below (Table 3).

Table 3. Conditional-null strategies for studying essential genes. Every different approach is exampled by a study in which DT40 or human cells are utilized.

Description	Level of conditionality	Advantage	Disadvantage
LoxP-Cre recombination (Arakawa et al., 2001; Brown and Baltimore, 2003)	Presence of gene	Individual 'floxed' cell eventually loses all protein	Asynchronous process, inefficient at some loci
Tet-on/off (Gossen and Bujardt, 1992)	Gene expression	Uniform response in a population	'Leaky' expression when 'off'
Degron-tag (Nishimura et al., 2009)	**Presence of protein**	**Degradation of protein gives a fast response**	**Need to tag the protein of interest (it might affect protein function)**
Protein inhibitor (Zachos et al., 2005; Zhang et al., 2009)	Protein activity	Acts quickly on wild-type protein	Not easy to find a specific inhibitor
Chemical-genetic (Hochegger et al., 2007)	Protein activity	Acts quickly, well established methods for kinases	Need to mutate your protein (it might affect protein function)

Consequently, our study was carried out using a conditional-null cell line designed in chicken DT40 cells based on the degron-tag strategy (further details explained in section below.). The selection of the degron-tag system was based on previous results from our laboratory (Eykelenboom et al., 2013).

1.5.1 Auxin inducible degron (AID) approach

Auxin represents a family of plant hormones including 3-Indolyacetic acid (IAA; a natural auxin) and 1-naphthaleneacetic acid (NAA; a synthetic auxin). Auxins have an essential role in many aspects of growth and development, being responsible for the control of gene expression in these processes (Teale et al., 2006).

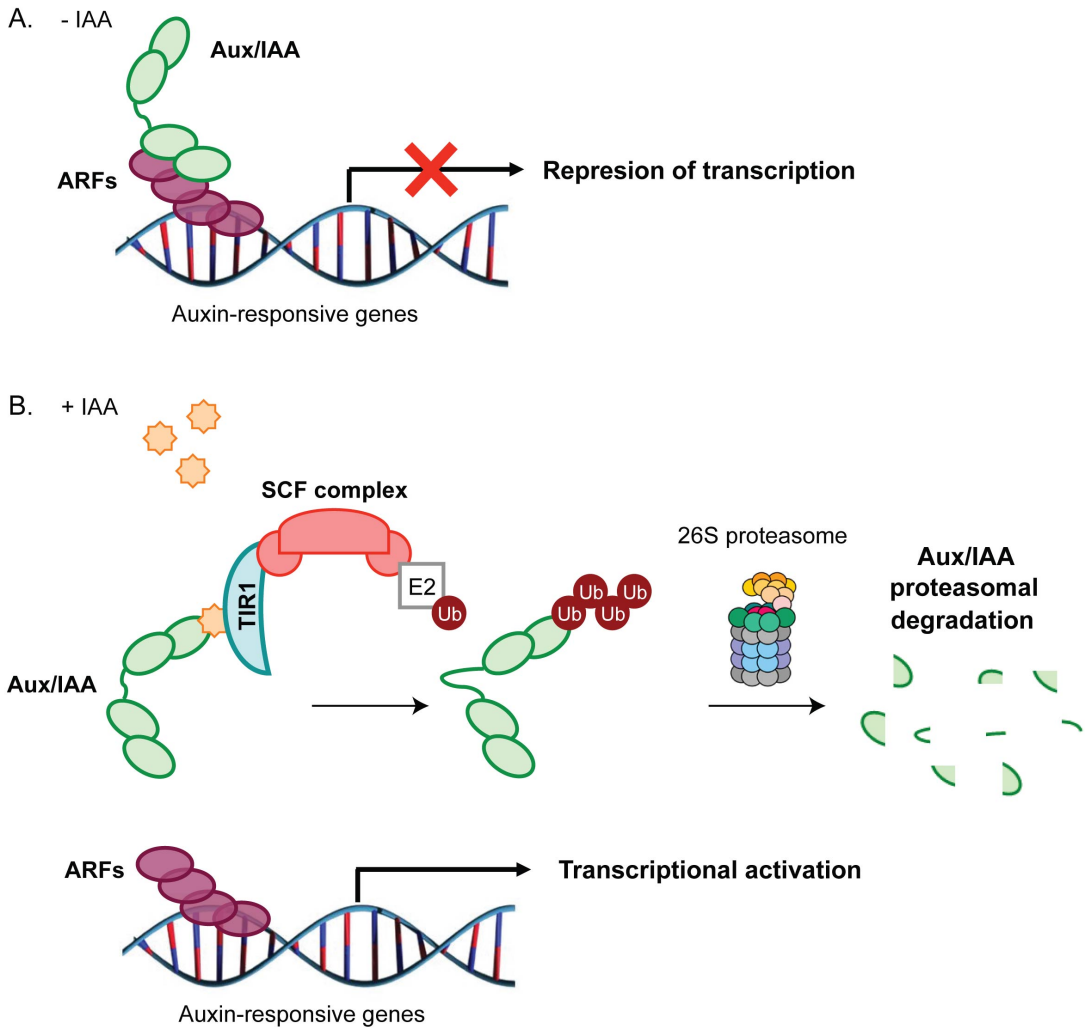


Figure 15. Diagram of auxin dependent gene expression, adapted from (Santner et al., 2009).

A. Expression of the Aux/IAA transcription factors downregulates the expression of the AUX responsive genes by inhibiting ARFs function.

B. In the presence of AUX, Aux/IAA proteins are targeted for degradation by the proteasome, indirectly stimulating ARFs function and allowing transcription of the AUX responsive genes.

Auxin regulates transcription through the action of two large families of transcription factors called the auxin/3-indolyacetic acid (Aux/IAA) proteins and the Auxin Response Factors (ARFs) (Santner et al., 2009). ARFs are directly bound to DNA and, when low auxin levels are present, the Aux/IAA proteins repress their function and consequently transcription (Figure 15A). However, when auxin is present in a high enough level, the auxin signaling pathway is activated and transcription stimulated: Auxin binds to the F-box Transport Inhibitor Response 1 (TIR1) protein, which is part of an SCF ubiquitin-ligase complex (consisting on a Skp1 protein, a Cullin, an F-box protein, and RBX1), and allows its interaction with the Aux/IAA proteins (Kepinski and Leyser, 2005). This interaction leads to polyubiquitination of the Aux/IAA transcription repressors and their degradation by the proteasome (Figure 15B) (Teale et al., 2006).

All eukaryotes have multiple forms of SCF in which an F-box protein determines substrate specificity, but orthologues of TIR1 and AUX/IAAs are only found in plant species (Nishimura et al., 2009). Based on the auxin pathway signaling and following the idea of Nishimura and colleagues, we wanted to recreate the SCF-TIR1 complex in DT40 cells so that a target protein fused to an Aux/IAA protein (hereafter called 'auxin inducible degron', AID) would be degraded in an auxin-dependent manner (Nishimura et al., 2009).

In previous work in the laboratory, a DT40 cell line was developed that expressed functional AID-tagged Atr from a single cDNA expression cassette that had been targeted to the *Ovalbumin* locus. In addition, the endogenous alleles of *Atr* in this cell line had been disrupted (by deletion of the kinase domains). Finally, the plant F-box protein gene TIR1 had been randomly integrated into the genome of this cell line as part of another expression cassette. Further experiments showed that following two hours treatment of this cell line with AUX there was complete loss of AID-Atr (and this was dependent on the presence of the AID tag and the TIR1 protein; Figure 16) (Eykelenboom et al., 2013).

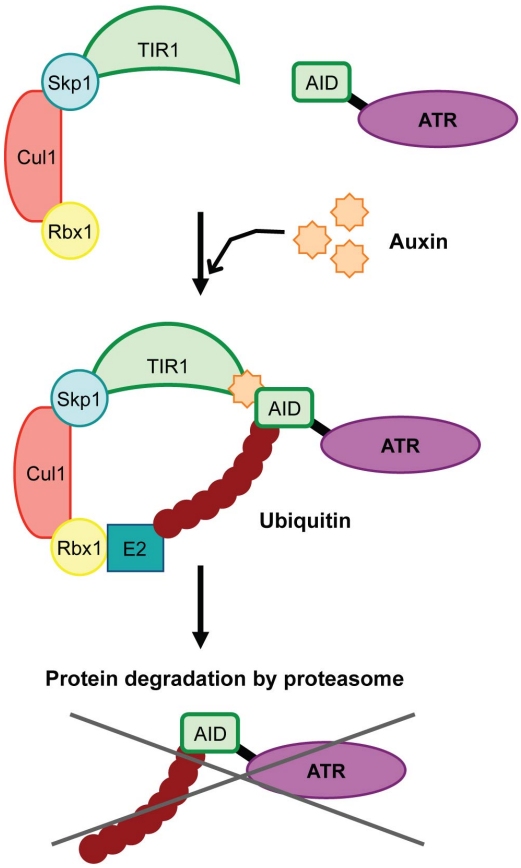


Figure 16. AID-Atr degron system, adapted from (Nishimura et al., 2009). The diagram shows the AID-Atr AUX-dependent degradation. The components that come from plants are shown in green; while the rest are normally expressed in higher cells, in this case chicken DT40. Once AUX is added to the system, AID-Atr interacts with the SCF-TIR1 complex in an AUX-dependent manner, which leads to AID-Atr polyubiquitination and degradation by the proteasome

1.6 Introduction to the present project

Despite ATR having a key role both in the DNA damage response and during a normal cell cycle, the molecular nature of its function remains ill-defined. Little is known about the regulation of its enzymatic activity, its substrate specificity and its specific interactors. Additionally, very few details are known about the structure of this protein. In particular, the non-kinase portion of ATR (about 85% of its structure) is made up of tandem helical repeats, called HEAT repeats, most of which remain uncharacterized. Based on studies in other HEAT-rich proteins, we speculate that HEAT repeats mediate protein-protein interactions. As part of this study, we investigate the structure-function relationship of the ATR protein (Chapter 3). The goal of this project was to

correlate specific functions of ATR to specific regions of the protein. To do this, we chose to generate a set of *ATR* mutants, called 'window' or internal deletion mutants, in which individual or groups of HEAT repeats were deleted. A parallel study for *Mec1* was also carried out in the laboratory (De Castro Abreu, 2012).

We also investigated the effect of several *ATR* disease-related mutations on its function. As explained in section 1.3.5 (and in more detail in Chapter 4), many *ATR* mutations have been described in the literature, with particular focus on hypomorphic mutations that cause Seckel syndrome. As part of this project, we set to study two novel *ATR* mutations that were found in unrelated Seckel sufferers (Chapter 4), as well as a novel dominant *ATR* mutation, which affected an evolutionary conserved residue and caused oropharyngeal cancer predisposition (Chapter 5). With this approach we plan to gain further insights into ATR function, but also widen the understanding of Seckel syndrome and cancer predisposition.

In order to achieve our objectives, we have mainly used the genetically tractable chicken DT40 system (Chapters 3, 4 and 5), as well as budding yeast (Chapter 3) and some human cell lines (Chapter 4). Using the conditional-null DT40 system described above (section 1.5.1), we generated different stable cell lines individually expressing the *Atr* window deletions, the Seckel and the cancer-related mutants. These *Atr* mutant versions were tested for the essential and checkpoint roles of Atr. In order to do this, we used auxin to deplete the AID-tagged form of Atr in these cell lines and monitored the ability of the mutant forms to support cell growth or to instigate a DNA damage checkpoint. This was accomplished using cellular and biochemical techniques such as growth analyses and checkpoint assays. Other biochemical and molecular techniques, including tandem purification, electron microscopy, splicing assays and karyotypic analyses, were employed to address the functionality of these mutants.

Deepening the knowledge of ATR structure, as well as its functions and activity regulation is vital to understand the general regulation of the DNA Damage Response. This is also fundamental to fully understand the role of ATR in unchallenged cells, which will help to elucidate the deficiencies underlying certain human pathologies and will allow us to generate better therapies.

CHAPTER 2

Materials and methods

2.1 Materials

In the tables below are shown the names, composition and use of common buffers and solutions (Table 4), as well as the reagents, such as drugs and inhibitors (Table 5), antibodies (Table 6), plasmids (Table 7) and cell lines (Table 8), used in this study.

Table 4. Common buffers and solutions. Reagents used were generally supplied by Sigma and, unless indicated, all the solutions were made with ddH_2O.

Buffers and solutions	Composition	Use
Blocking solution	5% milk in PBS-T	Western blot analysis
1x Ca^{2+} buffer	10mM HEPES pH 7.5, 140mM NaCl, 2.5mM $CaCl_2$	Flow cytometry analysis
Electrophoresis buffer	1xTG, 0.1% SDS	Running SDS-PAGE gels
Fixative	Methanol/Acetic acid (3:1). Prepared fresh for each experiment.	Metaphase spreads
Hybridization buffer	DIG Easy Hyb granules (Roche) reconstituted in 64ml ddH_2O, as manufacturer recommends.	Southern blot analysis
Hypotonic solution	75mM KCl	Metaphase spreads
5x Laemli sample buffer	30mM Tris-HCl ph 6.8, 10% glycerol, 3% SDS, 0.05% bromophenol blue, 2% ß-mercaptoethanol	Western blot analysis
LB	10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl in 1litre dH_2O. Sterilised by autoclaving.	*E.coli* transformation
LB/Agar	15g agar in 1l LB	*E.coli* transformation
Lysis buffer	50mM Tris, pH 7.5, 150mM NaCl, 0.5% NP40, 1mM EDTA, 5-10% glycerol,	Protein extraction

	protease and phosphatase inhibitors	
1x Phosphate-buffered saline (PBS)	137mM NaCl, 2.7mM KCl, 8mM Na_2HPO_4, and 2mM KH_2PO_4, pH to 7.4. Sterilised by autoclaving.	Cell culture
PBS-Tween	1x PBS with 0.1% Tween20	Washes and dilutions of the antibody in western blot analysis
50x Phosphatase inhibitors (PPIs)	0.105g NaF, 0.54g ß-glycerphosphate, 0.092g Na_3VO_4, 0.951g EGTA, 5.579g sodium pyrophosphate in 50ml dH_2O	Protein extraction
Ponceau-S solution	0.5% Ponceau-S, 5% Acetic Acid	Stain proteins on the nitrocellulose membranes
100x Protease inhibitors (PIs)	1.42mg leupeptin, 6.85mg pepstatin A, 0.85g PMSF, 1.65g benzamidine, 6.25mg antipain, 4mg chymostatin (dissolved in DMSO), EtOH to 50ml	Protein extraction
20x SSC buffer	3M NaCl, 300mM sodium citrate, pH 7.0	Southern blot analysis
TAIL buffer	50mM Tris pH 8.8, 100mM EDTA, 100mM NaCl, 1% SDS, adding 0.5mg/ml proteinase K (working concentration) before digesting the genomic DNA	Genomic DNA extraction
1xTAE	40mM Tris-Acetate pH 8.0, 1mM EDTA	Preparation and running of agarose gels
1xTBE	90mM Tris, 90mM Borate, 3mM EDTA, pH 8.3	Preparation and running of agarose gels
1xTE	10mM Tris pH 7.5, 1mM EDTA	Genomic DNA extraction
10x Tris – Glycine (TG)	0.25M Tris, 1.92M glycine	Protein gel running and transfer buffers
1x Tris-buffered saline (TBS) –	100mM Tris, 150mM NaCl, pH 7.5. With 0.1% Tween20	Washes and dilutions of the phospho

Tween		antibody in western blot analysis
Wet transfer buffer	1xTG, 20% Methanol, 0.1% SDS	Transfer SDS-PAGE gels onto nitrocellulose membranes

Table 5. Drugs and inhibitors.

Drugs	Composition	Use
Ampicillin	50mg/ml stock in ddH_2O. Used at 100µg/ml.	Selection of *E.coli* transformants
Bromodeoxyuridine (BrdU)	20mM stock in ddH_2O. Used at 30µM.	Proliferation analysis (BrdU pulse-chase)
3-Indolyacetic acid (IAA)	Referred to as 'AUX' in this work. From Sigma. Prepared fresh for each experiment at 0.5M in EtOH. Used at a concentration of 0.5mM.	DT40 cell culture
Hydroxyurea (HU)	From Sigma. Prepared at 0.66M in ddH_2O. Used at a concentration of 1mM.	DT40 cell culture
Nocodazole	From Sigma. Prepared at 2mg/ml in DMSO. Used at 0.25µg/ml.	Mitotic index analysis
N-deacetyl-N-methylcolchicine (Colcemid)	From Sigma. 1mg/ml stock in PBS. Used at 0.1µg/ml.	Metaphase spreads
Puromycin	From Invitrogen. 0.5mg/ml stock in ddH_2O. Used at 0.5µg/ml (for DT40) or 1µg/ml (for human cells).	Selection of DT40 and human transfected cells
Rad51 inhibitor	From Axon. 10mM stock in DMSO. Used at 10µM.	CRISPR transfections

Table 6. Antibodies used for western blotting (IB) and flow cytometry (FC).

Target	Reference no.	Host species	Source	Dilution for IB or FC	Conjugation
Primary antibodies					
ATR	sc-1887	Goat polyclonal	Santa Cruz	1:2,000	
BRCA2	ab123491	Rabbit polyclonal	Abcam	1:2,000	
BrdU	347580	Mouse monoclonal	Becton Dickinson	1:20	
Chk1	sc7898	Rabbit polyclonal	Santa Cruz	1:2,000	
Chk1 – phospho S345	2341	Rabbit polyclonal	Cell Signaling	1:1,000	
Histone H3 – phospho S10	06-570	Rabbit polyclonal	Millipore	1:50	
Secondary antibodies					
Goat IgG	sc-2020	Donkey	Santa Cruz	1:5,000	HRP
Mouse IgG	115-096-062	Goat	Jackson Immunoresearch	1:100	FITC
Rabbit IgG	111-096-045	Goat	Jackson Immunoresearch	1:100	FITC
Rabbit IgG	31460	Goat	Thermo Scientific	1:5,000	HRP

Table 7. Plasmids used in this study.

Plasmids	Details	Source
pLC4	pExpress containing the full length *Atr* cDNA under the control of a β-actin promoter and followed by a polyA tail. Used to generate the deletion, replacement and point *Atr* mutations.	(Eykelenboom et al., 2013)

pJE28	Plasmid containing Ovalbumin homology arms and a Puro^R cassette under the control of a Cmv promoter. Used in transfections to target the *Atr* mutants to the *Ovalbumin* locus of DT40 cells.	(Eykelenboom et al., 2013)
pTB NF1-29	Minigene plasmid containing exon 29 of the *NF1* gene and flanking intronic sequences. Used as a control in the *in vitro* splicing assay.	(Raponi et al., 2009)
pX335	Human codon-optimized SpCas9 nickase and chimeric guide RNA expression plasmid. Used for CRISPR targeting.	Addgene, originally from the Zhang laboratory (Cong et al., 2013).

Table 8. Source of cells lines used in this study

Cell lines	Source
DT40	Derived from WT, Clone 18 (Buerstedde et al., 1990)
HCT116	Supplied by Dr. B. Vogelstein (Bunz et al., 1998)
HEK293T	Obtained from ATCC.
HeLa	Obtained from ATCC.
hTERT-RPE1	Obtained from ATCC.
Patient-derived LCLs	Supplied by collaborator Prof. G. Stewart (Birmingham, UK).

2.2 Basic DNA methods

2.2.1 Plasmid DNA preparation and Enzymatic reactions

RBC and Thermo Scientific plasmid purification kits were used for DNA minipreps and Qiagen and Machery-Nagel kits for midipreps, respectively, and performed as per the manufacturer's recommendations. All cloning steps and preparation of plasmid DNA required before the transfection of DT40 and human cells were carried out using TOP10 competent *Escherichia coli* cells.

Restriction digests and ligation reactions were performed with buffers and protocols supplied with the enzymes. Restriction enzymes and T4 DNA ligase were from NEB.

2.2.2 Agarose gels electrophoresis and DNA gel extraction

Unless otherwise indicated, 1% agarose gels were prepared using Sigma electrophoresis grade agarose in 1xTAE buffer. Ethidium bromide was added at a final concentration of 0.5µg/ml. 6x Loading Buffer (Sigma) was used to load the samples at a final 1x dilution. The markers used were 1kb and 1kb Plus DNA Ladder from Fermentas. Gels were run in 1x TAE in Owl mini (50ml gel) or wide mini (100ml gel) cells at 80-100V until the required separation was achieved. The analysis of the gel was carried out using a transilluminator with UV light and photographs were taken with a digital camera (ChemiImager 5500, Alpha Innotech). If necessary, after the separation by electrophoresis and visualization by UV imaging, DNA bands were excised from gel and purified using QIAquik Gel Extraction kit (Qiagen or Machery-Nagel) according to the manufacturer's protocol.

2.2.3 Polymerase chain reaction (PCR)

PCR was used for two main purposes; KOD polymerase (Novagen) was used for cloning purposes and PCR analyses; while *Taq* Crimson polymerase (NEB) was used for DIG labelling PCR (for generating Southern blotting probes). Reactions were setup following manufacturers' instructions and amplification was

carried out using a DNA engine thermocycler (GRI). Examples of programmes used are shown in Tables 9 and 10.

Table 9. PCR conditions used for cloning and PCR analyses.

KOD PCR			
1. Initial denaturation		95ºC 2min	
2.	Denaturation	(i) 95ºC 20sec	x 30
	Annealing	(ii) 55-58ºC 10sec	
	Extension	(iii) 70ºC 5-60sec	
3. Final extension		70ºC 2min	
4. PCR finished. Keep product stable.		4ºC ∞	

Table 10. PCR conditions used in Southern probe DIG labelling.

***Taq* Crimson PCR**			
1. Initial denaturation		95ºC 5min	
2.	Denaturation	(i) 95ºC 30sec	x 30
	Annealing	(ii) 58ºC 15sec	
	Extension	(iii) 72ºC 1min 15sec	
3. Final extension		72ºC 10min	
4. PCR finished. Keep product stable.		4ºC ∞	

2.2.4 DNA sequencing

Plasmid DNA was sent to LGC genomics or Eurofins MWG for sequencing. 100ng of plasmid DNA was sent and sequencing was carried out using either a T7 forward primer provided by the sequencing company or specific primers that were mixed and sent together with the DNA.

2.3 DT40 techniques

2.3.1 DT40 medium and culture conditions

Chicken DT40 cells were grown at 39.5ºC in a humidified atmosphere of 5% CO_2. The medium used was RPMI 1640 with L-Glutamine (Gibco or Lonza) supplemented with 10% Foetal Bovine Serum (FBS; Sigma or Lonza), 1% Chicken Serum (CS; Sigma) and 1% Antibiotic penicillin/streptomycin mixture (Pen/Strep; Sigma or Lonza).

2.3.2 Generation of *Atr* mutants

Generation of *Atr*-window deletions, *Atr*-Seckel/Cancer-related point mutations and *Atr*-replacement mutants was achieved by fusion PCR. Primers used in each step are shown in Table S3. External primers used in the PCR reactions contained unique restriction sites, which were used for further cloning steps.

The resulting *Atr* mutant constructs were then inserted into a plasmid containing full-length *Atr* cDNA (pLC4; (Eykelenboom et al., 2013)) using unique restriction sites. Once cloned, constructs were preceded by the *β-actin* promoter and followed by a poly-A signal. The whole construct was flanked by two SpeI restriction sites. After cloning, mutant *Atr* cDNAs were fully sequenced to ensure no extra mutations were introduced during the PCR reaction (primer details in Table S4).

2.3.3 Cloning of *Atr* mutant constructs into an *Ovalbumin* targeting vector

To introduce the different *Atr* mutants into DT40 and generate stable cell lines, the *Atr* cDNAs were then cloned into an *Ovalbumin* targeting vector (pJE28; (Eykelenboom et al., 2013)). This vector included a puromycin resistance cassette flanked by *Ovalbumin* homology arms that allow targeting of the intervening sequence to the specific locus in DT40 cells by homologous recombination.

Cloning of the *Atr* mutants into the *Ovalbumin* targeting vector was carried out by two parallel restriction reactions: The plasmids containing the *Atr* cDNA mutations were digested with SpeI to release the *Atr* cDNA flanked by promoter and polyA signal; while the *Ovalbumin* targeting vector was digested with NheI (found between the *Ovalbumin* homology arms). After the digests, a ligation step was carried out and clones were screened by digestion, selecting those where the promoter for *Atr* was in the opposite orientation to the promoter of the resistance cassette.

2.3.4 Preparation of plasmid DNA for DT40 transfection

5–15µg of purified plasmid DNA were linearized using the Pvu1 restriction enzyme (NEB). The restriction reaction was carried out for 4h at 37ºC. DNA was then precipitated using an ethanol precipitation.

2.3.5 Stable transfection of DT40 cells

$1x10^7$ DT40 cells were transfected by electroporation with 5-15µg of linearized plasmids using 0.4cm Gene Pulsar cuvettes and Gene Pulsar electroporation apparatus (Bio-Rad) at 550V and 25µF. 16-24h after transfection, cells were placed in 96-well plates in fresh media containing puromycin selection (Invitrogen; final concentration 0.5µg/ml). After up to 10 days when clones were visible and large enough, single clones were picked and resuspended in fresh medium without selection drug in 24-well dishes. Clones were expanded until confluent in 12-well dishes, after which half of the culture was frozen and the other half was used for genomic DNA extraction (see following section 2.3.6).

2.3.6 Genomic DNA extraction and Southern blotting

Genomic DNA was prepared from cultured DT40 cells in order to screen colonies for targeted integration of genomic constructs by Southern blotting. $1.5x10^6$ cells were collected and lysed in 500µl TAIL buffer. After an overnight incubation at 37ºC, the genomic DNA was precipitated with saturated NaCl (6M)

and isopropanol. DNA pellet was allowed to air-dry before resuspension in 60µl of TE buffer. Samples were left at room temperature or overnight at 4ºC to allow resuspension of the DNA pellet and were then used for downstream applications or stored at -20ºC.

Approximately 10µg of genomic DNA were digested overnight at 37ºC by the restriction enzyme BamHI (NEB). Digested DNA was separated by electrophoresis on a 0.8% agarose gel. DNA was transferred to a nylon membrane (Amersham Hybond-N) in 2xSSC by upward capillary transfer. After transfer, the membrane was crosslinked by UV (3000Jm^{-2} for 1min) using the UV Stratalinker 2400 (Stratagene). A non-radioactive DIG-labeled probe was then hybridized and detected as per manufacturer's instructions (Roche). The probe used to screen clones from the *Ovalbumin* locus targeting was amplified from DT40 genomic DNA using the primers Ova_PF3 and Ova_PR3 primers (Table S4), and then cloned into the pGEM-T easy vector (Promega). The probe was then amplified from the pGEM-T-probe vector and labeled with digoxigenin by PCR (PCR DIG probe synthesis kit, Roche).

2.3.7 DT40 cell proliferation analysis

DT40 cell lines to be analysed were plated in 6-well dishes at a starting concentration of 0.1x10^6 cells/ml. Cells were allowed to grow in the incubator as normal and the cell number was counted every 24 hours for 96 hours using a haemocytometer. Exponential growth was maintained (between 0.1 and 1x10^6 cells/ml) by appropriate dilution in fresh medium every 24 hours. Any dilutions carried out were accounted for when plotting cell numbers. To construct a growth curve the cell number (x10^6) per ml of culture (on the Y axis) was plotted against time (on the X axis) for each cell line analysed.

2.3.8 BrdU pulse chase

Cells were grown to a density of approximately 0.5x10^6 cells/ml and then pulse treated with 30µM BrdU for 30min before washing twice with 5ml of PBS and splitting into two separate dishes. To one dish, 0.5mM AUX was added.

Samples were taken and fixed at time intervals following the pulse with BrdU as indicated (0, 2, 4, 6, 8 and 10h). Live single cells in G1 and early S phase were measured as a readout of BrdU incorporation for the indicated cells grown in the presence or absence of AUX.

2.3.9 Mitotic accumulation assay

Cells were grown to a density of approximately $0.5x10^6$ cells/ml and then treated split into two dishes, adding 0.5mM AUX to one of them. This was followed by the addition of 0.25µg/ml nocodazole and the collection and fixation of samples at time intervals as indicated (0, 2, 4, 6, 8 and 10h). Histone H3S10ph accumulation in the presence of nocodazole was plotted for the indicated cells grown in the presence or absence of AUX.

2.2.10 Metaphase spreads

Before fixation, between $1–5x10^6$ exponentially growing cells were treated with 0.1µg/ml colcemid for 1-2h. Cells were collected and resuspended in 5ml of hypotonic solution and incubated for 10min at 37ºC. Cells were then resuspended in 5ml of cold fixative and incubated for 20min at 37ºC. This fixation step was repeated and cells resuspended in approximately 200µl of fixative solution. Cells were stored at -20ºC. For spreading, 50µl of fixed cells were dropped from a 40cm height onto pre-wet (50% ethanol) polysine slides and air-dried. Chromosomes were stained using 1µg/ml DAPI. A light microscope with a 100x objective was used to visualize chromosome spreads.

2.4 Yeast techniques

2.4.1 Yeast medium and culture conditions

Yeast cells were grown in YPD liquid (1% (w/v) yeast extract (Difco), 2% (w/v) bactopeptone (Difco), 2% (w/v) glucose (Difco)) or solid media containing 2% (w/v) agar (Difco). Cells were grown in conical flasks with liquid media at 30ºC in a shaker incubator at 170 rpm.

Minimal media required for viability assay is based in YNB medium (Yeast Nitrogen Base without amino acids, Difco) with 2% agar (w/v) (Difco) and 2% glucose (w/v, Difco) in which the required amino acids were added to a final concentration of 20µl/ml (histidine, adenine, uracil and tryptophan, but no leucine).

2.4.2 Generation of yeast replacement mutants

All yeast strains used in this study are made in a W303-1a *sml1Δ* background. Fusion PCR was used to design *mec1* chimeric constructs. Primers used are listed in Table S5. Generation of the yeast replacement mutants was achieved by *Delitto Perfetto* (adapted from (Storici and Resnick, 2006)). Detailed methodology and selection conditions can be found in (De Castro Abreu, 2012).

Briefly, the *Delitto Perfetto* strategy (meaning 'perfect crime' in Italian) is based on the integration of a *COunterselectable REporter* (CORE) cassette into a desired genomic locus using standard targeting procedures, followed by the transformation with the appropriate targeting DNA fragments such that the CORE cassette is lost and the desired change is generated (Figure 17). Generation of *mec1* CORE strains was performed by Dr. Carla Abreu and use of these strains allows for the introduction of any desired modification (such as deletion, replacement, point mutation, etc.) within the *mec1* gene.

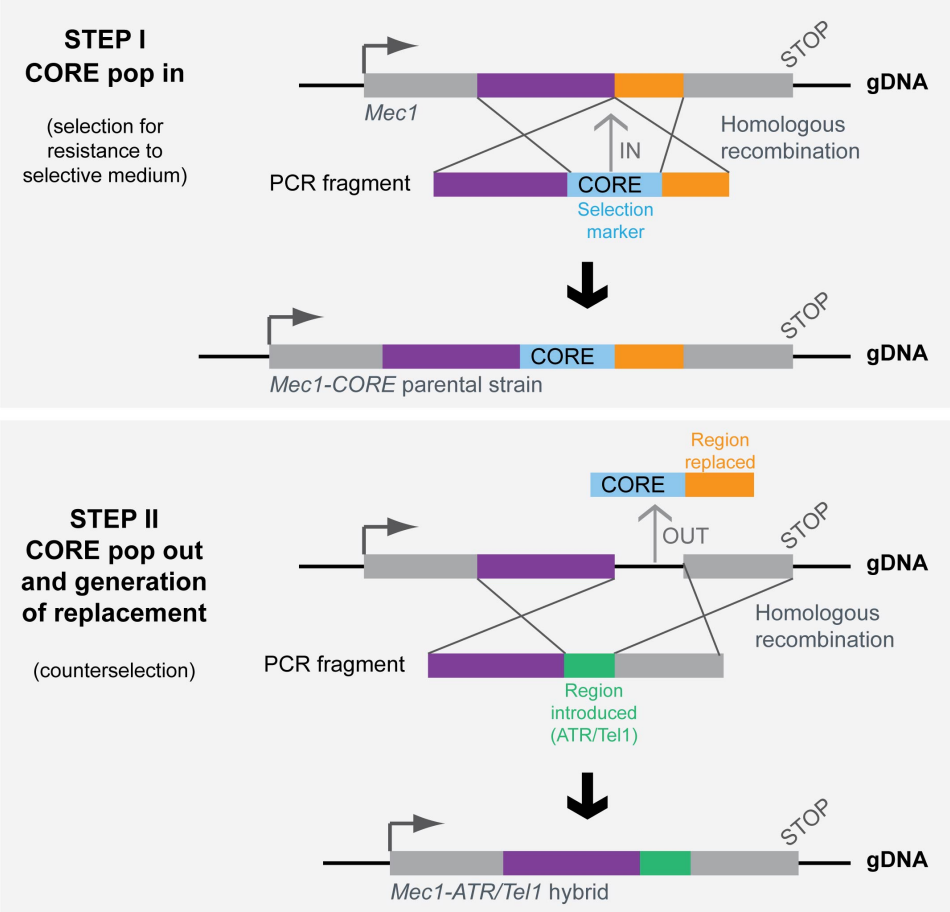


Figure 17. Schematic of *Delitto Perfetto* strategy. Generation of a parental cell line containing a CORE (COunterselectable REporter) cassette within the desired gene allows for the introduction of any desired mutations. In our case, certain *Mec1* regions were replaced by the equivalent *ATR* or *Tel1* sequences and resulting hybrids were identified by counterselection, or loss to the resistance given by the CORE. The *Mec1* gene is represented in grey, with regions highlighted in purple and orange, which are part of the homology arms that allow the introduction of the CORE cassette. Replacement region from *ATR* or *Tel1* is shown in green.

2.4.3 Viability assay

Viability assay was performed by transforming exponentially growing cells with 300ng of *SML1* centromeric plasmid (provided by R. Rothstein, (Zhao et al., 1998)), using the lithium acetate method described previously (Wach et al., 1994). $1x10^8$ cells were plated onto selective medium lacking leucine. Plates were then incubated at 30ºC and colonies were counted after two days of incubation (or until colonies were formed).

2.5 Human cell techniques

2.5.1 Human cell medium and culture conditions

HCT116, hTERT-RPE1, HEK293T and HeLa cells were grown at 37ºC in a humidified atmosphere of 5% CO_2. The medium used was DMEM (Lonza), except for hTERT-RPE1 cells, which were grown in DMEM-F12 (Lonza), supplemented with 10% FBS (Sigma or Lonza) and 1% P/S antibiotics (Sigma or Lonza).

Patient-derived lymphoblastoid cell lines (LCLs) were cultured at 37ºC in a humidified atmosphere of 5% CO_2. The medium used was RPMI 1640 with L-Glutamine (Gibco or Lonza) supplemented with 10% Foetal Bovine Serum (FBS; Sigma-Aldrich or Lonza), 1% Glutamine (Sigma-Aldrich) and 1% Pen/Strep (Sigma-Aldrich or Lonza).

2.5.2 Minigene splicing assay

2.5.2.1 Generation of *ATR* constructs and cloning into *NF1* minigene plasmid

To generate the hybrid minigene constructs, human genomic DNA was amplified to generate two fragments containing *ATR* exons 18 and 28, respectively, along with 230 to 300bp of the intronic flanking sequences. Primers used in these reactions carry an NdeI site in their 5'-ends. This restriction site was used to clone the *ATR* constructs into the NdeI site of an *NF1* minigene vector (pTB NF1-29) obtained from Diana Baralle's laboratory (Raponi et al., 2009).

In this vector, *NF1* exon 29 had been cloned into a modified version of the *α-globin-fibronectin extra type III homology B* or *extra domain B* (*EDB*) minigene, in which the alternatively spliced *EDB* exon has been removed to generate a site for the insertion of the genomic sequence under study (Pagani et al., 2000).

Site-directed mutagenesis was then performed to introduce the ATR^{M1159I} and ATR^{K1665N} point mutations. Sequencing of amplified regions was carried out to ensure absence of undesired mutations. Primers used to generate *ATR* minigene constructs and used for site-directed mutagenesis are shown in Table S6.

2.5.2.2 Transient DNA transfections and minigene analysis

80-90% confluent HEK293T or HeLa cells were transfected with 1µg *NF1* or *ATR* minigene plasmids and 4µl of Lipofectamine (Invitrogen) in 0.5ml of Opti-MEM I (Invitrogen) per 35mm dish. Transfection reagents were added to cells in combination with 2ml DMEM without antibiotics, for a total volume of 2.5ml. 4-6h post transfection, media was replaced with complete DMEM media. Cells were collected 24h later for RNA extraction and RT-PCR analyses.

RNA extraction was performed using TRI reagent (Sigma), while retro-transcription reaction was done following manufacturers' instructions for High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). PCRs to study the minigene splicing were carried out using KOD polymerase (Novagen) and primers complementary to sequences in the minigene vector flanking the *ATR* /*NF1* constructs (Table S6).

2.5.3 CRISPR techniques

pX335 CRISPR plasmid (expressing nickase version of Cas9; Cas9-D10A mutant) used in this study was obtained from Addgene. Sequences for CRISPR gRNA oligos are shown in Table S6.

2.5.3.1 Oligo annealing and cloning into pX335 vector

25µl of forward and reverse gRNA oligos were mixed and incubated at 95°C for 5min. To allow oligo annealing, samples were cooled down slowly until 22-25°C. Oligos were treated with T4 Polynucleotide Kinase (PNK4, from NEB) and re-annealed using the same procedure.

2-4µg of pX335 vector were linearized by BbsI digestion, CIP (Calf Intestinal Alkaline Phosphatase, from NEB) treated and purified using Thermo Scientific kit for PCR clean up. A 1:200 dilution of annealed oligos was used for ligation into 50ng of purified vector. Resulting gRNA/Cas9 plasmids were sequenced to confirm presence of correct sequences with 'insert seq rev' (Table S6).

2.5.3.2 CRISPR transfections

CRISPR transfections were performed following the method established by Davis and Maizels, in which siRNA transfections (to knockdown *BRCA2*) or Rad51 inhibition were followed by DNA transfections (to introduce the Cas9 and gRNAs into cells) (Davis and Maizels, 2014).

Two separate transfections were carried out in an attempt to introduce the ATR^{Q2144R} mutation into human cells (Figure 18). The first transfection was done using RPE1 cells, which were treated separately with control or BRCA2 siRNA. In the second transfection, we used HCT116 cells and compared treatment with BRCA2 siRNA and Rad51 inhibitor (Rad51i). In both cases, cells were transfected with either gRNA1 only (upstream of the site to mutate) or a combination of gRNA1 and 2 (cutting upstream and downstream of the site to mutate).

Schematic of the two transfections and concentration of reagents used are summarized below in Figure 18 and Table 11, respectively; while full details of transfections are shown in the next sections.

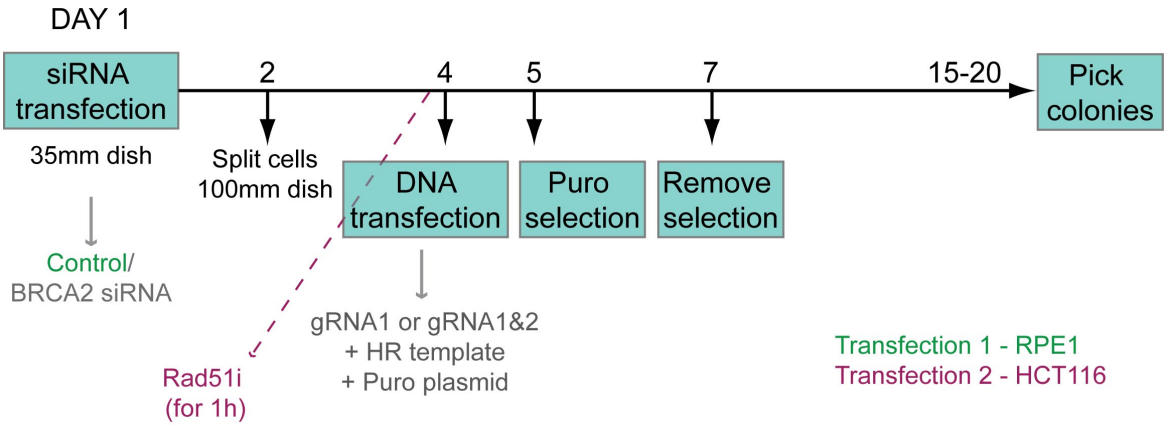


Figure 18. Schematic of CRISPR transfections. Both transfections were performed following the same timeline and protocol, except for the cell lines employed and the use of control siRNA or Rad51 inhibitor. These differences are marked in green for transfection 1 and purple for transfection 2. As a result, we had four different conditions in each transfection: 1) gRNA1 + siBRCA2, 2) gRNA1&2 + siBRCA2, 3) gRNA1 + siControl/Rad51i and 4) gRNA1&2 + siControl/Rad51i.

Table 11. Amounts of DNA, siRNA and HR template used in the CRISPR transfections.

Reagents	Concentration
gRNA/Cas9 plasmid	4µg gRNA1 or 2µg gRNA1 and 2µg gRNA2
pLox Puro plasmid	0.4µg
HR template	12.4nM of a 233nt ssDNA oligo
Control/BRCA2 siRNA	40nM
Rad51i	10µM

a) siRNA transfections

Control siRNA was obtained from Ambion (Silencer Negative Control #1 siRNA – AM4611), while BRCA2 siRNA was ordered from Dharmacon as a premixed pool of four siRNAs. BRCA2 siRNA pool was comprised of sequences: GAAACGGACUUGCUAUUUA, GGUAUCAGAUGCUUCAUUA, GAAGAAUGCAGGUUUAAUA and UAAGGAACGUCAAGAGAUA.

$2x10^5$ RPE1 or HCT116 cells were transfected with 40nM of siRNA (Ambion or Dharmacon) and 6µl of Oligofectamine (Invitrogen) in 0.8ml of Opti-MEM I (Invitrogen) per 35mm dish. 3h post transfection, 0.5ml of DMEM (without penicillin and streptomycin) supplemented with 30% FBS and 4mM L-glutamine was added to the media. After 24hr, 1ml of DMEM with 10% FBS and 1% penicillin-streptomycin was added.

If knockdown efficiency was analysed by western, 100µl of lysis buffer were added to each well and cells were scraped off the dish. A Bradford assay was performed to estimate protein concentration. Lysates were kept at -80°C until use.

b) Cas9/gRNA transfections

The day after siRNA transfections, cells were split into a 100mm dish. 24h later, 70-90% confluent RPE1 or HCT116 cells were co-transfected with the appropriate Cas9/gRNA plasmid/s, a puromycin selection plasmid (pLox Puro) and the HR template using the Lipofectamine reagent and incubated 24h at 37°C.

If Rad51 inhibition was carried out, cells were treated with 10μM Rad51i (Axon) for 1h before CRISPR transfection.

The following day, serial dilutions were performed and cells were plated in selection media (1μg/ml puromycin). After 48h, media was changed by complete DMEM F-12 with no puromycin selection. Cells were grown at 37°C until single colonies appeared.

Single colonies were picked from a 100mm dish using cloning disks (Sigma) and put into 24-well dishes in 1ml media. When confluent, clones were split into 12-well dishes in 2ml media and later expanded into 6-well dishes in 3ml media. Two confluent wells in a 6-well dish were saved for each clone: one for freezing the cells and a second one for gDNA extraction and further clone screening.

2.5.3.3 Clone screening

Selected clones were screened for the presence of the ATR^{Q2144R} mutation. gDNA was extracted using a TAIL-prep method (as in section 2.3.6) and PCR-amplified using primers F1.fw.noMreI and F2.rv.noMreI (Table S6). PCR products were subjected to diagnostic digests using both MfeI and BtgZI restriction enzymes (NEB). Positive clones identified by digest were then sequenced with primer Screening.rv (Table S6) to confirm the presence of the ATR^{Q2144R} mutation. PCR products were also cloned into pGEM-T easy to further identify by sequencing the mutations in each allele of *ATR*. Schematics of screening strategy can be found in Figure 46A.

2.6 Basic Protein Methods

2.6.1 Protein extraction

Protein extracts were prepared by lysing cells in lysis buffer for 45 min at 4ºC. Typically, ~1-2x10^7 DT40 cells were used. The cell pellet was resuspended in 50-100μl lysis buffer (depending on the pellet size). Lysates were spun down at 14,000rpm for 10-30 min at 4ºC. Protein concentration was determined by the

Bradford method and the supernatant was then used for downstream applications. If using extracts for western blotting, samples were stored in Laemmli buffer.

2.6.2 Western blotting

Samples were separated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) in 1xTG buffer with 0.1% SDS. ATR and BRCA2 were resolved using 6.5%, 80:1 acrylamide/bis-acrylamide gels; whist CHK1 was resolved using 12%, 37.5:1 acrylamide/bis-acrylamide gels. 50µg of protein extract and 5-10µl molecular weight markers (See Blue Protein Ladder, Invitrogen) were usually loaded per lane. Samples were transferred to a nitrocellulose membrane at 0.35mA for 2h on a BioRad apparatus at 4ºC in transfer buffer. Membranes were incubated in blocking buffer for 45min at room temperature. Primary antibody incubation was performed overnight at 4ºC following specific conditions in Table 6. Membranes were incubated with the relevant horseradish peroxidase-linked secondary antibody at 1:5000 dilution in 5% milk-PBS with 0.1% Tween20 for 1h at room temperature. Chemiluminescent detection was performed using an ECL kit according to the manufacturer's instructions (Amersham). When using phospho-specific antibodies, PBS was substituted by TBS to avoid phosphate from the buffer interfering with our results.

2.6.3 Protein purification

For protein purification, typically 1L (approximately $1x10^9$ cells) of HFSC-Atr (HA-FLAG-2xSTREP-CBP tagged Atr) DT40 cells were grown (Pessina and Lowndes, 2014). For this amount of cells, 5mL lysis buffer were used to lyse cells yielding to ~75mg of total protein.

The HFSC-Atr protein was purified using FLAG immunoprecipitation (IP), followed by STREP-IP. Prior to the purification, total cell extracts (TCE) were incubated with 120µl protein G sepharose beads (GE Healthcare) during a pre-clearing step to remove any unspecific interactors. The supernatant from the pre-clearing step (P-C) was then used in a FLAG-IP.

P-C extracts were incubated with 120µl FLAG beads (M2 FLAG beads, Sigma) for 2-3h at 4ºC and then washed 3x times (only the first wash was analysed by western – Wash). A supernatant sample (SN1) from the FLAG-IP was kept for analysis. HFSC-Atr was then eluted with 100µl lysis buffer containing a 100µg/ml FLAG peptide (3x FLAG peptide, Sigma) that serves as a competitor, displacing HFSC-Atr. The elution step was done three consecutive times to elute as much Atr as possible from the beads (E1-3).

FLAG elutions were then pulled and loaded into a STREP column (0.2ml Gravity flow Strep-Tactin sepharose columns, IBA). Flow through (FT2) was re-loaded onto the column to avoid losing any HFSC-Atr (FT3). Column was then washed 3x times (only the first wash was analysed by western – Wash) and protein eluted using 6x100µl consecutive rounds of elution (E1-6). Elution buffer contained desthiobiotin (IBA), which competes with the STREP tag on HFSC-Atr and releases it.

Purified HFSC-Atr protein was kept at -80ºC in elution or sample buffer, for microscopy analyses or western blotting, respectively.

2.6.4 Preparation of grids for negative staining electron microscopy (EM)

Once it was confirmed by western blotting that HFSC-Atr purification was successful, samples for microscopy were prepared. 5-10µl purified HFSC-Atr protein were applied to formvar/carbon-coated microscopy grids (Agar Scientific). Liquid was blotted off after 30-45sec and samples were stained with 2% uranyl acetate solution (negative staining) and incubated for 1min. Liquid was blotted off and grids were washed twice with milliQ H_2O. Samples were allowed to dry for a few minutes and were analysed under a transmission electron microscope (TEM).

Alternatively, microscopy grids were prepared using the ‘grid blotting’ technique described by (Knispel et al., 2012). Briefly, this is a method that allows the direct transfer of proteins separated by native electrophoresis to electron microscopy grids. HFSC-Atr samples were loaded twice (‘reference’ and ‘blotting’ lanes) on a native gel (NuPAGE Tris-Acetate gel run with Novex® Tris-Glycine Native Running Buffer, Invitrogen). The ‘reference’ lane was stained

with Silver staining (Sigma). HFSC-Atr protein was then blotted off from the 'blotting' lane directly onto the grid by comparison with the 'reference' lane. Samples were stained with uranyl acetate and analysed as previously described by TEM.

2.7 Flow cytometry Methods

2.7.1 Annexin V / PI staining

24h before analysis, cells were seeded at a density of $0.1x10^6$ cells/ml in 6-well dishes and, the next day, were treated with the relevant drugs. About $2x10^6$ cells were resuspended in 50µl $1xCa^{2+}$ buffer containing 1µl Annexin V – fluorescein isothiocyanate (FITC) peptide and incubated in the dark at room temperature for 15min. 300µl $1xCa^{2+}$ buffer containing 660ng/ml propidium iodide (PI) were added and the samples were measured by flow cytometry immediately. The readings were taken using a FACS Canto II flow cytometer and analysed using BD FACSDiva software, representing the data as a dotplot with the PI signal (DNA staining), versus the FITC channel (Annexin V staining). In order to represent the data easily we calculated the percentage of Annexin V positive cells represented on the dot plot. An example of the raw data can be found in Figure 25A (Chapter 3).

2.7.2 Fixed cell flow cytometry

In all cases, approximately $1x10^6$ cells were fixed in cold PBS and 70% ethanol and stored at 4ºC. For analysis, cells were first washed in 1ml PBS containing 1% BSA, then resuspended in 50µl PBS with 1% BSA and 0.1% Triton X-100 containing anti-H3S10ph or anti-BrdU antibodies, and incubated for 1–2h at room temperature. Cells were then washed twice in 1ml of PBS containing 1% BSA, then resuspended in 50µl PBS with 1% BSA and 0.1% Triton X-100 containing appropriate secondary antibodies, and then left for 1h on ice in the dark. Cells were then washed twice in 1ml of PBS containing 1% BSA and resuspended in 300-500µl of PBS containing 40µg/ml propidium iodide (Sigma) and 250µg/ml RNaseA (Qiagen). Samples were then measured using a

FACS Canto II flow cytometer and analyzed with BD FACSDiva software. Examples of the raw data can be found in Figure 44A (BrdU/PI) and B (H3S10ph/PI) (Chapter 5).

CHAPTER 3

Structure-function studies of Atr reveal its functions are dependent upon the integrity of its HEAT repeats

Running title: Structure-function studies of Atr

Keywords: ATR, Auxin Inducible Degron tag, DNA Damage, DT40, HEAT repeats, Mec1, PIKK, 'window' deletions

3.1 Introduction

3.1.1 Relation of PIK kinases function and structure

The principal regulators of the DNA damage response pathway (DDR) are large proteins, which belong to the phosphatidylinositol 3-kinase-like protein kinase (PIKK) family and share analogous structures (Baretić and Williams, 2014; Lovejoy and Cortez, 2009; Rivera-calzada et al., 2015) (Figure 19). Within the PIKK family, the C-terminal domain is highly conserved and consists of a relatively small kinase domain. On the other hand, the N-terminal regions are highly variable in sequence; nevertheless, a novel sequence algorithm determined that the non-kinase regions in the PIKK family proteins are composed of conserved helical HEAT repeats, suggesting that all family members have a related underlying structure (Perry and Kleckner, 2003).

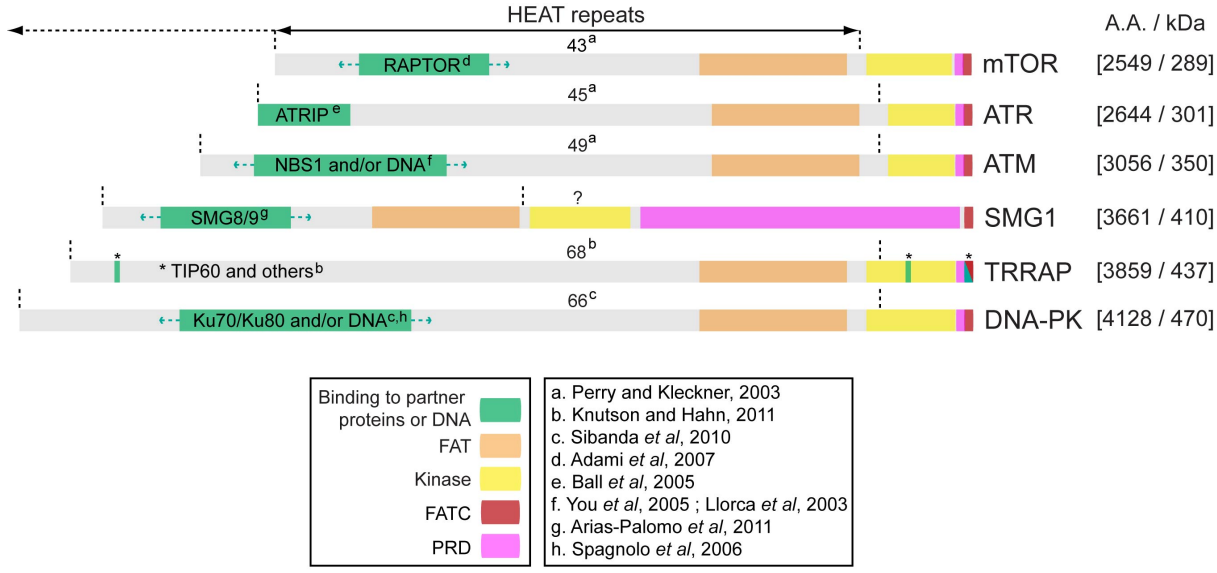


Figure 19. Structural features shared amongst PIKKs, adapted from (Baretić and Williams, 2014). PIKK domains are marked in different colours: kinase (yellow), FAT (orange), FATC (red), PRD (pink) and domains known to interact with partner proteins (green). Green arrows indicate that domains could comprise bigger or smaller HEAT-rich regions. Units that make up the HEAT repeat domains are indicated for each kinase (except for SMG1, for which information was not found in literature), together with its total length in amino acids and kDa. All PIKKs have analogous structures, except for SMG1, which contains a much larger PRD. mTOR, ATR and ATM are slightly smaller proteins, in the order of 300kDa; while SMG1, TRRAP and DNA-PK are larger than 400kDa. References regarding HEAT repeat number and partner interactors are marked with small letters.

Although PIKKs are involved in multiple key cellular processes, relatively little is known about their configuration. In particular, the major portion of the HEAT repeat-containing regions of these proteins still remains uncharacterized. The non-kinase part of ATR is made up of 45 HEAT repeats, seven of which are located at the N-terminus and contain the ATRIP interacting region (Ball et al., 2005). However, no role has been attributed to the remaining 38, of which three are specific to ATR homologues. During their studies, Perry and Kleckner identified common HEAT repeat units between ATR, ATM and mTOR, while they also found HEATs that were unique to each of the families. Regarding the ATR kinase, HEAT repeats 11 to 13 are only present in ATR and homologue proteins, but not in other PIKKs (Perry and Kleckner, 2003).

3.1.1.1 HEAT repeats as mediators of protein-protein interactions

HEAT repeat motifs are thought to be important for protein-protein interactions (Chook and Blobel, 1999; Cingolani et al., 1999), which could be crucial for PIKK function. Alternatively, HEAT repeats could be purely structural, allowing proper orientation of the functional parts of the protein (Grinthal et al., 2010).

Apart from structural analyses that show HEAT repeats mediate interaction between some proteins and the plasma membrane (Kunz et al., 2000; Vastiau et al., 2005), there are also studies that have investigated HEAT repeat clusters in PIKKs that mediate protein-protein interactions. In *S. cerevisiae* it has been shown by two-hybrid assays that Mec1-Ddc2 interaction is dependent on the first 500 amino acids of Mec1 N-terminus (Wakayama et al., 2001). Although this process was not attributed directly to the presence of HEAT repeats, this region of Mec1 corresponds to HEATs number 1-10 and it contains sequences required for interaction with Ddc2, the ATRIP homologue of yeast. Similarly, in human cells, the N-terminal region of ATR comprising amino acids 30-346 (HEATs 1-7) is involved in binding to ATRIP (Ball et al., 2005).

As further evidence of HEAT repeats mediating protein-protein interactions, studies of the PIKK Tel1 in *S. pombe* show that HEAT repeats 17, 18, 21 and 22 mediate its interaction with Nbs1 (You et al., 2005), which is part of

the MRN complex and is involved in DNA double-strand break repair (Rupnik et al., 2008). Hence, HEAT repeats are likely to have crucial functions in both mediating different protein-protein interactions as well as being required for the three-dimensional shape of the protein. Protein interactions and conformation might not be mutually exclusive and could continue to regulate different functions of the PIKK proteins. However, further experimental data on the structure and functions of these PIKK proteins, in particular for ATR, will be needed to clarify this matter.

Some groups have already carried out some structure-function analyses by deleting or replacing specific regions of the PIKKs (though not necessarily specific HEAT regions), in attempt to identify a particular region of the protein with a precise function. For example, the ATR FATC domain has been replaced by the corresponding region of ATM, resulting in a kinase-dead mutant (Mordes et al., 2008). Surprisingly, the opposite experiment, replacement of ATM FATC domain with the equivalent domain of ATR, DNA-PK or TRRAP, resulted in a functional protein. In particular, the FATC domain of ATM has been reported to mediate its interaction with TIP60, allowing its activation and further acetylation and autophosphorylation. ATM hybrids carrying the FATC domain of ATR, DNA-PK or TRRAP were still able to interact with TIP60 and retained an active kinase in response to damage (Jiang et al., 2006).

In other analyses, the ATRIP interacting domain of ATR was attached to the N-terminus of otherwise intact ATM ($ATM^{ATRIP\text{-}BD}$). The fusion protein was able to bind ATRIP and associate with RPA-coated ssDNA. It also gained the ability to localize efficiently to stalled replication forks as well as double strand breaks. However, although attaching the ATRIP binding domain of ATR to full-length ATM resulted in a gain of function, this chimeric protein did not substitute for either ATM or ATR roles. The addition of this domain onto ATM may block NBS1 access, disrupting the checkpoint activation and thereby CHK2 phosphorylation in response to damage. Furthermore, cells expressing the $ATM^{ATRIP\text{-}BD}$ protein were not able to phosphorylate CHK1 or support cell viability in the absence of ATR. This could be partly explained because the kinase activity of $ATM^{ATRIP\text{-}BD}$ chimera cannot be stimulated by TOPBP1 (Chen et al., 2007).

In this project we are interested in studying the role of HEAT repeats in ATR function and recruitment of downstream targets of ATR.

3.1.2 Deletion mutants

Over the years, different strategies have been developed in order to study protein function. In particular, to investigate the role of particular motifs in a protein, internal or 'window' deletions can be generated in order to associate the loss of a specific function with a particular deleted region. This methodology of creating internal deletions in a protein is widely used in basic research to gain knowledge of protein function and interactions.

For example, a recent study in *S. pombe* has used this methodology to analyse the importance of conserved Chk1 domains. Using internal deletion mutants, they have been able to show the importance of different domains in Chk1 regulation by mediating the interaction with Crb2 (53BP1's homologue) (Caparelli and O'Connell, 2013). Other studies have also used this technique to examine the role of different proteins; such as REV1, RAD17 and the NEK1 kinase, in the ATR-dependent DNA damage response pathway (DeStephanis et al., 2015; Lee et al., 2007; Liu et al., 2013).

In 2010, an American group performed a very detailed structure-function study of Tra1, the budding yeast homologue of TRRAP, another PIK kinase involved in development. They designed 44 internal deletions spanning Tra1 length, 35 of which were located in HEAT-rich regions. Using this approach, they were able to show that some HEAT repeats were involved in the interaction with proteins such as the TIP60 homologue. Besides, most deletion mutants were non-viable (Knutson and Hahn, 2011), supporting the idea that HEATs might not only be essential for mediating certain protein-protein interactions, but also to hold PIKK structure.

Using a similar strategy, wherein a set of 'window' deletion mutants was generated by deleting individual or groups of different HEAT repeats, we were able to show, for the first time, that the ATR protein does not tolerate deletions in its structure. Parallel studies in our lab have also shown equivalent results for ATR's yeast homologue, Mec1 (De Castro Abreu, 2012). Hence, as for other

PIKKs, we have been able to demonstrate that the function of ATR is also dependent on the integrity of the HEAT repeat motifs. While the deletion of single or groups of HEAT repeats leads to loss of all ATR functions; the replacement of these by equivalent motifs, at least between human ATR and chicken Atr, allows for proper protein function.

3.2 Results

3.2.1 Comparison of *Homo sapiens* and *Gallus gallus* ATR proteins

To study the role of HEAT repeats on ATR structure, we used the genetically tractable chicken DT40 vertebrate model system. As explained in the introduction, we will refer to the *H. sapiens* genes/proteins in capital letters (ie. *ATR*/ATR), while the *G. gallus* homologues will only have the starting letter in uppercase (ie. *Atr*/Atr).

Protein alignments have shown that the full length ATR protein is 77.2% identical between the human and chicken species (Figure 20A), within which the identity is highest in the C-terminal conserved kinase domain. We have also compared the ATR/Atr proteins in terms of their HEAT repeat composition (Figure 20B). In general terms, the N-terminus is well conserved, albeit not as well conserved as the kinase domain. Conservation typically exceeds 70% identity, except for HEAT repeats 3 to 14, which include the ATR-specific HEAT repeats (11-13). Repeats 11-13, whose function is not well understood, are specific for ATR and its homologues, but are not conserved in other proteins from the PIKK family such as ATM or TOR1 (Perry and Kleckner, 2003). From HEAT 15 onwards, conservation goes from 70 up to 100% for some of the HEAT repeats within the FAT domain, consistent with the C-terminal two thirds of the ATR/Atr proteins being more conserved than the N-terminal third. This may be due to the proximity of these HEAT repeats to the kinase domain, for instance, they could be important for holding overall structure of the protein or to support the catalytic function of ATR.

The high identity between the human and chicken proteins supports the assumption that DT40 is a good model to study ATR function and structure.

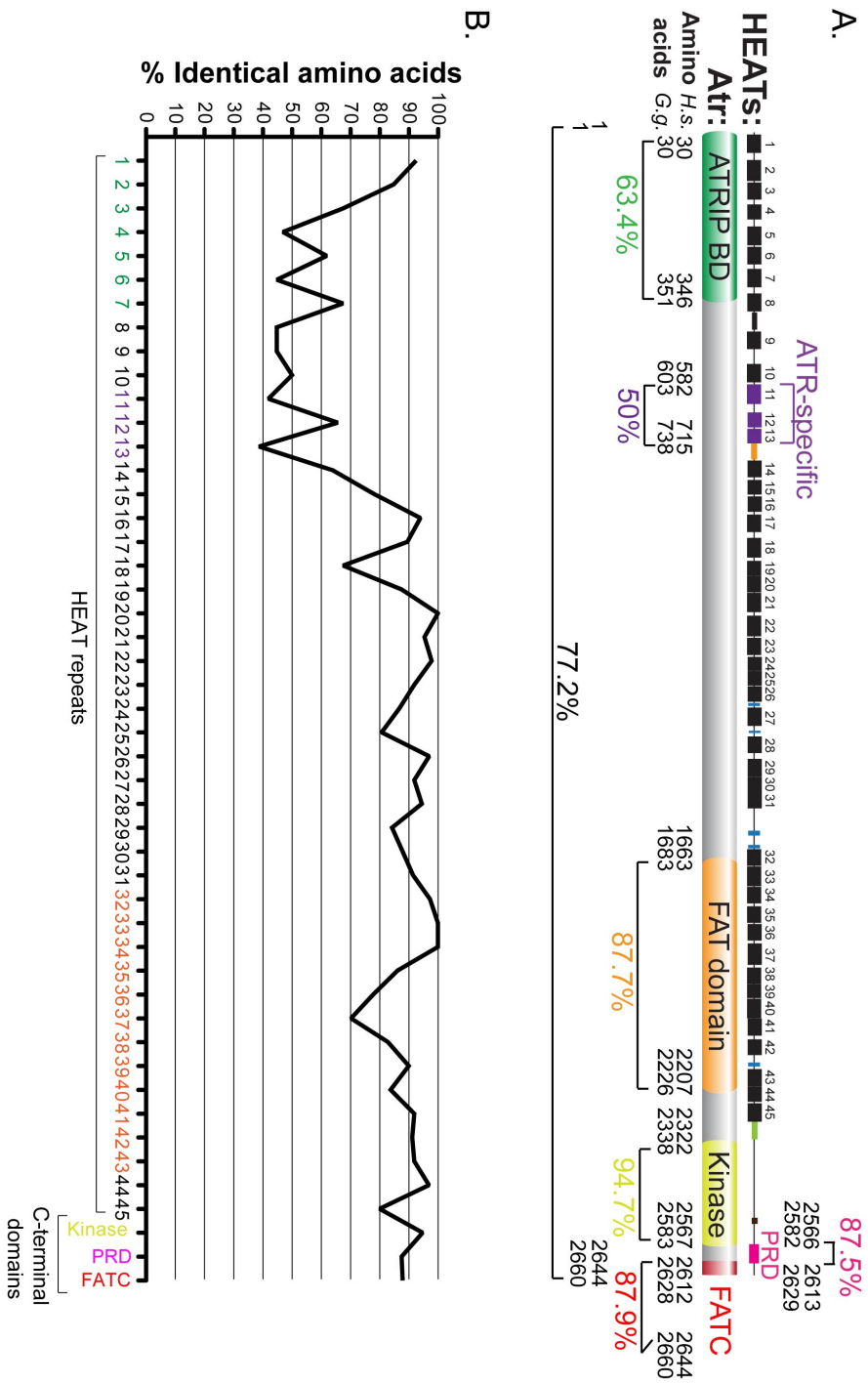


Figure 20. Comparison of human ATR and chicken Atr.

A. Overall identity between ATR/Atr proteins (amino acid level), as well as identity for each domain of the ATR protein was calculated using the SIM alignment tool. Different domains and important features of ATR are colour-coded as follows: kinase (yellow), FAT (orange), FATC (red), PRD (pink), ATRIP-interacting (green) and ATR-specific HEAT repeats (purple).

B. Detailed comparison of ATR/Atr proteins motifs. Identity for each individual HEAT repeat is shown in the graph, as well as for the C-terminal domains of the protein. Identity was also calculated by aligning each individual HEAT motif/domain with the SIM alignment tool.

3.2.2 Generation of window deletion mutants in DT40 cells

3.2.2.1 Parental and control cell lines

As *ATR* is an essential gene, we used a conditional-null cell line to analyse the effect of the window deletion mutants on ATR's functions. This system was previously established in our laboratory and it is based on the auxin degron system developed by Nishimura and colleagues (Nishimura et al., 2009). Upon AUX addition, the AID-tagged Atr protein (AID-Atr) is targeted for degradation by the proteasome (Eykelenboom et al., 2013).

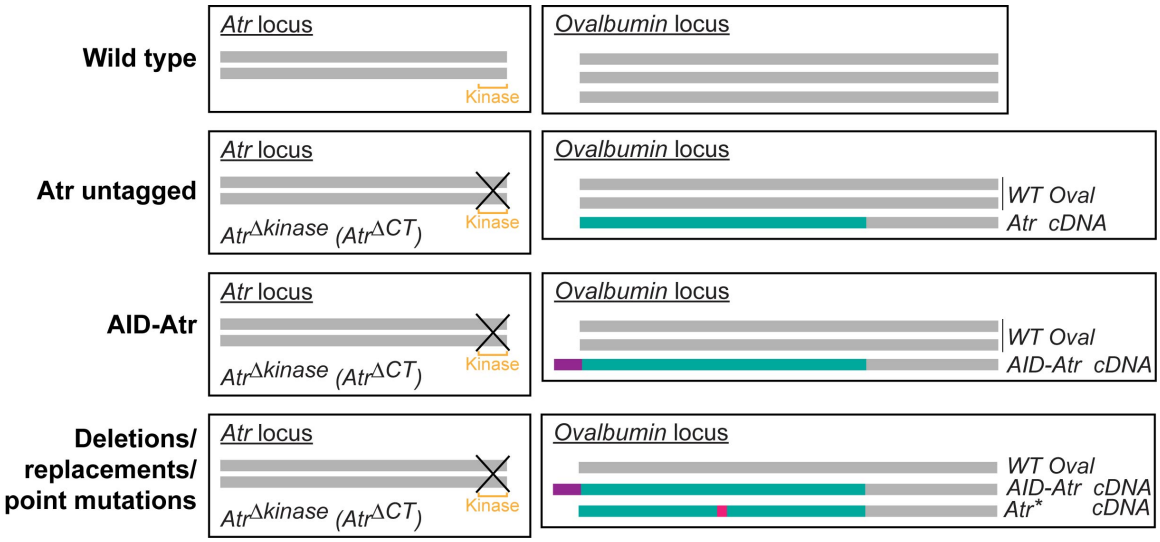


Figure 21. Diagram illustrating the generation of the DT40 cell lines used in this study. The endogenous copies of *Atr* and *Ovalbumin* loci are shown in grey, while the *Atr* cDNAs targeted to the *Ova* locus are shown in light blue. Other features are highlighted as follows: AID tag (purple), *Atr* mutations (pink) and the Atr kinase domain (yellow).

To generate the AID-Atr cell line, a former postdoctoral researcher first introduced an untagged *Atr* cDNA into one of the copies of the *Ovalbumin* locus, which is located in the trisomic chromosome two of DT40 cells. This was followed by the deletion of the kinase domain of both *Atr* endogenous loci, giving rise to an Atr untagged cell line. Generation of the AID-Atr cell line was then achieved by replacing the *Atr* cDNA in the *Ova* locus by the *AID-tagged Atr* cDNA (Eykelenboom et al., 2013). Last, the AID-Atr system was used as a parental cell line to generate all DT40 cell lines presented in this and following chapters. This was performed by introducing a copy of the corresponding mutant

Atr cDNA into a second copy of the *Ova* locus. Generation of all mutant Atr cell lines used in this study is illustrated in Figure 21.

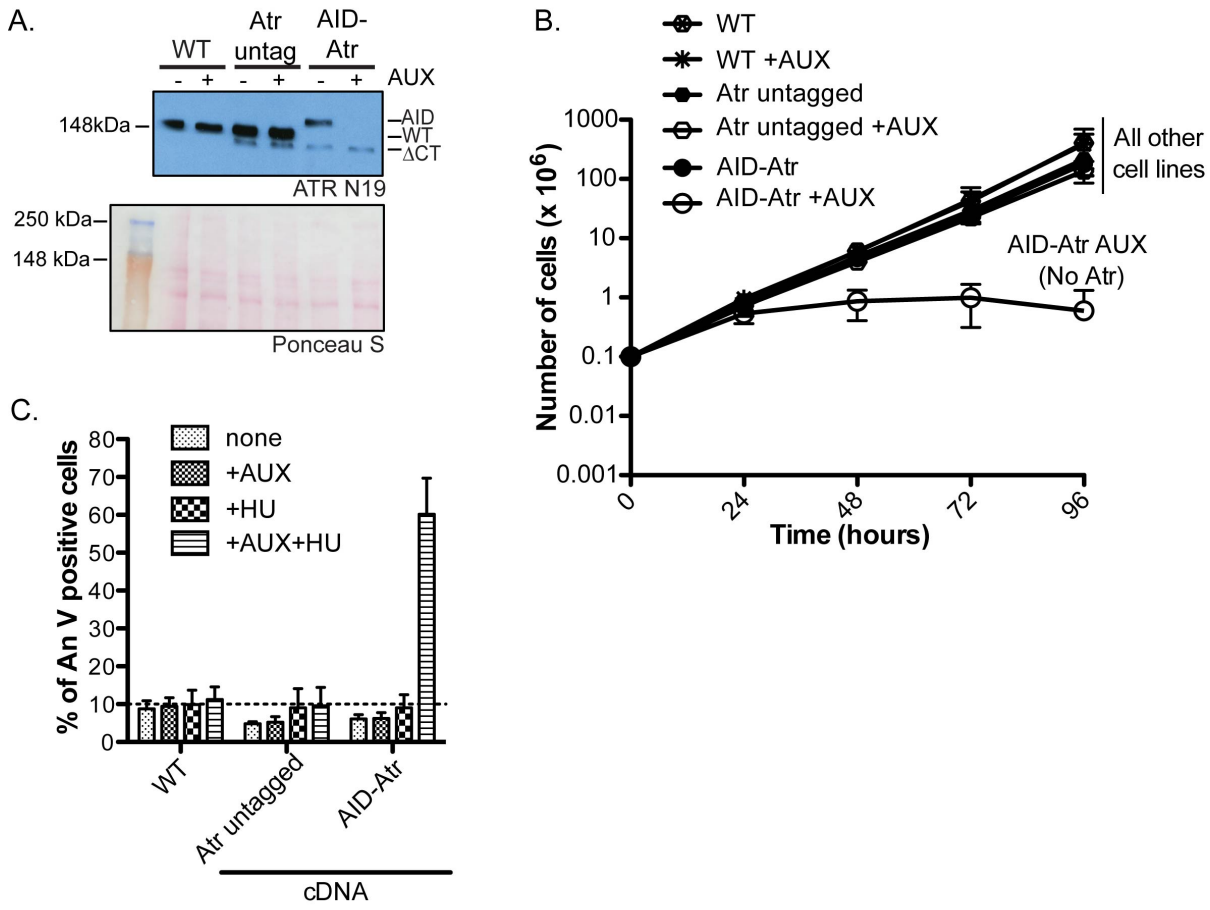


Figure 22. Comparison of Atr untagged (used as a 'WT' control), AID-Atr (used as a negative 'no Atr' control) and WT DT40 cell lines, regarding the expression of the Atr protein, cell proliferation and apoptosis in response to HU treatment.

A. Western blotting analysis of total cell extracts prepared from Atr untagged, AID-Atr and WT cells using the ATR-N19 antibody (Santa Cruz).

B. Growth analysis of indicated cell lines following treatment with AUX. Cell numbers were recorded every 24h for 5 days and are corrected for culture dilution. Error bars represent the SD from at least three experiments.

C. HU induced apoptosis. Annexin V/propidium iodide (PI) flow cytometry analysis after treatment with HU (4h) following growth in the presence or absence of AUX (2h). Quantification was done by measuring Annexin V positive cells. Error bars represent the SD from at least three experiments.

As the Atr untagged cell line expresses no depletable Atr, it was used as a wild type control in our experiments. As seen in Figure 22A, AUX only induces

the degradation of AID-Atr, while it has no effect on Atr expression in WT or Atr untagged cell lines. In the absence of AUX, Atr expression is similar between the three cell lines, although a loading control is needed to clarify this matter. Besides, as indicated in Figure 21, all cell lines generated still carry a C-terminally truncated form of Atr, which is detectable by western as a faster migrating band for Atr untagged and AID-Atr cell lines (Figure 22A). This Atr isoform is not able to support Atr viability or checkpoint functions (Eykelenboom et al., 2013), but it remains to be tested if it retains the ability of interacting with any ATR partners.

As previously confirmed in the laboratory, the AID-Atr cell line grows at the same rate as WT cells in the absence of AUX. In addition, both the Atr untagged and WT cell lines grow equally in the absence or presence of AUX (Figure 22B). Atr untagged also behaves similarly to WT cells in response to HU, being able to activate the checkpoint upon treatment with this DNA damaging agent, as confirmed by analysis of apoptosis (Annexin V/PI staining; see further details on this assay in section 3.2.5). On the other hand, the AID-Atr cells are not able to cope with the HU-induced damage after AUX treatment (Figure 22C). This confirms that the Atr untagged cell line can be used equivalent to a 'wild type' control and that AUX on its own does not affect cell viability (Eykelenboom et al., 2013).

3.2.2.2 Window deletion mutants ($Atr^{\Delta HEATs}$)

To obtain the desired stable cell lines in DT40 cells, we separately targeted five different $Atr^{\Delta HEATs}$ cDNAs to the *Ovalbumin* locus of the AID-Atr DT40 cell line. These cDNAs contained several deletions, which spanned different sets of HEAT repeats, as indicated in Figure 23 (see information on amino acids, kDa and HEAT repeat composition). As a result, the individual final cell lines contain copies of both *AID-Atr* and the $Atr^{\Delta HEATs}$. To study the effect of the Atr HEAT repeat deletions, AID-Atr is depleted by the addition of AUX. As previously mentioned, cells still express a C-terminally truncated version of the endogenous copies of the *Atr* gene, which lacks the kinase domain and is not able to support Atr roles (Eykelenboom et al., 2013).

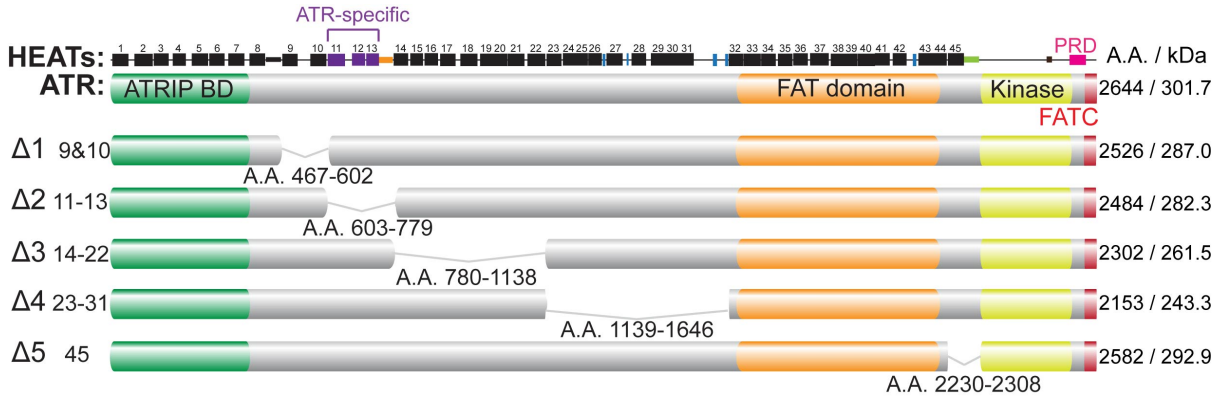


Figure 23. Schematic of window deletions. Different sets of HEAT repeats were deleted in the Atr protein. HEAT repeats and amino acids deleted in each mutant are shown in the figure, as well as predicted protein sizes, both in amino acids and kDa.

The screening of positive clones was performed by Southern blotting of genomic DNA after a BamHI digest. The probe used in the analysis recognises a specific region of exon eight of the *Ovalbumin* gene (Figure 24A). The AID-Atr parental cell line used for transfection contains two copies of the non-targeted *Ovalbumin* genomic locus and one of the *AID-Atr* targeted *Ovalbumin* locus (Figure 24B: lane 2). Therefore, since there are three copies of this gene, positive clones are expected to have three different BamHI restriction patterns due to targeting of the $Atr^{\Delta HEATs}$ cDNAs (Figure 24B: lane 3-7).

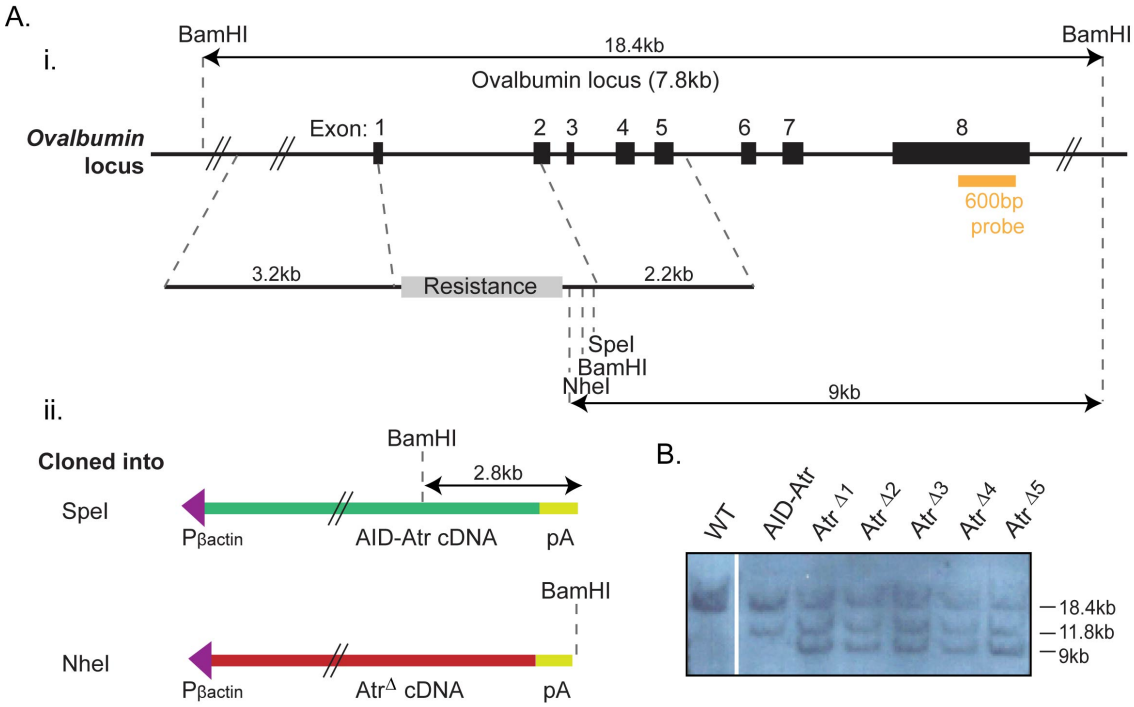


Figure 24. Strategy for the generation of Atr$^{\Delta HEATs}$ mutant cell lines.

A. Schematic of the *G. gallus Ovalbumin* locus and the targeting vector (i) and *AID-Atr* and *Atr$^{\Delta}$* expression cassettes (ii). Relevant restriction sites (as grey dotted lines) and Southern probe binding sites (in orange) are indicated. Relevant BamHI site and subsequent fragment expected following targeting of the construct to the *Ovalbumin* locus are also shown in the figure. For a non-targeted *Ovalbumin* genomic locus, the expected band is 18.4kb. For *Atr$^{\Delta HEATs}$* cDNAs targeting, the expected band is 9kb; while for the *AID-Atr* cDNA targeting, already present in parental cell line, the expected band is 11.8 kb (9 + 2.8 kb). The *Atr$^{\Delta HEATs}$* cDNAs were cloned into the NheI site of the targeting vector, whereas the *AID-Atr* cDNA was introduced in the SpeI restriction site. That difference in the cloning site corresponds to a different pattern after the BamHI digest.

B. Southern screening of Atr$^{\Delta HEATs}$ clones for *Ovalbumin* targeting. The first lane represents wild type cells; the second lane corresponds to the AID-Atr parental cell line and lanes three to seven represent targeting to the *Ovalbumin* locus of *Atr$^{\Delta 1}$ to Atr$^{\Delta 5}$* cDNAs.

3.2.3 Atr window deletion mutants are stably expressed in DT40 cells

Once the Atr$^{\Delta HEATs}$ mutant cell lines were generated, a preliminary characterisation of *Atr* expression was carried out. Previous studies in this laboratory, using the budding yeast (*S. cerevisiae*) model system, in which the ATR homologue is Mec1, showed that similar *mec1* deletion mutants led to expression of unstable proteins (De Castro Abreu, 2012). Therefore, it was necessary to confirm that these mutant versions of *Atr* were being stably expressed in the DT40 cell lines.

To check for stable expression, we performed western blotting using an antibody raised against the N-terminus of the ATR protein (ATR-N19 obtained from Santa Cruz). As shown in Figure 25, the AID tagged form of Atr is expressed in all mutant and AID-Atr parental cell lines, as expected. During the SDS-PAGE, the AID-Atr protein migrates more slowly than the WT form of Atr present in the Atr untagged cell line. In the cell lines that express a truncated form of Atr ($Atr^{\Delta CT}$) in place of WT Atr due to disruption of both loci, the WT protein is replaced by the C-terminally truncated form lacking its kinase domain. The $Atr^{\Delta CT}$ form can easily be distinguished in the Atr untagged and AID-Atr cell lines as the fastest migrating band. The third band corresponding to the specific window deletion mutants ($Atr^{\Delta HEATs}$; pointed out by the white arrowheads) migrates according to the overall size of the mutant protein (Figure 25A).

Despite the fact some of our deletions were in the N-terminus of the protein, and that the epitope recognised by the antibody was not known, in each mutant cell line we were able to detect stable truncated $Atr^{\Delta HEATs}$ proteins that correspond to the expected size. Thus, contrary to the results obtained in yeast, we could detect by western blotting the Atr-window deletion forms for all the five mutants that were introduced in the chicken DT40 cell line.

As mentioned above, it should also be noted that in each mutant cell line (expressing both AID-Atr and the $Atr^{\Delta HEATs}$ forms), a C-terminally truncated Atr form lacking the kinase domain is also expressed from the endogenous alleles ($Atr^{\Delta CT}$). In three of the mutants ($Atr^{\Delta 1}$, $Atr^{\Delta 2}$ and $Atr^{\Delta 5}$) the endogenous truncated form migrates at or close to the expected size for the window deletion generated. In these cases we have not been able to achieve sufficient resolution to separate these two forms. However, the intensity of the band increases in these three lanes, which is consistent with two proteins migrating through the gel at the same rate. Therefore, we believe all the window deletion forms of Atr are stably expressed at least at levels equivalent to $Atr^{\Delta CT}$.

We attempted to generate an antibody raised against the C-terminus of Atr to confirm the stable expression of the window deletion mutants, but we were unsuccessful (data not shown).

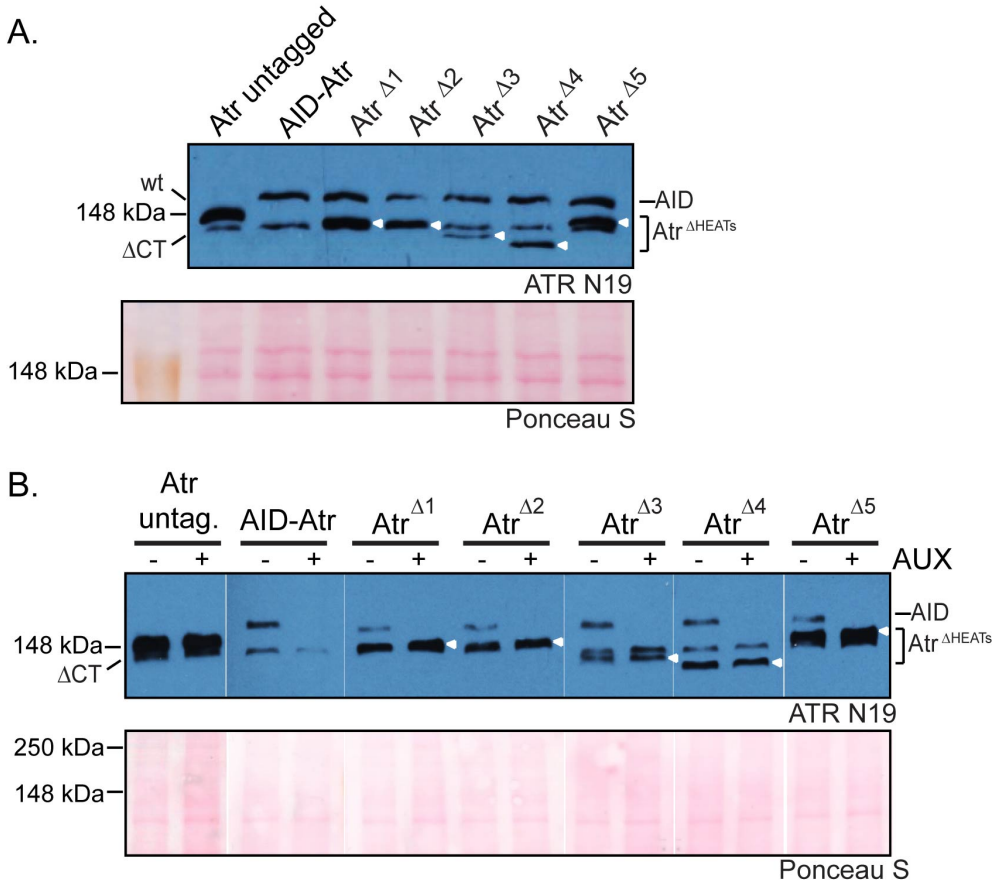


Figure 25. Expression of Atr$^{\Delta HEATs}$ proteins.

A. Western blotting analysis of total cell extracts prepared from Atr untagged, AID-Atr and Atr$^{\Delta HEATs}$ cells using the ATR-N19 antibody (Santa Cruz).

B. Western blotting analysis of the indicated cell lines in the absence and presence of AUX (0.5mM AUX treatment for 2h).

Atr isoforms corresponding to window deletions are marked with white arrowheads in both panels. The protein from the endogenous Atr loci, which lacks the kinase domain (Atr$^{\Delta CT}$), serves as a loading control in both westerns. Note that deletion mutant forms Atr$^{\Delta 1}$, Atr$^{\Delta 2}$ and Atr$^{\Delta 5}$ co-migrate with the Atr$^{\Delta CT}$ form. For this reason, Ponceau S staining is also shown in the figure.

The Atr window deletion expression tests were also performed in the absence and presence of AUX (Figure 25B). The equivalent *Mec1* mutations showed that the window deletion mutant proteins were not stably expressed in yeast (De Castro Abreu, 2012). However, the yeast cell line used did not express wild type Mec1. *Mec1* is an essential gene, disruption of which causes lethality (Kato and Ogawa, 1994); nevertheless, this lethality can be overcome if the gene *sml1* is also disrupted (Zhao et al., 1998). Hence, the background in which the *Mec1* mutants were analysed was Δ*sml1* and Δ*mec1*.

In our case, the Atr cell lines harbouring window deletions Δ1-5 have a functional Atr (AID-Atr) that, possibly, given the yeast results, may lead to improved stability of the mutant form. Consequently, to confirm that AID-Atr was not contributing to the Atr mutants' stability, we also had to formally prove that these proteins were still detectable after the depletion of AID-Atr following AUX treatment. Our results show that, addition of AUX to the media leads to the degradation of AID-Atr; nevertheless, loss of AID-Atr does not affect the untagged Atr or the $Atr^{\Delta HEATs}$ proteins. In the absence of AID-Atr, the Atr window deletion forms seem to remain stable and can be detected by western blotting.

3.2.4 Atr window deletion mutants cannot support viability

To test the ability of the different $Atr^{\Delta HEATs}$ mutants to substitute for the essential role of Atr, growth of the cell lines was measured in the presence or absence of AUX. When no AUX is added to the medium, all cell lines grow as control cells (AID-Atr). On the other hand, following AUX treatment, the AID-Atr cell line and all the *Atr* window deletion mutants displayed a gradual decline in growth, culminating in almost no survival after 96 hours of treatment (Figure 26).

We noticed that there were some mutants ($Atr^{\Delta1}$, $Atr^{\Delta2}$ and $Atr^{\Delta3}$, corresponding to deletions immediately N-terminal, including the ATR-specific HEAT repeats) that seemed to be more severely affected by AUX than the parental AID-Atr cell line. While we would need to confirm these differences by analysing independently generated clones, a possible explanation for these differences in sensitivity to AUX is that these particular deletions within Atr could have an extra negative effect. Therefore, it is possible that, after addition of AUX, some of the mutant cell lines die quicker than the AID-Atr cell line in the presence of an Atr window deletion, particularly $Atr^{\Delta1}$, $Atr^{\Delta2}$ and $Atr^{\Delta3}$.

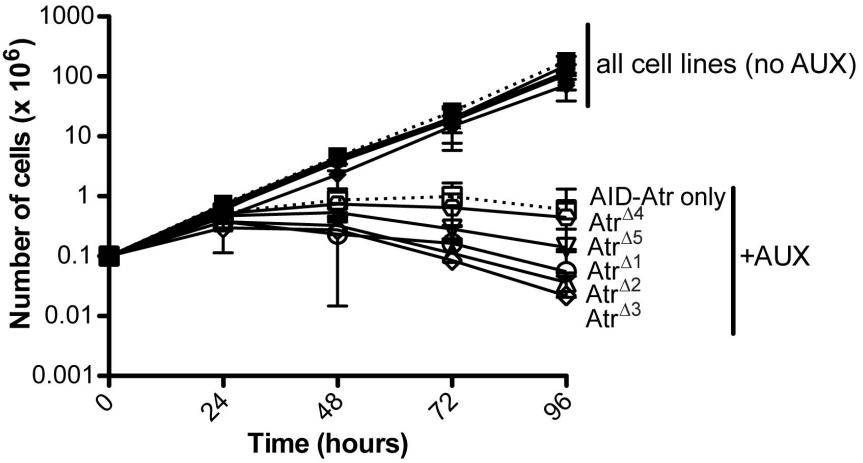


Figure 26. Growth analysis of Atr window deletion mutant cell lines following treatment with AUX. Cell numbers were recorded every 24h for 5 days and are corrected for culture dilution. Error bars represent the SD from at least three experiments.

3.2.5 Atr window deletion mutants are not able to stimulate the S-phase checkpoint in response to HU

It has been suggested that one of the essential roles of ATR could be the monitoring of DNA replication to avoid the onset of mitosis in the presence of unreplicated DNA (Eykelenboom et al., 2013; Zhou and Elledge, 2000). However, it is not well established whether the essential role of ATR is related or not to its checkpoint roles. For this reason, we wanted to test if any of the $Atr^{\Delta HEATs}$ mutants, which cannot perform the essential function of Atr (section 3.2.4), might be able to perform the checkpoint functions of Atr. Similar separation-of-function mutations in the *ATR/Mec1* genes have been reported in the past (Luzwick et al., 2014; Nam et al., 2011a; Paciotti et al., 2001) (see general introduction).

Therefore, to further characterize the Atr^{HEATs} mutants, we carried out an investigation of their ability to stimulate the S-phase checkpoint in response to DNA damage. In our experiments, we used hydroxyurea treatment to study the checkpoint functions of Atr. HU inhibits ribonucleotide reductase, causing the depletion of dNTPs and leading to replication fork arrest (Håkansson et al., 2006). We measured apoptosis by flow cytometry analysis (in particular, Annexin V/PI staining) and Chk1 phosphorylation by western blotting as readouts of Atr checkpoint function (Figure 27).

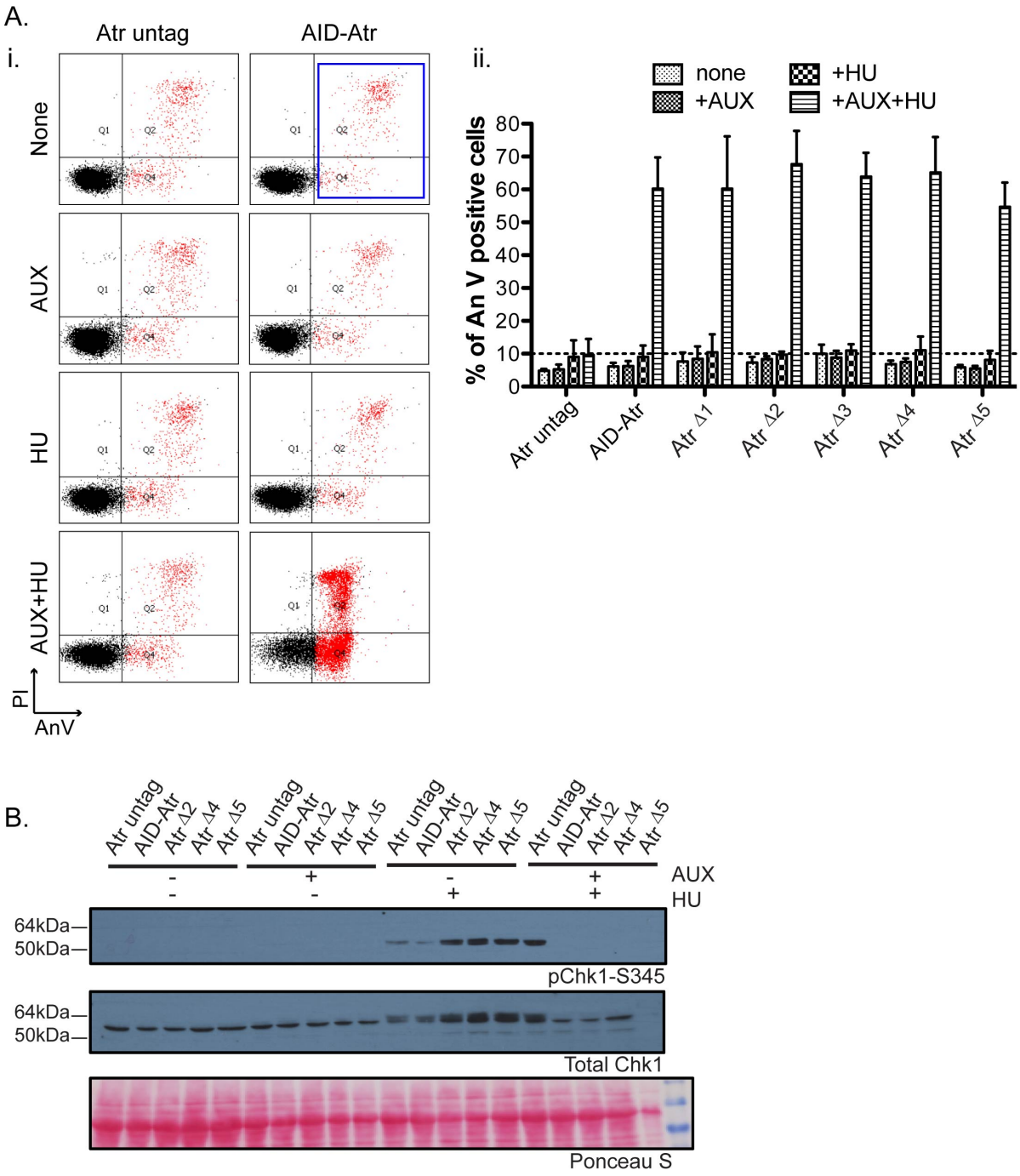


Figure 27. Checkpoint activation in response to HU in the Atr window deletion mutant cell lines.

A. Apoptosis assay. Annexin V/propidium iodide (PI) were analysed by flow cytometry following treatment of cells with 1mM HU (4 hr) following growth in the presence or absence of AUX (2 hr). i. Examples of flow cytometry raw data for the AID-Atr and Atr untagged samples are shown in the figure. The $Atr^{\Delta HEATs}$ mutants behave similarly to the AID-Atr cell line. By representing the data as a dot plot of PI signal versus Annexin V signal (FITC labeled), it was possible to distinguish between: Q1, PI + AnV - cells. There should not be any cell in this quadrant, since every cell whose membrane allows passage of PI would allow passage of Annexin V-FITC; Q2, PI + AnV + cells. This quadrant corresponds to the late apoptotic or dead cells; Q3, PI - AnV -. These are the cells that are alive; Q4, PI - AnV + cells. These are the early apoptotic cells. Quantification of total

apoptosis was performed adding Q2 and Q4 values, which correspond to the AnV positive cells (blue quadrant). ii. Quantification of HU-induced apoptosis. Percentage of Annexin V positive cells was plotted for the cell lines indicated. Error bars represent the SD from at least three experiments.

B. Western analysis of Chk1-S345 phosphorylation upon HU treatment (4h) following growth in the presence or absence of AUX (2h). Western blotting of protein extracts harvested from cell lines as indicated. A Chk1 antibody that recognises phosphorylation on S345 (Cell Signalling) was used to analyse checkpoint proficiency. An antibody against total Chk1 (FL-476 Santa Cruz) was used as a loading control.

Our flow cytometry analysis shows that in untreated and AUX treated cells, there is a basal level of apoptosis/cell death that reaches between 5 and 10%. When cells are treated with HU, there is a slight increase in cell death. However, in this case, AID-Atr has not been depleted, and so all the cell lines have a functional checkpoint. After the depletion of AID tagged Atr and the induction of damage (+AUX +HU samples), there is a significant increase in cell death (up to 55-70%) in both the AID-Atr and mutant cell lines (Figure 27Aii). This phenotype is not observed in the wild type cells (Atr untagged) where Atr is not depleted in response to AUX. In this case, cell death is only about 10% since this cell line has a functional checkpoint. These results indicate that, since the $Atr^{\Delta HEATs}$ mutants behave in the same way as the conditional-null cell line, they are not capable of substituting the checkpoint role of Atr.

Another way to study checkpoint functionality of the Atr-window deletion mutants is to analyse downstream Chk1 phosphorylation following treatment with HU. In this assay, we detected Chk1 phosphorylation after HU treatment, but not in the untreated or AUX treated samples, as expected. However, when AID-Atr is depleted prior to the addition of HU, this phosphorylation event is no longer evident in the AID-Atr cell line nor in any of the mutant cell lines (Figure 27B). This suggests that none of the $Atr^{\Delta HEATs}$ mutants tested are able to phosphorylate Chk1 in response to replication fork stalling. As expected, our wild type control cell line with non-depletable Atr is able to phosphorylate Chk1 in response to HU treatment.

This result is coherent with the outcome of the Annexin V/PI assay (Figure 27A), in which we detected a significant increase in apoptotic or dead cells in the mutant cell lines, when replication is inhibited after the depletion of Atr.

Probably, the increase in Annexin V positive cells might be a direct consequence of the cells inability to phosphorylate Chk1 and to activate the checkpoint. As a consequence, cells accumulate damage, which would not be repaired, leading to cell death.

As seen in Figure 27B, only three of the mutants were studied for proficient Chk1 phosphorylation ($Atr^{\Delta 2}$, $Atr^{\Delta 4}$ and $Atr^{\Delta 5}$). We expect all mutants to behave similarly as they did in all previous analyses. Specially, all mutants yielded to the same result in the Annexin V/PI assay, in which the aim of the experiment was, as in this western, to analyse checkpoint activation after HU treatment.

Note that both of these checkpoint experiments were performed a short time after the depletion of AID-Atr (6h) in comparison to the study of the essential role that was carried out for 5 days (Figure 26). We have shown in section 3.2.4 that all Atr window deletion mutants displayed a gradual increase in cell death, which began after about 24 hours of AUX treatment and culminated at 96 hours with almost no survival. Due to this delay, the activation of the checkpoint machinery can be analysed in a small window when damage is induced but before the cells start dying due to loss of Atr (Eykelenboom et al., 2013).

3.2.6 Generation of HEAT replacement mutants in both DT40 cells and budding yeast

We hypothesised that different groups of HEAT repeats could mediate particular Atr interactions. However, since all window deletion mutants were unable to fulfil any roles of Atr we took a different approach by testing whether the lack of functionality could be related to protein folding. It is possible that a minimum length of amino acids is required to ensure Atr assumes the correct 3D shape and consequent ability to bind to damaged DNA. Perhaps none of the mutants tested can fold properly to perform Atr functions.

Instead of deleting unique or groups of HEAT repeats, we decided to replace them with other HEAT-rich regions. Replacing HEAT repeats with

similar structures should at least maintain the overall size of the protein so we can analyse whether there are any functions gained or lost. We pursued this strategy both in chicken DT40 cells and budding yeast (Figure 28).

In the case of DT40 cells, we replaced the Atr HEAT repeats 32 and 33 with the equivalent human HEATs ($Atr^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$, Figure 28A). Generation of the replacement cell line was achieved using the same approach as in section 3.2.2. HEAT repeats 32 and 33 are two of the most conserved HEATs between chicken and human ATR (Figure 20B), being this a conservative replacement in order to validate the overall size hypothesis. Furthermore, these repeats were chosen to be significantly far away in the primary sequence from the kinase domain to avoid interfering with its activity.

When using budding yeast, we used the *Delitto Perfetto* strategy (Materials and Methods, section 2.4.2) to generate the different replacement mutants (Figure 28B). In this case, HEAT repeats 32 and 33 from Mec1 were also replaced with the equivalent HEATs from human ATR ($Mec1^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$). Besides, we also replaced Mec1 specific HEATs with the ATR equivalent repeats ($Mec1^{H11\text{-}13\Delta>ATR\ H11\text{-}13}$); in order to confirm if the function/structure of the ATR-specific HEAT repeats is conserved from yeast to human.

In addition, we also performed other replacements in which certain Mec1 HEATs were replaced with similar domains from Tel1, the ATM homologue (Perry and Kleckner, 2003). In particular, we replaced Mec1 specific HEATs with the Tel1 specific units ($Mec1^{H11\text{-}13\Delta>Tel1\ H13\text{-}14}$) and also Mec1 HEATs 14-21 with Tel1 HEATs 17-23 ($Mec1^{H14\text{-}21\Delta>Tel1\ H17\text{-}23}$), four of which have been shown to mediate Tel1 interaction with Nbs1 (You et al., 2005). If these replacements were functional, we would expect the structure to be maintained, but we can also check for the gain/loss of specific Mec1/Tel1 functions.

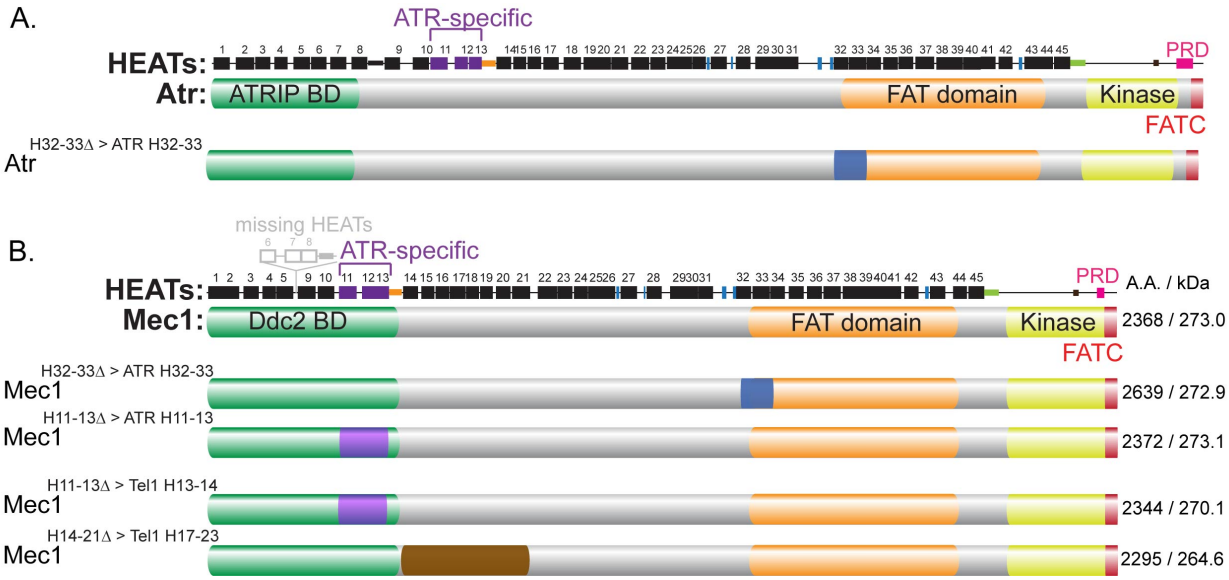


Figure 28. HEAT repeat replacement mutants in *G. gallus* and *S. cerevisiae*.

A. Diagram of replacement mutant generated in DT40 cells ($Atr^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$). HEAT repeats 32 and 33 in Atr were replaced by the equivalent HEAT repeats from ATR.

B. Diagram of replacement mutants generated in budding yeast ($Mec1^{H11\text{-}13\Delta>ATR\ H11\text{-}13}$, $Mec1^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$, $Mec1^{H11\text{-}13\Delta>Tel1\ H13\text{-}14}$ and $Mec1^{H14\text{-}21\Delta>Tel1\ H17\text{-}23}$). Different sets of Mec1 HEAT repeats were replaced by the equivalent HEATs from either ATR or Tel1. Sizes of the replacement mutants are shown in the figure.

The design of these replacement mutants was performed according to Perry and Kleckner's data on HEAT repeat alignments (Perry and Kleckner, 2003). When possible, we tried to precisely maintain protein length, taking into consideration the size of the different linker regions between HEATs.

3.2.7 HEAT repeat replacement mutants are viable in chicken but not yeast

To investigate whether the Atr replacement mutant was functional in DT40 cells, we examined, as before, cell proliferation and checkpoint activation (see section 3.2.2). Replacing part of the FAT domain of chicken Atr with the equivalent ATR sequence results in a viable mutant. $Atr^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$ mutant protein was normally expressed and its stability was not affected by the addition of AUX to the media (Figure 29A). Proliferation pattern (Figure 29B) and response to HU (Figure 30) were also equivalent between $Atr^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$ and our control cell line. These results indicate that the structure/function of HEAT

repeats 32 and 33 are conserved from human to chicken. Furthermore, Atr$^{H32-33\Delta>ATR\ H32-33}$ serves as a positive control for the window deletion mutants previously analysed.

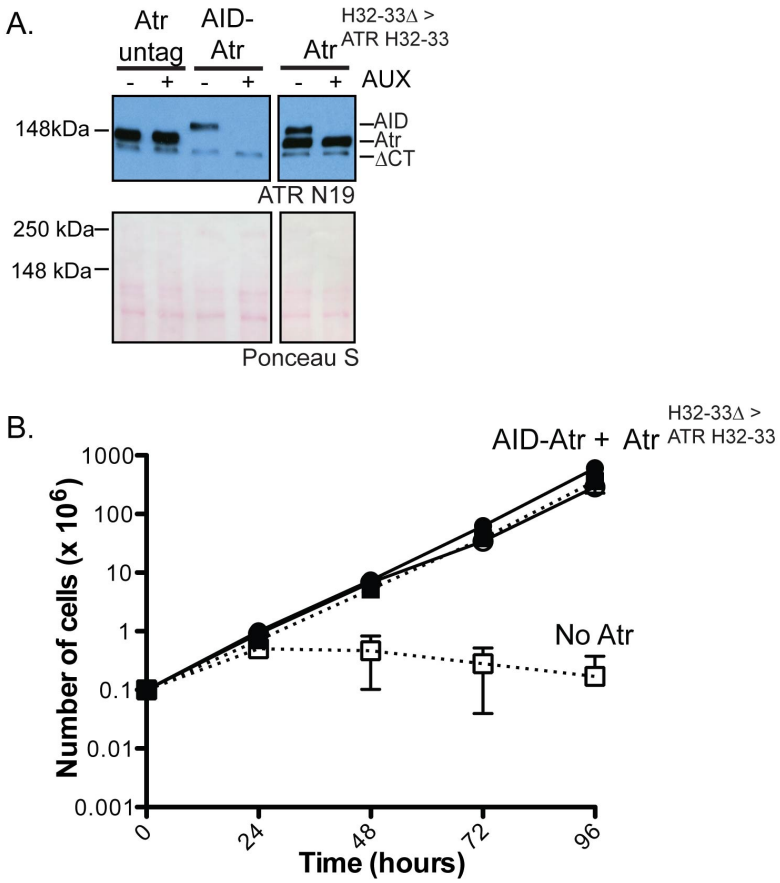


Figure 29. Expression and viability of Atr$^{H32-33\Delta>ATRH32-33}$ replacement mutant in chicken DT40 cells.

A. Western blotting analysis of total cell extracts prepared from Atr untagged, AID-Atr and Rep1 cells using the ATR-N19 antibody (Santa Cruz). Atr$^{\Delta CT}$ was used as a loading control.

B. Growth analysis of the indicated cell lines following treatment with AUX. Cell numbers were recorded every 24h for 5 days and are corrected for culture dilution. Error bars represent the SD from at least three experiments.

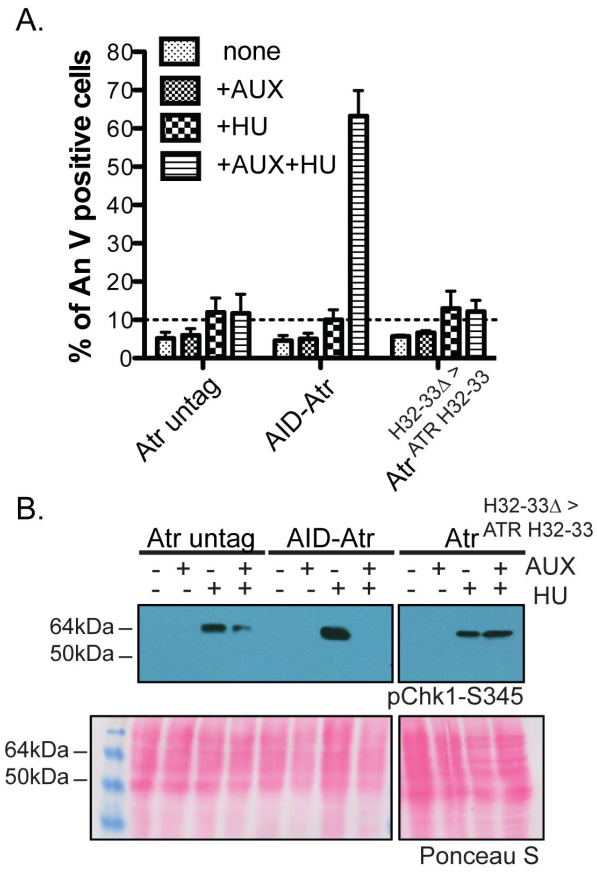


Figure 30. Checkpoint activation of Atr$^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$ replacement mutant in chicken DT40 cells.

A. HU induced apoptosis in the indicated cell lines. Annexin V/PI flow cytometry analysis after treatment with HU (4h) following growth in the presence or absence of IAA (2h). Annexin V positive cells were plotted in the graph. Error bars represent the SD from at least three experiments.

B. Western analysis of Chk1-S345 phosphorylation upon HU treatment. Western blotting of protein extracts from cell lines and conditions indicated was performed using a Chk1-S345ph antibody (Cell Signalling). Loading control could not be carried out due to technical problems with the total Chk1 antibody (FL-476 Santa Cruz). As a consequence, Ponceau S staining is shown to indicate protein loading. Atr$^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$ replacement mutant samples are slightly overloaded compared to controls.

However, this is not the case for the replacement mutants generated in yeast. As, these mutants were generated in a *sml1*Δ background, cells were transformed with a *SML1* plasmid containing a *LEU2* gene for selection to test whether Mec1 replacement mutants were functional. Therefore, cells were plated in minimal media lacking leucine following transformation. Our results show that, after addition of *Sml1* back into the cells, only WT and a positive control (replacement cell line in which HEAT repeats 32 and 33 from Mec1 were

swapped by the same sequence: Mec1$^{H32\text{-}33\Delta>Mec1\ H32\text{-}33}$) were able to grow on minimal Leu$^-$ media. None of the four replacement mutants were able to grow on this media (Figure 31), similar to the 'no Mec1' control (*sml1*Δ *mec1*Δ mutant). These results suggest that the Mec1 replacement mutants are not functional and cannot support Mec1 functions.

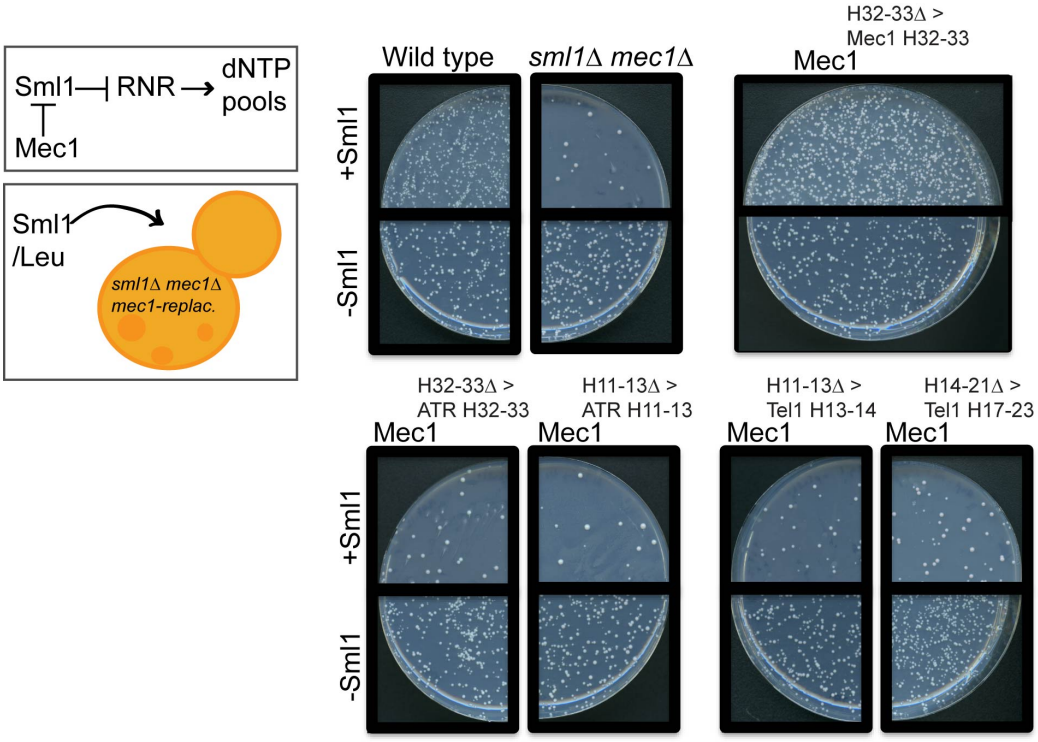


Figure 31. Viability of Mec1/ATR&Tel1 replacements in budding yeast. Cells were transformed with a *Sml1/Leu*$^+$ plasmid, plated in minimal media lacking leucine and surviving colonies counted. Representative images are shown in the figure. Transformation with a *Leu*$^+$ only plasmid served as a transformation control (-Sml1 panels).

Thus, replacing HEAT repeats 11-13 or 32-33 of Mec1 with their human equivalents does not result in a functional Mec1 protein. Similarly, swapping HEAT repeats 11-13 or 14-21 of Mec1 with HEAT repeats 13-14 or 17-23 of Tel1 does not result in a functional protein either. It is likely that the evolutionary gap between yeast Mec1 and human ATR is too great for the protein function to be restored by the equivalent HEAT repeats, let alone those from Tel1. Even if the Mec1 structure were maintained by the replaced motifs, the sequences between ATR and Mec1 are so divergent (section 8.5) that it is possible important residues for proper Mec1 activities are lost.

3.2.8 Atr may have a ring-like structure, similar to other PIKKs

Our results are consistent with HEAT repeat motifs being important to maintain proper Atr structure. We attempted to examine the structure of Atr by electron microscopy (EM).

A previous postdoctoral researcher in our laboratory generated a DT40 cell line expressing tagged Atr, from now on referred to as HFSC-Atr (HA-FLAG-2xSTREP-CBP tagged Atr; Figure 32A) (Pessina and Lowndes, 2014). This cell line was useful for a highly efficient purification of the Atr protein using a tandem affinity purification protocol. We used sequential FLAG followed by STREP immunnoprecipitation to purify HFSC-Atr (Figure 32B). Preliminary negative staining-EM was then performed using the WT HFSC-Atr purified protein (Figure 33). Success with this approach would then be followed by HFSC-tagging and purification of the deletion and replacement mutants.

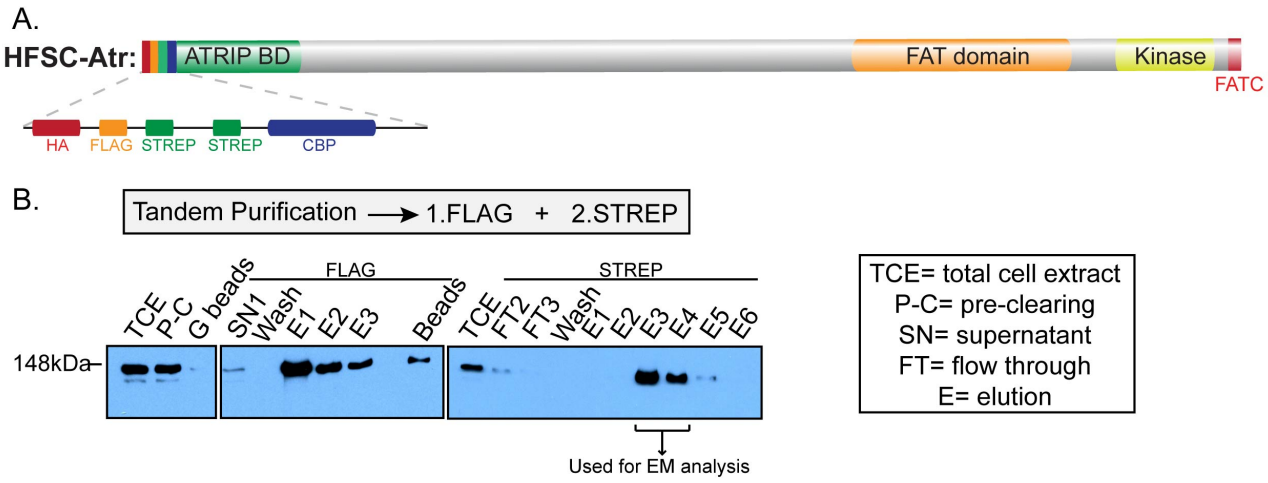


Figure 32. HFSC-Atr purification.

A. Diagram of the HFSC-Atr protein. The HFSC-Atr cell line expresses Atr tagged with four different motifs, including HA, FLAG, 2xSTREP and CBP. This was designed by Dr. John Eykelenboom and Dr. Fabio Pessina, two former researchers at this laboratory (Pessina and Lowndes, 2014).

B. Example of tandem purification for EM analysis. The HFSC-Atr protein was purified using FLAG immunoprecipitation (IP), followed by STREP-IP. HFSC-Atr was always observed to peak in STREP elutions 3 and 4. Further details of purification protocol can be found in Material and Methods, section 2.6.3.

The three-dimensional configuration of other PIKKs has been previously studied by EM and crystallography analyses (Llorca et al., 2003; Sibanda et al., 2010). The resolution of ATM and DNA-PK structures reveals that PIKKs may fold into large structures of 10-15nm (120-160Å). We expect that ATR would have a similar, although slightly smaller, size given its molecular weight (300kDa versus 350 for ATM and 460 for DNA-PK). Therefore, in the electron micrographs we are looking for negatively stained particles of around 10nm in size.

Different techniques were tested to prepare our EM samples. First, we tried to add a drop of the purified protein directly onto an electron microscopy grid (Figure 33A). With this approach we were able to detect negatively stained particles that were dependent on the addition of the purified sample (see 'empty grid' and 'buffer-only' controls) to the grid. However, during the analysis, we found it hard to identify which of the different size particles corresponded to HFSC-Atr and encountered several other problems. For instance, protein density was low within the grid; while numerous particles seemed to be more concentrated at the edges (Figure 33A; 'grid centre' vs. 'grid edges' samples). However, with increased magnification to study these specimens, we could not achieve proper focus and were not able to analyse them. Besides, many artefacts were detected in the process, such as the presence of 'halo structures' or small positively stained deposits (Figure 33C).

Using the same method of adding the protein onto the grid, but in an attempt to increase the protein density, we analysed the Atr samples directly after the staining step, without washing afterwards. With this methodology, we successfully increased the protein density (Figure 33A; 'washing' vs. 'no washing' samples). When no washing step is performed, potential HFSC-Atr particles are detected in the sample (~10nm). However, resolution was not good at this high magnification (150,000x) to discern whether these specimens corresponded to HFSC-Atr, resulting in no satisfactory results.

In an effort to solve these problems, we used an alternative method to prepare our samples (Figure 33B), known as grid blotting (Knispel et al., 2012). Briefly, this is a simple method for the direct transfer of proteins separated by native electrophoresis to electron microscopy grids (further details can be found in

the Materials and Methods, section 2.6.4). When preparing our samples for grid blotting, we found two potential HFSC-Atr bands in the reference lane when compared to a native ladder by silver staining. These were of bigger size than expected, but this could be due to the different running conditions used in this case (ie. native gel). Therefore, we decided to study both proteins using this methodology.

After using two separate grids for these analyses, we performed negative staining-EM. We still encountered some of the previous problems, such as positively stained particles and halo structures (Figure 33C). Furthermore, we could now detect the presence of some gel pieces only in the grids used for grid blotting and not in the 'empty grid' control (Figure 33B; 'empty grid' vs 'gel pieces' samples). Within those gel pieces, we could also find some artefacts that were absorbed in those particular regions (Figure 33B; 'gel pieces' sample). Although it was still difficult to identify the HFSC-Atr protein on the grids and concentration was still an issue, we obtained some interesting preliminary results for the lower band detected on the gel. We were able to distinguish several ring-link structures of 10-15nm (Figure 33B; 'ring structures' sample), which could fit with the expected Atr size. These results suggest that Atr might have a similar 'ring-like' configuration to other PIK kinases, such as DNA-PK (Sibanda et al., 2010). Note that we have sent the HFSC-Atr cell line to Laurence Pearl's laboratory in Sussex, a laboratory with expertise in cryo-EM.

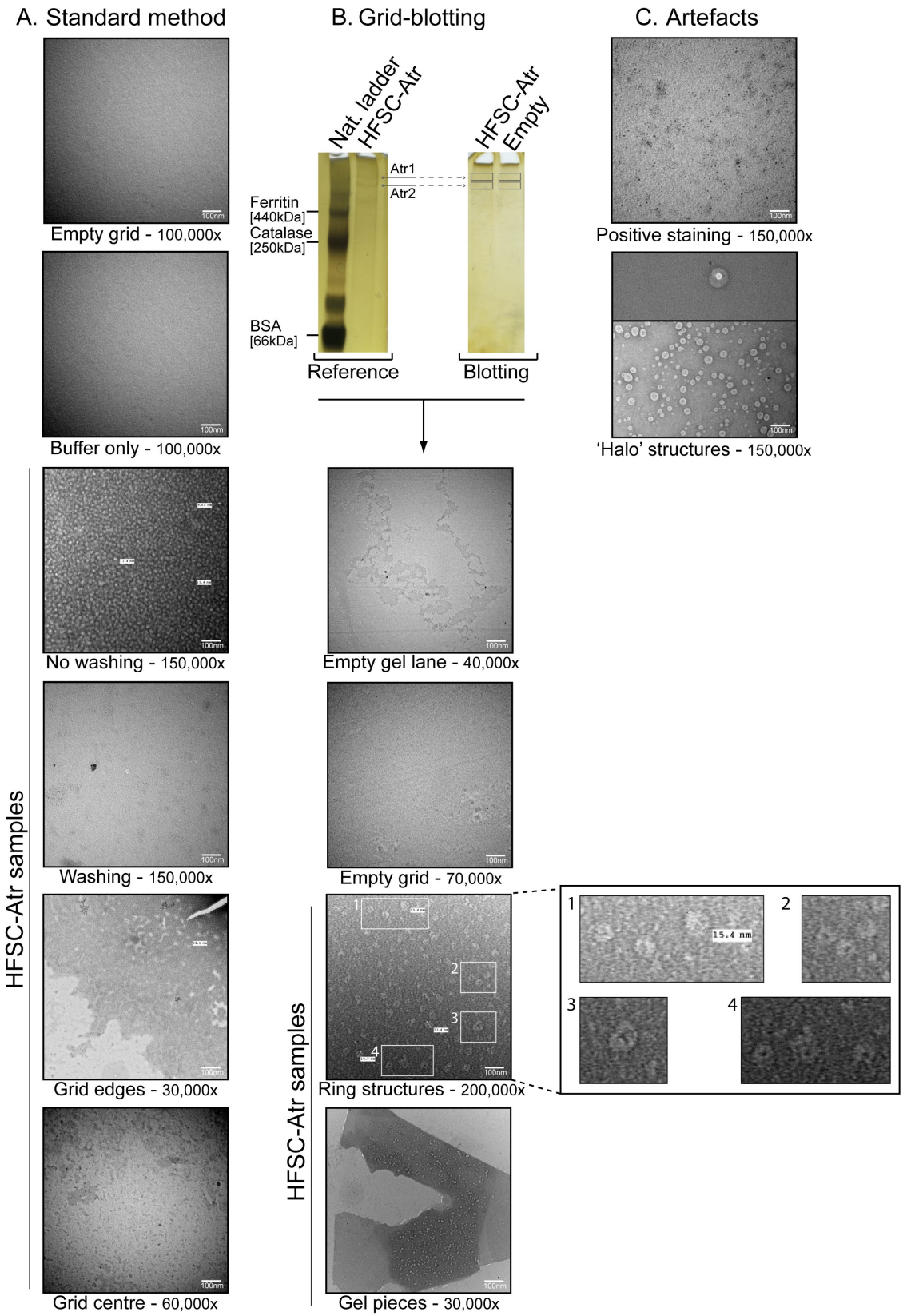


Figure 33. Analysis of Atr configuration by negative staining - EM.

A. Results for EM samples prepared using the standard method (drop-wise). Empty grid and buffer only controls are shown in the figure, along with examples of the HFSC-Atr

protein density at the centre and edges of the grids. Magnification is indicated at the bottom of each picture.

B. Results for negative staining EM samples prepared using the grid-blotting technique. Example of grid-blotting gel is shown in the figure after silver staining. Potential HFSC-Atr bands were analysed from blotting lanes and compared to an empty lane. Representative images for 'empty grid' and 'empty lane' controls, as well as for HFSC-Atr samples are shown in the figure. Examples of gel pieces found when using grid blotting are also disclosed. Within those gel pieces, there were some promising 'ring-like' structures of the expected size for HFSC-Atr (~10-15nm, expanded views in 1-4).

C. Examples of the artefacts detected in both standard and grid blotting techniques are shown, which include 'halo' structures and positively stained particles of unknown nature.

3.3 Discussion

The ATR kinase is both a fundamental player of the DNA damage checkpoints and in the monitoring of DNA replication during a normal cell cycle. Therefore, ATR is considered essential for maintaining genome integrity (Cimprich and Cortez, 2008; Friedel et al., 2009). ATR belongs to the PIK kinase family, which share analogous configurations, including the presence of α-helical repeats, known as HEAT repeats, throughout the non-kinase portions of their structures (Baretić and Williams, 2014; Perry and Kleckner, 2003). As HEAT repeat are structural units that have been proposed to act like platforms for protein-protein and DNA-protein interactions in many systems (Arias-Palomo et al., 2011; Kunz et al., 2000; Llorca et al., 2003; Ruediger et al., 1994; Sibanda et al., 2010), we set as our starting hypothesis that HEAT repeats may play a key role in mediating PIKKs interactions with other proteins and regulating the assembly of multiprotein complexes.

To test this hypothesis and in an attempt to correlate single or sets of HEAT repeats with specific functions of Atr, we began a structure-function study of Atr using a DT40 model system. We have analysed five *Atr* mutants in which different groups of internal HEAT repeats were individually deleted ($Atr^{\Delta HEATs}$). To study the essential role of the mutants, growth analyses were performed in the absence and presence of AUX; while for examining the checkpoint role, the checkpoint activation was monitored after HU treatment. We found that every HEAT repeat deletion mutant tested showed an inability to support cell viability and checkpoint functions of Atr.

During the growth studies, it was noticeable that some of the mutants seemed to be more sensitive to AUX than the AID-Atr cell line. $Atr^{\Delta 1}$, $Atr^{\Delta 2}$ and $Atr^{\Delta 3}$ appeared to die slightly more quickly than the conditional cell line. There are two possibilities that could account for this. First, this could be a consequence of a dominant negative effect carried by the particular mutation and could possibly mean that the deleted regions in those cases are involved in more important Atr interactions. Secondly, this could be a specific effect seen in these particular DT40 clones, meaning that some secondary mutation could have arisen post-introduction of the mutant *Atr* cDNA. To distinguish between these

possibilities, it would be important to analyse independently-generated clones for each mutant.

Taken together, these data suggest that all window HEAT repeat deletions generated in this study seem to impact on all Atr functions, although they do not affect the protein expression levels. The results obtained reflect those seen in Mec1 studies carried out in this laboratory (De Castro Abreu, 2012). In these experiments the *mec1* window deletion mutants could not perform the essential or checkpoint roles, showing sensitivity to several DNA damaging agents. This suggests that Atr/Mec1 kinases cannot tolerate the deletion of these combinations of HEAT repeats.

Given that the deletions generated varied from a single to multiple HEAT repeats in different regions of Atr, these results were unexpected. We had predicted that we would be able to observe different phenotypes, consistent with the idea that the different HEAT repeats studied would mediate different Atr interactions or functions. As several groups have reported separation of function mutations in *ATR* (Laurençon et al., 2003; Nam et al., 2011a; Paciotti et al., 2001), it was reasonable to expect that some of the mutants would be able to support cell viability but not the checkpoint roles, or vice versa. Further mechanistic analyses could be performed on the deletion mutants. These would be essential to check, for example, their ability to interact with ATRIP, co-localize to ssDNA coated by RPA or to phosphorylate known substrates *in vitro*, to test if the kinase remains active. However, these experiments rely on reagents such antibodies, which may not work in DT40 cells. Tagging the $Atr^{\Delta HEATs}$ mutants and/or known interactors, as well as the use of mammalian cells, would facilitate these studies.

It has been previously shown that HEAT repeats 1-7 in ATR mediate the interaction with ATRIP (Ball et al., 2005). Generation of a deletion mutant spanning these HEAT repeats would have been a good control to check for the loss of a known ATR function. However, this mutant was not generated due to technical difficulties. A different approach that could be used to study the different domains of ATR, could be the generation of GST-fusion proteins. ATR could be divided into smaller HEAT repeat containing fragments and fused to GST to check for binding with ATRIP, TOPBP1 or other known interactors.

As none of the mutants were functional, we cannot rule out the possibility that multiple essential functions are distributed throughout the protein, meaning that all HEAT repeats studied could be involved in essential interactions for Atr activity. However, it seems more likely that the problem is related with protein folding. In other words, the lack of functionality might not only be caused by the loss of protein-protein interactions, but the insufficient length of the protein that could alter the overall structural integrity.

The structure of other PIK kinases, such as SMG-1, ATM, DNA-PK and mTOR, has been solved using high-resolution techniques (Arias-Palomo et al., 2011; Llorca et al., 2003; Sibanda et al., 2010; Yang et al., 2013). These studies indicate that PIKKs share similar 3D configurations, in which the kinase domain forms a 'head-like' domain encircled by the FAT domain, while HEAT repeats are shaped as a 'ring-like' structure, facilitating the interactions with the different substrates (Baretić and Williams, 2014; Rivera-calzada et al., 2015). Furthermore, a study, in which a similar 'internal deletion' approach was taken to study SMG-1 structure-function relationship, has shown that the SMG-1 catalytic activity is lost after disrupting continuity of the HEAT-rich region when deleting various regions at its N-terminus (Morita et al., 2007). Another report focusing on TRRAP structure-function relationship has also shown that deletion of single or small HEAT repeat groups leads to inviable mutants (Knutson and Hahn, 2011). This supports the idea that PIKK function is tightly coupled to the integrity of HEAT repeats.

Interestingly, it was also shown recently that HEAT repeats seem to act as elastic connectors that respond to mechanical stress. This was reported for both PR65, a HEAT-rich subunit of the protein phosphatase 2 (PP2A), but also Mec1. PP2A undergoes structural changes in response to forces to allow proper orientation of its regulatory and catalytic subunits impacting on its activity (Grinthal et al., 2010). At the same time, Mec1 was found to regulate the collision of transcription and replication forks, in response to mechanical forces, allowing the release of transcribed genes that are trapped in nuclear pores (Kumar et al., 2014). These functions are thought to rely on the helical nature of HEAT repeats, which would provide flexibility to the structure of PIKKs, and other HEAT-

containing proteins, allowing them to undergo conformational changes that result in a structurally active form of these proteins.

To further investigate if protein folding was dependent on HEAT repeat structural organization, a more conservative approach was followed in which Atr and Mec1 hybrids were generated. We replaced specific groups of HEAT repeats with the equivalent motifs from the orthologous ATR kinase or the related Tel1 kinase (ATM yeast homologue). Our prediction was that replacing HEAT repeats with such equivalent structures should at least maintain the protein length, therefore restoring Atr/Mec1 functions. Additionally, replacement mutants might gain specific functions associated with the PIKK from which the substitutions were obtained.

Our analyses showed that the $Atr^{H32\text{-}33\Delta>ATR\ H32\text{-}33}$ replacement mutant generated in DT40 cells was viable and able to activate the checkpoint in response to HU, supporting the idea that HEAT repeat integrity is crucial for Atr activity. In particular, HEAT repeats 32 and 33 seem to have a conserved function between human and chicken. This is not surprising, considering that they are two of the most conserved HEAT repeats between the two species. On the other hand, all yeast replacements behaved as a *mec1* null. It is possible that the replaced Mec1 HEAT repeats are either responsible for mediating essential protein-protein interactions not conserved in ATR or Tel1, or that the new HEAT repeats introduced in Mec1 result in a structural arrangement that is not compatible with its function. It is also worth noting that, although function and structure might be conserved between Mec1 and ATR proteins, the underlining sequences are very divergent, probably due to a high evolutionary distance.

Note that western blot analyses were performed for the Atr replacement mutant, but they remain to be carried out in these Mec1 hybrids to exclude the possibility that the replacing HEATs are affecting Mec1 expression. Besides, even though the resulting mutant is not functional, it would have also been interesting to test if the $Mec1^{H14\text{-}21\Delta>Tel1\ H17\text{-}23}$ hybrid had gained the ability to interact with Nbs1, an interaction that is known to occur through Tel1 HEAT repeats, 17, 18, 21 and 22 (You et al., 2005).

Analysis of the molecular architecture of the deletion and replacement mutants relative to the WT Atr and Mec1 proteins would be critical to confirm or rule out the possibility of structural abnormalities in these mutants. Besides, this would also provide valuable information to the DDR and PIKK field, since ATR three-dimensional structure has not been resolved yet. We speculated that ATR might have a similar highly ordered structure, with a circular or ring-like shape made of HEATs, the dimensions of which might be crucial to interact with DNA or other proteins, allowing its proper recruitment to DNA lesions and activation. Additionally, HEAT repeats could also be responsible for a conformational change that may be essential for ATR catalytic activity.

To address this, within this study, we have also attempted to resolve Atr structure using electron microscopy. A very preliminary analysis was performed using purified tagged WT Atr and our results, though not conclusive, are very promising, suggesting that Atr might indeed follow the trend seen for other PIKKs. It is likely that Atr could also have a 'ring-like' shape, although further structural analyses using high-resolution techniques are necessary to clarify the nature of Atr structure.

Together, it is clear that ATR functions are strongly dependent on its HEAT repeat composition. Remarkably, although separation of functions could not be achieved using our structure-function strategy, it is noteworthy that Atr (and Mec1), and possibly other PIKKs, are very sensitive to changes in their primary structure. Consistent with the idea of HEATs as elastic connectors, our results indicate that PIKKs require these helical motifs for their functions.

CHAPTER 4

Modelling two novel missense *ATR* Seckel mutations

Running title: Novel missense *ATR* Seckel mutations

Keywords: *ATR*, exon skipping Seckel syndrome, splicing

4.1 Introduction

4.1.1 Syndromes resulting from defects in DDR genes

Alterations in genes related to the DNA Damage Response (DDR) can give rise to several human diseases. Examples include disorders in which repair or signaling are defective due to the malfunction or absence of one or more DDR key players. The study of these human disorders has led to a better understanding of these pathways and will ultimately help in the development of new therapies (Driscoll, 2013; Jackson and Bartek, 2009).

Some of the best-studied DDR syndromes are caused by defects in genes that encode central players of the double strand break repair pathway. It is known that mutations in the *ATM* gene result in Ataxia-telangiectasia (A-T) (Frappart and McKinnon, 2006; Lavin, 2008), a disorder characterised by progressive neurodegeneration, elevated radiosensitivity and high incidence of lymphoid tumours, amongst other clinical signs (Gatti et al., 1991). Mutations in the *NBS1* gene, which encodes another protein involved in DSB recognition, are also causative of a disorder known as the Nijmegen breakage syndrome (NBS) (Weemaes et al., 1981). This recessive disease is characterised by microcephaly, growth retardation, immunodeficiency, and predisposition to cancer. Although the phenotypical characteristics of A-T and NBS are different, these syndromes share multiple chromosomal rearrangements involving chromosomes 7 and 14 (Kleier et al., 2000).

Mutations in *ATR*, another major DDR player that responds to the presence of single stranded DNA, have been also shown to cause a rare human disorder; Seckel syndrome (Alderton et al., 2004; O'Driscoll et al., 2003). Initially, Seckel syndrome was described as an autosomal recessive condition caused by a hypomorphic mutation in the *ATR* gene. Individuals with this disease suffer from proportionately short stature, severe microcephaly and mental retardation, and a typical 'bird-head' facial appearance.

Although Seckel syndrome is normally associated with alterations in the *ATR* gene and, therefore, defective ATR-dependent DNA damage signalling, mutations in other genes also lead to development of this disease. It has been

shown that Seckel patients can carry mutations in the *ATRIP* gene (encoding for ATR's partner in the DDR) (Ogi et al., 2012), but also in genes coding for some centrosomal proteins such as pericentrin (Griffith et al., 2008; Willems et al., 2010), CEP152 (Kalay et al., 2011) and CEP63 (Marjanović et al., 2015).

These findings point towards a possible connection between centrosomal proteins and the ATR-dependent DDR. Centrosomes have a central role in regulating cell cycle progression and, it has been shown recently in our laboratory that, ATR is also involved in this process. During normal cell proliferation, ATR ensures that DNA replication is completed before cell division occurs to avoid mitotic failure (Eykelenboom et al., 2013).

4.1.2 Classic *ATR* Seckel mutation

Until 2012, all known *ATR* Seckel patients belonged to one of two related Pakistani families (Goodship et al., 2000), which harbour the same identical mutation in *ATR* corresponding to a point mutation (A2101G) that results in low levels of expression of functional ATR (O'Driscoll et al., 2003).

This hypomorphic mutation, which is located in exon 9 of *ATR* and does not cause an amino acid substitution (silent change), is known as the classic Seckel mutation and was shown to affect splicing. In the presence of this mutation, exon 9 is skipped, leading to a frameshift and the introduction of a premature stop codon. Low levels of correctly spliced transcript resulting in residual expression of full length ATR protein were detected in the patients. This indicates that the mutation does not fully abolish correct splicing, which is consistent with *ATR* being an essential gene (O'Driscoll et al., 2003).

The mechanism by which exon skipping occurs is not very well understood, but it has been speculated to affect downregulation of an enhancer or upregulation of a silencer of splicing (O'Driscoll et al., 2003). These elements are present in both introns and exons and regulate the splicing process (Wang and Burge, 2008).

Murine models of this disease have been generated to further study ATR deficiency (O'Driscoll, 2009). Murga and colleagues created a mouse model

expressing the classic Seckel mutation (ATR^{A2101G}) that recapitulated the splicing defects seen previously. They were able to generate this model by swapping the human genomic region containing exons 8 to 10, including the intervening introns, into the equivalent mouse locus (Murga et al., 2009). The incorporation of the intronic regions was shown to be essential, when a different group found that engineering the ATR^{A2101G} mutation alone did not affect *ATR* splicing (Ragland et al., 2009). Brown's group have also have reproduced the Seckel phenotype in adult mice. Generation of an adult conditional *ATR*-knockdown mouse model led to defects in tissue homeostasis and the rapid appearance of age-related phenotypes (Ruzankina et al., 2007). Differing from other genomic instability disorders that display cancer predisposition, Seckel murine models do not develop spontaneous tumours. In fact, low levels of ATR have been shown to prevent the growth of some tumours (Murga et al., 2011; Schoppy et al., 2012).

4.1.3 Novel *ATR* Seckel mutations

After the discovery of the classic *ATR* Seckel mutation in 2003 (a homozygous hypomorphic mutation), it was only in 2012 when additional mutations in *ATR* were linked to Seckel syndrome.

Mokrani-Benhelli and colleagues reported the case of a French patient with primary microcephaly who harboured compound heterozygous mutations in *ATR* (Mokrani-Benhelli et al., 2013). The patient inherited from her father an allele that lacked the *ATR* gene due to a 540kb deletion on chromosome 3 encompassing this gene, amongst others. The second allele, which was inherited from her mother, carried a point mutation (ATR^{G5635T}) located within exon 33 that caused a missense mutation (D1879Y). As a consequence of both *ATR* mutations, there was a profound decrease in ATR expression that caused the development of Seckel syndrome (Mokrani-Benhelli et al., 2013).

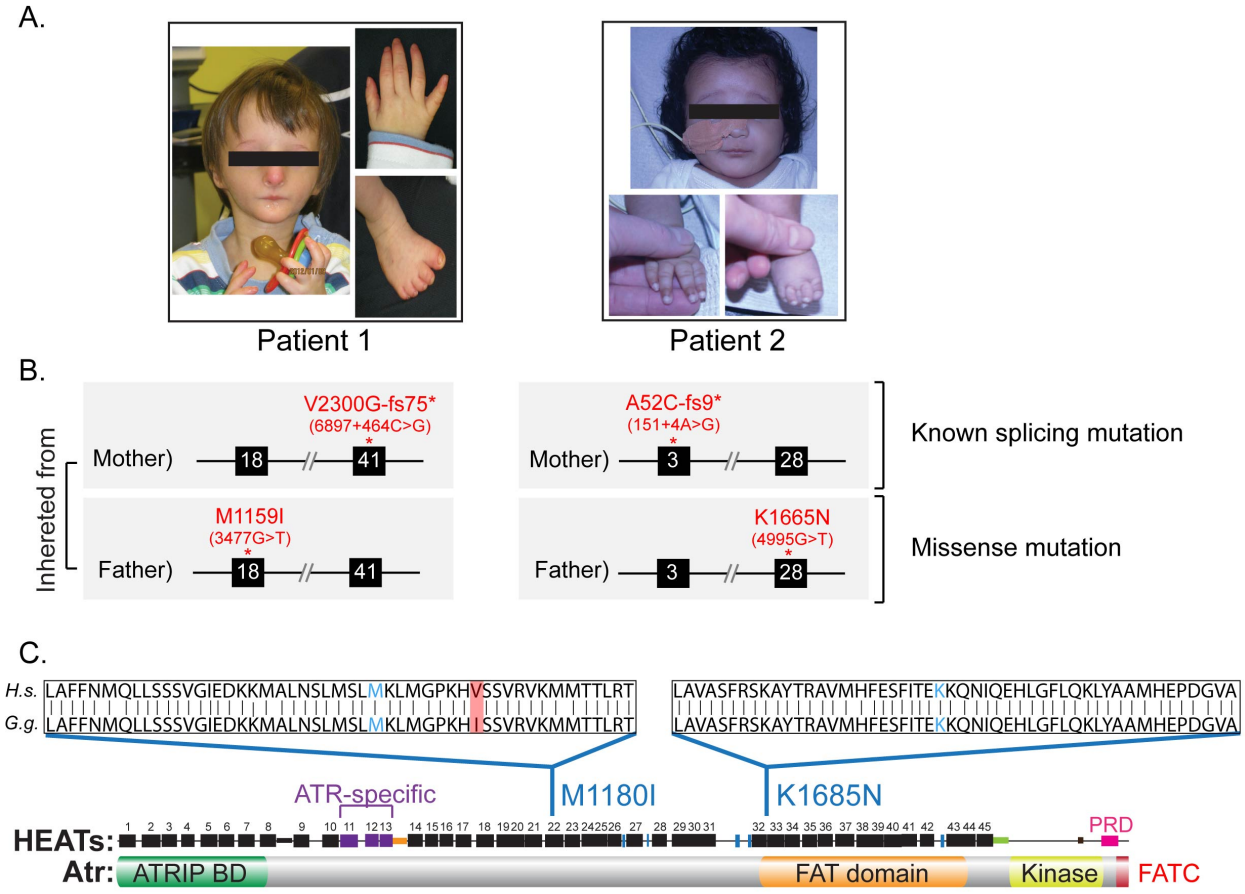


Figure 34. Novel *ATR*^{M1159I} *and ATR*^{K1665N} Seckel mutations.

A. Representative patients harbouring novel Seckel mutations in *ATR*. Both patients show typical Seckel features, including microcephaly, mental retardation and growth abnormalities, such as 'bird-head' facial appearance or defects in their limbs (images provided by Prof. Grant Stewart).

B. Details of *ATR* alleles. Both patients are heterozygotes and carry a known splicing mutation in one allele, plus a novel missense mutation of unknown effects in the other allele. In particular, the allele inherited from the mother carries the splicing mutation, being *ATR*$^{6897+464C>G}$ for patient 1 and *ATR*$^{151+4A>G}$ for patient 2, which translate into the following changes: ATR$^{V2300G\text{-}fs75*}$ and ATR$^{A52C\text{-}fs9*}$, respectively. This nomenclature, used in the human genetics field, shows that the mutations occur in intronic regions; specifically, the C>G and A>G changes occur 464 and 4 nucleotides downstream of the indicated position in the coding sequence (6897 and 151), respectively. While at protein level, these are detected as frameshifts at positions 2300 and 52 in the ATR protein, with a stop codon being created a few amino acids downstream (marked with a *). The second allele, which is inherited from the father, harbours a point mutation, being *ATR*$^{3477G>T}$ for patient 1 and *ATR*$^{4955G>T}$ for patient 2. These result in the *ATR*^{M1159I} and *ATR*^{K1665N} missense mutations, respectively. Patients with these mutations suffer from Seckel syndrome and, as seen in the pictures, they present microcephaly, growth abnormalities in their limbs and also mental retardation. These pictures were kindly provided by Grant Stewart.

C. Structure and localisation of missense mutations. The structure of chicken Atr is shown in the figure, zooming in the HEAT-rich regions that are affected by the Seckel mutations. The M1180I mutation is located within the 22nd HEAT repeat (exon 18 of the *ATR* gene), which lies in the uncharacterised UME domain found in all ATR homologues and which is predicted to be required for protein-protein interactions (Rounds and Larsen, 2008). The K1685N mutation is located in the linker between HEAT repeats 32 and 33 (exon 28 of the *ATR* gene), which is also part of the conserved FAT domain.

The same year, other novel mutations in the *ATR* gene were linked to Seckel syndrome. Our collaborator Prof. Grant Stewart, along with other groups, was involved in the detection of different *ATR* mutations in two unrelated Seckel sufferers native to the UK, one of which was reported in the literature (Ogi et al., 2012). Mutations found in the second patient remain unpublished and a brief description of their effects on ATR function is enclosed in this project.

Prof. Grant Stewart's findings show that both patients suffer from Seckel syndrome and harbour compound heterozygous mutations (Figure 34); in other words, they contain different mutations in the two *ATR* alleles.

The first *ATR* allele harbours a known splicing mutation, which is different for the two patients. In particular, patient 1 carries a mutation within intron 40 of *ATR* ($ATR^{6897+464C>G}$; see Figure 34B, for further information on nomenclature). This mutation resulted in the activation of a cryptic splice site and the consequent introduction of part of an Alu repeat into the mRNA. At the protein level, this was observed as a change in the reading frame at residue 2300 and the creation of a premature stop codon 75 amino acids downstream ($ATR^{V2300G\text{-}fs75*}$). The splicing mutation carried by patient 2 was located in intron 2 of *ATR* ($ATR^{151+4A>G}$) and had similar effects to the previous one. In this case, a frameshift was detected at position 52 in the ATR protein, with a stop codon being created 9 amino acids downstream ($ATR^{A52C\text{-}fs9*}$).

Schematics of the splicing defect caused by $ATR^{6897+464C>G}$ mutation in patient 1 are shown in Figure 35. The $ATR^{151+4A>G}$ mutation found in patient 2 causes a similar defect, but the specific mechanism by which this occurs is not known.

Figure 35. Splicing defects caused by the *ATR*$^{6897+464C>G}$ mutation (Ogi et al., 2012). The *ATR*$^{6897+464C>G}$ mutation affects one of the *ATR* alleles in patient 1. This mutation is located in intron 40 of *ATR*, specifically 464bp downstream of position 6897 on the coding sequence, which corresponds to the end of exon 40. This C>G change within an intronic region is thought to create new a splice site and results in the introduction of part of an Alu repeat into the mRNA (intron retention). As a consequence, a 'new exon 41' is generated giving rise to the V2300G change and the introduction of a stop codon 75 amino acids downstream due to a frameshift.

On the other hand, the second *ATR* allele in these patients contains a point mutation, which is located in exon 18 for patient 1 (*ATR*$^{3477G>T}$) and exon 28 for patient 2 (*ATR*$^{4955G>T}$). These mutations cause the M1159I and K1665N amino acid substitutions, respectively. While the effect of the splicing mutations is known, there is no evidence of the effects of the missense mutations in these patients.

In this study we investigate any effects that the *ATR*^{M1159I} and *ATR*^{K1665N} Seckel missense substitutions have on ATR function. We addressed this by using both chicken DT40 cells and human cells derived from Seckel patients. We are able to demonstrate that, although these mutations cause amino acid changes, they do not directly affect ATR function but, instead, they are involved in splicing regulation.

4.2 Results

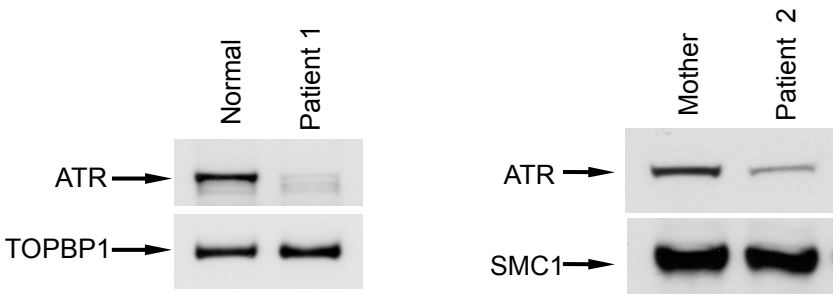
4.2.1 Description of M1159I and K1665N patient cell lines: Previous data on DDR defects

In this study we aimed to analyse the effect of two independent missense mutations on ATR function (*ATR*^{M1159I} and *ATR*^{K1665N}). These mutations were identified in unrelated Seckel patients by our collaborator Prof. Grant Stewart and colleagues, one of which has been previously published (Ogi et al., 2012).

Stewart and collaborators had previously observed that these patient cells express very low levels of ATR (Figure 36A) and show various defects in their response to HU and UV lesions (Figure 36B). They described that cells derived from patient 1 contain the mutations described above and are unable to activate the UV-induced G2/M checkpoint arrest. Additionally, these mutations cause impaired phosphorylation of a range of ATR substrates following exposure to 0.5mM HU, such as pSMC1, pCHK1, p53BP1, pNBS1 and γH2AX (Ogi et al., 2012).

Unpublished data from Prof. Grant Stewart and collaborators also suggests that the ATR signalling pathway is altered in cells derived from patient 2. However, in this case, mutations do not seem to cause such a strong Seckel phenotype. They confirmed that, in response to 20J/m^2 of UV, phosphorylation of pSMC1 and pNBS1 is normal, while Chk1 phosphorylation seems to decrease only slightly. Besides, there is an up-regulation in RPA2 and H2AX phosphorylation.

A.



B.

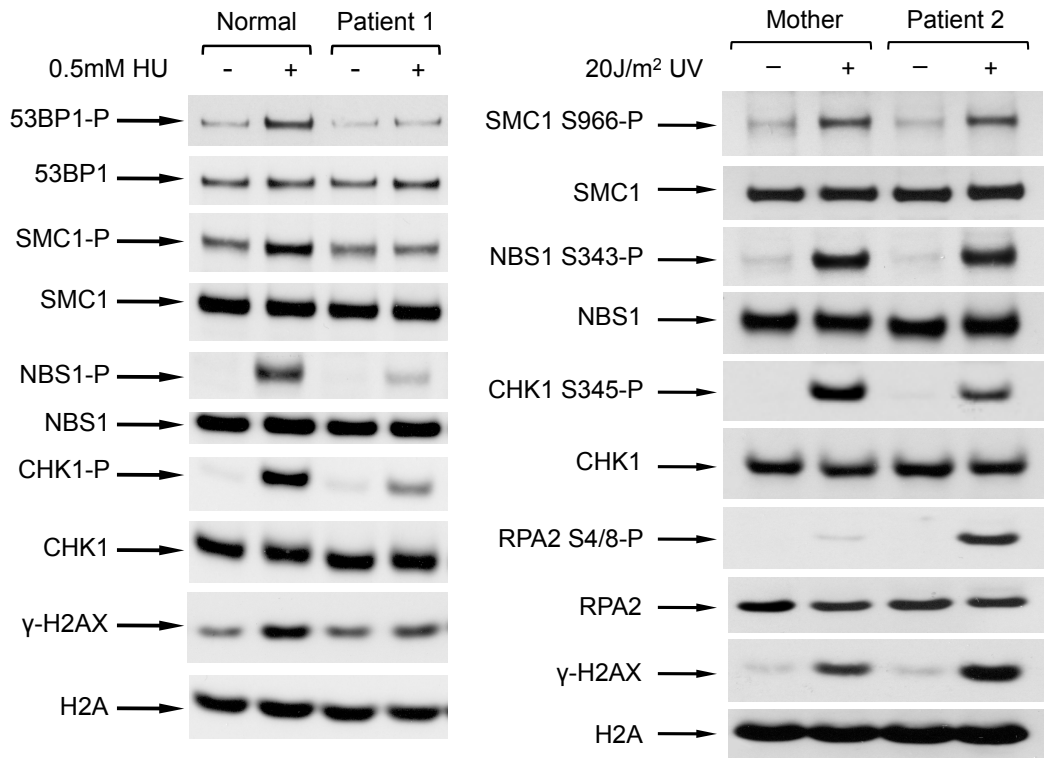


Figure 36. Altered DNA damage signalling in patient cell lines (data provided by Prof. Grant Stewart).

A. ATR expression levels in patients 1 and 2. ATR levels are analysed using the ATR-N19 antibody (Santa Cruz). TOPBP1 and SMC1 are used as loading controls.

B. Phosphorylation pattern of several ATR targets. Patient cells were examined for their ability to phosphorylate the indicated ATR substrates, including the classical CHK1-S345 and H2AX, but also SMC1, NBS1, 53BP1 and RPA2, at 1h following exposure to 0.5mM HU (Patient 1) or 20J/m^2 UV (Patient 2). The total form of these proteins is used as loading control in each case.

These results indicate that the described mutations have an effect on ATR function; however, they might be affecting different parts of the signalling cascade. Since, in both patients, one of the alleles carries a mutation that results in aberrant splicing, we hypothesised that the missense mutations could be directly responsible for the defects on ATR signalling by affecting protein function rather than expression. Therefore, we used our DT40 model system to address this possibility.

As seen in Figure 34C, the equivalent amino acids of the mutated human residues correspond to positions M1180 and K1685 in the chicken sequence. Both of these *ATR* mutations are located in conserved regions between human and chicken species, supporting the idea that they could be important residues for ATR function.

4.2.2 Generation of Seckel mutants in DT40 cells: M1180I and K1865N

To generate the two stable Seckel cell lines in DT40 cells, we targeted Atr^{M1180I} and Atr^{K1685N} cDNAs, in separate experiments, to the Ovalbumin locus of the AID-Atr cell line. The screening of positive clones was performed by Southern blotting of genomic DNA after a BamHI digest (Figure 37), as explained in Chapter 3. As a result, the final cell line contains both AID-Atr and Seckel-Atr proteins. As explained in the previous chapter, to study the effect of the missense Seckel ATR mutations, AID-Atr is depleted by the addition of auxin to the system. After auxin addition, the only Atr copy that is functional in the cells is the Seckel-Atr protein.

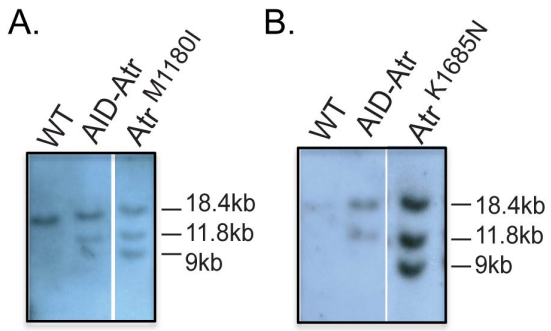


Figure 37. Generation of Atr^{Seckel} mutant cell lines. Southern screening of Atr^{Seckel} clones for *Ovalbumin* targeting. In both panels, the first lane represents wild type cells; the second lane corresponds to the AID-Atr parental cell line, while the third lane represents targeting to the *Ovalbumin* locus of Atr^{M1180I} (A) *to* Atr^{K1685N} (B) cDNAs.

4.2.3 Point mutations have no effect on Atr function in DT40 cells

Since Seckel mutations previously described are caused by hypomorphic mutations, which result in very low levels of ATR expression (O'Driscoll et al., 2003; Ogi et al., 2012), we first examined whether Atr^{M1180I} and Atr^{K1685N}

mutations cause a reduction in protein levels. However, the missense *Atr*^{Seckel} mutations have no effect on protein expression in our DT40 system, both being expressed equivalently to the WT control (Figure 38A).

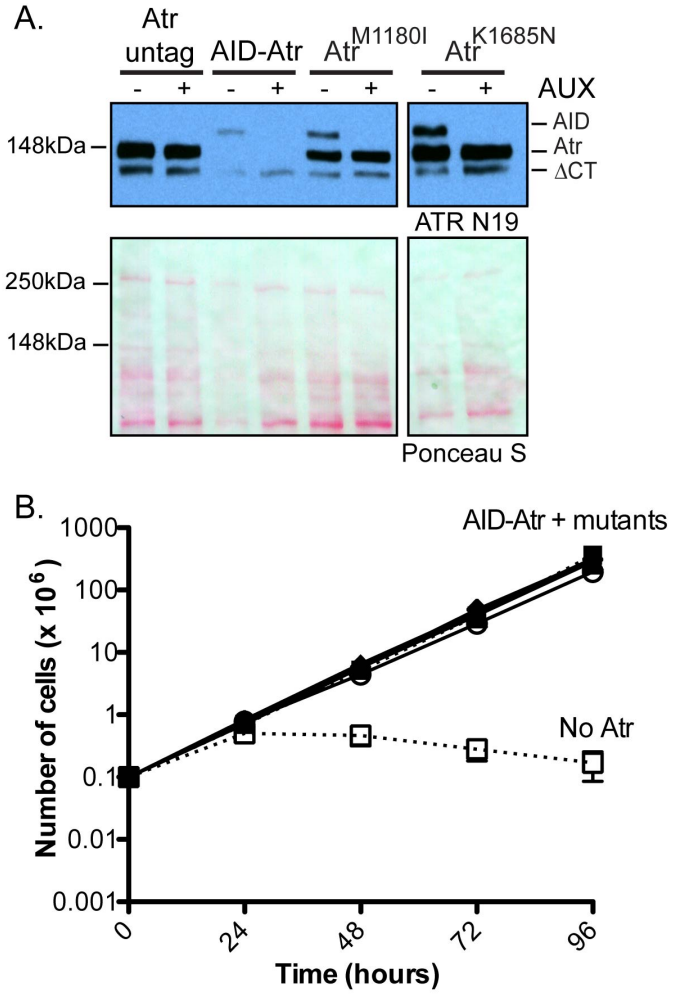


Figure 38. Expression and viability of novel Seckel mutants in chicken DT40 cells.

A. Western blotting analysis of total cell extracts prepared from Atr untagged, AID-Atr, Atr^{M1180I} and Atr^{K1685N} cells using the ATR-N19 antibody (Santa Cruz). This analysis was performed in the absence and presence of auxin (0.5mM AUX treatment for 2h). The Atr$^{\Delta CT}$ protein from the endogenous Atr loci serves as a loading control. Note that AID-Atr untreated sample is slightly under loaded. This can also be seen in the Ponceau S staining.

B. Growth curve analysis of the Atr^{M1180I} and Atr^{K1685N} cell lines following treatment with AUX. Cell numbers were recorded every 24h for 5 days and are corrected for culture dilution. Error bars represent the SD from at least three experiments.

Next, we investigated if the *Atr*^{Seckel} mutations affect proliferation. We analysed proliferation by counting cells every 24h in the presence and absence of AUX in the media. As seen in Figure 38B, these mutations have no effect on cell growth; mutant cell lines and control cells divide every 8 hours.

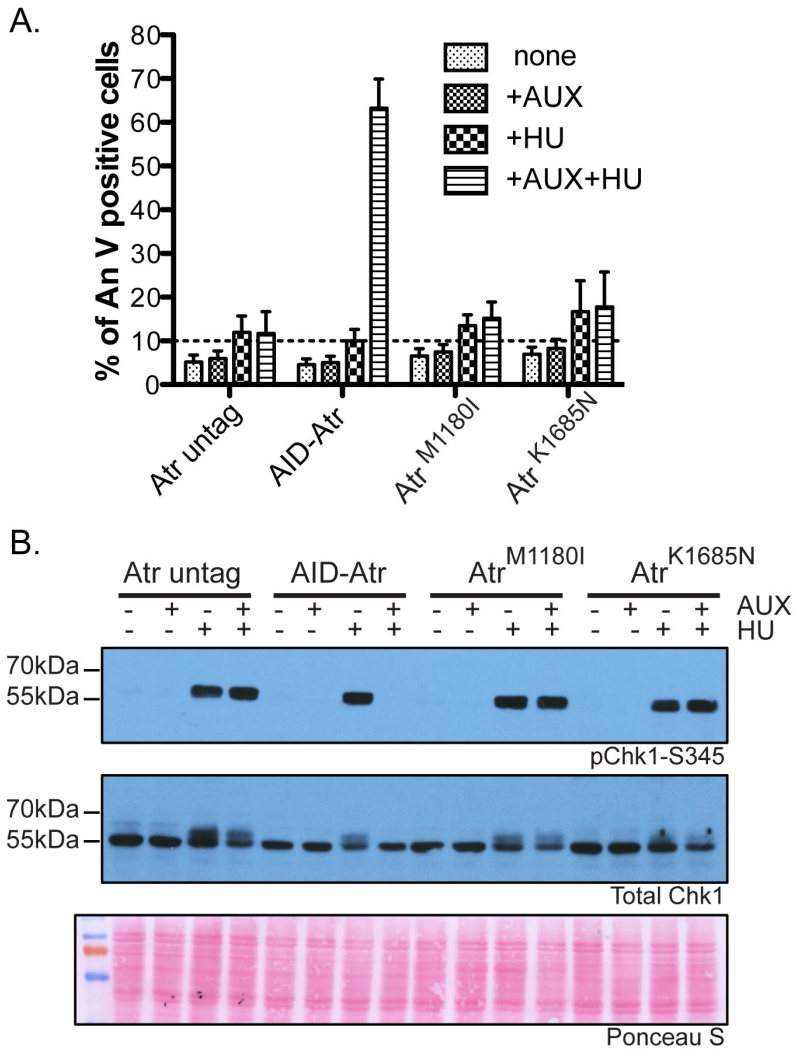


Figure 39. Checkpoint activation of novel Seckel mutants in chicken DT40 cells.

A. HU induced apoptosis in the indicated cell lines. Annexin V/ PI flow cytometry analysis after treatment with HU (4h) following growth in the presence or absence of IAA (2h). Annexin V positive cells were plotted in the graph. Error bars represent the SD from at least three experiments.

B. Western analysis of Chk1-S345 phosphorylation upon HU treatment. Western blotting of protein extracts from cell lines and conditions indicated; performed using a Chk1 antibody that recognises phosphorylation on S345 (Cell Signalling). An antibody against total Chk1 (FL-476 Santa Cruz) was used as a loading control.

To more carefully explore the effect of these mutations in the DT40 system, we assessed whether mutant cell lines were able to activate the intra S-phase checkpoint in response to hydroxyurea (HU) (Figure 39). HU indirectly causes replication fork stalling, so ATR activation can be measured by assessing the phosphorylation of its main substrate, Chk1. Our results show that Atr^{M1180I} and Atr^{K1685N} are still able to phosphorylate Chk1 at serine 345 (Figure 39B). We have also confirmed this indirectly, by analysing the percentage of apoptotic cells upon AUX and HU treatment. This assay demonstrates that the Atr^{M1180I} and Atr^{K1685N} proteins are able to activate the checkpoint after depletion of AID-Atr and induction of replication fork stalling (+AUX +HU) (Figure 39A).

Our data indicates that ATR^{M1159I} and ATR^{K1665N} Seckel missense mutations do not affect the function of the Atr protein. This suggests that the underlying pathology of those missense mutations is not related to protein function.

4.2.4 Point mutations cause exon skipping in patient cell lines

As the novel *ATR* Seckel mutations show no effect on Atr function in DT40 cells, we hypothesise that they may affect splicing by abrogating a splice enhancer or generating a splice silencer within exonic sequences. Thus, both mutations could potentially result in exon skipping, similarly to that described for the classic Seckel mutation (O'Driscoll, 2009; O'Driscoll et al., 2003). Since our DT40 system is cDNA based, it is not suitable for studying splicing defects. Therefore, to confirm if splicing defects occur in the novel Seckel mutants, we obtained human lymphocyte cell lines (LCLs) derived from patients (kindly donated by Prof. Grant Stewart).

In order to investigate if ATR^{M1159I} and ATR^{K1665N} mutations cause exon 18 and 28 skipping, respectively, we extracted RNA and used RT-PCR followed by standard PCR. We designed primers to the flanking exons and expected to see smaller PCR products in the mutant cell lines as a consequence of exon skipping. As shown in Figure 40A and B, shorter splice variants of the expected sizes were detected in both patient samples. In addition, we could still detect normally spliced ATR, which would correspond to the transcript from the second ATR allele. Since these PCRs were performed in a semi-quantitative way (few PCR

cycles), we believe that each of the splice variants could contribute to 50% of the mRNA produced. We have attempted to study the abundance of the different splice variants by qPCR, but we have not been able to distinguish the WT and mutant variants using this method. The use of Taqman probes might be helpful to study this in more detail.

To confirm whether the smaller PCR products actually corresponded to exon-skipped variants, we sequenced these PCR products (Figure 40C and D). Our data shows that exon 18 and 28 are skipped in the aberrant splice variants (Figure 40E). In these transcripts, exons 17&19 and 27&29 are spliced together and we predict this would cause frameshifts and result in the introduction of a stop codon (Figure 41). These putative ATR truncated forms were not detected by westerns, thereby, we would expect that the resulting protein, in Seckel patients, may not be expressed or may be very unstable. Alternatively, the aberrant transcript lacking exon18/28 could also be eliminated by non-sense mediated decay.

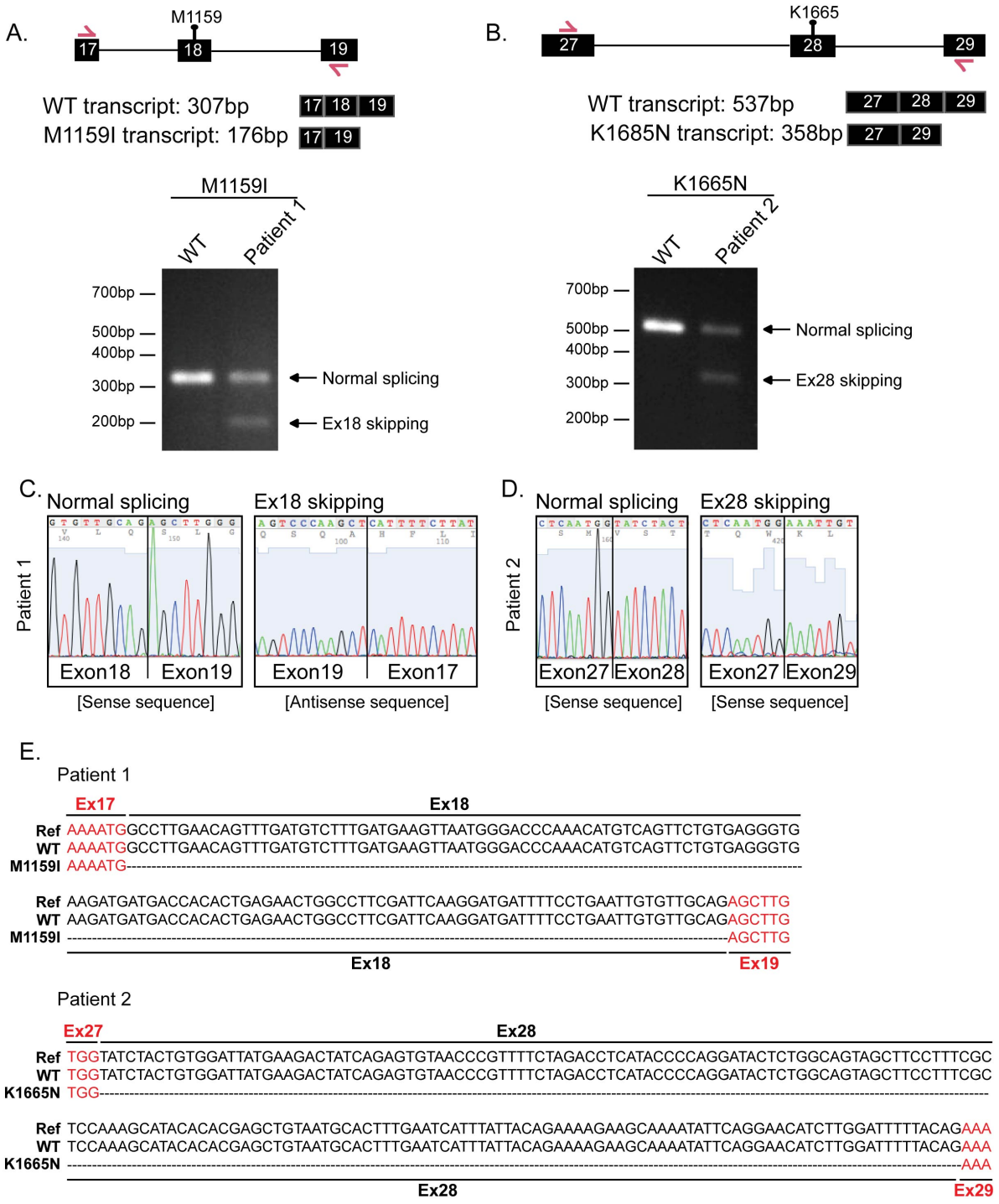


Figure 40. Study of exon skipping in patient-derived cell lines.

A. PCR analysis of WT and M1159I cDNAs. Primers against exon 17 and 19 were used to confirm exon 18 skipping in the presence of the M1159I mutation. Results shown in the figure correspond to a PCR setup with 27 amplification cycles.

B. PCR analysis of WT and K1665N cDNAs. Primers against exon 27 and 29 were used to confirm exon 28 skipping in the presence of the K1665N mutation. Results shown in the figure correspond to a PCR setup with 29 amplification cycles.

C. Sequencing results for patient 1 (M1159I mutation). Upper and lower band (from panel A) corresponding to patient 1 transcripts were sequenced using the forward and reverse PCR primers, respectively. Chromatogram illustrates exon 18 skipping in shorter mRNA variant. Sequencing of this shorter variant using forward primer also confirmed this result, however, since the primer is too close to the start of exon 18, it was difficult to visualize the exon17/19 transition.

D. Sequencing results for patient 2 (K1665N mutation). Upper and lower band (from panel B) corresponding to patient 2 transcripts were sequenced using the forward PCR primer. Chromatogram illustrates exon 28 skipping in shorter mRNA variant.

E. Alignment of WT (upper PCR band = normal splicing) and mutant (lower PCR band = exon skipping) sequences. WT and mutant sequences obtained in sequencing reactions were aligned with a reference ('Ref') encompassing part of *ATR* exons 17 to 19 or 27 to 29.

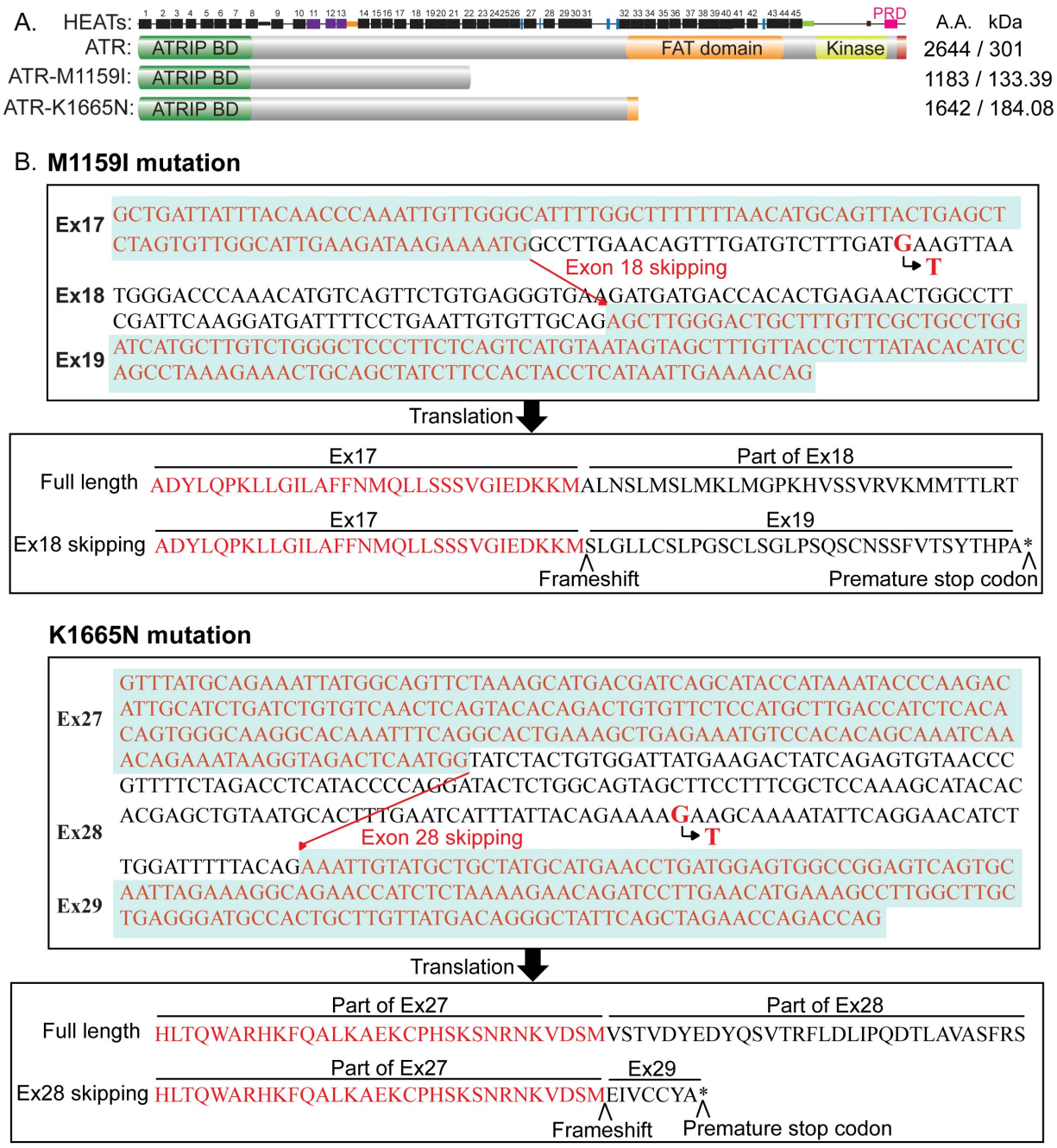


Figure 41. *In silico* effects of atypical Seckel mutations on the *ATR* transcript and protein.

A. Predicted ATR truncations in the presence of the M1159I and K1665N mutations. Predicted size of the mutant proteins is shown in both amino acids and kDa.

B. Effect of ATR mutations at DNA and protein level: Sequences of exons 17 to 19 and 27 to 29 are shown in the figure along with part of the corresponding translation for those exons. Presence of M1159I and K1665N mutations causes exon skipping, which leads to a change in the reading frame and the introduction of a premature stop codon.

4.2.5 Confirmation of exon skipping using a minigene system

To confirm that exon skipping was not an artefact seen only in patient cells, we used a minigene system to obtain independent results. This is a established technology to study splicing, which has also been applied to investigate splicing regulation of other DDR genes, such as *ATM* and *BRCA1* (Lewandowska et al., 2005; Pagani et al., 2002; Raponi et al., 2014).

This method consists of an *in vivo* splicing assay, in which a plasmid containing the *β-globin* gene under a strong promoter is used. The genomic region of interest is inserted within exon 3 of the *α-globin* gene and this plasmid is then transiently transfected into higher cells to examine if splicing is affected *in vivo*.

We performed our experiments with the control *Neurofibromin* (*NF1*) minigene obtained from Diana Baralle's laboratory. This plasmid was originally used to assess whether a mutation in the *NF1* gene affected its splicing (Raponi et al., 2009). To analyse the effect of ATR^{M1159I} and ATR^{K1665N} mutations on exon 18 and 28 respectively, we cloned these exons and flanking intronic regions into the NdeI sites of the *NF1* plasmid (Figure 42A). Site-directed mutagenesis was then performed to introduce the ATR^{M1159I} and ATR^{K1665N} point mutations. WT and mutant plasmids were used to transfect both HeLa and HEK293T cells, which was followed by RNA extraction and RT-PCR. Standard PCR analysis with primers annealing to the minigene vector (pink arrows in panel A) were then used to assess the splicing possibilities.

Using this approach, we were able to reproduce the results from the Baralle's group, regarding splicing of exon 29 of the *NF1* gene. Study of *ATR-ex18/28* constructs confirmed the results obtained in patient cells. Shorter splice variants corresponding to exon 18 and 28 skipping were also generated in the presence of M1159I and K1665N mutations when analysing splicing of these minigenes *in vivo*, both in HeLa and HEK293T cells (Figure 42Bi and ii, respectively). This was also confirmed by sequencing (Figure 42C).

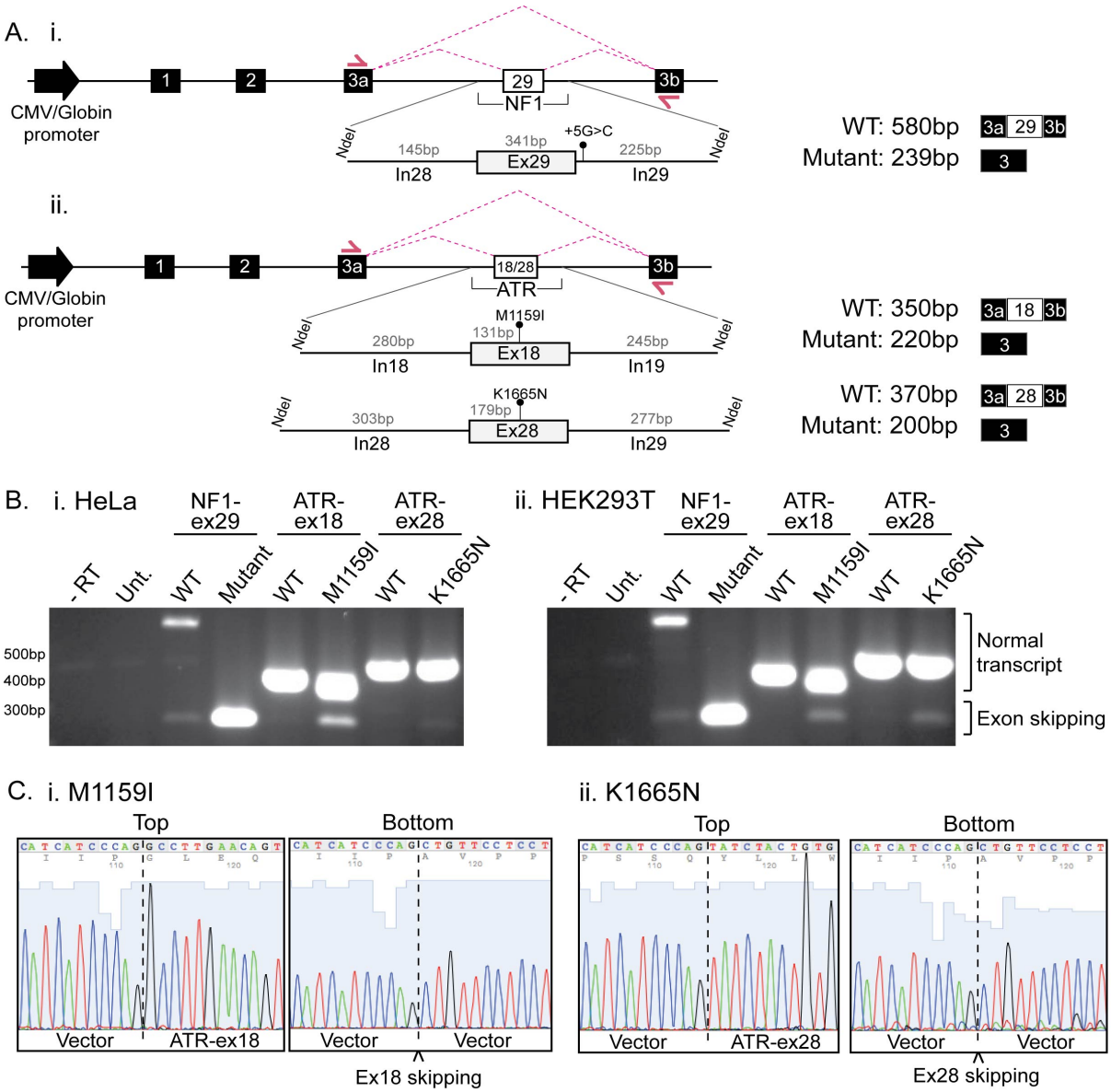


Figure 42. Minigene system transfection assay.

A. Schematic representation of *NF1* and *ATR* minigenes used in transient transfection splicing assays. Diagram shows the *NF1/ATR* genomic regions inserted within exon 3 of the globin gene, together with the location of the mutations studied. Minigene exons and introns are indicated as boxes and lines, respectively. Pink arrows indicate primers used to check splicing efficiency by PCR, while dotted lines show the two splicing variants that are expected (normal transcript / exon skipping variant). Length of these PCR products is also shown in the figure.

B. PCR results of the minigene transfections in HeLa (i) and HEK293T (ii) cells. Controls for DNA contamination (minus retrotranscriptase) and minigene specificity (untransfected cells) were included in the analysis, along with the WT and mutant *NF1* minigenes, which were used as splicing controls (*NF1-ex29*). Analysis of *ATR-ex18/28* constructs is shown in the figure. RNA splicing variants corresponding to exon 18 and 28 inclusion (upper band) and exclusion (lower band) are displayed for mutant M1159I and K1665N, respectively. A higher band can also be detected in the mutants, which seems to be unspecific as it also appears in the *WT-NF1* construct.

C. Sequencing results for *ATR-ex18* (i) and *ATR-ex28* (ii) mutant minigenes. Upper and lower band for minigenes containing the M1159I/K1665N mutations (from panel Bi) were sequenced using a primer annealing to the vector. Chromatogram illustrates exon 18 and 28 skipping in shorter mRNA variant.

4.2.6 Point mutations could potentially disrupt binding sites for splice factors

To assess whether the ATR^{M1159I} and ATR^{K1665N} mutations could affect an exonic splicing enhancer or silencer (ESE/ESS) or any other splicing regulatory sequence, Human Splicing Finder (HSF) (Desmet et al., 2009) and SFmap (Paz et al., 2010) online tools were used. These bioinformatic tools predicted that both missense mutations disrupt a 9G8 exonic splicing enhancer in *ATR* exons 18/28, while different silencer sites are created by the presence of the mutations (Table 12).

Table 12. Predicted binding sites in the presence of the ATR^{M1159I} and ATR^{K1665N} mutations. Part of *ATR* exon 18 and 28 WT and mutant sequences were analysed using the HSF and SFmap bioinformatic tools.

```
WT:      TGTCTTTGATGAAGTTAATG
M1159I:  TGTCTTTGATTAAGTTAATG
```

Sites gained	Sites disrupted
hnRNPA1	9G8

```
WT:      TACAGAAAAGAAGCAAAATA
K1665N:  TACAGAAAACAAGCAAAATA
```

Sites gained	Sites disrupted
Tra2 Beta	9G8
YB1	

Further studies would be necessary to confirm these hypotheses and to study in more detail binding to these regulatory sites. An electrophoretic mobility shift assay (EMSA) would be ideal to corroborate if these proteins regulate *ATR* splicing. This would give new insights on the understanding of Seckel syndrome but also into the general splicing mechanisms.

4.3 Discussion

Seckel syndrome is a rare genetic disease caused by mutations in several genes, including the *ATR* gene, which are known to affect splicing (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012). Patients with this syndrome suffer from microcephaly and developmental defects. Our collaborator, Prof. Grant Stewart, recently found two novel missense mutations (ATR^{M1159I} and ATR^{K1665N}) of unknown effects in unrelated Seckel patients. As these mutations caused amino acid substitutions, we hypothesized they might directly affect ATR function. Therefore, in this study, we set to model those new Seckel mutations using a DT40 model system.

To test this assumption we evaluated if the mutants were capable of supporting cell viability and activating the checkpoint in response to replication fork stalling. To study this, growth analysis were performed in the absence and presence of AUX, in parallel with the monitoring of checkpoint activation upon HU treatment. Surprisingly, we found that both Seckel point mutations, corresponding to Atr^{M1180I} and Atr^{K1685N} substitutions in chicken, were fully functional in DT40 cells. This suggests that Seckel defects seen in patients are not a direct effect of ATR protein malfunction. An alternative hypothesis is that these mutations could be affecting splicing by disrupting regulatory elements such as splicing enhancers or silencers located within exonic regions of the *ATR* gene (Wang and Burge, 2008).

It has previously been shown that the classic *ATR* Seckel mutation (ATR^{A2101G}) causes a defect in splicing. Even though this mutation was located in an exonic region (exon 9) and did not cause an amino acid substitution (silent change), it was shown to cause exon skipping (O'Driscoll et al., 2003). The mechanism behind this is unclear, but it was suggested that the mutation could affect a regulatory motif important for splicing. A similar mechanism was proposed to cause Seckel syndrome in a French patient, who harboured a mutation in exon 33 of *ATR* that was predicted to disrupt a splicing enhancer (Mokrani-Benhelli et al., 2013).

Since our DT40 model system was cDNA based, we were not able to test the splicing hypothesis. Therefore, to further investigate if the missense ATR^{M1159I}

and ATR^{K1665N} mutations affected splicing, we obtained human cells derived from patients. Using these cells, we were able to study splicing of exons 18 and 28 of the *ATR* locus. Our results show that, in the presence of the ATR^{M1159I} and ATR^{K1665N} mutations, a shorter splice variant is produced, which corresponds to either exon 18 or 28 skipping, respectively. These results were also confirmed using an *in vivo* splicing assay widely used in the splicing field. We used a minigene system to determine that exon skipping also occurred when the minigene was transiently transfected into standard human cell lines such as HEK293T or HeLa cells.

Taken together these results indicate that the novel Seckel missense mutations, although altering the amino acid sequence, they do not directly affect ATR function, rather they dysregulate splicing. Nowadays, there is growing evidence that multiple human diseases caused by exonic mutations are, in fact, the consequence of splicing defects instead of loss of protein function (Baralle and Baralle, 2005; Lopez-Bigas et al., 2005). This is true for missense, but also silent mutations, which are usually not considered as disruptive since the amino acid sequence is kept intact, but they still have the potential to, for example, drive cancer development (Supek et al., 2014; Zheng et al., 2014). Even nonsense mutations, which are commonly assumed to only disrupt protein function by generating truncated isoforms, have been shown to cause splicing defects instead (Cartegni et al., 2002; Mazoyer et al., 1998).

Splicing mutations within introns or exons can create novel splice sites, alter binding sites for core transcription factors, but also disrupt regulatory elements such as splicing enhancers or silencers. As a result, exon skipping, intron retention or the use of alternative splice sites can occur (Baralle and Baralle, 2005; Faustino and Cooper, 2003). A possible outcome of these splicing mutations is the introduction of premature termination codons (PTCs), which usually trigger nonsense-mediated mRNA decay (NMD). This is an mRNA surveillance mechanism that limits the amount of transcripts containing PTCs and, consequently, of the corresponding truncated protein.

The missense ATR^{M1159I} and ATR^{K1665N} mutations are predicted to cause a frameshift and the introduction of a premature stop codon. As the truncated proteins were not detected by western, we could speculate that they either had a

low stability or perhaps the mRNA production was regulated by NMD. It would be interesting to test if NMD affects the transcripts for the exon-skipped variants. For this, we could inhibit the NMD machinery and study if the aberrant variants accumulate over time by both PCR and western.

Examples that illustrate different scenarios of splicing defects include mutations that lead to development of Cystic Fibrosis, Ataxia-Telangiectasia or BRCA1-dependent ovarian and breast cancer. For instance, an intronic mutation within *Cystic Fibrosis transmembrane regulator* (*CFTR*) gene causes an imbalance between enhancer and silencer sequences, leading to exon skipping and disease development (Pagani et al., 2000). Another intronic mutation in *ATM* results in Ataxia-Telangiectasia, due to incorrect intron processing and its insertion into the transcript (intron retention) (Pagani et al., 2002). On the other hand, missense and nonsense exonic mutations in the *BRCA1* gene have been found in patients suffering ovarian and/or breast cancer. These mutations cause exon skipping, possibly due to a disruption of regulatory elements, such enhancers, or splice sites (Mazoyer et al., 1998; Rouleau et al., 2010).

These data suggest that splicing mutations may have a bigger impact than previously thought in human disorders (Baralle et al., 2009). New diagnostic techniques, in which genomic screens are accompanied by analyses of transcripts, are necessary. This will avoid the misclassification of missense, silent or nonsense mutations and allow for proper evaluation of how many human syndromes are caused by splicing mutations, which currently might be underestimated.

In order to analyse in more detail the effects of the ATR^{M1159I} and ATR^{K1665N} mutations on splicing, HSF (Human Splice Finder) and SFmap online tools were used to predict if specific regulatory motifs were affected by these mutations. In particular, a site for 9G8 (also known as SFRS7; Splicing Factor Arginine/Serine-Rich 7) was disrupted by both ATR^{M1159I} and ATR^{K1665N} mutations. This protein belongs to the serine/arginine (SR) family, which is comprised of several splice factors known to play significant roles in constitutive and alternative splicing. In addition, they are also involved in post-splicing activities, such as mRNA nuclear export, nonsense-mediated decay and mRNA translation (Huang and Steitz, 2001; Long and Caceres, 2009). Although the SR proteins usually act as splicing activators by binding to exonic enhancers

(Graveley, 2000; Shen et al., 2004), other studies show for example that the 9G8 factor strongly promotes exon 10 skipping of the Tau protein (Gao et al., 2007).

On the other hand, a site for hnRNP A1 was created when the ATR^{M1159I} mutation was present, while sites for Tra2Beta and YB1 were generated by the ATR^{K1665N} mutation. hnRNP A1 belongs to the hnRNP (heterogeneous nuclear ribonucleoproteins) family, which is a very diverse group of RNA binding proteins that display multiple functions in the regulation of transcriptional events. In particular, hnRNP A1 it is known to be a key regulator of alternative splicing, which competes with the SR proteins for common binding sites and antagonizes their function. The balance between hnRNP A1 and SR factors determines the ratio of exon inclusion/exclusion. However, apart from inducing transcriptional repression, hnRNP A1 has also been shown to function as a splicing activator in certain cases (Jean-Philippe et al., 2013).

Tra2Beta (Transformer 2 Beta Homolog, previously named SFRS10) is also part of the SR family and it is known to participate in splicing control. For instance, Tra2Beta has been also shown to bind splicing enhancer motifs and regulate exon inclusion in different genes (Kondo et al., 2004; Raponi et al., 2014; Tacke et al., 1998). Finally, YB-1 (Y box binding protein 1) belongs to a family of DNA/RNA binding proteins, which are conserved from bacteria to humans (Wolffe et al., 1992). It has been shown that Y-box proteins are essential for transcription and translation regulation (Matsumoto and Wolffe, 1998), as well as splicing control. In particular, YB-1 can regulate splicing by direct binding to exon enhancers (Stickeler et al., 2001) or via interaction with proteins from the SR family (Li et al., 2003).

In the context of the ATR Seckel mutations studied in this chapter, it is possible that both ATR^{M1159I} and ATR^{K1665N} mutations disrupt an ESE recognized by the 9G8 factor. We could speculate that splicing activated by 9G8 could be downregulated by the presence of these mutations. On the contrary, or perhaps in combination, these changes could potentially create ESS sites recognized by the hnRNPA1, Tra2Beta and YB1 proteins. These would counteract the 9G8 function by playing a role as splicing silencers/inhibitors and, thus, favouring exon skipping. To confirm these hypotheses, it would be interesting to perform specific binding assays, for example using an electrophoretic mobility shift assay (EMSA).

Over the years, most disease-related mutations were assumed to directly affect protein function thus leading to the development of the syndrome. However, there is growing evidence that splicing mutations may play a big role in human disease (Li et al., 2016; Padgett, 2012). It is clear that all known *ATR* Seckel mutations have deleterious effects on splicing regulation instead of causing a direct loss of protein function. Further studies are necessary to have a detailed understanding of the splicing process behind these disorders.

CHAPTER 5

Modelling A Non-Seckel Cancer Predisposing *ATR* Mutation

Running title: Non-Seckel cancer predisposing *ATR* mutation

Keywords: ATR, ATR signalling, CHK1 phosphorylation, Oropharyngeal cancer

5.1 Introduction

5.1.1 Link between DDR mutations, cancer and therapy

The DNA damage response (DDR) allows for the detection and repair of DNA lesions to ensure preservation of genomic stability and cell viability. Activation of this pathway may limit early tumourigenesis by acting as a barrier to the propagation of aberrant cells (Bartkova et al., 2005). Therefore, it is not surprising that anomalies in DDR players have been identified in relation to cancer progression and cancer biology. Somatic mutations in different DDR genes, including *ATR*, *ATM*, *CHK1*, *CHK2* and *BRCA1* and *BRCA2*, have been observed in multiple human tumours (Ahmed and Rahman, 2006; Antoniou et al., 2003; Lewis et al., 2007; Meijers-Heijboer et al., 2003; Menoyo et al., 2001; Poehlmann and Roessner, 2010; Zighelboim et al., 2009).

Such mutations may give cancer cells a growth advantage, allowing them to overcome the proliferation barrier set by the DDR and to survive even in the presence of replication stress and genomic instability. In fact, genomic instability is a common feature of human cancers, as it leads to increased genetic alterations that drive tumour development (Negrini et al., 2010).

In addition, as many cancers have defects in some components of the DDR, they become highly dependent on the remaining DDR pathways for survival. Therefore, therapies directed to target the remaining DDR pathways may be a selective treatment for tumours (Weber and Ryan, 2014). For instance, inhibition of ATR is preferentially toxic for cells expressing *MYC* oncogene (Murga et al., 2011), lacking tumour suppressors such as ATM or p53 (Reaper et al., 2011) or cells with deficiencies in nucleotide excision repair due to the lack of ERCC1 (Mohni et al., 2014).

There are many examples of small molecules that have been developed to specifically inhibit ATR and have been used in the development of new cancer treatments (Fokas et al., 2012; Foote et al., 2013; Prevo et al., 2012; Toledo et al., 2011). Besides, the combination of inhibitors that target both ATR and other DDR players may have a more effective outcome in cancer treatment (Ruiz et al., 2016; Sanjiv et al., 2015).

It has been recently shown that CDC25 expression dictates the sensitivity of tumours to ATR inhibitors. When cells face replication stress in the absence of ATR, there is a premature mitotic entry followed by subsequent accumulation of irreparable DNA breaks that leads to cell death. However, cancer cells deficient in CDC25 are resistant to ATR inhibition, as now CDK1 is not available to promote cell cycle progression. A way to overcome this resistance is to force mitotic entry with WEE1 inhibitors (Ruiz et al., 2016).

WEE1 phosphorylates CDK1, inhibiting its function in cell cycle progression. Therefore, WEE1 inhibition would allow the activation of CDK1 and entry into mitosis. In addition, WEE1 has been shown to play a role in the repair of replication fork stalling through its interaction with Mus81/Eme1 endonuclease (Domínguez-Kelly et al., 2011). Therefore its inhibition would also produce a more acute problem in the detection and repair of stalled forks in the absence of ATR, which would sensitize cancer cells.

Another alternative to increase cancer sensitivity would be to combine ATR and CHK1 inhibitors, as CHK1 functions upstream of both CDC25 and WEE1. Combined treatment with these inhibitors leads to replication catastrophe and specific cell death in cancer cells (Sanjiv et al., 2015).

5.1.2 *ATR* mutations causative oropharyngeal cancer

Inherited mutations in *ATR* have been associated with Seckel syndrome (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012). However, considering the major role of ATR in regulating proper chromosome segregation, fork stability and maintaining genome integrity (Cimprich and Cortez, 2008; Errico and Costanzo, 2012), it is perhaps surprising that the Seckel patients do not develop cancer.

In 2012, a new mutation in the *ATR* gene (6431A>G) was implicated for the first time in oropharyngeal cancer predisposition. Tanaka and collaborators found a heterozygous missense mutation in the ATR protein (Q2144R), located in the FAT domain, that segregated with the disease in an American family (Tanaka et al., 2012). The serine adjacent to the mutated glutamine (S2143) is a putative SQ phosphorylation site, which could potentially regulate ATR activity.

In contrast to Seckel syndrome, which is a recessive genetic disorder, this mutation translates into an autosomal dominant inherited disease. Besides, Seckel patients suffer from premature aging but they usually do not develop cancer (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012). Although ATR is considered a genome safeguard (Shiloh, 2003), the low ATR expression in Seckel patients does not correlate with a higher cancer predisposition. This is not the case for the new *ATR* mutation discovered by Tanaka and colleagues.

Their results reveal that the Q2144R mutation does not lead to a reduction in ATR expression, but fibroblasts derived from patients show lower p53 levels after HU treatment. Additionally, loss of heterozygosity was shown for the *ATR* locus in the cancerous tissue, linking ATR to the pathophysiology of oropharyngeal cancer and pointing towards a tumour suppressor role for ATR.

In recent years, the link between ATR and oro/nasopharyngeal cancer is more noticeable. It has been shown that the ATM and ATR-dependent checkpoints are compromised in nasopharyngeal carcinomas (NPC) (Poon, 2014). Dysregulation of these pathways suggests that these proteins could be possible targets for cancer treatment. In fact, NPC cells depend on CHK1 and WEE1 activity for growth and the combined inhibition of these kinases may serve as potential therapeutics for this type of cancer (Mak et al., 2015).

In this project we have focused on the study of the ATR^{Q2144R} mutation involved in oropharyngeal cancer. We show that this mutation affects both cell proliferation and ATR-dependent signalling. Our results highlight the importance that this residue might have on ATR activity and the DDR pathway.

5.2 Results

5.2.1 Description of Q2144R mutation and its conservation in evolution

In this study we intend to analyse the effect of the Atr^{Q2144R} mutation on protein function. This mutation was identified in 24 individuals spanning generations of an American family (example patient in Figure 43A), who have a higher predisposition for oropharyngeal cancer (Tanaka et al., 2012). To do this, we used both human and DT40 cells. Residue Q2144 in human corresponds to Q2163 in chicken (Figure 43B).

The Q2144 residue is present in a very conserved region amongst vertebrates (Figure 43B). This residue is also conserved between some of the most used animal models; such as *Arabidopsis*, *Drosophila* and both budding and fission yeast. When looking at the SQ motif that includes the mutated glutamine, we can observe that it is conserved in all organisms studied, except for *Arabidopsis* and *Aspergillus*. In budding yeast, it is converted to a TQ, which is still a potential substrate for PIKK phosphorylation.

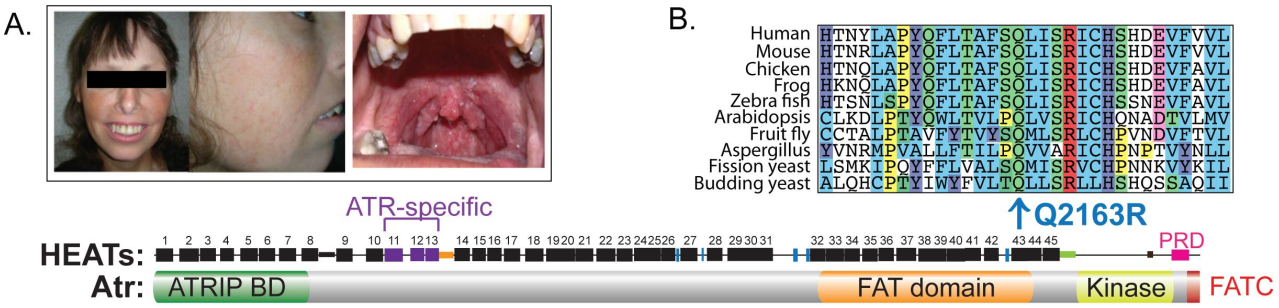


Figure 43. Details of novel *ATR* Seckel mutations.

A. Representative patients harbouring ATR^{Q2144R} mutation (taken from Tanaka et al., 2012). Patients carrying this mutation suffer from skin telangiectasias and oropharyngeal cancer predisposition.

B. Localisation and conservation of the Atr^{Q2164R} mutation. The structure of chicken Atr is shown in the figure, zooming in the region that is affected by the mutation. Alignment for multiple species shows that the Q2163 residue is well conserved through evolution.

5.2.2 Generation of Atr^{Q2163R} cell line in DT40 cells

To generate a stable cell line in DT40 cells, we targeted *Atr*^{Q2163R} cDNA to the *Ovalbumin* locus of the AID-Atr cell line, as previously carried out for the *Atr* window deletion and Seckel mutants. Southern blotting was carried out to identify positive clones (Figure 44), as explained in chapter 3. As previously, as a result, the final cell line contains both AID-Atr and Atr^{Q2163R} proteins. The effect of the *Atr*^{Q2163R} mutation was studied in the same fashion as in previous chapters; protein function and cell survival were studied after auxin addition and depletion of AID-Atr protein.

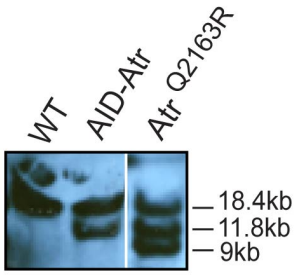


Figure 44. Generation of Atr^{Q2163R} mutant cell line. Southern screening of Atr^{Q2163R} clones for *Ovalbumin* targeting. The first lane represents wild type cells; the second lane corresponds to the AID-Atr parental cell line and the last lane represents targeting to the *Ovalbumin* locus of *Atr*^{Q2163R} cDNA.

5.2.3 *Atr*^{Q2163R} mutation causes a proliferation defect in DT40 cells

Initially, we looked at Atr expression levels in the mutants and saw that the *Atr*^{Q2163R} mutation had no effect on protein expression in our DT40 system (Figure 45A). This was consistent with the phenotype described in patients carrying the *ATR*^{Q2144R} mutation by Tanaka and colleagues (Tanaka et al., 2012).

We next analysed if the cancer-related mutation affects cell growth. As in previous chapters, we performed growth curves and counted cells every 24h in the presence and absence of auxin in the media. As seen in Figure 45B, wild type cells divide every 8 hours; while, in the absence of Atr, cells are able to grow normally up to 24h, but they start dying after this point. However, the *Atr*^{Q2163R} mutation has a profound effect on proliferation. Cells expressing this mutant protein seem to be able to divide only every 24h. Furthermore, the mutant protein shows a dominant effect over AID-Atr, which behaves as wild type, and cells expressing both of them divide every 11h.

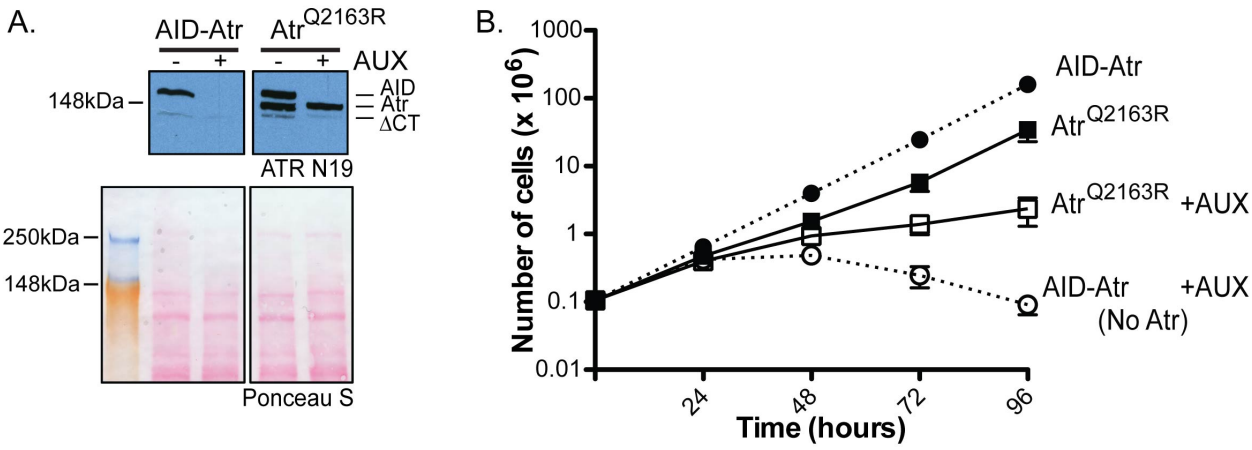


Figure 45. Modelling *Atr*^{Q2163R} mutation in chicken DT40 cells.

A. Western blotting analysis of total cell extracts prepared from AID-Atr and Atr^{Q2163R} cells using the ATR-N19 antibody (Santa Cruz). This analysis was performed in the absence and presence of auxin (0.5mM AUX treatment for 2h). Since intensity of the Atr$^{\Delta CT}$ protein from the endogenous *Atr* locus varies slightly between samples, the Ponceau S staining has also been included as a loading control.

B. Growth curve analysis of the AID-Atr and Atr^{Q2163R} cell lines following treatment with AUX. Cell numbers were recorded every 24h for 5 days and are corrected for culture dilution. Error bars represent the SD from at least three experiments.

We also studied cell proliferation in a smaller window after depleting AID-Atr in the mutant cells. We looked at BrdU incorporation and mitotic entry in the first cell cycle after depletion of AID-Atr and confirmed that mutant cells are progressing more slowly through the cell cycle (Figure 46). This was independent of the cell cycle phase analysed (S phase or mitosis). Although these experiments have only been performed once and further repeats are needed to confirm these results, we have reproduced the defects seen in the absence of Atr previously seen in the laboratory (Eykelenboom et al., 2013).

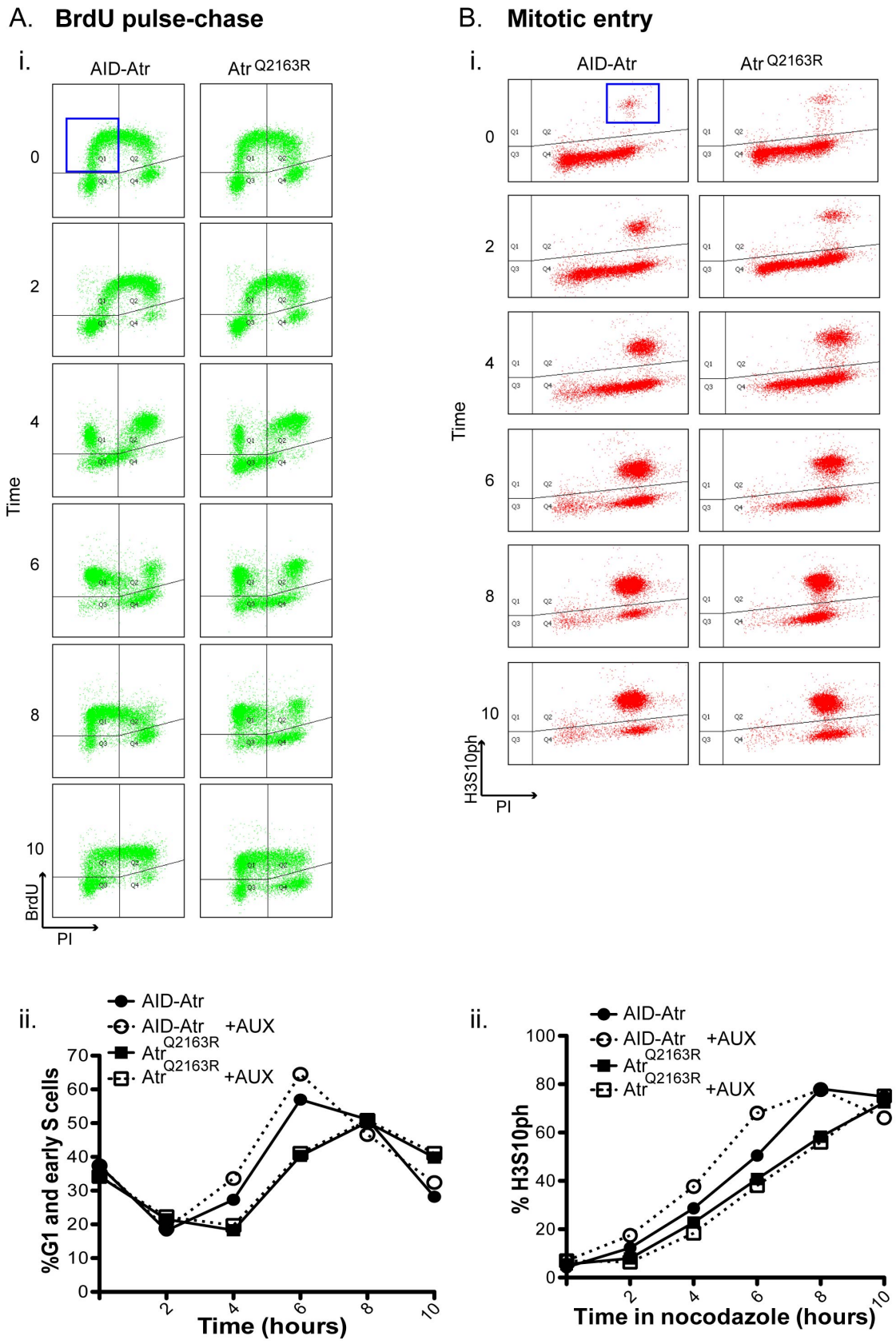


Figure 46. Cell cycle progression of Atr^{Q2163R} in the first cell cycle after depleting AID-Atr.

A. Proliferation profile following a BrdU pulse-chase. AID-Atr and Atr^{Q2163R} cell lines were grown in the presence or absence of AUX and were analysed by flow cytometry (BrdU/PI staining) for cell-cycle position at 0, 2, 4, 6, 8 and 10h following BrdU pulse-chase. i. Examples of raw data for *AID-Atr* and Atr^{Q2163R} cell lines are shown in the figure. By plotting BrdU versus PI we were able to distinguish cells in G1 (Q3), in early (Q1) and late (Q2) S-phase and in G2/M (Q4). Quantification was done by plotting cells in Q1. As cells progress through the cell cycle and divide, BrdU -/+ populations get closer over time; therefore, Q1 includes cells in early S-phase, but also some cells in G1. ii. Plot representing percentage of cells in G1/early S for the cell lines indicated. Values correspond to only one repetition; thus, no error bars are included.

B. Mitotic index. AID-Atr and Atr^{Q2163R} cell lines were grown in the presence or absence of AUX and were analysed by flow cytometry (H3S10ph/PI staining) for mitotic entry at 0, 2, 4, 6, 8 and 10h following addition of nocodazole. i. Examples of raw data for *AID-Atr* and Atr^{Q2163R} cell lines are shown in the figure. Cells in Q2 were quantified (blue quadrant) representing cells positive for H3S10 phosphorylation (mitotic marker). ii. Plot representing H3S10ph accumulation in the presence of nocodazole for the cell lines indicated. Values correspond to only one repetition; thus, no error bars are included.

5.2.4 Increased cell death and chromosome breakages in Atr^{Q2163R} cells in the absence of exogenous damage

To discern whether proliferation defects are due to an actual slower cell division or to an increase in cell death, we briefly looked at the percentage of apoptotic/dead cells in the absence of damage over a lapse of 5 days (Figure 47). These results have only been carried out once, but they show that Atr^{Q2163R} mutant cells do indeed have a higher basal level of cell death (from 20% in day 1 to about 50% in the later days). However, this is only true for cells expressing only the Atr^{Q2163R} protein. In the presence of AID-Atr, cells have the same percentage of death as wild type cells, which is only about 5% cell death in a healthy wild type culture. Besides, when treating mutant cells with AUX for a longer period of time (>1 week), DT40 mutant cells die. This suggests that the Atr^{Q2163R} protein might only be able to fulfil some Atr roles, causing an accumulation of damage over time that could trigger cell death. However, further repeats are necessary to confirm these results.

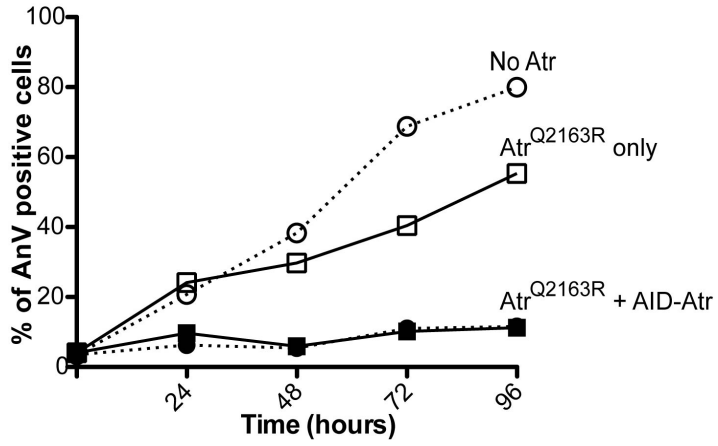


Figure 47. Cell death in the absence of damage for AID-Atr and Atr^{Q2163R} cell lines. Cells were counted every 24h for 5 days and 1x10^6 cells were collected each day and stained with Annexin V/ PI. Flow cytometry analyses were performed in the absence of any exogenous damage to check for basal levels of cell death. Annexin V positive cells were plotted in the graph. Values correspond to only one repetition; thus, no error bars are included.

Previous studies in the laboratory have shown that, in the absence of Atr, there is an increase in chromosome loss and chromatid gaps/breaks (Eykelenboom et al., 2013). Therefore, we measured the impact that the *Atr^{Q2163R}* mutation had upon the chicken macro-chromosomes' structure. To do this, we analysed cells at metaphase; a cell cycle stage when chromosomes are fully condensed and chromosome abnormalities can be easily spotted. The chicken macro-chromosomes 1, 2 (triploid in DT40), 3, 4 and Z, which represent about half of the genetic material of DT40 cells, can be distinguished in Figure 48A.

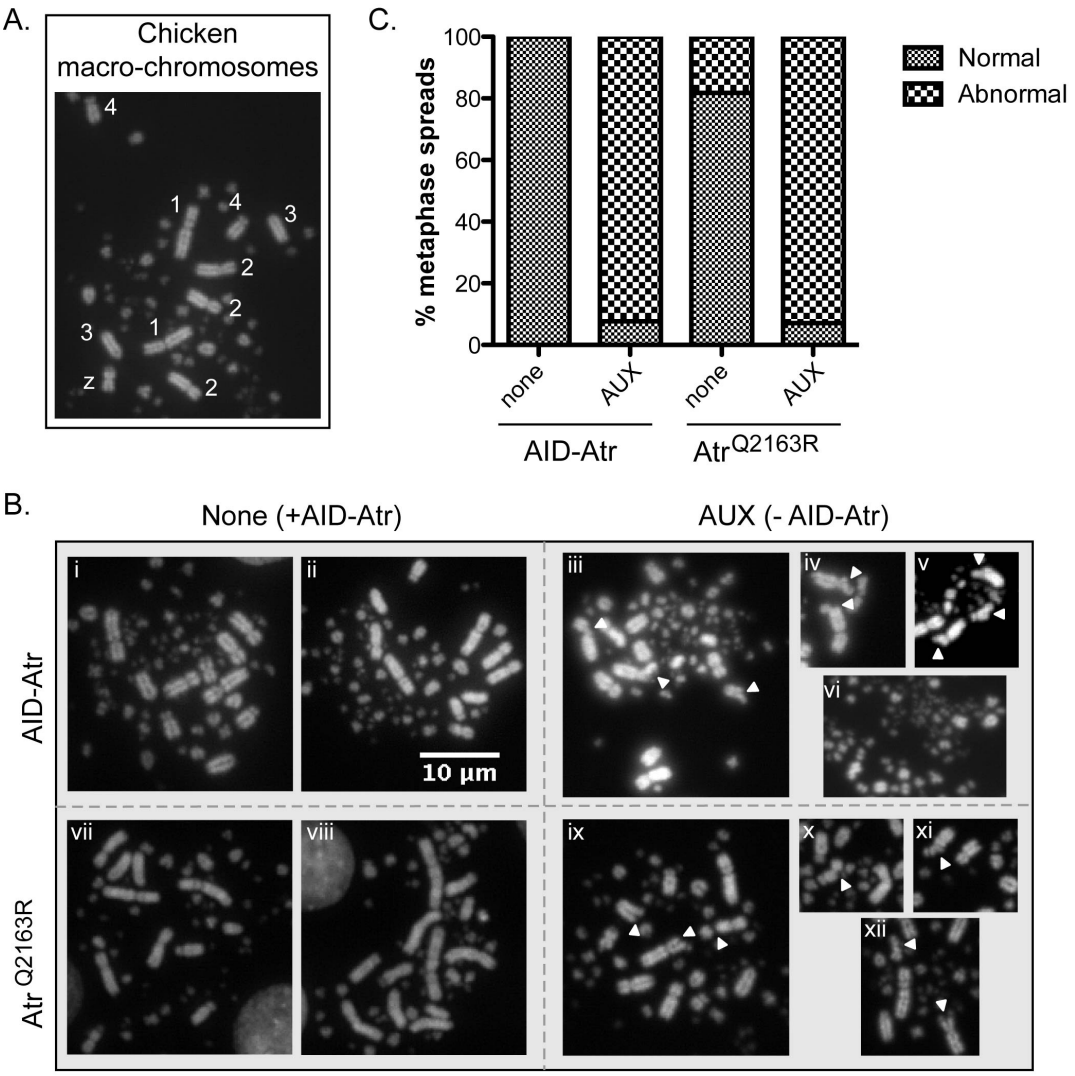


Figure 48. Increased chromatid gaps and missing chromosomes in the presence of the *Atr*^{Q2163R} mutation.

A. Identification of chicken macro-chromosomes. Chromosome 1 corresponds to the biggest submetacentric chromosome, chromosome 2 is trisomic in DT40; chromosome 3 is telocentric, chromosome 4 is acrocentric, and chromosome z is monosomic, being one of the sex chromosomes. In *G. gallus*, zz animals are male, while zw animals are female. Therefore, the DT40 cells that we are studying come from a female bird.

B. Examples of normal spreads, in the absence of AUX, are shown for AID-Atr (i, ii) and Atr^{Q2163R} cell lines (vii and viii). Examples of chromatid breaks and missing chromosomes, in the presence of AUX, are given for both cell lines (ii-vi and ix-xii, respectively), including an example of chromosome pulverization or shattering (vi). Chromosome gaps are indicated with white arrowheads. Expanded view of representative chromatid gaps are included for the AUX treated samples.

C. Quantification of normal versus abnormal spreads in AID-Atr and Atr^{Q2163R} cell lines in the presence and absence of AUX. Abnormal spreads include those with chromatid breaks and/or missing chromosomes. Only 15 spreads were scored per condition. Replicates are necessary to confirm these results and include error bars.

Chromosome spreads were prepared from AID-Atr and Atr^{Q2163R} cells grown in the presence or absence of AUX for 24h (Figure 48B). In cells containing AID-Atr (absence of AUX), the majority of spreads looked normal for the AID-Atr cell line (i and ii). However, although most spreads were also normal for the Atr^{Q2163R} cell line (vii and viii), there seemed to be a slightly higher frequency of abnormal spreads (up to 20%) even in the presence of AID-Atr. After 24h of AUX treatment, spreads for Atr^{Q2163R} cells showed multiple defects, such as chromatid breaks and/or missing chromosomes (ix to xii). The presence of abnormal spreads, including chromatid gaps, missing chromosomes or both, increased up to about 90% (Figure 48C). These defects were previously reported in cells expressing no functional Atr by Eykelenboom et al. and were reproduced in our experiments (AID-Atr + AUX control, iii to vi). These results indicate that the mutant protein behaves as the null cell line, however, this experiment was only performed once and very low cell numbers were scored in each case. Hence, repetition is needed to obtain accurate numbers and proper quantification on chromosome loss and breakage in the presence of the Atr^{Q2163R} mutation.

When scoring the chromosomal abnormalities, we noticed several spreads with too many abnormalities for accurate scoring. The presence of multiple shattered chromosomes made the identification of the different chicken macro-chromosomes difficult. For these reasons, these spreads were not included in the analysis (Figure 48B, vi).

5.2.5 Atr^{Q2163R} mutation abrogates Atr signalling in response to HU even in the presence of AID-Atr

To explore if checkpoint functions are intact in the Atr^{Q2163R} mutant, we then looked at the activation of Chk1 in response to HU and other damaging agents. Our results show that mutant cells are only able to phosphorylate Chk1 upon HU treatment in the presence of AID-Atr. However, after depletion of AID-Atr with AUX (0.5mM) followed by induction of replication fork stalling with HU (1mM), phosphorylation of Chk1 is fully abrogated in mutant cells (Figure 49A). This indicates that the Atr^{Q2163R} mutant behaves as a null cell line and is unable to activate Atr signalling in response to HU.

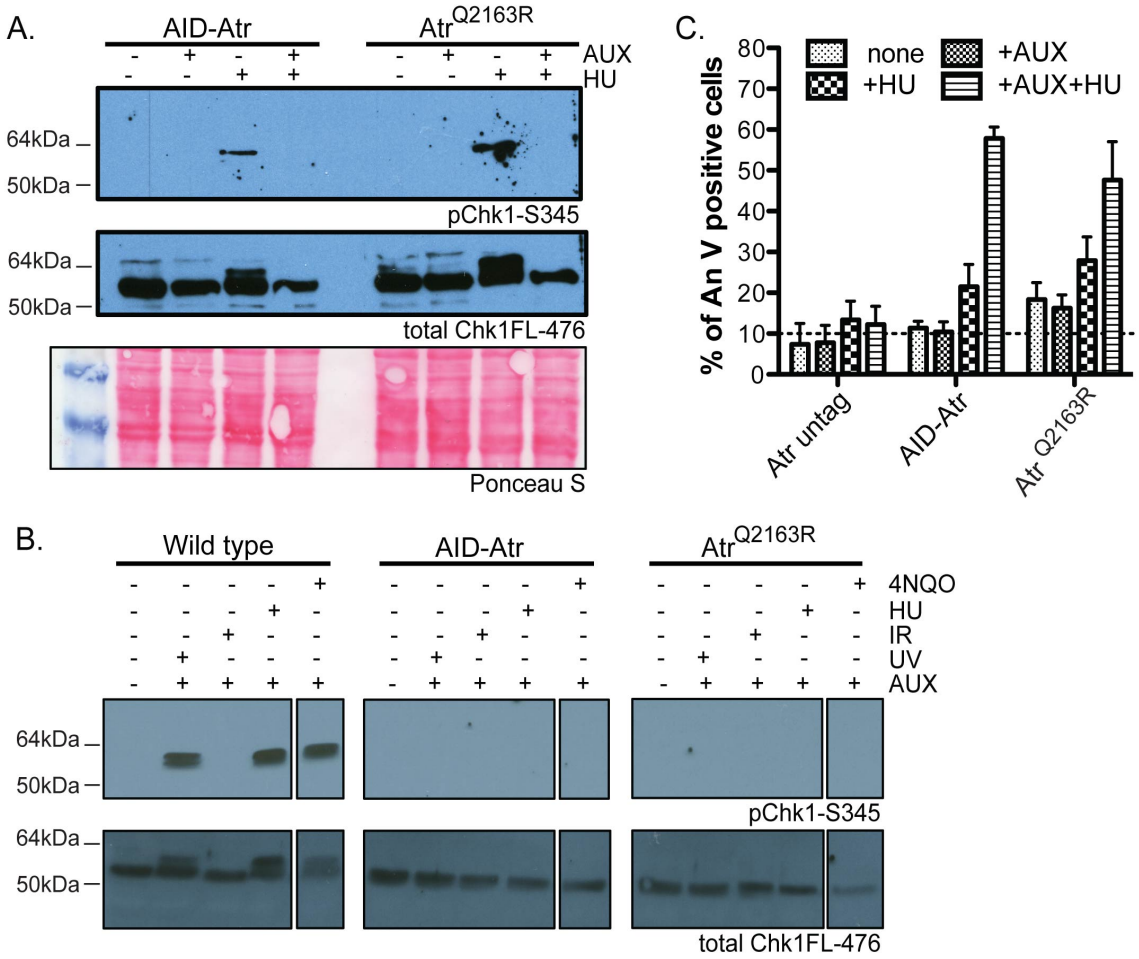


Figure 49. Checkpoint activation in response to HU in the AID-Atr and Atr^{Q2163R} cell lines.

A. Western analysis of Chk1-S345 phosphorylation upon HU treatment (4h) following growth in the presence or absence of AUX (2h). Western blotting of protein extracts harvested from cell lines as indicated. A Chk1 antibody that recognises phosphorylation on S345 (Cell Signalling) was used to analyse checkpoint proficiency. An antibody against total Chk1 (FL-476 Santa Cruz) was used as a loading control.

B. Further study of Chk1 phosphorylation in response to different DNA damage, including UV (10J/m^2), IR (3Gy), HU (1mM) and 4NQO (0.2µg/ml). In this case, all samples treated with damage had been previously treated with AUX for 2h to deplete AID-Atr. No damage or AUX only controls were added in this case.

C. HU induced apoptosis. Annexin V/propidium iodide (PI) flow cytometry analysis after treatment with HU (4h) following growth in the presence or absence of AUX (2h). Quantification was done by measuring Annexin V positive cells. Error bars represent the SD from at least three experiments.

This is also the case when cells are treated with other types of damage (Figure 49B), such as UV (10J/m^2) and 4NQO (0.2µg/ml). Chk1 phosphorylation

after IR treatment (3Gy) was not detected in WT cells, perhaps the timepoint was too early for an Atr-dependent response (30min).

We have also confirmed that Atr^{Q2163R} cells are not able to activate the checkpoint by examining the percentage of apoptotic cells upon auxin and HU treatment. An increase in apoptosis shows that Atr^{Q2163R} cells are incapable of activating the checkpoint after both HU and auxin treatment (Figure 49C).

5.2.6 Generation of ATR^{Q2144R} cell line in human cells

Chicken DT40 cells are very convenient due to their high efficiency in gene targeting experiments; however, these cells are not ideal when reagents such as antibodies are needed. A lot of these reagents, such as total RPA, ATRIP and phospho-Chk2 antibodies (data not shown), do not work in DT40 cells, making it difficult to extract information in some cases.

For this reason, to dissect the part of the ATR signalling pathway that the Q2144R mutation affects, we attempted to generate a human cell line containing this mutation. With the discovery of the CRISPR-Cas9 technology, targeting in human cells became more approachable (Cong et al., 2013; Mali et al., 2013). This technology, which derived from the adaptive immune system of bacteria and archaea, exploits the bacterial DNA nuclease Cas9 to cut sequence specifically when guided with an RNA directed to a target region (Cho et al., 2013; Mali et al., 2013). Therefore, we used this technology for generating an ATR^{Q2144R} cell line.

Cancer patients are heterozygotes for this locus; however, heterozygosity is lost in the cancerous tissue. Therefore, our interest was to generate both the homozygote and heterozygote cell lines. Heterozygotes would allow us to understand what happens in the patients, while the homozygotes would show a clearer effect of this mutation on ATR function and could give new insights on a potential key residue in ATR signalling.

5.2.6.1 CRISPR strategy

To generate a cell line containing the ATR^{Q2144R} mutation, hTERT-RPE1 (hereafter RPE1) and HCT116 cells were used in parallel experiments. The approach that we followed was developed by Davis and Maizels (Figure 50) and

relies on an alternative route of homologous repair that is specific for nicks in DNA (Davis and Maizels, 2014). Thus, to introduce the Q2144R point mutation, we used the nickase version of Cas9 (D10A). This enzyme was used in combination with two gRNAs (marked in turquoise in Figure 50B), which target the same strand (transcribed strand). These gRNAs are 88 and 94bp away from the site to mutate; however, they have higher GC content than other potential and closer gRNAs (in an AT-rich region) and therefore we predict that they will be more specific. gRNA quality was also confirmed with the MIT CRISPR design tool (http://crispr.mit.edu/). This online tool identifies all possible gRNAs in a given sequence and scores them depending on the likelihood of off-target binding. gRNAs were classified as "high quality", with scores of 90 and 73 for gRNA1 and 2, respectively.

The Cas9 enzyme is directed by the gRNAs to cleave the DNA three base pairs upstream of the protospacer adjacent motif (PAM) recognition sequence (5'-NGG-3') (Jinek et al., 2012). The D10A Cas9 is engineered to cleave only one strand of DNA, opposite to the recognition site (Ran et al., 2013a). Once cut, we provide a template that includes the desired mutation, which will be used for homologous directed repair and will mediate the introduction of the point mutation. This template also includes extra mutations in the regions targeted by the gRNAs that allow the template to elude Cas9 cleavage (marked in red in Figure 50B). In particular, the template used for HDR at nicks was a ssDNA oligo (233nt) corresponding to exon 38, which covered the nicked region (Figure 50B).

The use of two gRNAs (rather than a single gRNA as in Davis and Maizels report) is expected to mediate a quicker/easier repair. After these gRNAs lead Cas9 to the target region, the DNA would be nicked both upstream and downstream of the site to be mutated. In this situation, when the transcription machinery localises to the break site, we hypothesise that the cut strand would be completely peeled off and this would allow the ssDNA oligo to come in directly, allowing for repair (Figure 50C). To test this assumption, control transfections using only the upstream gRNA were also performed.

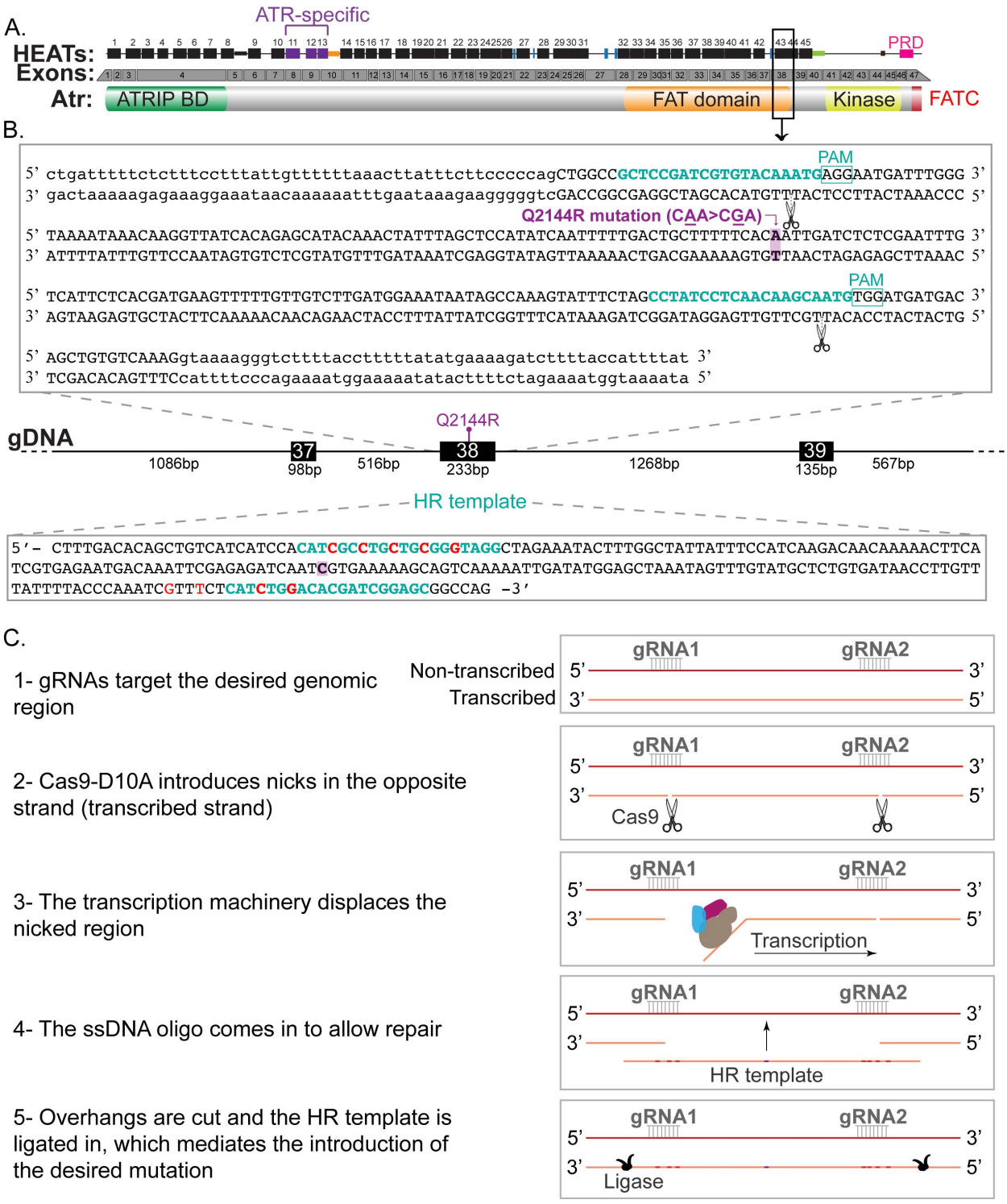


Figure 50. CRISPR strategy to generate a human cell line containing the *ATR*^{Q2144R} mutation.

A. Alignment of ATR exons with ATR protein structure and HEAT repeat region. Residue Q2144 is located in exon 38, which spans HEAT repeat 43 and falls within the end of the FAT domain.

B. Details of targeted region. Exon 38 and flanking intronic sequences are shown in the figure along with the details of the gRNAs (in turquoise) and the HR template (purple marks the Q2144R mutation and red the extra mutations that avoid Cas9 cleavage).

gRNA1 and gRNA2 stimulate Cas9 cleavage 88bp and 94bp away from the site to mutate, while the HR template spans the whole exon 38 (233nt).

C. Diagram of alternative HDR at nicks, in combination with the use of CRISPR/Cas9 technology for introduction of the Q2144R mutation. gRNAs target the desired genomic region, guiding the Cas9 nickase to cut upstream and downstream of the site to mutate. These nicks are introduced on the transcribed strand, which allows the nicked region to be displaced by the transcription machinery. When a ssDNA donor is provided, it can be used as a template mediating the introduction of the desired mutations. This allows for the mutations to be incorporated in one allele; the generation of a mutant homozygote cell line would depend on mismatch repair.

Finally, Davis and Maizels have shown that this HDR repair pathway requires BRCA1; however, other proteins such as BRCA2 and Rad51, which are also essential for the canonical HDR pathway (Prakash et al., 2015), inhibit the alternative path (Davis and Maizels, 2014). For this reason, to direct the repair towards the alternative route, we also used either a Rad51 inhibitor or a BRCA2 siRNA to deplete this protein prior to our CRISPR transfections. Efficiency of BRCA2 siRNA was tested in RPE1 cells (Figure 51). Knockdown was approximately 95% efficient at 24-48h, with a slight increase in expression starting at 72-96h. During CRISPR experiments, cells were transfected with gRNA/Cas9 plasmids 48h after knocking down BRCA2.

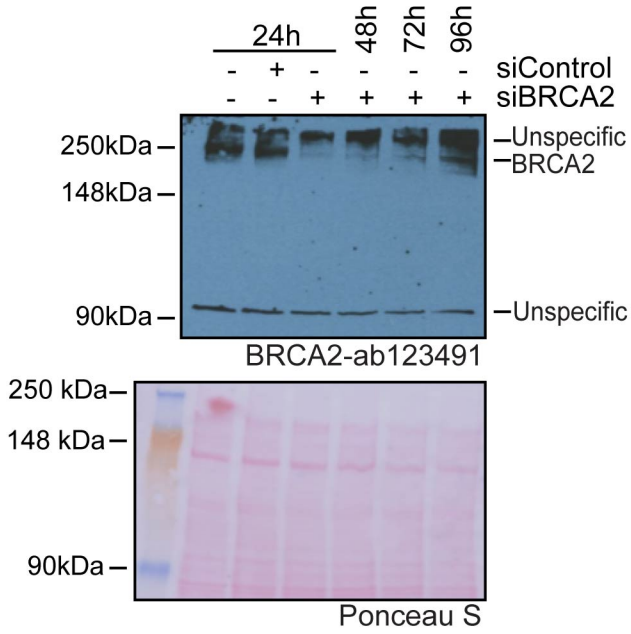


Figure 51. Efficiency of BRCA2 siRNA knockdown. RPE1 cells were transfected with siBRCA2 and collected 24, 48, 72 and 96h later to check knockdown efficiency. Untransfected and siControl samples were collected together with the 24h timepoint. Total cell extracts were analysed by western blotting using the BRCA2-ab123491 antibody (Abcam). Two unspecific bands were detected when using this antibody, with the lower band being used as a loading control.

5.2.6.2 CRISPR results

The introduction of the desired point mutation disrupted an MfeI restriction site. Additionally, in the HR template we also included a mutation to avoid Cas9 cleavage at gRNA2 targeting site, which created a new restriction site for BtgZI enzyme. Therefore, to screen both hTERT-RPE1 and HCT116 clones, we performed a PCR followed by a diagnostic digest consisting of two separate MfeI and BtgZI digests (Figure 52A). Using this screening method, we observed that the strategy, and also the cell lines used in the CRISPR experiments had a big impact on the number of clones and targeting results that were obtained (Figure 52B).

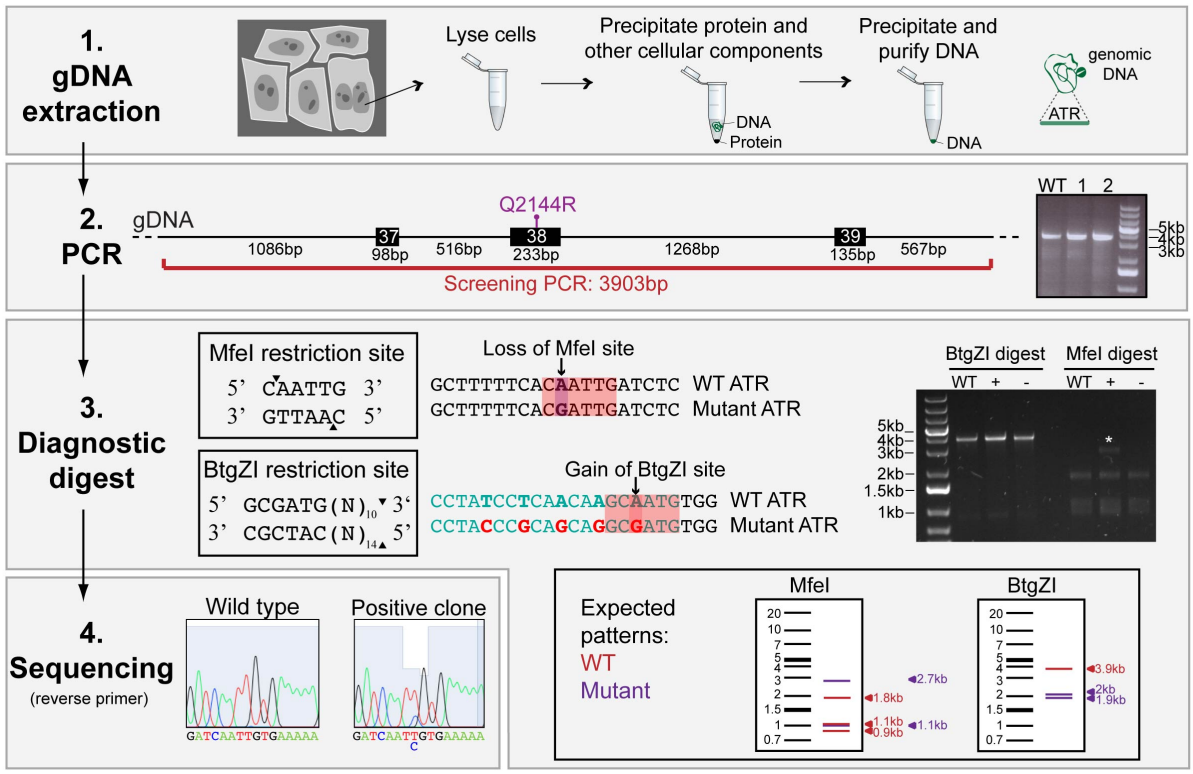
hTERT-RPE1 is a diploid immortalized cell line derived from retinal pigment epithelium cells by constitutive expression of human telomerase, while HCT116 is a near-diploid transformed cell line derived from human hereditary colon cancer. Although both RPE1 and HCT116 cells are considered chromosomally stable cell lines (Thompson and Compton, 2008), HCT116 cells show some extent of microsatellite instability (Lengauer et al., 1997; Masramon et al., 2000). Therefore, we initially used RPE1 cells due to their more normal karyotype. However, this cell line is not ideal for single cell isolation experiments (data not shown). As a consequence, fewer RPE1 clones were obtained, all of which were negative in that they did not contain the desired mutation. In contrast, as HCT116 cells might have higher potential for rearrangements, we obtained more clones in these transfections, although we could only detect the desired mutation in two of them (Figure 52B). These two clones originated from a transfection in which we only used one gRNA (gRNA1) in combination with BRCA2 siRNA. In the diagnostic digests, we could identify the positive clones by the loss of MfeI restriction site, as marked by the asterisk in Figure 52A, box 3. In those clones, both WT and mutant patterns co-existed, meaning that they were possibly heterozygotes. Sequencing confirmed the existence of a double peak, although the peak for the mutant nucleotide was smaller (Figure 52A, box 4). Besides, we also noticed that the pattern for BtgZI digest did not change in any of the positive clones (Figure 52A, box 3).

To clarify the mutations incorporated in each *ATR* allele in the two positive clones, the 3.9kb PCR product used as a template in the diagnostic digest

was cloned into a pGEM-T easy vector. Several colonies were sent for sequencing and results were aligned to both wild type and the expected mutant sequence (HR template; Figure 52C). When analysing the results, we could detect more than two alleles (up to 4 different ones). This suggests that clones isolated using cloning disks were a mixed population of 2 different genetic backgrounds. This explains the presence of a smaller peak for the mutant nucleotide (C) versus wild type (T) at position 6431 (Figure 52A, box 4).

Within the four alleles detected, we could see a number of the desired point mutations, including the 6431A>G change and some of the extra mutations to avoid Cas9 cleavage, as well as other spontaneous mutations that might have occurred during the repair process. As mentioned previously, we were unable to detect the gain of a BtgZI restriction site by diagnostic digest. This was also confirmed by sequencing and shows that none of the positive clones incorporated the extra mutations to elude Cas9 cleavage at gRNA2 targeting site, which were included in the ssDNA oligo provided as HR template. This suggests that HR directed by the ssDNA oligo may have occurred after the desired point mutation (in purple) but before the extra mutations at gRNA2 (in red). Both HCT116 positive clones showed the same digest and sequencing patterns, although only clone 1 appears in Figure 52C as an example.

A.



B.

Strategy	Cell line	HR template	gRNAs	Rad51i/ siBRCA2	Total number of clones		Positive clones
Davis & Maizels strategy	hTERT-RPE1	ssDNA oligo: 233nt	gRNA1	si control	21	97	-
				siBRCA2	21		
			gRNA1&2	si control	36		
				siBRCA2	19		
	HCT116		gRNA1	Rad51i	59	199	-
				siBRCA2	48		2
			gRNA1&2	Rad51i	45		-
				siBRCA2	47		-

C.

```
      1                                                                                                  100
WT    CTGGCCGCTCCGATCGTGTACAAATGAGGAATGATTTGGGTAAAATAAACAAGGTTATCACAGAGCATACAAACTATTTAGCTCCATATCAATTTTTGAC
Mutant CTGGCCGCTCCGATCGTGTCCAGATGAGAAACGATTTGGGTAAAATAAACAAGGTTATCACAGAGCATACAAACTATTTAGCTCCATATCAATTTTTGAC
c1    CTGGCCGCTCCGATCGTGTACAAATGAGGAATGATTTGGGTAAAATAAACAAGGTTATCACAGAGCATACAAACTATTTAGCTCCATATCAATTTTTGAC
c2    CTGGCCGCTCCGATCGTGTCCAGATGAGAAATGATTTGGGTAAAATAAACAAGGTTATCACAGAGCATACAAACTATTTAGCTCCATATCAATTTTTGAC
c3    CTGGCCGCTCCGATCGTGTACAAATGAGGAATGATTTGGGTAAAATAAACAAGGTTATCACAGAGCATACAAACTATTTAGCTCCATATCAATTTTTGAC
c4    CTGGCCGCTCCGATCGTGTCCAGATGAGAAATGATTTGGGTAAAATAAACAAAGTTATCACAGAGCATACAAACTATTTAGCTCCATATCAATTTTTGAC

      101                                                                                                200
WT    TGCTTTTTCACAATTGATCTCTCGAATTTGTCATTCTCACGATGAAGTTTTTGTTGTCTTGATGGAAATAATAGCCAAAGTATTTCTAGCCTATCCTCAA
Mutant TGCTTTTTCACGATTGATCTCTCGAATTTGTCATTCTCACGATGAAGTTTTTGTTGTCTTGATGGAAATAATAGCCAAAGTATTTCTAGCCTACCGCAG
c1    TGCTTTTTCACAATTGATCTCTCGAATTTGTCATTCTCACGATGAAGTTTTTGTTGTCTTGATGGAAATAATAGCCAAAGTATTTCTAGCCTATCCTCAA
c2    TGCTTTTTCACAATTGATCTCTCGAATTTGTCATTCTCACGATGAAGTTTTTGTTGTCTTGATGGAAATAATAGCCAAAGTATTTCTAGCCTATCCTCAA
c3    TGCTTTTTCACAATTGATCTCTCGAATTTGTCATTCTCACGATGAAGTTTTTGTTGTCTTGATGGAAATAATAGCCAAAGTATTTCTAGCCTATCCTCAA
c4    -GCTTTTTCACGATTGATCTCTCGAATTTGTCATTCTCACGATGAAGTTTTTGTTGTCTTGATGGAAATAATAGCCAAAGTATTTCTAGCCTATCCTCAA

      201                              233
WT    CAAGCAATGTGGATGATGACAGCTGTGTCAAAG
Mutant CAGGCGATGTGGATGATGACAGCTGTGTCAAAG
c1    CAAGCAATGTGGATGATGACAGCTGTGTCAAAG
c2    CAAGCAATGTGGATGATGACAGCTGTGTCAAAG
c3    CAAGCAATGTGAATGATGACAGCTGTGTCAAAG
c4    CAAGCAATGTGGATGATGACAGCTGTGTCAAAG
```

Figure 52. Results of CRISPR attempt to generate a human cell line containing the *ATR*^{Q2144R} mutation.

A. Screening strategy. gDNA was extracted from isolated clones (1), followed by PCR (2) and diagnostic digest with both MfeI and BtgZI enzymes (3). Clones positively identified by digest were also confirmed by sequencing (4). Examples of these four steps are shown in the figure. The asterisk (3) marks the appearance of an undigested band in positive clones due to the loss of an MfeI restriction site.

B. Summary of clones obtained in each strategy. Table shows the total and positive number of clones obtained in each transfection. Details of cell lines used, gRNAs, HR template and siRNAs/inhibitors added are also provided.

C. Examples of HCT116 targeted alleles. Exon 38 sequence of four different pGEM-T easy colonies (c1-4) was aligned with wild type and mutant controls: c1) wild type; c2) an allele containing three of the four mutations to avoid Cas9 cleavage at gRNA1 targeting site; c3) an allele containing none of the desired mutations, but an unexpected nucleotide change downstream of gRNA2 targeting site; and c4) an allele containing three of the four mutations to avoid Cas9 cleavage at gRNA1 targeting site, plus an unexpected nucleotide change downstream of gRNA1 targeting site, and an extra single nucleotide deletion a few base pairs before the desired 6431A>G (Q2144R) mutation. gRNA sequences are highlighted in turquoise, while mutations are shown in red.

Unfortunately, the alleles that contained the 6431A>G (Q2144R) mutation, also contained an extra single nucleotide deletion that caused a change in the frame. As a consequence, we were unable to obtain neither heterozygous nor homozygous cell lines using the CRISPR methodology.

Further optimization of the CRISPR strategy might be necessary for the successful generation of a human cell line expressing the ATR^{Q2144R} mutation. This would be essential to gain additional insights on the importance of this residue in supporting ATR function.

5.3 Discussion

Defects in ATR have been reported in several cancers (Fang et al., 2004; Menoyo et al., 2001; Zighelboim et al., 2009). In particular, a non-Seckel mutation in the ATR FAT domain (ATR^{Q2144R}) was recently reported to cause oropharyngeal cancer predisposition (Tanaka et al., 2012). As the FAT domain is conserved amongst PIKKs and it is thought to regulate kinase activity (Bosotti et al., 2000; Lempiäinen and Halazonetis, 2009), it is conceivable that this new *ATR* mutation has a direct impact on protein function leading to the development of cancer. To evaluate this possibility, we decided to model the ATR^{Q2144R} mutation (Atr^{Q2163R} in chicken) using a DT40 model system previously established in the laboratory.

In order to test if the Atr^{Q2163R} mutant was able to perform the essential function of Atr, we carried out growth analyses in the presence and absence of AUX. Remarkably, when cells expressed only the Atr^{Q2163R} protein, they showed a partial growth defect. That is, cells were not as sick as in the complete absence of Atr, but they were still slower growing than WT, showing an intermediate phenotype. Moreover, even in the presence of AID-Atr, cells expressing the Atr^{Q2163R} protein, did not fully recover from the proliferation defect. Mutant protein seems to have a dominant negative effect on the 'WT' form of Atr (AID-Atr).

To further investigate the proliferation defects, we were intrigued to know if these were due to an actual slower growth or perhaps a higher basal level of cell death. To discern between these possibilities, we analysed the percentage of apoptotic cells in the absence of damage, in parallel to a karyotypic analysis in which the integrity of chicken macro-chromosomes was scored. The results indicate that there is a higher basal level of cell death and chromosomal abnormalities in cells expressing only the Atr^{Q2163R} protein. Multiple chromosomes breaks and missing chromosomes were detected, indicating that the Atr^{Q2163R} mutant was not able to support faithful cell division and chromosome segregation. These defects could potentially be leading to the increase detected in the apoptotic cell population over time, after AID-Atr is depleted.

It is possible that the Atr^{Q2163R} protein might be able to fulfil only for some roles of Atr and, after a few cell cycles, cells are not able to live with only this

copy of Atr. As indicated, these cells would be kept alive during a longer time than when cells completely lack Atr. The dominant negative effect seen in the presence of AID-Atr could perhaps be explained if we speculate that mutant Atr^{Q2163R} could still bind to substrates, but was unable to activate effector proteins. This would be affecting some cells in the population that would grow more slowly until the pathway is activated by functional copies of AID-Atr.

As ATR's role in the DNA damage response is well established, we also wanted to study if the Atr^{Q2163R} mutant was able to respond to HU damage. Apoptosis (using Annexin V/PI staining) and Chk1 phosphorylation analyses were performed upon HU treatment. Results indicate that the Atr^{Q2163R} mutant causes a defect in checkpoint activation. In particular, in the presence of the Atr^{Q2163R} mutation, there is a complete abrogation of Chk1 phosphorylation, together with the accumulation of apoptotic cells upon HU treatment and depletion of AID-Atr.

Taken together, these data indicate that a single mutation in the *Atr* gene (Atr^{Q2163R}) has a profound effect on protein activity in chicken DT40 cells. However, it may be particularly difficult to extrapolate the proliferation defects displayed by a chicken lymphoma cell line to oropharyngeal cancer predisposition observed in patients carrying the ATR^{Q2144R} mutation. It is currently unknown why patients harbouring this mutation should be specifically predisposed to oropharyngeal and not other types of cancer.

Mutations in the *ATR* gene are also known to cause Seckel syndrome. Patients suffering from this disease express very low levels of ATR due to defects in splicing (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012). This syndrome is characterized by microcephaly, developmental defects and accelerated aging. Considering the established role for ATR in preserving genomic integrity, it is intriguing that Seckel patients do not develop cancer, an effect that has been reproduced using mouse models (Murga et al., 2009; Ruzankina et al., 2007).

It has been previously shown for other genes that mutations that affect protein expression leading to different degrees of residual protein activity, can have very different outcomes. For instance, different mutations in the *XPF* gene (*Xeroderma Pigmentosum complementation group F*, also known as *Excision Repair Cross-Complementation Group 4, ERCC4*), which encodes a protein

involved in nucleotide excision repair, can cause either cancer predisposition or a syndrome characterized by accelerated aging. These effects are thought to be a direct consequence of the different levels of remaining ERCC4 functional protein in each situation (Niedernhofer et al., 2006).

In the same fashion, Murga and colleagues already proposed that the amount of residual ATR might be crucial to determine the outcome of diminished ATR function, that is cancer development or Seckel syndrome (Murga et al., 2009). If ATR activity is reduced below a certain level, the accumulation of irreparable genomic lesions might be incompatible with the uncontrolled growth of a potential cancerous cell. Seemingly, above this threshold there might be an activity window where genomic instability can potentially stimulate malignant transformation.

This assumption would explain the different outcome of the ATR^{Q2144R} mutation in contrast to other *ATR* mutations related to Seckel syndrome. As Seckel mutations affect splicing, the residual level of ATR expressed in the cells might generate a setting that is not compatible with cancer development. In contrast, the ATR^{Q2144R} mutation seems to disturb ATR function directly, but the levels of expressed protein would remain above the threshold necessary to direct genomic instability towards cancer progression.

It is important to mention that residue Q2144 is located next to a serine in ATR structure, thus being part of a potential SQ site. As this SQ motif is well conserved through evolution, we hypothesize that S2143 could be phosphorylated. This event could be key for ATR signaling. Further phosphorylation analysis, using a specific phospho antibody or mass spectrometry, would be necessary to clarify this matter.

In this study, we propose a model in order to explain the importance of residue ATR^{Q2144} (and neighbouring ATR^{S2143}) in the ATR-dependent response to stalled replication forks. In this model, this SQ motif could be important for either stimulating or activating the catalytic activity of ATR or to mediate ATR recruitment to sites of damage (Figure 53). It is possible that this residue is located in a region that directly mediates ATR interaction with partner proteins, such as TOBP1, RAD9 (from the 9-1-1 complex), etc. Alternatively, as this

residue is located in the FAT domain, it could potentially affect proper folding of ATR structure and thus its catalytic activity.

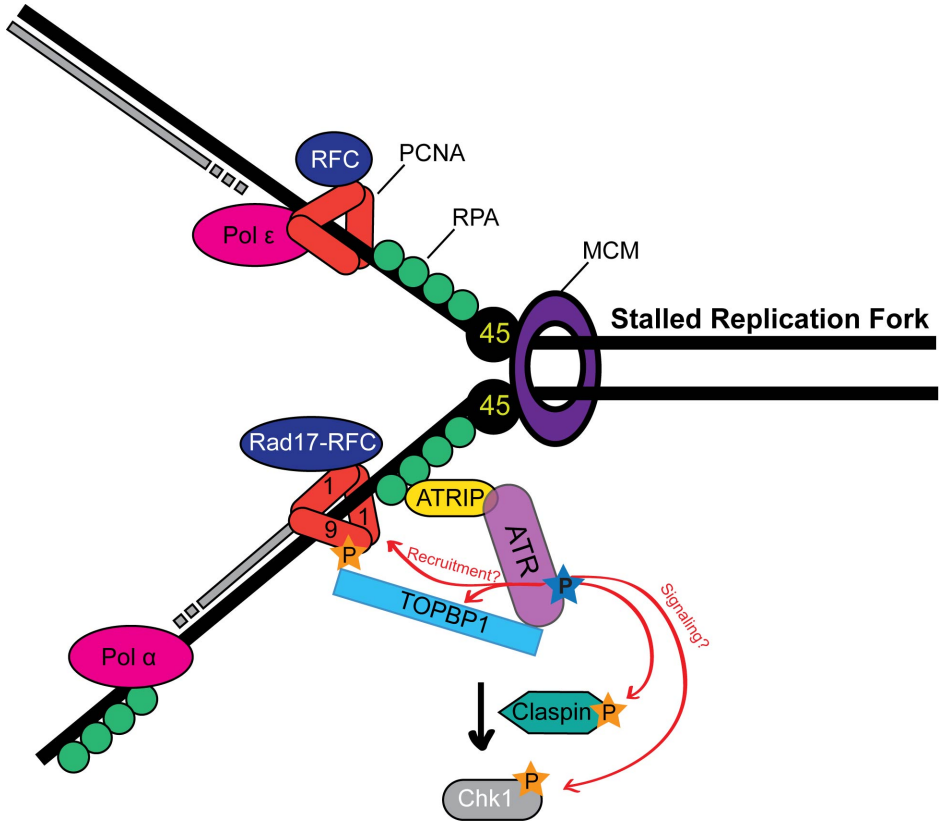


Figure 53. Model for ATR activation. Importance of residue Q2144 and neighbouring S2143 is illustrated in the model. This putative SQ site could be important for mediating the recruitment of ATR to ssDNA or perhaps is essential for the catalytic activity of the protein.

In a recent paper in which the structure of part of mTOR was resolved by X-ray crystallography, it was determined that the FAT domain encircles the kinase domain and these domains make contact through hydrogen bonds. These interactions are mediated by glutamine and arginine residues, most which are conserved among PIKK family members (Yang et al., 2013). In fact, residue Q2144 in ATR corresponds to residue Q1941 in mTOR (Figure 54), one of the amino acids located in the FAT domain that mediates binding to the mTOR kinase domain. For this reason, it is understandable that the ATR^{Q2144R} mutation could lead to a defective kinase.

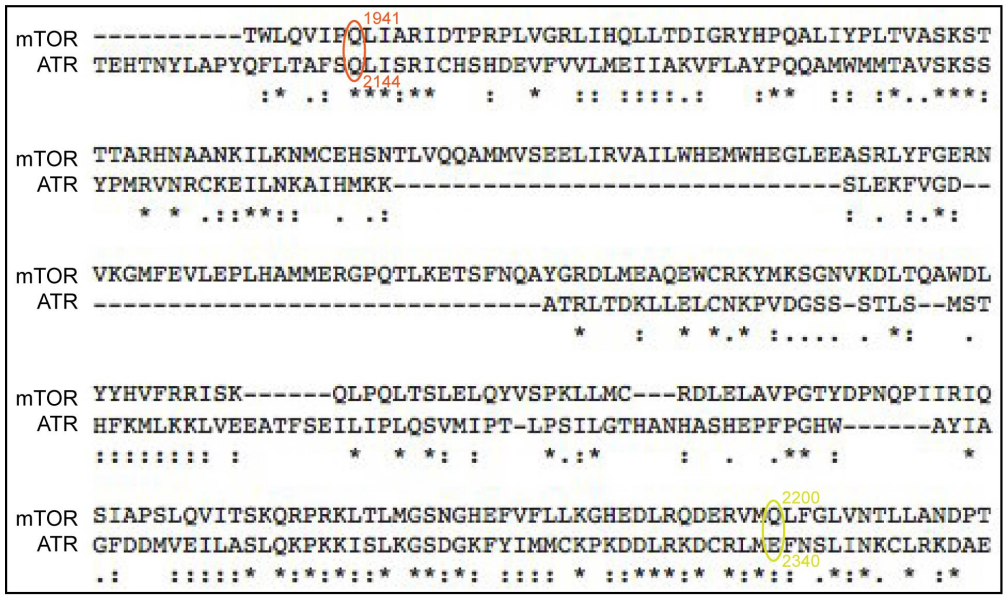


Figure 54. Alignment of part of mTOR and ATR C-terminus. mTOR residues that mediate the interaction between the FAT (in orange) and kinase (in yellow) domains are encircled. ATR equivalent residues are shown in the alignment. Identity between mTOR and ATR C-terminus is shown at the bottom of each line, as calculated using the Tcoffee alignment tool.

Recently, it has also been shown that the FATC domain is necessary for localization of Tel1 to DNA ends (Ogi et al., 2015). If this function were conserved in humans and both FAT and FATC domains work perhaps as a group, the ATR^{Q2144R} mutation could also disrupt this interaction.

To explore which of the possibilities conceived by our model was true, we aimed to generate a human cell line containing this mutation by using the CRISPR Cas9 technology. We decided to use a CRISPR strategy based on alternative homology-directed repair discovered by Davis and Maizels (Davis and Maizels, 2014). This alternative HR route is specific for nicks, the repair of which is mediated by ssDNA. Therefore, in our CRISPR experiments we used the nickase version of Cas9 in combination with two gRNAs and a ssDNA template. Unfortunately, while we were able to generate a human cell line containing the ATR^{Q2144R} mutation with a low frequency, this change was also accompanied by inactivating indel mutations.

Our results suggest that this methodology has very low efficiency (~1% targeting) and allows for extra undesired mutations to accumulate. In Davis and Maizels publication, they report a 3% targeting efficiency when using a similar CRISPR/Cas9 strategy, meaning that in general this might not be a very effective

approach. Therefore, other techniques might be used to improve the targeting efficiency and to reduce the off-target effects. Although the nickase Cas9 version is known to be less efficient than the nuclease, it has been preferentially used in recent years because it caused fewer off-targets effects (Ran et al., 2013b). However, new nuclease Cas9 enzymes have been engineered to be more specific, with off-target effects being reduced (Kleinstiver et al., 2016).

Apart from enzyme specificity, proximity of the Cas9 cut site and the nucleotide to mutate has been shown to be extremely important for efficient targeting (Paquet et al., 2016; Ran et al., 2013b). For instance, in the recent report by Paquet and colleagues, they have been unable to detect targeting using gRNAs located more than 45bp away to the site of interest. In this study, we attempted to use gRNAs with a small cut-to-mutation distance; however, we were unable to design close gRNAs with good scores, due to the A-T richness of *ATR* exon 38. As a result, the gRNAs used in this study were located 88 and 94bp away from the site to mutate; therefore, it is not surprising that the observed targeting efficiency was very low. Reduction of the cut-to-mutation distance is essential for future attempts in generating a human cell line containing the ATR^{Q2144R} mutation. New Cas9 enzymes that recognise different PAM sequences (for example NGA instead of NGG) are now available and its use would be ideal for this purpose (Kleinstiver et al., 2015).

New Cas9 enzymes that do not introduce double strand breaks (DSBs), but directly mediate the conversion of one target DNA base into another, have also been developed this year. Researchers have created a Cas9 and cytidine deaminase fusion enzyme, which retains the ability to target specific DNA bases (with a gRNA), without the induction of DSBs, but mediating the direct conversion of a C>T (or G>A) (Komor et al., 2016). This is a valuable technique for introducing point mutations; however, it can only be used for mutations that require a cytidine to uridine change. In our case we would require an enzyme that mediates the reverse reaction (A>G, or T>C in the reverse strand), since the ATR^{Q2144R} mutation converts CAA (glutamine, Q) into CGA (arginine, R).

Another possibility to improve the CRISPR design would be to use a dsDNA template cloned into a vector containing the selection marker. The fact that we obtained 199 clones in the CRISPR transfection but targeting was only detected in two (heterozygote) clones indicates that the design could be improved.

This suggests that we are selecting for cells that incorporated the plasmid containing the selection marker (puromycin) but did not incorporate the desired mutation provided in the ssDNA template. Cloning the template and selection marker into one single vector would eliminate one of the factors in the co-transfection. Besides, using a longer template might also increase the efficiency of homologous recombination.

A different explanation for the low targeting efficiency is that the mutation were toxic or deleterious for the cells. As a consequence, we would be selecting for intact or heterozygote clones. In this study, we have been able to generate a DT40 cell line expressing the Atr^{Q2163R} mutation; however, it is a conditional-null cell line in which WT Atr (AID-Atr) is present in the cells and could be compensating for the defects that appear when the WT protein is depleted.

If generation of human cell line containing the ATR^{Q2144R} mutation using CRISPR/Cas9 technology were achieved, the priority would be to study if the mutant is kinase active. It would also be exciting to test if this mutation affects ATR recruitment to stalled forks for example by immunofluorescence or immunoprecipitation studies. Besides, it would also be important to confirm if this mutation affects binding between the FAT and kinase domains. This could be achieved by fusing the FAT and kinase domain of ATR to separate tags in order to perform *in vitro* interaction studies. Analysis of neighbouring serine (ATR^{S2143}) would also be important to confirm if this putative SQ site is phosphorylated. Generation of phospho mimetic mutants (D and E) would be essential for this purpose, but also to study if the FAT-kinase interaction is altered.

In summary, we could hypothesize that residue Q2144 in ATR is essential for a conformational change necessary for the activation of the kinase activity. For this reason, the ATR^{Q2144R} mutation may result in a partially defective kinase. In a normal cell cycle, cells carrying this mutation may encounter problems during replication that cannot be resolved. Therefore, although cells might be able to survive longer that in the complete absence of Atr, they would ultimately accumulate irreparable damage leading to cell death. This damage would be generated due to a lack of checkpoints that would lead to cell cycle progression in the presence of, for example, un-replicated DNA resulting in chromosome breaks or aberrant segregation.

Further studies on the importance of this residue in ATR function will bring light into unknown ATR activation mechanisms. The resolution of ATR three-dimensional structure will be essential to fully understand this pathway.

CHAPTER 6

Conclusions and Future Perspectives

The study of ATR kinase function and regulation has been the focus of several investigations over the past decades. Significant knowledge has been gained regarding its role in the DNA damage response and, more recently, the basics of its essential role in regulating replication. However, there still remains much to learn, particularly concerning ATR configuration. One of the aims of this project has been to characterise the structure-function relationship of ATR using the genetic tractable DT40 cells as a model system.

ATR belongs to the PIKK family, a group of large proteins that share homologous structures (Baretić and Williams, 2014; Rivera-calzada et al., 2015). In particular, their C-terminus is comprised of a conserved kinase together with FAT and FATC domains; while their non-kinase portions, with the exception of the FACT domain, are mostly made up of helical motifs named HEAT repeats (Perry and Kleckner, 2003). The role of most of these repeats in PIKK function is unknown. For this reason, we set to study their function by creating a unique set of window deletion mutants spanning different HEAT-rich regions of the ATR protein. We hypothesized that HEAT repeats mediate specific interactions contributing to different functions of ATR and, therefore, we predicted that we would be able to assign specific functions to particular regions of ATR. However, the five window deletion mutants generated exhibited a null phenotype, equivalent to the complete loss of ATR. The phenotypes studied were proliferation and checkpoint response to replication fork stalling. These unexpected results were confirmed by equivalent studies performed in *S. cerevisiae* expressing similar window deletions of Mec1 (De Castro Abreu, 2012).

One explanation of these results is that all combinations of HEAT repeats that we deleted included motifs required for an essential function. This would be consistent with protein-protein interactions necessary for ATR essential function being distributed throughout the entire protein length. This possibility could be tested by the generation of smaller deletions, such as individual HEAT repeat deletions, as previously performed for the TRRAP protein (Knutson and Hahn, 2011). This study has shown that loss of some individual HEAT repeats did not result in a null phenotype. To further investigate particular ATR interactions, it would also be interesting to fuse different domains of ATR to GST for pull down experiments to study binding to known interactors with more detail.

On the other hand, it is also likely that loss of function in the deletion mutants is due to a problem in ATR folding. In other words, as HEAT repeats have been shown to act as solenoids or springs, which may be important to sense mechanical forces (Kumar et al., 2014), it is conceivable that loss of even a single HEAT repeat could disrupt ATR structure and thus affect its function. We tested this hypothesis by generating a replacement mutant in DT40 cells, in which the equivalent HEAT repeats of human ATR were introduced into chicken Atr. This strategy was intended to maintain ATR configuration intact. As the replacement mutant was fully functional in DT40, we understand that HEAT repeats are key for holding ATR structure, although it is still likely that in addition they are also required for important protein-protein interactions. In fact, when replacement mutants were generated in yeast by using the equivalent HEAT repeats from the orthologous and related PIKKs, ATR and Tel1 respectively, the Mec1-ATR and Mec1-Tel1 hybrids still displayed null phenotypes. As budding yeast is a more distant species from human than chicken, we cannot exclude the possibility that particular interactions vital for Mec1 function were lost using this approach.

The resolution of the three-dimensional architecture of ATR would be vital to get a better understanding of how HEAT repeats contribute to ATR function. In this project we have also analysed the ATR structure using electron microscopy and, although these results are very preliminary, our results suggest that ATR may have a similar configuration to other PIK kinases. The study of mTOR, SMG-1, ATM and DNA-PK structures has shown that PIKKs fold in a 'ring-like' shape, in which HEAT repeats act as an arm that wraps around DNA or binds to other substrate proteins (Adami et al., 2007; Arias-Palomo et al., 2011; Llorca et al., 2003; Sibanda et al., 2010).

Mutations in the *ATR* gene have been found in several cancers (Fang et al., 2004; Menoyo et al., 2001; Zighelboim et al., 2009) as well as in patients that suffer from Seckel syndrome (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003; Ogi et al., 2012), a rare disease characterized by premature aging. As part of this work, we have also researched the effect that three novel point mutations had on ATR function, two of which were associated with Seckel syndrome (ATR^{M1159I} and ATR^{K1665N}), while the third one was found in a family with cancer predisposition (ATR^{Q2144R}) (Tanaka et al., 2012). We focused our studies on these

mutations as we had hypothesized that they would have a direct effect on the ATR protein function and their study would help understand the cause behind those diseases.

The analysis of these disease-related mutations was also carried out using DT40 cells. Surprisingly, the ATR^{M1159I} and ATR^{K1665N} substitutions, which were found in unrelated Seckel patients, had no effect on ATR function in this model system. Both mutant cell lines generated showed a normal proliferation pattern and a functional checkpoint in response to replication fork stalling. As Seckel mutations had been previously reported to affect ATR splicing (Mokrani-Benhelli et al., 2013; O'Driscoll et al., 2003), we then investigated this possibility using human cell lines derived from patients. Accordingly, we were able to demonstrate that the ATR^{M1159I} and ATR^{K1665N} mutations were responsible for aberrant *ATR* splicing, resulting in exon 18 and 28 skipping, respectively. These defects, which were also confirmed using a minigene assay, are predicted to cause a frameshift resulting in the incorporation of a premature stop codon.

Nowadays, there are multiple reports in the literature that support the idea that many human diseases may be caused by point mutations that affect splicing (Li et al., 2016; Lopez-Bigas et al., 2005). The removal of introns, or splicing, is a complex process regulated by many components. Enhancer and silencer motifs play an important role in this process by recruiting positive or negative regulators of splicing. As these sequences can be located in either exons or introns, silent, missense or even non-sense point mutations located in exonic sequences do not necessarily affect protein function directly, but may disrupt or create one of these regulator motifs. For this reason, splicing mutations may be currently underestimated, making apparent the need for new diagnostic techniques that include transcript analyses when identifying the mutations causative of a disease.

Accordingly, it is possible that the ATR^{M1159I} and ATR^{K1665N} mutations have an impact on *ATR* splicing by altering enhancer or silencer motifs. To further investigate this possibility, potential splicing regulatory motifs that could have been affected by these modifications were analysed using bioinformatic online tools. These tools predicted that both missense mutations disrupt a 9G8 exonic splicing enhancer, while different silencer sites are created for hnRNPA1, Tra2Beta and YB1 splice factors. It would be exciting to perform binding studies

such as EMSA (Electrophoretic Mobility Shift Assay) to confirm if binding to the predicted splice factors is in fact disrupted/stimulated *in vitro*.

On the contrary, the *ATR*^{Q2144R} mutation, which was linked to cancer predisposition, was not found to alter splicing but to directly affect protein function. This mutation alters an evolutionarily conserved residue positioned in the ATR FAT domain. Besides, this glutamine is part of a putative SQ site and it is, therefore, conceivable that its mutation may disrupt a key motif for ATR activity. In the analyses carried out in our DT40 model system, it was clear that this mutation (*Atr*^{Q2163R} in DT40) caused a proliferation defect. Cells expressing Atr^{Q2163R} displayed an intermediate phenotype, showing a slower growth than WT, but higher survival than in the complete absence of Atr. When analysing the mutant response to replication fork stalling, we noticed a striking effect on ATR signaling. CHK1 phosphorylation was abolished in the presence of the *Atr*^{Q2163R} mutation. In addition, cells displayed a higher basal level of cell death and many chromosomal defects, including chromosome breaks and missing chromosome copies.

We speculate that abrogation of checkpoint signaling in Atr^{Q2163R} cells may lead to cell cycle progression even in the presence of unreplicated DNA or damage, giving rise to genomic instability. As a consequence, cells would accumulate karyotypic abnormalities, which would ultimately result in cell death. This could result in the 'slower' growth phenotype we detected in the proliferation assays. However, with regards to the development of cancer in patients, this hypothesis needs to be examined further. It is possible that defective checkpoint control triggered by the *ATR*^{Q2144R} mutation allowed for the accumulation of a specific set of additional mutations in patients, which would have given the cells a growth advantage over their neighbors. Therefore, the *ATR*^{Q2144R} mutation would be the driver mutation leading to such disease. Study of other mutations that may be associated with the disease would be important to confirm this idea. In the first report of the *ATR*^{Q2144R} mutation, other mutations were found, although they did not seem to correlate with the disease. However, they could potentially contribute to the unrestrained cell growth and cancer development in the oropharynx. If the 'driver mutation' theory were true, we would expect to see an accumulation of mutations in the DT40 cell line generated.

The precise event that leads to a defect in ATR signaling still remains to be clarified. One could consider that the ATR^{Q2144R} mutation may disrupt a motif necessary for the recruitment of ATR to ssDNA. It is possible that Q2144 may be a key residue for an essential ATR interaction with for example TOPBP1, ATRIP, RPA or other proteins that mediate ATR localization to forks. Alternatively, another possibility is that the mutation ablates the kinase activity of ATR, thus resulting in a kinase dead protein. It has been recently shown by crystallography that the FAT domain of mTOR encircles the kinase domain. These domains interact through conserved resides located in the FAT domain and, in fact, Q2144 is the equivalent ATR residue to Q1941 in mTOR, which mediates one of the key interactions with the kinase domain (Yang et al., 2013). Therefore, it is likely that the mutation observed in cancer patients results in a defective kinase. Kinase assays and interaction studies will be necessary to better understand the importance of this residue.

Additionally, phosphorylation of the serine (S2143) adjacent to the mutated glutamine (Q2144) could play a role in any of these scenarios. As previously mentioned, SQ/TQ sites are preferred phosphorylation sites for PIK kinases (Traven and Heierhorst, 2005). Therefore, S2143-Q2144 in ATR could presumably be phosphorylated in an ATR dependent or independent manner, contributing to ATR activation. Confirmation of this phosphorylation event by mass spectrometry or using a phospho-specific antibody would be crucial to understand the function of this residue. Mutation of S2143 to E or D (phospho-mimetic mutants) would be interesting to investigate its phosphorylation, but also to test if they affect the potential interaction between the FAT and kinase domains.

To study these theories, as DT40 cells have their limitations, it would be ideal obtain cell lines derived from patients. However, as this possibility was not available, we attempted to generate a human cell line containing the ATR^{Q2144R} mutation using the CRISPR/Cas9 technology. Our strategy, based in homology directed repair at nicks (Davis and Maizels, 2014), was found to be inefficient and prone to introducing other undesired mutations. Therefore, as our attempt was unsuccessful, the use of new strategies and improved Cas9 enzymes with more specificity and capability of recognising new PAM sequences would be required

to generate a cell line with this mutation (Kleinstiver et al., 2015, 2016; Paquet et al., 2016).

It is noteworthy that mutations in *ATR* can instigate very different and opposing outcomes, such as Seckel syndrome, a disorder associated with accelerated aging but also cancer progression, which main feature is unlimited growth and self-renewal. It is thought that the development of one disease or another is dependent on ATR dosage (Murga et al., 2009). As ATR^{Seckel} and ATR^{Q2144R} mutations have different effects on ATR expression, the degree of residual protein activity might be crucial to determine if the genomic lesions accumulated are overwhelming for the repair machinery (as for Seckel sufferers) or if, in contrast, they are able to stimulate uncontrolled cellular growth and malignant transformation (as in cancer patients). Further studies are necessary to investigate the fine line where this threshold lies.

CHAPTER 7

References

Abraham, R.T. (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev. *15*, 2177–2196.

Abraham, R.T. (2004). PI 3-kinase related kinases: "big" players in stress-induced signaling pathways. DNA Repair *3*, 883–887.

Adami, A., Garcia-Alvarez, B., Arias-Palomo, E., Barford, D., and Llorca, O. (2007). Structure of TOR and Its Complex with KOG1. Mol. Cell *27*, 509–516.

Agarwal, M.L., Agarwal, A., Taylor, W.R., and Stark, G.R. (1995). p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts. Proc. Natl. Acad. Sci. U. S. A. *92*, 8493–8497.

Aguilera, A. (2005). Cotranscriptional mRNP assembly: From the DNA to the nuclear pore. Curr. Opin. Cell Biol. *17*, 242–250.

Ahmed, M., and Rahman, N. (2006). ATM and breast cancer susceptibility. Oncogene *25*, 5906–5911.

Alberts, Bray, Hopkin, Johnson, Lewis, Raff, Roberts, and Walter (2009). Essential Cell Biology, 3rd Edition (Garland Science).

Alderton, G.K., Joenje, H., Varon, R., Borglum, A.D., Jeggo, P.A., and O'Driscoll, M. (2004). Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway. Hum. Mol. Genet. *13*, 3127–3138.

Andrade, M.A., and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. Nature *11*, 115–116.

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Muller, C.W., and Bork, P. (2001a). Comparison of ARM and HEAT protein repeats. J. Mol. Biol. *309*, 1–18.

Andrade, M.A., Perez-Iratxeta, C., and Ponting, C.P. (2001b). Protein Repeats: Structures, Functions, and Evolution. J. Struct. Biol. *134*, 117–131.

Antoniou, A., Pharoah, P.D.P., Narod, S., Risch, H.A., Eyfjord, J.E., Hopper, J.L., Loman, N., Olsson, H., Johannsson, O., Borg, A., et al. (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am. J. Hum. Genet. *72*, 1117–1130.

Arakawa, H., Lodygin, D., and Buerstedde, J.M. (2001). Mutant loxP vectors for selectable marker recycle and conditional knock-outs. BMC Biotechnol. *1*, 7.

Arias-Palomo, E., Yamashita, A., Fernández, I.S., Núñez-Ramírez, R., Bamba, Y., Izumi, N., Ohno, S., and Llorca, O. (2011). The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. Genes Dev. *25*, 153–164.

Baba, T.W., Giroir, B.P., and Humphries, E.H. (1985). Cell lines derived from avian lymphomas exhibit two distinct phenotypes. Virology *144*, 139–151.

Bahassi, E.M., Ovesen, J.L., Riesenberg, a L., Bernstein, W.Z., Hasty, P.E., and Stambrook, P.J. (2008). The checkpoint kinases Chk1 and Chk2 regulate the functional associations between hBRCA2 and Rad51 in response to DNA damage. Oncogene *27*, 3977–3985.

Bakkenist, C.J., and Kastan, M.B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature *421*, 499–506.

Bakr, A., Köcher, S., Volquardsen, J., Reimer, R., Borgamnn, K., Dikomey, E., Rothkamm, K., and Mansour, W.Y. (2015). Functional crosstalk between DNA damage response proteins 53BP1 and BRCA1 regulates double strand break repair choice. Radiother. Oncol. 1–6.

Ball, H.L., Myers, J.S., and Cortez, D. (2005). ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation. Mol. Biol. Cell *16*, 2372–2381.

Ball, H.L., Ehrhardt, M.R., Mordes, D.A., Glick, G.G., Chazin, W.J., and Cortez, D. (2007). Function of a conserved checkpoint recruitment domain in ATRIP proteins. Mol. Cell. Biol. *27*, 3367–3377.

Banaszynski, L.A., and Wandless, T.J. (2006). Conditional control of protein function. Chem. Biol. *13*, 11–21.

Baralle, D., and Baralle, M. (2005). Splicing in action: Assessing disease causing sequence changes. J. Med. Genet. *42*, 737–748.

Baralle, D., Lucassen, A., and Buratti, E. (2009). Missed threads: The impact of pre-mRNA splicing defects on clinical practice. EMBO Rep. *10*, 810–816.

Baralle, M., Baralle, D., De Conti, L., Mattocks, C., Whittaker, J., Knezevich, a, Ffrench-

Constant, C., and Baralle, F.E. (2003). Identification of a mutation that perturbs NF1 agene splicing using genomic DNA samples and a minigene assay. J. Med. Genet. *40*, 220–222.

Baretić, D., and Williams, R.L. (2014). PIKKs — the solenoid nest where partners and kinases meet. Curr. Opin. Struct. Biol. *29*, 134–142.

Barnes, J.W., Tischkau, S.A., Barnes, J.A., Mitchell, J.W., Burgoon, P.W., Hickok, J.R., and Gillette, M.U. (2003). Requirement of Mammalian Timeless for Circadian Rhythmicity. Science *302*, 439–442.

Bartek, J., and Lukas, J. (2001). Pathways governing G1/S transition and their response to DNA damage. FEBS Lett. *490*, 117–122.

Bartek, J., and Lukas, J. (2007). DNA damage checkpoints: from initiation to recovery or adaptation. Curr. Opin. Cell Biol. *19*, 238–245.

Bartek, J., Lukas, C., and Lukas, J. (2004). Checking on DNA damage in S phase. Nat. Rev. Mol. Cell Biol. *5*, 792–804.

Bartkova, J., Horejsí, Z., Koed, K., Krämer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., et al. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature *434*, 864–870.

Bensimon, A., Aebersold, R., and Shiloh, Y. (2011). Beyond ATM: The protein kinase landscape of the DNA damage response. FEBS Lett. *585*, 1625–1639.

Bermejo, R., Capra, T., Jossen, R., Colosio, A., Frattini, C., Carotenuto, W., Cocito, A., Doksani, Y., Klein, H., Gómez-González, B., et al. (2011). The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. Cell *146*, 233–246.

Bermudez, V.P., Lindsey-Boltz, L. a, Cesare, A.J., Maniwa, Y., Griffith, J.D., Hurwitz, J., and Sancar, A. (2003). Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro. Proc. Natl. Acad. Sci. U. S. A. *100*, 1633–1638.

Binz, S.K., Sheehan, A.M., and Wold, M.S. (2004). Replication Protein A phosphorylation and the cellular response to DNA damage. DNA Repair (Amst). *3*, 1015–1024.

Blankley, R.T., and Lydall, D. (2004). A domain of Rad9 specifically required for

activation of Chk1 in budding yeast. J. Cell Sci. *117*, 601–608.

Boskovic, J., Rivera-Calzada, A., Maman, J.D., Chac??n, P., Willison, K.R., Pearl, L.H., and Llorca, O. (2003). Visualization of DNA-induced conformational changes in the DNA repair kinase DNA-PKcs. EMBO J. *22*, 5875–5882.

Bosotti, R., Isacchi, A., and Sonnhammer, E.L. (2000). FAT: a novel domain in PIK-related kinases. Trends Biochem. Sci. *25*, 225–227.

Brown, E.J., and Baltimore, D. (2000). ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev. *14*, 397–402.

Brown, E.J., and Baltimore, D. (2003). Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. Genes Dev. *17*, 615–628.

Brown, A.D., Sager, B.W., Gorthi, A., Tonapi, S.S., Brown, E.J., and Bishop, A.J.R. (2014). ATR Suppresses Endogenous DNA Damage and Allows Completion of Homologous Recombination Repair. PLoS One *9*, e91222.

Brown, W.R.A., Hubbard, S.J., Tickle, C., and Wilson, S.A. (2003). The chicken as a model for large-scale analysis of vertebrate gene function. Nat. Rev. Genet. *4*, 87–98.

Buerstedde, J.-M., and Takeda, S. (2006). Reviews and Protocols in DT40 Research (Springer).

Buerstedde, J.M., Reynaud, C.A., Humphries, E.H., Olson, W., Ewert, D.L., and Weill, J.C. (1990). Light chain gene conversion continues at high rate in an ALV-induced cell line. EMBO J. *9*, 921–927.

Buisson, R., Boisvert, J.L., Benes, C.H., and Zou, L. (2015). Distinct but Concerted Roles of ATR, DNA-PK, and Chk1 in Countering Replication Stress during S Phase. Mol. Cell 1011–1024.

Bunting, S.F., Callén, E., Wong, N., Chen, H., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo, O., Cao, L., et al. (2010). 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell *141*, 243–254.

Bunz, F., Dutriaux, A., Lengauer, C., Waldman, T., Zhou, S., Brown, J.P., Sedivy, J.M., Kinzler, K.W., and Vogelstein, B. (1998). Requirement for p53 and p21 to Sustain G2 Arrest After DNA Damage. Science *282*, 1497–1502.

Burrows, A.E., and Elledge, S.J. (2008). How ATR turns on: TopBP1 goes on ATRIP with ATR. Genes Dev. *22*, 1416–1421.

Byun, T.S., Pacek, M., Yee, M.C., Walter, J.C., and Cimprich, K.A. (2005). Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes Dev. *19*, 1040–1052.

Caparelli, M.L., and O'Connell, M.J. (2013). Regulatory motifs in Chk1. Cell Cycle *12*, 916–922.

Cara, L., Baitemirova, M., Duong, F., Larios-Sanz, M., and Ribes-Zamora, A. (2014). SCDFinder, a Web-based tool for the identification of putative novel ATM and ATR targets. Bioinformatics *30*, 3394–3395.

Cara, L., Baitemirova, M., Follis, J., Larios-Sanz, M., and Ribes-Zamora, A. (2016). The ATM- and ATR-related SCD domain is over-represented in proteins involved in nervous system development. Sci. Rep. *6*, 19050.

Carballo, J.A., and Cha, R.S. (2007). Meiotic roles of Mec1, a budding yeast homolog of mammalian ATR/ATM. Chromosom. Res. *15*, 539–550.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat. Rev. Genet. *3*, 285–298.

De Castro Abreu, C.M. (2012). Mec1 and Rad9 functions in the DNA Damage Response. National University of Ireland, Galway.

Cha, R.S., and Kleckner, N. (2002). ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. Science *297*, 602–606.

Chen, X., Zhao, R., Glick, G.G., and Cortez, D. (2007). Function of the ATR N-terminal domain revealed by an ATM/ATR chimera. Exp. Cell Res. *313*, 1667–1674.

Chen, Y., Chen, P.L., Chen, C.F., Jiang, X., and Riley, D.J. (2008). Never-in-mitosis related kinase 1 functions in DNA damage response and checkpoint control. Cell Cycle *7*, 3194–3201.

Chen, Y., Chen, C.-F., Chiang, H.-C., Pena, M., Polci, R., Wei, R.L., Edwards, R.A., Hansel, D.E., Chen, P.-L., and Riley, D.J. (2011). Mutation of NIMA-related kinase 1 (NEK1) leads to chromosome instability. Mol. Cancer *10*, 5.

Chittenden, T. (2002). BH3 domains: Intracellular death-ligands critical for initiating apoptosis. Cancer Cell *2*, 165–166.

Chiu, C.Y., Cary, R.B., Chen, D.J., Peterson, S.R., and Stewart, P.L. (1998). Cryo-EM imaging of the catalytic subunit of the DNA-dependent protein kinase. J. Mol. Biol. *284*, 1075–1081.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232.

Chook, Y.M., and Blobel, G. (1999). Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. Nature *399*, 230–237.

Chu, W.K., and Hickson, I.D. (2009). RecQ helicases: Multifunctional genome caretakers. Nat. Rev. Cancer *9*, 644–654.

Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: Making it safe to play with knives. Mol. Cell *40*, 179–204.

Cimprich, K.A., and Cortez, D. (2008). ATR: An essential regulator of genome integrity. Nat. Rev. Mol. Cell Biol. *9*, 616–627.

Cingolani, G., Petosa, C., Weis, K., and Muller, C.W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. Nature *399*, 221–229.

Cliby, W.A., Roberts, C.J., Cimprich, K.A., Stringer, C.M., Lamb, J.R., Schreiber, S.L., and Friend, S.H. (1998). Overexpression of a kinase-inactive ATR protein causes sensitivity to DNA-damaging agents and defects in cell cycle checkpoints. EMBO J. *17*, 159–169.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Cooper, G.M. (2000). The cell: A molecular approach, 2nd Edition (Sunderland, Mass. Bethesda, MD: Sinauer Associates; NCBI).

Cortez, D., Guntuku, S., Qin, J., and Elledge, S.J. (2001). ATR and ATRIP: Partners in checkpoint signaling. Science *294*, 1713–1716.

Cotta-Ramusino, C., McDonald, E.R., Hurov, K., Sowa, M.E., Harper, J.W., and Elledge,

S.J. (2011). A DNA damage response screen identifies RHINO, a 9-1-1 and TopBP1 interacting protein required for ATR signaling. Science *332*, 1313–1317.

Couch, F.B., Bansbach, C.E., Driscoll, R., Luzwick, J.W., Glick, G.G., Bétous, R., Carroll, C.M., Jung, S.Y., Qin, J., Cimprich, K. a, et al. (2013). ATR phosphorylates SMARCAL1 to prevent replication fork collapse. Genes Dev. *27*, 1610–1623.

Daley, J.M., and Sung, P. (2014). 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. Mol. Cell. Biol. *34*, 1380–1388.

Daub, H., Olsen, J. V, Bairlein, M., Gnad, F., Oppermann, F.S., Korner, R., Greff, Z., Keri, G., Stemmann, O., and Mann, M. (2008). Kinase-Selective Enrichment Enables Quantitative Phosphoproteomics of the Kinome across the Cell Cycle. Mol. Cell *31*, 438–448.

Davies, S.L., North, P.S., and Hickson, I.D. (2007). Role for BLM in replication-fork restart and suppression of origin firing after replicative stress. Nat. Struct. Mol. Biol. *14*, 677–679.

Davis, L., and Maizels, N. (2014). Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. Proc. Natl. Acad. Sci. U. S. A. *111*, E924–E932.

Dephoure, N., Zhou, C., Villen, J., Beausoleil, S.A., Bakalarski, C.E., Elledge, S.J., and Gygi, S.P. (2008). A quantitative atlas of mitotic phosphorylation. Proc. Natl. Acad. Sci. U. S. A. *105*, 10762–10767.

Desany, B.A., Alcasabas, A.A., Bachant, J.B., and Elledge, S.J. (1998). Recovery from DNA replicational stress is the essential function of the S-phase checkpoint pathway. Genes Dev. *12*, 2956–2970.

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., and Beroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. *37*, e67–e67.

DeStephanis, D., McLeod, M., and Yan, S. (2015). REV1 is important for the ATR-Chk1 DNA damage response pathway in Xenopus egg extracts. Biochem. Biophys. Res. Commun. *460*, 609–615.

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D.H., Pepperkok, R.,

Ellenberg, J., Panier, S., Durocher, D., Bartek, J., et al. (2009). RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. Cell *136*, 435–446.

Domínguez-Kelly, R., Martín, Y., Koundrioukoff, S., Tanenbaum, M.E., Smits, V.A.J., Medema, R.H., Debatisse, M., and Freire, R. (2011). Wee1 controls genomic stability during replication by regulating the Mus81-Eme1 endonuclease. J. Cell Biol. *194*, 567–579.

Driscoll, M.O. (2013). Diseases Associated with Defective Responses to DNA Damage. Cold Spring Harb. Perspect. Biol. 1–26.

Duina, A.A., Miller, M.E., and Keeney, J.B. (2014). Budding yeast for budding geneticists: A primer on the Saccharomyces cerevisiae model system. Genetics *197*, 33–48.

Duursma, A.M., Driscoll, R., Elias, J.E., and Cimprich, K. a (2013). A role for the MRN complex in ATR activation via TOPBP1 recruitment. Mol. Cell *50*, 116–122.

Errico, A., and Costanzo, V. (2012). Mechanisms of replication fork protection: a safeguard for genome stability. Crit. Rev. Biochem. Mol. Biol. *47*, 222–235.

Eykelenboom, J.K., Harte, E.C., Canavan, L., Pastor-Peidro, A., Calvo-Asensio, I., Llorens-Agost, M., and Lowndes, N.F. (2013). ATR Activates the S-M Checkpoint during Unperturbed Growth to Ensure Sufficient Replication Prior to Mitotic Onset. Cell Rep. *5*, 1095–1107.

Falck, J., Mailand, N., Syljuåsen, R.G., Bartek, J., and Lukas, J. (2001). The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. Nature *410*, 842–847.

Falck, J., Petrini, J.H.J., Williams, B.R., Lukas, J., and Bartek, J. (2002). The DNA damage-dependent intra-S phase checkpoint is regulated by parallel pathways. Nat. Genet. *30*, 290–294.

Fang, Y., Tsao, C.-C., Goodman, B.K., Furumai, R., Tirado, C.A., Abraham, R.T., and Wang, X.-F. (2004). ATR functions as a gene dosage-dependent tumor suppressor on a mismatch repair-deficient background. EMBO J. *23*, 3164–3174.

Fant, X., Samejima, K., Carvalho, A., Ogawa, H., Xu, Z., Yue, Z., Earnshaw, W.C., and Ruchaud, S. (2010). Use of DT40 conditional-knockout cell lines to study

chromosomal passenger protein function. Biochem. Soc. Trans. *38*, 1655–1659.

Faustino, N.A., and Cooper, T.A. (2003). Pre-mRNA splicing and human disease. Genes Dev. *17*, 419–437.

Feijoo, C., Hall-Jackson, C., Wu, R., Jenkins, D., Leitch, J., Gilbert, D.M., and Smythe, C. (2001). Activation of mammalian Chk1 during DNA replication arrest: a role for Chk1 in the intra-S phase checkpoint monitoring replication origin firing. J. Cell Biol. *154*, 913–923.

Flynn, R.L., and Zou, L. (2011). ATR: a master conductor of cellular responses to DNA replication stress. Trends Biochem. Sci. *36*, 133–140.

Fokas, E., Prevo, R., Pollard, J.R., Reaper, P.M., Charlton, P.A., Cornelissen, B., Vallis, K.A., Hammond, E.M., Olcina, M.M., Gillies McKenna, W., et al. (2012). Targeting ATR in vivo using the novel inhibitor VE-822 results in selective sensitization of pancreatic tumors to radiation. Cell Death Dis. *3*, e441.

Foote, K.M., Blades, K., Cronin, A., Fillery, S., Guichard, S.S., Hassall, L., Hickson, I., Jacq, X., Jewsbury, P.J., Mcguire, T.M., et al. (2013). Discovery of 4‐{4-[(3R)‐3-Methylmorpholin-4-yl]-6-[1-(methylsulfonyl)cyclopropyl]pyrimidin-2-yl}‐1H‐indole (AZ20): A Potent and Selective Inhibitor of ATR Protein Kinase with Monotherapy in Vivo Antitumor Activity. J. Med. Chem. *56*, 2125–2138.

Fraenkel, P.G., Hospital, C., Israel, B., and Medical, D. (2005). Zebrafish as a Model for Human Diseases. Encycl. Life Sci. *43*, 107–120.

Frappart, P.-O., and McKinnon, P.J. (2006). Ataxia-Telangiectasia and Related Diseases. Neuromolecular Med. *8*, 175–190.

Friedel, A.M., Pike, B.L., and Gasser, S.M. (2009). ATR/Mec1: Coordinating fork stability and repair. Curr. Opin. Cell Biol. *21*, 237–244.

Gao, L., Wang, J., Wang, Y., and Andreadis, A. (2007). SR protein 9G8 modulates splicing of tau exon 10 via its proximal downstream intron, a clustering region for frontotemporal dementia mutations. Mol. Cell. Neurosci. *34*, 48–58.

Garcia, V., Furuya, K., and Carr, A.M. (2005). Identification and functional analysis of TopBP1 and its homologs. DNA Repair (Amst). *4*, 1227–1239.

Gatti, R., Boder, E., Vinters, H. V, Sparkes, R.S., Norman, A., and Lange, K. (1991).

Ataxia-telangiectasia: An interdisciplinary approach to pathogenesis. Medicine (Baltimore). *70*, 99–117.

Gilbert, C.S., Green, C.M., and Lowndes, N.F. (2001). Budding yeast Rad9 is an ATP-dependent Rad53 activating machine. Mol. Cell *8*, 129–136.

Goodarzi, A.A., Jonnalagadda, J.C., Douglas, P., Young, D., Ye, R., Moorhead, G.B.G., Lees-Miller, S.P., and Khanna, K.K. (2004). Autophosphorylation of ataxia-telangiectasia mutated is regulated by protein phosphatase 2A. EMBO J. *23*, 4451–4461.

Goodship, J., Gill, H., Carter, J., Jackson, A., Splitt, M., and Wright, M. (2000). Autozygosity mapping of a seckel syndrome locus to chromosome 3q22. 1-q24. Am. J. Hum. Genet. *67*, 498–503.

Gossen, M., and Bujardt, H. (1992). Tight Control of Gene Expression in Mammalian Cells by Tetracycline-Responsive Promoters. Proc. Natl. Acad. Sci. U. S. A. *89*, 5547–5551.

Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. RNA *6*, 1197–1211.

Griffith, E., Walker, S., Martin, C., and Vagnarelli, P. (2008). Mutations in Pericentrin cause Seckel syndrome with defective ATR-dependent DNA damage signalling. Nat. Genet. *40*, 232–236.

Grinthal, A., Adamovic, I., Weiner, B., Karplus, M., and Kleckner, N. (2010). PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. Proc. Natl. Acad. Sci. U. S. A. *107*, 2467–2472.

de Groote, F.H., Jansen, J.G., Masuda, Y., Shah, D.M., Kamiya, K., de Wind, N., and Siegal, G. (2011). The Rev1 translesion synthesis polymerase has multiple distinct DNA binding modes. DNA Repair (Amst). *10*, 915–925.

Groves, M.R., and Barford, D. (1999). Topological characteristics of helical repeat proteins. Curr Opin Struct Biol *9*, 383–389.

Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A., and Barford, D. (1999). The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. Cell *96*, 99–110.

Håkansson, P., Hofer, A., and Thelander, L. (2006). Regulation of mammalian

ribonucleotide reduction and dNTP pools after DNA damage and in resting cells. J. Biol. Chem. *281*, 7834–7841.

Harper, J.W., and Elledge, S.J. (2007). The DNA damage response: Ten years after. Mol. Cell *28*, 739–745.

Harrison, J.C., and Haber, J.E. (2006). Surviving the breakup: The DNA damage checkpoint. Annu. Rev. Genet. *40*, 209–235.

Harsha, H.C., and Pandey, A. (2010). Phosphoproteomics in cancer. Mol. Oncol. *4*, 482–495.

Hastings, M.L., and Krainer, A.R. (2001). Pre-mRNA splicing in the new millennium. Curr.Opin.Cell Biol. *13*, 302–309.

Hilton, B.A., Li, Z., Musich, P.R., Wang, H., Cartwright, B.M., Serrano, M., Zhou, X.Z., Lu, K.P., and Zou, Y. (2015). ATR Plays a Direct Antiapoptotic Role at Mitochondria , which Is Regulated by Prolyl Isomerase Pin1. Mol. Cell *60*, 1–12.

Ho, G.P.H.G.P., Margossian, S.S., Taniguchi, T.T., and D'Andrea, A.D.A.D. (2006). Phosphorylation of FANCD2 on two novel sites is required for mitomycin C resistance. Mol. Cell. Biol. *26*, 7005–7015.

Hochegger, H., Dejsuphong, D., Sonoda, E., Saberi, A., Rajendra, E., Kirk, J., Hunt, T., and Takeda, S. (2007). An essential role for Cdk1 in S phase control is revealed via chemical genetics in vertebrate cells. J. Cell Biol. *178*, 257–268.

Hu, J., Sun, L., Shen, F., Chen, Y., Hua, Y., Liu, Y., Zhang, M., Hu, Y., Wang, Q., Xu, W., et al. (2012). The intra-S phase checkpoint targets Dna2 to prevent stalled replication forks from reversing. Cell *149*, 1221–1232.

Huang, Y., and Steitz, J.A. (2001). Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. Mol. Cell *7*, 899–905.

Iliakis, G., Wang, Y., Guan, J., and Wang, H. (2003). DNA damage checkpoint control in cells exposed to ionizing radiation. Oncogene *22*, 5834–5847.

Inanç, B., Pütz, M., Lalor, P., Dockery, P., Kuriyama, R., Gergely, F., and Morrison, C.G. (2013). Abnormal centrosomal structure and duplication in Cep135-deficient vertebrate cells. Mol. Biol. Cell *24*, 2645–2654.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and

disease. Nature *461*, 1071–1078.

Jarrett, S.G., Horrell, E.M.W., Christian, P.A., Vanover, J.C., Boulanger, M.C., Zou, Y., and D'Orazio, J.A. (2014). PKA-Mediated Phosphorylation of ATR Promotes Recruitment of XPA to UV-Induced DNA Damage. Mol. Cell *54*, 999–1011.

Jean-Philippe, J., Paz, S., and Caputi, M. (2013). hnRNP A1: The Swiss army knife of gene expression. Int. J. Mol. Sci. *14*, 18999–19024.

Jeong, S.Y., Kumagai, A., Lee, J., and Dunphy, W.G. (2003). Phosphorylated Claspin Interacts with a Phosphate-binding Site in the Kinase Domain of Chk1 during ATR-mediated Activation. J. Biol. Chem. *278*, 46782–46788.

Jiang, H., Reinhardt, H.C., Bartkova, J., Tommiska, J., Blomqvist, C., Nevanlinna, H., Bartek, J., Yaffe, M.B., and Hemann, M.T. (2009). The combined status of ATM and p53 link tumor development with therapeutic response. Genes Dev. *23*, 1895–1909.

Jiang, K., Pereira, E., Maxfield, M., Russell, B., Goudelock, D.M., and Sanchez, Y. (2003). Regulation of Chk1 Includes Chromatin Association and 14-3-3 Binding following Phosphorylation on Ser-345. J. Biol. Chem. *278*, 25207–25217.

Jiang, X., Sun, Y., Chen, S., Roy, K., and Price, B.D. (2006). The FATC domains of PIKK proteins are functionally equivalent and participate in the Tip60-dependent activation of DNA-PKcs and ATM. J. Biol. Chem. *281*, 15741–15746.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science *337*, 816–822.

Johnson, D.G., and Walker, C.L. (1999). Cyclins and cell cycle checkpoints. Annu. Rev. Pharmacol. Toxicol. *39*, 295–312.

Kaindl, A.M., Passemard, S., Kumar, P., Kraemer, N., Issa, L., Zwirner, A., Gerard, B., Verloes, A., Mani, S., and Gressens, P. (2010). Many roads lead to primary autosomal recessive microcephaly. Prog. Neurobiol. *90*, 363–383.

Kalay, E., Yigit, G., Aslan, Y., Brown, K.E., Pohl, E., Bicknell, L.S., Kayserili, H., Li, Y., Tüysüz, B., Nürnberg, G., et al. (2011). CEP152 is a genome maintenance protein disrupted in Seckel syndrome. Nat. Genet. *43*, 23–26.

Kato, R., and Ogawa, H. (1994). An essential gene, ESR1, is required for mitotic cell

growth, DNA repair and meiotic recombination in Saccharomyces cerevisiae. Nucleic Acids Res. *22*, 3104–3112.

Kemp, M.G., Akan, Z., Yilmaz, S., Grillo, M., Smith-Roe, S.L., Kang, T.H., Cordeiro-Stone, M., Kaufmann, W.K., Abraham, R.T., Sancar, A., et al. (2010). Tipin-replication protein A interaction mediates Chk1 phosphorylation by ATR in response to genotoxic stress. J. Biol. Chem. *285*, 16562–16571.

Kepinski, S., and Leyser, O. (2005). The Arabidopsis F-box protein TIR1 is an auxin receptor. Nature *435*, 446–451.

Kim, J.M., Kakusho, N., Yamada, M., Kanoh, Y., Takemoto, N., and Masai, H. (2008). Cdc7 kinase mediates Claspin phosphorylation in DNA replication checkpoint. Oncogene *27*, 3475–3482.

Kim, S., Lim, D., Christine, E., Kastan, M.B., and Canman, C.E. (1999). Substrate Specificities and Identification of Putative Substrates of ATM Kinase Family Members. J. Biol. Chem. *274*, 37538–37543.

Kleier, S., Herrmann, M., Wittwer, B., Varon, R., Reis, a, and Horst, J. (2000). Clinical presentation and mutation identification in the NBS1 gene in a boy with Nijmegen breakage syndrome. Clin. Genet. *57*, 384–387.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V. V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature *523*, 481–485.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Keith Joung, J. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature *000*, 1–6.

Knispel, R., Kofler, C., and Boicu, M. (2012). Blotting protein complexes from native gels to electron microscopy grids. Nature *9*, 182–185.

Knutson, B.A., and Hahn, S. (2011). Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. Mol. Cell. Biol. *31*, 818–831.

Kobayashi, M., Hirano, A., Kumano, T., Xiang, S.-L., Mihara, K., Haseda, Y., Matsui, O., Shimizu, H., and Yamamoto, K. (2004). Critical role for chicken Rad17 and Rad9 in the cellular response to DNA damage and stalled DNA replication. Genes

Cells *9*, 291–303.

Kobe, B., Gleichmann, T., Horne, J., Jennings, I.G., Scotney, P.D., and Teh, T. (1999). Turn up the HEAT. Structure *7*, R91–R97.

Köhler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nat. Rev. Mol. Cell Biol. *8*, 761–773.

Kolas, N.K., Chapman, J.R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F.D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T.M., et al. (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. Science *318*, 1637–1640.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 1–17.

Kondo, S., Yamamoto, N., Murakami, T., Okumura, M., Mayeda, A., and Imaizumi, K. (2004). Tra2 beta, SF2/ASF and SRp30c modulate the function of an exonic splicing enhancer in exon 10 of tau pre-mRNA. Genes to Cells *9*, 121–130.

Kumagai, A., and Dunphy, W.G. (2000). Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in Xenopus egg extracts. Mol. Cell *6*, 839–849.

Kumagai, A., Lee, J., Yoo, H.Y., and Dunphy, W.G. (2006). TopBP1 activates the ATR-ATRIP complex. Cell *124*, 943–955.

Kumar, A., Mazzanti, M., Mistrik, M., Kosar, M., Beznoussenko, G.V., Mironov, A.A., Garrè, M., Parazzoli, D., Shivashankar, G.V., Scita, G., et al. (2014). ATR Mediates a Checkpoint at the Nuclear Envelope in Response to Mechanical Stress. Cell *158*, 633–646.

Kunz, J., Schneider, U., Howald, I., Schmidt, A., and Hall, M.N. (2000). HEAT repeats mediate plasma membrane localization of Tor2p in yeast. J. Biol. Chem. *275*, 37011–37020.

Labib, K., and De Piccoli, G. (2011). Surviving chromosome replication: the many roles of the S-phase checkpoint pathway. Philos. Trans. R. Soc. B Biol. Sci. *366*, 3554–3561.

Laurençon, A., Purdy, A., Sekelsky, J., Hawley, R.S., and Su, T.T. (2003). Phenotypic analysis of separation-of-function alleles of MEI-41, Drosophila ATM/ATR.

Genetics *164*, 589–601.

Lavin, M.F. (2008). Ataxia-telangiectasia: From a rare disorder to a paradigm for cell signalling and cancer. Nat. Rev. Mol. Cell Biol. *9*, 759–769.

Le, Y., and Sauer, B. (2000). Conditional gene knockout using cre recombinase. Methods Mol. Biol. *136*, 477–485.

Lee, J., Kumagai, A., and Dunphy, W.G. (2007). The Rad9-Hus1-Rad1 checkpoint clamp regulates interaction of TopBP1 with ATR. J. Biol. Chem. *282*, 28036–28044.

Lees-Miller, S.P., Sakaguchi, K., Ullrich, S.J., Appella, E., and Anderson, C.W. (1992). Human DNA-activated protein kinase phosphorylates serines 15 and 37 in the amino-terminal transactivation domain of human p53. Mol. Cell. Biol. *12*, 5041–5049.

Lempiäinen, H., and Halazonetis, T.D. (2009). Emerging common themes in regulation of PIKKs and PI3Ks. EMBO J. *28*, 3067–3073.

Lengauer, C., Kinzler, K.W., and Vogelstein, B. (1997). Genetic instability in colorectal cancers. Nature *386*, 623–627.

Leuther, K.K., Hammarsten, O., Kornberg, R.D., and Chu, G. (1999). Structure of DNA-dependent protein kinase: implications for its regulation by DNA. EMBO J. *18*, 1114–1123.

Lewandowska, M. a., Stuani, C., Parvizpur, A., Baralle, F.E., and Pagani, F. (2005). Functional studies on the ATM intronic splicing processing element. Nucleic Acids Res. *33*, 4007–4015.

Lewis, K., Bakkum-Gamez, J., Loewen, R., French, A., Thibodeau, S., and Cliby, W. (2007). Mutations in the Ataxia Telangiectasia and Rad3-Related–Checkpoint Kinase 1 DNA Damage Response Axis in Colon Cancers. Genes. Chromosomes Cancer *46*, 1061–1068.

Li, J., Hawkins, I.C., Harvey, C.D., Jennings, J.L., Link, A.J., and Patton, J.G. (2003). Regulation of alternative splicing by SRrp86 and its interacting proteins. Mol. Cell. Biol. *23*, 7437–7447.

Li, Y.I., Geijn, B. Van De, Raj, A., Knowles, D.A., Petti, A.A., Golan, D., Gilad, Y., and Pritchard, J.K. (2016). RNA splicing is a primary link between genetic variation and disease. Science *352*, 600–604.

Lindqvist, A., Rodríguez-Bravo, V., and Medema, R.H. (2009). The decision to enter mitosis: feedback and redundancy in the mitotic entry network. J. Cell Biol. *185*, 193–202.

Lindsey-Boltz, L.A., Kemp, M.G., Capp, C., and Sancar, A. (2015). RHINO forms a stoichiometric complex with the 9-1-1 checkpoint clamp and mediates ATR-Chk1 signaling. Cell Cycle *14*, 99–108.

Liou, Y.C., Zhou, X.Z., and Lu, K.P. (2011). Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. Trends Biochem. Sci. *36*, 501–514.

Liu, Q.H., Guntuku, S., Cui, X.S., Matsuoka, S., Cortez, D., Tamai, K., Luo, G.B., Carattini-Rivera, S., DeMayo, F., Bradley, A., et al. (2000). Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. Genes Dev. *14*, 1448–1459.

Liu, S., Bekker-Jensen, S., Mailand, N., Lukas, C., Bartek, J., and Lukas, J. (2006). Claspin operates downstream of TopBP1 to direct ATR signaling towards Chk1 activation. Mol. Cell. Biol. *26*, 6056–6064.

Liu, S., Shiotani, B., Lahiri, M., Maréchal, A., Tse, A., Leung, C.C.Y., Glover, J.N.M., Yang, X.H., and Zou, L. (2011). ATR autophosphorylation as a molecular switch for checkpoint activation. Mol. Cell *43*, 192–202.

Liu, S., Ho, C.K., Ouyang, J., and Zou, L. (2013). Nek1 kinase associates with ATR-ATRIP and primes ATR for efficient DNA damage signaling. Proc. Natl. Acad. Sci. U. S. A. *110*, 2175–2180.

Llorca, O., Rivera-Calzada, A., Grantham, J., and Willison, K.R. (2003). Electron microscopy and 3D reconstructions reveal that human ATM kinase uses an arm-like domain to clamp around double-stranded DNA. Oncogene *22*, 3867–3874.

Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochem. J *417*, 15–27.

Lopez-Bigas, N., Audit, B., Ouzounis, C., Parra, G., and Guigo, R. (2005). Are splicing mutations the most frequent cause of hereditary disease? FEBS Lett. *579*, 1900–1903.

Lopez-Contreras, A.J., Specks, J., Barlow, J.H., Ambrogio, C., Desler, C., Vikingsson, S., Rodrigo-Perez, S., Green, H., Rasmussen, L.J., Murga, M., et al. (2015). Increased

Rrm2 gene dosage reduces fragile site breakage and prolongs survival of ATR mutant mice. Genes Dev. *29*, 690–695.

Lovejoy, C. a., and Cortez, D. (2009). Common mechanisms of PIKK regulation. DNA Repair (Amst). *8*, 1004–1008.

Lowndes, N., and Murguia, J. (2000). Sensing and responding to DNA damage. Curr. Opin. Genet. Dev. 17–25.

Lu, Z., and Hunter, T. (2014). Prolyl isomerase Pin1 in cancer. Cell Res. *24*, 1033–1049.

Lupardus, P.J., Byun, T., Yee, M.-C., Hekmat-Nejad, M., and Cimprich, K. a (2002). A requirement for replication in activation of the ATR-dependent DNA damage checkpoint. Genes Dev. *16*, 2327–2332.

Luzwick, J.W., Nam, E. a, Zhao, R., and Cortez, D. (2014). Mutation of serine 1333 in the ATR HEAT repeats creates a hyperactive kinase. PLoS One *9*, e99397.

Madhani, H.D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. Annu. Rev. Genet. *28*, 1–26.

Mailand, N., Falck, J., Lukas, C., Syljuasen, R.G., Welcker, M., Bartek, J., and Lukas, J. (2000). Rapid Destruction of Human Cdc25A in Response to DNA Damage. Science *288*, 1425–1429.

Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007). RNF8 Ubiquitylates Histones at DNA Double-Strand Breaks and Promotes Assembly of Repair Proteins. Cell *131*, 887–900.

Majka, J., Niedziela-Majka, A., and Burgers, P.J. (2006). The Checkpoint Clamp Activates Mec1 Kinase during Initiation of the DNA Damage Checkpoint. Mol. Cell *24*, 891–901.

Mak, J.P.Y., Man, W.Y., Chow, J.P.H., Ma, H.T., and Poon, R.Y.C. (2015). Pharmacological inactivation of CHK1 and WEE1 induces mitotic catastrophe in nasopharyngeal carcinoma cells. *6*, 1–11.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

Mamely, I., van Vugt, M.A., Smits, V.A., Semple, J.I., Lemmens, B., Perrakis, A.,

Medema, R.H., and Freire, R. (2006). Polo-like Kinase-1 Controls Proteasome-Dependent Degradation of Claspin during Checkpoint Recovery. Curr. Biol. *16*, 1950–1955.

Mantiero, D., Clerici, M., Lucchini, G., and Longhese, M.P. (2007). Dual role for Saccharomyces cerevisiae Tel1 in the checkpoint response to double-strand breaks. EMBO Rep. *8*, 380–387.

Marjanović, M., Sánchez-Huertas, C., Terré, B., Gómez, R., Scheel, J.F., Pacheco, S., Knobel, P. a., Martínez-Marchal, A., Aivio, S., Palenzuela, L., et al. (2015). CEP63 deficiency promotes p53-dependent microcephaly and reveals a role for the centrosome in meiotic recombination. Nat. Commun. *6*, 7676.

Martin, R.G., and Stein, S. (1976). Resting state in normal and simian virus 40 transformed Chinese hamster lung cells. Proc Natl Acad Sci U S A *73*, 1655–1659.

Masramon, L., Ribas, M., Cifuentes, P., Arribas, R., García, F., Egozcue, J., Peinado, M.A., and Miró, R. (2000). Cytogenetic characterization of two colon cell lines by using conventional G-banding, comparative genomic hybridization, and whole chromosome painting. Cancer Genet. Cytogenet. *121*, 17–21.

Matsumoto, K., and Wolffe, A.P. (1998). Gene regulation by Y-box proteins: Coupling control of transcription and translation. Trends Cell Biol. *8*, 318–323.

Matsuoka, S., Rotman, G., Ogawa, A., Shiloh, Y., Tamai, K., and Elledge, S.J. (2000). Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. Proc. Natl. Acad. Sci. U. S. A. *97*, 10389–10394.

Matsuoka, S., Ballif, B. a, Smogorzewska, A., McDonald, E.R., Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., et al. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science *316*, 1160–1166.

Mazoyer, S., Puget, N., Perrin-Vidoz, L., Lynch, H., Serova-Sinilnikova, O., and Lenoir, G.M. (1998). A BRCA1 nonsense mutation causes exon skipping. Am. J. Hum. Genet.

McGlynn, P., Savery, N.J., and Dillingham, M.S. (2012). The conflict between DNA replication and transcription. Mol. Microbiol. *85*, 12–20.

Meijers-Heijboer, H., Wijnen, J., Vasen, H., Wasielewski, M., Wagner, A., Hollestelle,

A., Elstrodt, F., van den Bos, R., de Snoo, A., Fat, G.T. a, et al. (2003). The CHEK2 1100delC mutation identifies families with a hereditary breast and colorectal cancer phenotype. Am. J. Hum. Genet. *72*, 1308–1314.

Menoyo, A., Alazzouzi, H., Espín, E., Armengol, M., Yamamoto, H., and Schwartz S., J. (2001). Somatic mutations in the DNA damage-response genes ATR and CHK1 in sporadic stomach tumors with microsatellite instability. Cancer Res. *61*, 7727–7730.

Mohni, K.N., Kavanaugh, G.M., and Cortez, D. (2014). ATR Pathway Inhibition Is Synthetically Lethal in Cancer Cells with ERCC1 Deficiency. Cancer Res. *74*, 2835–2845.

Mokrani-Benhelli, H., Gaillard, L., Biasutto, P., Le Guen, T., Touzot, F., Vasquez, N., Komatsu, J., Conseiller, E., Pïcard, C., Gluckman, E., et al. (2013). Primary microcephaly, impaired DNA replication, and genomic instability caused by compound heterozygous ATR mutations. Hum. Mutat. *34*, 374–384.

Mordes, D.A., Glick, G.G., Zhao, R.X., and Cortez, D. (2008). TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. Genes Dev. *22*, 1478–1489.

Morgan, D.O. (1997). Cyclin-dependent kinases : Engines , clocks , and microprocessors. Annu. Rev. Cell Dev. Biol. *13*, 261–291.

Morita, T., Yamashita, A., Kashima, I., Ogata, K., Ishiura, S., and Ohno, S. (2007). Distant N- and C-terminal domains are required for intrinsic kinase activity of SMG-1, a critical component of nonsense-mediated mRNA decay. J. Biol. Chem. *282*, 7799–7808.

Murga, M., Bunting, S., Montaña, M.F., Soria, R., Mulero, F., Cañamero, M., Lee, Y., Mckinnon, P.J., and Nussenzweig, A. (2009). A mouse model of the ATR-Seckel Syndrome reveals that replicative stress during embryogenesis limits mammalian lifespan. Nat. Genet. *41*, 891–898.

Murga, M., Campaner, S., Lopez-Contreras, A.J., Toledo, L.I., Soria, R., Montaña, M.F., D'Artista, L., Schleker, T., Guerra, C., Garcia, E., et al. (2011). Exploiting oncogene-induced replicative stress for the selective killing of Myc-driven tumors. Nat. Struct. Mol. Biol. *18*, 1331–1335.

Nam, E.A., Zhao, R., and Cortez, D. (2011a). Analysis of mutations that dissociate G(2) and essential S phase functions of human ataxia telangiectasia-mutated and Rad3-

related (ATR) protein kinase. J. Biol. Chem. *286*, 37320–37327.

Nam, E.A., Zhao, R.X., Glick, G.G., Bansbach, C.E., Friedman, D.B., and Cortez, D. (2011b). Thr-1989 Phosphorylation Is a Marker of Active Ataxia Telangiectasia-mutated and Rad3-related (ATR) Kinase. J. Biol. Chem. *286*, 28707–28714.

Nasmyth, K., Gunther, A., Lydall, D., and Seddon, A. (1990). The identification of a second cell cycle control on the HO promoter in yeast: Cell cycle regulation of SWI5 nuclear entry. Cell *62*, 631–647.

Navadgi-Patil, V.M., and Burgers, P.M. (2009). A tale of two tails: activation of DNA damage checkpoint kinase Mec1/ATR by the 9-1-1 clamp and by Dpb11/TopBP1. DNA Repair (Amst). *8*, 996–1003.

Navadgi-Patil, V.M., and Burgers, P.M. (2011). Cell-cycle-specific activators of the Mec1/ATR checkpoint kinase. Biochem. Soc. Trans. *39*, 600–605.

Negrini, S., Gorgoulis, V.G., and Halazonetis, T.D. (2010). Genomic instability--an evolving hallmark of cancer. Nat. Rev. Mol. Cell Biol. *11*, 220–228.

Niedernhofer, L.J., Garinis, G. a, Raams, A., Lalai, A.S., Robinson, A.R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., et al. (2006). A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature *444*, 1038–1043.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods *6*, 917–922.

O'Driscoll, M. (2009). Mouse models for ATR deficiency. DNA Repair *8*, 1333–1337.

O'Driscoll, M., and Jeggo, P.A. (2003). Clinical impact of ATR checkpoint signalling failure in humans. Cell Cycle *2*, 194–195.

O'Driscoll, M., Ruiz-Perez, V.L., Woods, C.G., Jeggo, P. a, and Goodship, J. a (2003). A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat. Genet. *33*, 497–501.

Oestergaard, V.H., Pentzold, C., Pedersen, R.T., Iosif, S., Alpi, A., Bekker-Jensen, S., Mailand, N., and Lisby, M. (2012). RNF8 and RNF168 but not HERC2 are required for DNA damage-induced ubiquitylation in chicken DT40 cells. DNA Repair (Amst). *11*, 892–905.

Ogi, H., Goto, G.H., Ghosh, A., Zencir, S., Henry, E., and Sugimoto, K. (2015). Requirement of the FATC domain of protein kinase Tel1 for localization to DNA ends and target protein recognition. Mol. Biol. Cell *26*, 3480–3488.

Ogi, T., Walker, S., Stiff, T., Hobson, E., Limsirichaikul, S., Carpenter, G., Prescott, K., Suri, M., Byrd, P.J., Matsuse, M., et al. (2012). Identification of the first ATRIP-deficient patient and novel mutations in ATR define a clinical spectrum for ATR-ATRIP Seckel Syndrome. PLoS Genet. *8*, e1002945.

Orengo, J., Ward, A., and Cooper, T. (2011). Alternative splicing misregulation secondary to skeletal muscle regeneration. Ann. Neurol. *69*, 681–690.

Osmani, S.A., Pu, R.T., and Morris, N.R. (1988). Mitotic induction and maintenance by overexpression of a G2-specific gene that encodes a potential protein kinase. Cell *53*, 237–244.

Paciotti, V., Clerici, M., Scotti, M., Lucchini, G., and Longhese, M.P. (2001). Characterization of mec1 kinase-deficient mutants and of new hypomorphic mec1 alleles impairing subsets of the DNA damage response pathway. Mol. Cell. Biol. *21*, 3913–3925.

Padgett, R.A. (2012). New connections between splicing and human disease. Trends Genet. *28*, 147–154.

Pagani, F., Buratti, E., Stuani, C., Romano, M., Zuccato, E., Niksic, M., Giglio, L., Faraguna, D., and Baralle, F.E. (2000). Splicing Factors Induce Cystic Fibrosis Transmembrane Regulator Exon 9 Skipping through a Nonevolutionary Conserved Intronic Element. J. Biol. Chem. *275*, 21041–21047.

Pagani, F., Buratti, E., Stuani, C., Bendix, R., Dörk, T., and Baralle, F.E. (2002). A new type of mutation causes a splicing defect in ATM. Nat. Genet. *30*, 426–429.

Panier, S., and Boulton, S.J. (2014). Double-strand break repair: 53BP1 comes into focus. Nat. Rev. Mol. Cell Biol. *15*, 7–18.

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., and Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature *000*, 1–18.

Patro, B.S., Frohlich, R., Bohr, V. a., and Stevnsner, T. (2011). WRN helicase regulates the ATR-CHK1-induced S-phase checkpoint pathway in response to

topoisomerase-I-DNA covalent complexes. J. Cell Sci. *124*, 3967–3979.

Paull, T.T., Rogakou, E.P., Yamazaki, V., Kirchgessner, C.U., Gellert, M., and Bonner, W.M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. Curr. Biol. *10*, 886–895.

Paz, I., Akerman, M., Dror, I., Kosti, I., and Mandel-Gutfreund, Y. (2010). SFmap: A web server for motif analysis and prediction of splicing factor binding sites. Nucleic Acids Res. *38*, 281–285.

Pellicioli, A., Lucca, C., Liberi, G., Marini, F., Lopes, M., Plevani, P., Romano, A., Di Fiore, P.P., and Foiani, M. (1999). Activation of Rad53 kinase in response to DNA damage and its effect in modulating phosphorylation of the lagging strand DNA polymerase. EMBO J. *18*, 6561–6572.

Peng, C., Graves, P.R., Thoma, R.S., Wu, Z., Shaw, A.S., and Piwnica-worms, H. (1997). Mitotic and G2 Checkpoint Control: Regulation of 14-3-3 Protein Binding by Phosphorylation of Cdc25C on Serine-216. Science *277*, 1501–1505.

Perry, J., and Kleckner, N. (2003). The ATRs, ATMs, and TORs are giant HEAT repeat proteins. Cell *112*, 151–155.

Pessina, F., and Lowndes, N.F. (2014). The RSF1 Histone-Remodelling Factor Facilitates DNA Double-Strand Break Repair by Recruiting Centromeric and Fanconi Anaemia Proteins. PLoS Biol. *12*.

Petermann, E., Helleday, T., and Caldecott, K.W. (2008). Claspin promotes normal replication fork rates in human cells. Mol. Biol. Cell *19*, 2373–2378.

Pichierri, P., and Rosselli, F. (2004). The DNA crosslink-induced S-phase checkpoint depends on ATR-CHK1 and ATR-NBS1-FANCD2 pathways. EMBO J. *23*, 1178–1187.

Pichierri, P., Nicolai, S., Cignolo, L., Bignami, M., and Franchitto, A. (2011). The RAD9–RAD1–HUS1 (9.1.1) complex interacts with WRN and is crucial to regulate its response to replication fork stalling. Oncogene *31*, 2809–2823.

Pirzio, L.M., Pichierri, P., Bignami, M., and Franchitto, A. (2008). Werner syndrome helicase activity is essential in maintaining fragile site stability. J. Cell Biol. *180*, 305–314.

Poehlmann, A., and Roessner, A. (2010). Importance of DNA damage checkpoints in the

pathogenesis of human cancers. Pathol. Res. Pract. *206*, 591–601.

Poon, R.Y.C. (2014). DNA damage checkpoints in nasopharyngeal carcinoma. Oral Oncol. *50*, 339–344.

Prakash, R., Zhang, Y., Feng, W., and Jasin, M. (2015). Homologous recombination and human health: The roles of BRCA1, BRCA2, and associated proteins. Cold Spring Harb. Perspect. Biol. *7*, 1–37.

Prevo, R., Fokas, E., Reaper, P.M., Charlton, P.A., Pollard, J.R., McKenna, W.G., Muschel, R.J., and Brunner, T.B. (2012). The novel ATR inhibitor VE-821 increases sensitivity of pancreatic cancer cells to radiation and chemotherapy. Cancer Biol. Ther. *13*, 1072–1081.

Prosser, S.L., and Morrison, C.G. (2015). Centrin2 regulates CP110 removal in primary cilium formation. J. Cell Biol. *208*, 693–701.

Puddu, F., Piergiovanni, G., Plevani, P., and Muzi-Falconi, M. (2011). Sensing of replication stress and Mec1 activation act through two independent pathways involving the 9-1-1 complex and DNA polymerase ε. PLoS Genet. *7*, e1002022.

Putnam, C.D., Jaehnig, E.J., and Kolodner, R.D. (2010). Perspectives on the DNA damage and replication checkpoint responses in Saccharomyces cerevisiae. DNA Repair *8*, 974–982.

Ragland, R.L., Arlt, M.F., Hughes, E.D., Saunders, T.L., and Glover, T.W. (2009). Mice hypomorphic for Atr have increased DNA damage and abnormal checkpoint response. Mamm. Genome *20*, 375–385.

Rainey, M.D., Harhen, B., Wang, G.-N., Murphy, P. V, and Santocanale, C. (2013). Cdc7-dependent and -independent phosphorylation of Claspin in the induction of the DNA replication checkpoint. Cell Cycle *12*, 1560–1568.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D. a, and Zhang, F. (2013a). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D. a, Inoue, A., Matoba, S., Zhang, Y., et al. (2013b). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389.

Raponi, M., Buratti, E., Dassie, E., Upadhyaya, M., and Baralle, D. (2009). Low U1

snRNP dependence at the NF1 exon 29 donor splice site. FEBS J. *276*, 2060–2073.

Raponi, M., Smith, L.D., Silipo, M., Stuani, C., Buratti, E., and Baralle, D. (2014). BRCA1 exon 11 a model of long exon splicing regulation. RNA Biol. *11:4*, 351–359.

Reaper, P.M., Griffiths, M.R., Long, J.M., Charrier, J.-D., MacCormick, S., Charlton, P.A., Golec, J.M.C., and Pollard, J.R. (2011). Selective killing of ATM- or p53-deficient cancer cells through inhibition of ATR. Nat. Chem. Biol. *7*, 428–430.

Riches, L.C., Lynch, A.M., and Gooderham, N.J. (2008). Early events in the mammalian response to DNA double-strand breaks. Mutagenesis *23*, 331–339.

Rivera-Calzada, A., Maman, J.P., Spagnolo, L., Pearl, L.H., and Llorca, O. (2005). Three-dimensional structure and regulation of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). Structure *13*, 243–255.

Rivera-calzada, A., López-Perrote, A., Melero, R., and Boskovic, J. (2015). Structure and Assembly of the PI3K-like Protein Kinases (PIKKs) Revealed by Electron Microscopy. AIMS Biophys. *2*, 36–57.

Rizzardi, L.F., and Cook, J.G. (2012). Flipping the switch from G1 to S phase with E3 ubiquitin ligases. Genes Cancer *3*, 634–648.

Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139. J. Biol. Chem. *273*, 5858–5868.

Rogakou, E.P., Boon, C., Redon, C., and Bonner, W.M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. J. Cell Biol. *146*, 905–916.

Rouleau, E., Lefol, C., Moncoutier, V., Castera, L., Houdayer, C., Caputo, S., Bièche, I., Buisson, M., Mazoyer, S., Stoppa-Lyonnet, D., et al. (2010). A missense variant within BRCA1 exon 23 causing exon skipping. Cancer Genet. Cytogenet. *202*, 144–146.

Rounds, M.A., and Larsen, P.B. (2008). Aluminum-dependent root-growth inhibition in Arabidopsis results from AtATR-regulated cell-cycle arrest. Curr. Biol. *18*, 1495–1500.

Ruediger, R., Hentz, M., Fait, J., Mumby, M., and Walter, G. (1994). Molecular model of

the A subunit of protein phosphatase 2A: Interaction with other subunits and tumor antigens. J. Virol. *68*, 123–129.

Ruiz, S., Mayor-Ruiz, C., Lafarga, V., Murga, M., Vega-Sendino, M., Ortega, S., and Fernández-Capetillo, O. (2016). A Genome-wide CRISPR Screen Identifies CDC25A as a Determinant of Sensitivity to ATR Inhibitors. Mol. Cell *62*, 1–7.

Rupnik, A., Grenon, M., and Lowndes, N. (2008). The MRN complex. Curr. Biol. *18*, 455–457.

Ruzankina, Y., Pinzon-Guzman, C., Asare, A., Ong, T., Pontano, L., Cotsarelis, G., Zediak, V.P., Velez, M., Bhandoola, A., and Brown, E.J. (2007). Deletion of the Developmentally Essential Gene ATR in Adult Mice Leads to Age-Related Phenotypes and Stem Cell Loss. Cell Stem Cell *1*, 113–126.

Samejima, I., Spanos, C., De Lima Alves, F., Hori, T., Perpelescu, M., Zou, J., Rappsilber, J., Fukagawa, T., and Earnshaw, W.C. (2015). Whole-proteome genetic analysis of dependencies in assembly of a vertebrate kinetochore. J. Cell Biol. *211*, 1141–1156.

Sanchez, Y., Bachant, J., Wang, H., Hu, F., Liu, D., Tetzlaff, M., and Elledge, S.J. (1999). Control of the DNA damage checkpoint by chk1 and rad53 protein kinases through distinct mechanisms. Science *286*, 1166–1171.

Sanjiv, K., Hagenkort, A., Calderón-Montaño, J.M., Koolmeister, T., Reaper, P.M., Mortusewicz, O., Jacques, S.A., Kuiper, R.V., Schultz, N., Scobie, M., et al. (2015). Cancer-Specific Synthetic Lethality between ATR and CHK1 Kinase Activities. Cell Rep. 298–309.

Santner, A., Calderon-Villalobos, L.I., and Estelle, M. (2009). Plant hormones are versatile chemical regulators of plant growth. Nat. Chem. Biol. *5*, 301–307.

Santocanale, C., and Diffley, J.F. (1998). A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. Nature *395*, 615–618.

Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S.P. (2007). Human CtIP promotes DNA end resection. Nature *450*, 509–514.

Satir, P., and Christensen, S.T. (2007). Overview of structure and function of mammalian cilia. Annu. Rev. Physiol. *69*, 377–400.

Satyanarayana, A., and Kaldis, P. (2009). A dual role of Cdk2 in DNA damage response. Cell Div. *4*, 9.

Schlegel, B.P., Jodelka, F.M., and Nunez, R. (2006). BRCA1 promotes induction of ssDNA by ionizing radiation. Cancer Res. *66*, 5181–5189.

Schoppy, D.W., Ragland, R.L., Gilad, O., Shastri, N., Peters, A. a, Murga, M., Fernandez-capetillo, O., Diehl, J.A., and Brown, E.J. (2012). Oncogenic stress sensitizes murine cancers to hypomorphic suppression of ATR. *122*, 241–252.

Shen, H., Kan, J.L.C., and Green, M.R. (2004). Arginine-Serine-Rich Domains Bound at Splicing Enhancers Contact the Branchpoint to Promote Prespliceosome Assembly. Mol. Cell *13*, 367–376.

Sherr, C.J., and Roberts, J.M. (2004). Living with or without cyclins and cyclin-dependent kinases Living with or without cyclins and cyclin-dependent kinases. Genes Dev. *18*, 2699–2711.

Shiloh, Y. (2003). ATM and related protein kinases: safeguarding genome integrity. Nat. Rev. Cancer *3*, 155–168.

Shirahige, K., Hori, Y., Shiraishi, K., Yamashita, M., Takahashi, K., Obuse, C., Tsurimoto, T., and Yoshikawa, H. (1998). Regulation of DNA-replication origins during cell-cycle progression. Nature *395*, 618–621.

Sibanda, B.L., Chirgadze, D.Y., and Blundell, T.L. (2010). Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. Nature *463*, 118–121.

Smith, K.D., Fu, M. a, and Brown, E.J. (2009). Tim-Tipin dysfunction creates an indispensible reliance on the ATR-Chk1 pathway for continued DNA synthesis. J. Cell Biol. *187*, 15–23.

Smits, V.A.J., Reaper, P.M., and Jackson, S.P. (2006). Rapid PIKK-dependent release of Chk1 from chromatin promotes the DNA-damage checkpoint response. Curr. Biol. *16*, 150–159.

Sørensen, C.S., Hansen, L.T., Dziegielewski, J., Syljuåsen, R.G., Lundin, C., Bartek, J., and Helleday, T. (2005). The cell-cycle checkpoint kinase Chk1 is required for mammalian homologous recombination repair. Nat. Cell Biol. *7*, 195–201.

Spagnolo, L., Rivera-Calzada, A., Pearl, L.H., and Llorca, O. (2006). Three-Dimensional

Structure of the Human DNA-PKcs/Ku70/Ku80 Complex Assembled on DNA and Its Implications for DNA DSB Repair. Mol. Cell *22*, 511–519.

Stewart, G.S., Wang, B., Bignell, C.R., Taylor, A.M.R., and Elledge, S.J. (2003). MDC1 is a mediator of the mammalian DNA damage checkpoint. Nature *421*, 961–966.

Stickeler, E., Fraser, S.D., Hornig, A., Chen, A.L., Berget, S.M., and Cooper, T.A. (2001). The RNA binding protein YB-1 binds A/C-rich exon enhancers and stimulates splicing of the CD44 alternative exon v4. EMBO J. *20*, 3821–3830.

Stiff, T., Casar Tena, T., O'Driscoll, M., Jeggo, P.A., and Philipp, M. (2016). ATR promotes cilia signalling: links to developmental impacts. Hum. Mol. Genet. *25*, 1574–1587.

Storici, F., and Resnick, M.A. (2006). The Delitto Perfetto Approach to In Vivo Site-Directed Mutagenesis and Chromosome Rearrangements with Synthetic Oligonucleotides in Yeast. Methods Enzymol. *409*, 329–345.

Sun, Y., Jiang, X., Chen, S., Fernandes, N., and Price, B.D. (2005). A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. Proc. Natl. Acad. Sci. U. S. A. *102*, 13182–13187.

Supek, F., Minana, B., Valcarcel, J., Gabaldon, T., and Lehner, B. (2014). Synonymous mutations frequently act as driver mutations in human cancers. Cell *156*, 1324–1335.

Suzuki, K., Kodama, S., and Watanabe, M. (1999). Recruitment of ATM protein to double strand DNA irradiated with ionizing radiation. J. Biol. Chem. *274*, 25571–25575.

Sweeney, F.D., Yang, F., Chi, A., Shabanowitz, J., Hunt, D.F., and Durocher, D. (2005). Saccharomyces cerevisiae Rad9 acts as a Mec1 adaptor to allow Rad53 activation. Curr. Biol. *15*, 1364–1375.

Syljuåsen, R.G., Sørensen, C.S., Hansen, T., Fugger, K., Lundin, C., Helleday, T., Sehested, M., Lukas, J., Hansen, L.T., Johansson, F., et al. (2005). Inhibition of Human Chk1 Causes Increased Initiation of DNA Replication, Phosphorylation of ATR Targets, and DNA Breakage. Mol. Cell. Biol. *25*, 3553–3562.

Tacke, R., Tohyama, M., Ogawa, S., and Manley, J.L. (1998). Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. Cell *93*, 139–148.

Takisawa, H., Mimura, S., and Kubota, Y. (2000). Eukaryotic DNA replication: From pre-replication complex to initiation complex. Curr. Opin. Cell Biol. *12*, 690–696.

Tanackovic, G., Ransijn, A., Ayuso, C., Harper, S., Berson, E.L., and Rivolta, C. (2011). A missense mutation in PRPF6 causes impairment of pre-mRNA splicing and autosomal-dominant retinitis pigmentosa. Am. J. Hum. Genet. *88*, 643–649.

Tanaka, K. (2010). Multiple functions of the S-phase checkpoint mediator. Biosci. Biotechnol. Biochem. *74*, 2367–2373.

Tanaka, A., Weinel, S., Nagy, N., O'Driscoll, M., Lai-Cheong, J.E., Kulp-Shorten, C.L., Knable, A., Carpenter, G., Fisher, S. a, Hiragun, M., et al. (2012). Germline mutation in ATR in autosomal- dominant oropharyngeal cancer syndrome. Am. J. Hum. Genet. *90*, 511–517.

Tange, T., Damgaard, C.K., Guth, S., Valcarcel, J., and Kjems, J. (2001). The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. EMBO J. *20*, 5748–5758.

Taylor, W.R., and Stark, G.R. (2001). Regulation of the G2/M transition by p53. Oncogene *20*, 1803–1815.

Teale, W.D., Paponov, I.A., and Palme, K. (2006). Auxin in action: signalling, transport and the control of plant growth and development. Nat. Rev. Mol. Cell Biol. *7*, 847–859.

Thanasoula, M., Escandell, J.M., Suwaki, N., and Tarsounas, M. (2012). ATM/ATR checkpoint activation downregulates CDC25C to prevent mitotic entry with uncapped telomeres. EMBO J. *31*, 3398–3410.

Thompson, S.L., and Compton, D.A. (2008). Examining the link between chromosomal instability and aneuploidy in human cells. J. Cell Biol. *180*, 665–672.

Toledo, L.I., Murga, M., Zur, R., Soria, R., Rodriguez, A., Martinez, S., Oyarzabal, J., Pastor, J., Bischoff, J.R., and Fernandez-Capetillo, O. (2011). A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. Nat. Struct. Mol. Biol. *18*, 721–727.

Toledo, L.I., Altmeyer, M., Rask, M.-B., Lukas, C., Larsen, D.H., Povlsen, L.K., Bekker-Jensen, S., Mailand, N., Bartek, J., and Lukas, J. (2013). ATR Prohibits Replication Catastrophe by Preventing Global Exhaustion of RPA. Cell *155*, 1088–1103.

Traven, A., and Heierhorst, J. (2005). SQ/TQ cluster domains: Concentrated ATM/ATR kinase phosphorylation site regions in DNA-damage-response proteins. BioEssays *27*, 397–407.

Tuul, M., Kitao, H., Iimori, M., Matsuoka, K., Kiyonari, S., Saeki, H., Oki, E., Morita, M., and Maehara, Y. (2013). Rad9, Rad17, TopBP1 and Claspin Play Essential Roles in Heat-Induced Activation of ATR Kinase and Heat Tolerance. PLoS One *8*, e55361.

Ünsal-Kaçmaz, K., Mullen, T.E., and William, K. (2005). Coupling of Human Circadian and Cell Cycles by the Timeless Protein. Mol. Cell. Biol. *25*, 3109–3116.

Ünsal-Kaçmaz, K., Chastain, P.D., Qu, P.-P., Minoo, P., Cordeiro-Stone, M., Sancar, A., and Kaufmann, W.K. (2007). The human Tim/Tipin complex coordinates an Intra-S checkpoint response to UV that slows replication fork displacement. Mol. Cell. Biol. *27*, 3131–3142.

Vastiau, A., Cao, L., Jaspers, M., Owsianik, G., Janssens, V., Cuppens, H., Goris, J., Nilius, B., and Cassiman, J.J. (2005). Interaction of the protein phosphatase 2A with the regulatory domain of the cystic fibrosis transmembrane conductance regulator channel. FEBS Lett *579*, 3392–3396.

Van Vugt, M.A.T.M., Bras, A., and Medema, R.H. (2004). Polo-like kinase-1 controls recovery from a G2 DNA damage-induced arrest in mammalian cells. Mol. Cell *15*, 799–811.

Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793–1808.

Wakayama, T., Kondo, T., Ando, S., Matsumoto, K., and Sugimoto, K. (2001). Pie1, a protein interacting with Mec1, controls cell growth and checkpoint responses in Saccharomyces cerevisiae. Mol. Cell. Biol. *21*, 755–764.

Walker, M., Black, E.J., Oehler, V., Gillespie, D. a, and Scott, M.T. (2009). Chk1 C-terminal regulatory phosphorylation mediates checkpoint activation by de-repression of Chk1 catalytic activity. Oncogene *28*, 2314–2323.

Wang, Z., and Burge, C.B. (2008). Splicing regulation : From a parts list of regulatory elements to an integrated splicing code. RNA *14*, 802–813.

Wang, B., Matsuoka, S., Ballif, B.A., Zhang, D., Smogorzewska, A., Gygi, S.P., and Elledge, S.J. (2007). Abraxas and RAP80 Form a BRCA1 Protein Complex Required for the DNA Damage Response. Science *316*, 1194–1198.

Wang, H., Zhang, X., Geng, L., Teng, L., and Legerski, R.J. (2009). Artemis Regulates Cell Cycle Recovery from the S Phase Checkpoint by Promoting Degradation of Cyclin E. J. Biol. Chem. *284*, 18236–18243.

Wang, X., Zou, L., Lu, T., Bao, S., Hurov, K.E., Hittelman, W.N., Elledge, S.J., and Li, L. (2006). Rad17 Phosphorylation Is Required for Claspin Recruitment and Chk1 Activation in Response to Replication Stress. Mol. Cell *23*, 331–341.

Weber, A.M., and Ryan, A.J. (2014). ATM and ATR as therapeutic targets in cancer. Pharmacol. Ther. *149*, 124–138.

Weemaes, C., Hustinx, T., Scheres, J., van Munster, P., Bakkeren, J., and Taalman, R. (1981). A new chromosomal instability disorder: the Nijmegen breakage syndrome. Acta Paediatr. Scand. *70*, 557–564.

Willems, M., Geneviève, D., Borck, G., Baumann, C., Baujat, G., Bieth, E., Edery, P., Farra, C., Gerard, M., Héron, D., et al. (2010). Molecular analysis of pericentrin gene (PCNT) in a series of 24 Seckel/microcephalic osteodysplastic primordial dwarfism type II (MOPD II) families. J. Med. Genet. *47*, 797–802.

Winding, P., and Berchtold, M.W. (2001). The chicken B cell line DT40: a novel tool for gene disruption experiments. J. Immunol. Methods *249*, 1–16.

Wolffe, A., Tafuri, S., Ranjan, M., and Familari, M. (1992). The Y-box factors: a family of nucleic acid binding proteins conserved from Escherichia coli to man. New Biol. *4*, 290–298.

Wu, J., and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell *75*, 1061–1070.

Xu, B., Kim, S., Kastan, M.B., and Xu, B.O. (2001). Involvement of Brca1 in S-Phase and G 2 -Phase Checkpoints after Ionizing Irradiation. Mol. Cell. Biol. *21*, 3445–3450.

Xu, X., Vaithiyalingam, S., Glick, G.G., Mordes, D.A., Chazin, W.J., and Cortez, D. (2008). The Basic Cleft of RPA70N Binds Multiple Checkpoint Proteins, Including RAD9, To Regulate ATR Signaling. Mol. Cell. Biol. *28*, 7345–7353.

Yan, J., and Jetten, A.M. (2008). RAP80 and RNF8, key players in the recruitment of repair proteins to DNA damage sites. Cancer Lett. *271*, 179–190.

Yang, H., Rudge, D.G., Koos, J.D., Vaidialingam, B., Yang, H.J., and Pavletich, N.P. (2013). mTOR kinase structure, mechanism and regulation. Nature *497*, 217–223.

Yang, S., Kuo, C., Bisi, J.E., and Kim, M.K. (2002). PML-dependent apoptosis after DNA damage is regulated by the checkpoint kinase hCds1/Chk2. Nat. Cell Biol. *4*, 865–870.

Yarden, R.I., Pardo-Reoyo, S., Sgagias, M., Cowan, K.H., and Brody, L.C. (2002). BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage. Nat. Genet. *30*, 285–289.

Yazdi, P.T., Wang, Y., Zhao, S., Patel, N., Lee, E.Y.H.P., and Qin, J. (2002). SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint. Genes Dev. *16*, 571–582.

You, Z., Chahwan, C., Bailis, J., Hunter, T., and Russell, P. (2005). ATM activation and its recruitment to damaged DNA require binding to the C terminus of Nbs1. Mol. Cell. Biol. *25*, 5363–5379.

Yu, X., and Chen, J. (2004). DNA Damage-Induced Cell Cycle Checkpoint Control Requires CtIP , a Phosphorylation-Dependent Binding Partner of BRCA1 C-Terminal Domains DNA Damage-Induced Cell Cycle Checkpoint Control Requires CtIP , a Phosphorylation-Dependent Binding Partner of BRCA1. Mol. Cell. Biol. *24*, 9478–9486.

Zachos, G., Rainey, M.D., and Gillespie, D. a F. (2003). Chk1-deficient tumour cells are viable but exhibit multiple checkpoint and survival defects. EMBO J. *22*, 713–723.

Zachos, G., Rainey, M.D., and Gillespie, D.A. (2005). Chk1-dependent S-M checkpoint delay in vertebrate cells is linked to maintenance of viable replication structures. Mol. Cell. Biol. *25*, 563–574.

Zegerman, P., and Diffley, J.F.X. (2010). Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. Nature *467*, 474–478.

Zhang, J., Yang, P.L., and Gray, N.S. (2009). Targeting cancer with small molecule kinase inhibitors. Nat. Rev. Cancer *9*, 28–39.

Zhao, H., and Piwnica-Worms, H. (2001). ATR-Mediated Checkpoint Pathways Regulate

Phosphorylation and Activation of ATR-Mediated Checkpoint Pathways Regulate Phosphorylation and Activation of Human Chk1. Mol. Cell. Biol. *21*, 4129–4139.

Zhao, X., Muller, E.G., and Rothstein, R. (1998). A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Mol. Cell *2*, 329–340.

Zheng, S., Kim, H., and Verhaak, R.G.W. (2014). Silent mutations make some noise. Cell *156*, 1129–1131.

Zhou, B.S.B., and Elledge, S.J. (2000). The DNA damage response: putting checkpoints in perspective. Nature *408*, 433–439.

Zhou, Z., Licklider, L.J., Gygi, S.P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. Nature *419*, 182–185.

Zighelboim, I., Schmidt, a. P., Gao, F., Thaker, P.H., Powell, M. a., Rader, J.S., Gibb, R.K., Mutch, D.G., and Goodfellow, P.J. (2009). ATR mutation in endometrioid endometrial cancer is associated with poor clinical outcomes. J. Clin. Oncol. *27*, 3091.

Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542–1548.

CHAPTER 8

Appendixes

8.1 Cell lines used in this study

Table S 1. DT40 cell lines. Genomic details of DT40 cell lines used in this study.

Cell line name	Genotype
WT	WT, Clone 18 (Buerstedde et al., 1990) Ova $^{+/+/+}$ Atr $^{+/+}$
Atr untagged	Ova $^{+/+/wt\ Atr\ cDNA\ (PuroR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
AID-Atr	Ova $^{+/+/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
HFSC-Atr	Ova $^{+/+/HFSC\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$
Atr $^{\Delta 1}$	Ova $^{+/\ \Delta 467\text{-}602\ Atr\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr $^{\Delta 2}$	Ova $^{+/\ \Delta 603\text{-}779\ Atr\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr $^{\Delta 3}$	Ova $^{+/\ \Delta 780\text{-}1138\ Atr\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr $^{\Delta 4}$	Ova $^{+/\ \Delta 1139\text{-}1646\ Atr\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr $^{\Delta 5}$	Ova $^{+/\ \Delta 2230\text{-}2308\ Atr\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr $^{H32\text{-}33\Delta > ATRH32\text{-}33}$	Ova $^{+/\ Atr\ H32\text{-}33\Delta >\ ATR\ H32\text{-}33\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr ^{M1180I}	Ova $^{+/\ Atr\ M1180I\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr ^{K1685N}	Ova $^{+/\ Atr\ K1685N\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$
Atr ^{Q2163R}	Ova $^{+/\ Atr\ Q2163R\ cDNA\ (PuroR)/AID\text{-}Atr\ cDNA\ (HistoR)}$ Atr $^{\Delta CT\ (HygroR)\ /\ \Delta CT\ (BsdR)}$ Pcmv OsTIR1 $^{(G418R)}$

Table S 2. Yeast cell lines. Genomic details of yeast cell lines generated in this study.

Cell line name	Genotype
WT	W303-1a: *MATα ade2-1; his3-11,15; leu2-3,112; trp1-1; ura3; can1-100* (Nasmyth et al., 1990)
$Mec1^{H32\text{-}33\Delta>ATRH32\text{-}33}$	*MATα sml1::TRP1* $mec1^{H32\text{-}33\Delta>ATR^{H32\text{-}33}}$
$Mec1^{H11\text{-}13\Delta>Tel1H13\text{-}14}$	*MATα sml1::TRP1* $mec1^{H11\text{-}13\Delta>Tel1^{H13\text{-}14}}$
$Mec1^{H14\text{-}21\Delta>Tel1H17\text{-}23}$	*MATα sml1::TRP1* $mec1^{H14\text{-}21\Delta>Tel1^{H17\text{-}23}}$

8.2 List of primers used in this study

Generation of *Atr* and *Mec1* mutant constructs was carried out using primers below (Table S3 and S5, respectively). Fusion PCR of the indicated fragments allowed for the generation of the different constructs.

DT40 constructs were then cloned into an *Atr* cDNA vector (pLC4; plasmid original from (Eykelenboom et al., 2013)) by double digest, as indicated in Table S3. Once in a pLC4 background, *Atr* constructs were sequenced using primers in Table S4. Constructs were then released from the *Atr* cDNA vector by a SpeI digest and subsequently cloned into an NheI site in the *Ovalbumin* targeting vector, which was used for transfection and generation of stable cell lines.

Mec1 PCR products were directly used in the generation of stable cell lines by *Delitto Perfetto*. Primers and parental CORE strains, generated by Dr. Carla Abreu, used in each case are indicated in Table S5.

Primers used for splicing assays and CRISPR experiments using human cells are also shown below (Table S6).

Table S 3. DT40 primers used to generate *Atr* mutants: PCR and cloning details.

PCR	Primers	Sequence (5'>3')	Details
Fragment 1 for deletion mutant 1: Δ468-602	Del_T3pr_F1	GCGCGAAATTAACCCTCACTAAAG	Cloning into pLC4 was carried out by SfiI/StuI digest.
	Del_468-602_R1	CTGATGATTCAG-CTTCAATTTCATGTTTACCGGGC	
Fragment 2 for deletion mutant 1: Δ468-602	Del_468-602_F2	ATGAAATTGAAG-CTGAATCATCAGATTCTTGATAAT	
	Del_KpnI_R2	CAGATGGGTTTCTTGTACTGACTG	
Fragment 1 for deletion mutant 2: Δ603-779	Del_T3pr_F1	GCGCGAAATTAACCCTCACTAAAG	Cloning into pLC4 was carried out by SfiI/StuI digest.
	Del_603-779_R1	ACAGGCTTGCAC-TGGAATTTTTATTTGTGCATACAGC	
Fragment 2 for deletion mutant 2: Δ603-779	Del_603-779_F2	ATAAAAATTCCA-GTGCAAGCCTGTGTTTTTAAGCCA	
	Del_KpnI_R2	CAGATGGGTTTCTTGTACTGACTG	
Fragment 1 for deletion mutant 3: Δ780-1138	Del_SfiI_F1	CTAGAGCCTCTGCTAACCATGTTC	Cloning into pLC4 was carried out by SfiI/StuI digest.
	Del_780-1138_R1	ATCAGCCATTTG-ACAGCAAGTTTTTTCATGAGCAGA	
Fragment 2 for deletion mutant 3: Δ780-1138	Del_780-1138_F2	AAAACTTGCTGT-CAAATGGCTGATTATCTGCAGCCT	
	Del_StuI_R2	TGTTCAATTCAGAAATCCATTCAG	
Fragment 1 for deletion mutant 4: Δ1139-1646	Del_SfiI_F1	CTAGAGCCTCTGCTAACCATGTTC	Cloning into pLC4 was carried out by SfiI/StuI digest.
	Del_1139-1646_R1	TGTCACATTCTG-CTCAGATGACGTTATTTCCCTAGG	
Fragment 2 for deletion mutant 4: Δ1139-1646	Del_1139-1646_F2	ACGTCATCTGAG-CAGAATGTGACACGCTTTTTAGACC	
	Del_StuI_R2	TGTTCAATTCAGAAATCCATTCAG	
Fragment 1 for deletion mutant 5:	Del_SfiI_F1	CTAGAGCCTCTGCTAACCATGTTC	Cloning into pLC4 was carried

Δ2230-2308	Del_2230-2308_R1	AAAGCCTGCAAT-TTTTCCTAAGGACTCCTTCATGTTG	out by StuI/NotI digest.
Fragment 2 for deletion mutant 5: Δ2230-2308	Del_2230-2308_F2	TCCTTAGGAAAA-ATTGCAGGCTTTGATGATACGGTAG	
	Del_T7pr_R2	CGTAATACGACTCACTATAGGGCG	
Fragment 1 for Replacement mutant (homology arm1): Atr ΔH32-33 > ATRH32-33	Del_SfiI_F1	CTAGAGCCTCTGCTAACCATGTTC	Cloning into pLC4 was carried out by SfiI/NotI digest.
	Rep_32-33_arm1_R2	Human tail: GGTTACACTCTG Chicken: ATATTCTCCATAGTCCTCACTTGA	
Fragment 2 for Replacement mutant: Atr ΔH32-33 > ATRH32-33	Rep_32-33_F1	Chicken tail: TATGGAGAATAT Human: CAGAGTGTAACCCGTTTTCTAGAC	
	Rep_32-33_R2	Chicken tail: GCCAATACTTTC Human: ATGTTCAAGGATCTGTTCTTTTAG	
Fragment 3 for Replacement mutant (homology arm2): Atr ΔH32-33 > ATRH32-33	Rep_32-33_arm2_F1	Human tail: ATCCTTGAACAT Chicken: GAAAGTATTGGCTTGTTAAGAGAT	
	Del_T7pr_R2	CGTAATACGACTCACTATAGGGCG	
Fragment 1 for Seckel mutant 1: M1180I	Del_SfiI_F1	CTAGAGCCTCTGCTAACCATGTTC	Cloning into pLC4 was carried out by SfiI/StuI digest.
	Point_M1159I_R2 [higher GC%]	GGGACCCATAAGCTTGATTAAAGACATAAG	
Fragment 2 for Seckel mutant 1: M1180I	Point_ M1159I_F1 [higher GC%]	CTTATGTCTTTAATCAAGCTTATGGGTCCC	
	Del_StuI_R2	TGTTCAATTCAGAAATCCATTCAG	
Fragment 1 for Seckel mutant 2: K1685N	Del_SfiI_F1	CTAGAGCCTCTGCTAACCATGTTC	Cloning into pLC4 was carried out by SfiI/StuI digest.
	Point_K1665N_R2 [higher GC%]	GCTCCTGAATATTTTGTTTGTTCTCTGTGATG	
Fragment 2 for Seckel mutant 2: K1685N	Point_K1665N_F1 [higher GC%]	CATCACAGAGAACAAACAAAATATTCAGGAGC	
	Del_T7pr_R2	CGTAATACGACTCACTATAGGGCG	
Fragment 1 for	Del_StuI_F1	AACCATCTCTCAAAGAACAAAT	Cloning into

Cancer-related mutant: Q2163R		CC	pLC4 was carried out by SfiI/NotI digest.
	Point_Q2144R_R2	CGGGAGATTAACCGAGAAAAAGCTGTCAGAAACTG	
Fragment 2 for Cancer-related mutant: Q2163R	Point_Q2144R_F1	GACAGCTTTTTCTCGGTTAATCTCCCGAATCTGCCATTC	
	Del_T7pr_R2	CGTAATACGACTCACTATAGGGCG	

Table S 4. DT40 primers used for sequencing and Southern blotting.

Primers	Sequence (5'>3')	Details
Ova_PF3	TTTATGGGGGAAAAATGCAG	Southern probe generation
Ova_PR3	CAGATGAGTTGTCCCAGGTG	
Atr_seq_F1	CATTGCACCTGATGGGTATG	Sequencing of *Atr* constructs
Atr_seq_F2	AGTTTTTGTGAGCGTGCTTG	
Atr_seq_F3	ATGCTCCTTGGATTCTGTGG	
Atr_seq_F4	CCAAGTGTGCAGTCCTTCAG	
Atr_seq_F5	TATTGGAAGGGCATCGAAAG	
Atr_seq_F6	CGTATAGGAGAACACTATCAGCAAG	
Atr_seq_F7	TTGCAACTGTCCATGAGAGC	
Atr_seq_F8	CAGAATGGTCAGCAACTTGG	
Atr_seq_F9	TGTGACACGCTTTTTAGACCTG	
Atr_seq_F10	GAACAGATTGTGCCCCTTTC	

Table S 5. Yeast primers used to generate *mec1* hybrid mutants.

PCR	Primers	Sequence (5'>3')	Details
Fragment 1 for yeast mutant (homology arm1): Mec1 ΔH11-13 > ATRH11-13	T2	ATGGAATCACACGTCAAATA	Parental strain used for generation of mutant: *mec1::CORE2* strain (De Castro Abreu, 2012)
	B6R1	CAAATCAAAATATACAATCG	
Fragment 2 for yeast mutant: Mec1 ΔH11-13 > ATRH11-13	R1.2ATR-5'	Mec1 tail: GACTAATAAAGCTCTCTCGATT GTATATTTTGATTTG ATR: TTTGAAGATCATATCCTGGAAG	
	R1.2ATR-3'	Mec1 tail: GTTTGAATGAAGTATCCTGAAG ATTTCTGATTTACC ATR: AGTGGCTTTCAAGTTCCTAC	
Fragment 3 for yeast mutant (homology arm2): Mec1 ΔH11-13 > ATRH11-13	T5	GGTAAATCAGAAATCTTCAGG	
	B3	CCTTTGTCAAAAGTTGGAGCC	
Fragment 1 for yeast mutant (homology arm1): Mec1 ΔH32-33 > ATRH32-33	T14	AATAACTGGACGGATGATCAG	Parental strain used for generation of mutant: *mec1::CORE3* strain (De Castro Abreu, 2012)
	B9	AAGCATGTTAGTTGTTTTCGTA TC	
Fragment 2 for yeast mutant: Mec1 ΔH32-33 > ATRH32-33	R4.1 ATR 5'	Mec1 tail: CACGGCACTTTTATCATTAAAG ATACGAAAACAACTAACATGC TT ATR: CAGAGTGTAACCCGTTTTCTAG	
	R4.1 ATR 3'	Mec1 tail: GACATTAAAGCAGTCTTGTGCG AGTTTCCAGTTTTCAGAATATT G ATR: AAGGCTTTCATGTTCAAGGATC	
Fragment 3 for yeast mutant (homology arm2): Mec1 ΔH32-33 >	R4.1-T.F3	CAATATTCTGAAAACTGGAAAC	
	R4.1-B.F3	GTATTCAAACTTATCTTTCGC	

ATRH32-33			
Fragment 1 for yeast mutant (homology arm1): Mec1 ΔH11-13 >Tel1H13-14	T2	ATGGAATCACACGTCAAATA	Parental strain used for generation of mutant: *mec1::CORE2* strain (De Castro Abreu, 2012)
	B6R1	CAAATCAAAATATACAATCG	
Fragment 2 for yeast mutant: Mec1 ΔH11-13 >Tel1H13-14	R1.1tel1-5'	Mec1 tail: GACTAATAAAGCTCTCTCGATTGTATATTTTGATTTG Tel1: CTCACACGTTATAATGAAATTC	
	R1.1tel1-3'	Mec1 tail: GTTTGAATGAAGTATCCTGAAGATTTCTGATTTACC Tel1: TCCAATTGTGTACGGACTATTATTC	
Fragment 3 for yeast mutant (homology arm2): ΔH11-13 >Tel1H13-14	T5	GGTAAATCAGAAATCTTCAGG	
	B3	CCTTTGTCAAAAGTTGGAGCC	
Fragment 1 for yeast mutant (homology arm1): Mec1 ΔH14-21 > Tel1H17-23	T4	AAAAGAAGGTCCCTACCTGTTGAAGC	Parental strain used for generation of mutant: *mec1::CORE2* strain (De Castro Abreu, 2012)
	B4	CGCAGCCTCTGGTCTATTTGG	
Fragment 2 for yeast mutant: Mec1 ΔH14-21 > Tel1H17-23	R2.2tel1-5'	Mec1 tail: CGTGCGAATAAATCCAAATAGACCAGAGGCTGCG Tel1: AGTTCTTCCAATCAAGCTATG	
	R2.2tel1-3'	Mec1 tail: GGAACGCTTGTTCCTGATCATCCGTCCAGTTATT Tel1: AATACACAAGTAATAGCTGTTATC	
Fragment 3 for yeast mutant (homology arm2): Mec1 ΔH14-21 > Tel1H17-23	T14	AATAACTGGACGGATGATCAG	
	B9	AAGCATGTTAGTTGTTTTCGTATC	

Table S 6. Human primers used in CRISPR and minigene experiments

Primers	Sequence (5'>3')	Details
Ex17_fw	CATGCAGTTACTGAGCTCTAGTG	Study of splicing defects caused by *ATR*^{M1159I} mutation by PCR/Sequencing – Human cell lines
Ex19_rv	GAGGTAGTGGAAGATAGCTGCAG	
Ex27_fw	GCAGTTCTAAAGCATGACGATC	Study of splicing defects caused by *ATR*^{K1665N} mutation by PCR/Sequencing – Human cell lines
Ex29_rv	CTGGTCTGGTTCTAGCTGAATAG	
minigene_in17.fw	GACACATATGCTGTTCTAGATGAATACTGGGC	Generation of *ATR-ex18* construct – Minigene assay
minigene_in18.rv	CAGTCATATGGGGAAATAGGTATGTAGGTTC	
minigene_in27.fw	GACACATATGCACCAACAACGTTGTAAGGAAG	Generation of *ATR-ex28* construct – Minigene assay
minigene_in28.rv	CAGTCATATGCACCATCTCCAAGAAGCCATTC	
M1159I.fw.sitedm	CAGTTTGATGTCTTTGATTAAGTTAATGGGACCCAAAC	Introduction of M1159I mutation in *ATR-ex18* construct – Minigene assay
M1159I.rv.sitedm	GTTTGGGTCCCATTAACTTAATCAAAGACATCAAACTG	
K1665N.fw.sitedm	GAATCATTTATTACAGAAAATAAGCAAAATATTCAGG	Introduction of K1665N mutation in *ATR-ex28* construct – Minigene assay
K1665N.rv.sitedm	CCTGAATATTTTGCTTATTTTCTGTAATAAATGATTC	
minig_NF1.fw	CAACTTCAAGCTCCTAAGCCACTGC	Final analysis of minigene splicing by PCR
minig_NF1.rv	TAGGATCCGGTCACCAGGAAGTTGGTTAAATCA	
NF1_sequencing	CCACACAGCAAAGAGAAACATAG	Sequencing of *NF1*/*ATR* minigenes
gRNA1.fw	CACC – GCTCCGATCGTGTACAAATG	Generation of *Atr*^{Q2144R} point mutant by CRISPR
gRNA1.rv	AAAC – CATTTGTACACGATCGGAGC	
gRNA2.fw	CACC – CCTATCCTCAACAAGCAATG	
gRNA2.rv	AAAC – CATTGCTTGTTGAGGATAGG	
insert seq rev	GTCTGCAGAATTGGCGCACG	Sequencing of pX335 plasmid – gRNA cloning

F1.fw.noMreI	GGTACAACATAGATGAACCTCG	Screening of CRISPR clones (PCR)
F2.rv.noMreI	GATAACCATCTGCTTAGTTCCC	
Screening.rv	CTGCATTCCAGCCTGGGCAAC	Sequencing of CRISPR clones

8.3 ATR/CHK1 gel recipes

All volumes are in µl and are sufficient for one Biorad 1mm gel.

Table S 7. Preparation of an ATR (or BRCA2) SDS-PAGE resolving gel.

ATR resolving gel	
Gel %	**6.5**
Crosslinking ratio	**80:1**
dH2O	2290
40% Acrylamide	856
2% Bis-acrylamide	213
1M Tris pH 8.7	2000
10% SDS	33
10% Ammonium persulfate (APS)	30
TEMED	10

Table S 8. Preparation of a CHK1 SDS-PAGE resolving gel.

CHK1 resolving gel	
Gel %	**12**
Crosslinking ratio	**37.5:1**
dH2O	1186
30% Acrylamide/bis-acrylamide	2173
1M Tris pH 8.7	2000
10% SDS	33
10% Ammonium persulfate (APS)	30
TEMED	10

Table S 9. Preparation of SDS-PAGE stacking gel.

Stacking gel	
Gel %	**5**
dH2O	2250
40% Acrylamide	427
2% Bis-acrylamide	233
1M Tris pH 8.7	417
10% SDS	17
10% Ammonium persulfate	17

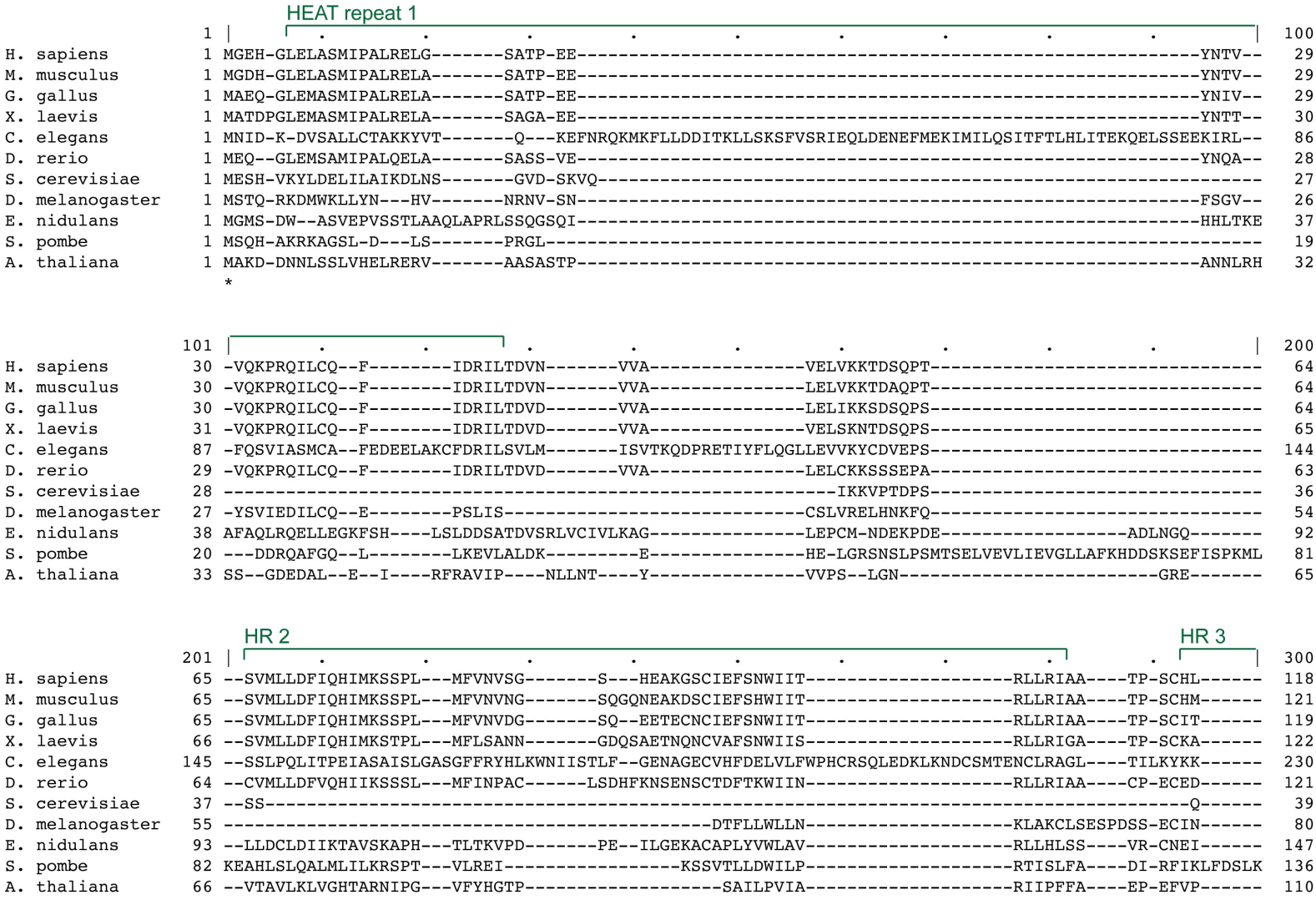
(APS)	
TEMED	7

8.4 Alignment of ATR orthologues

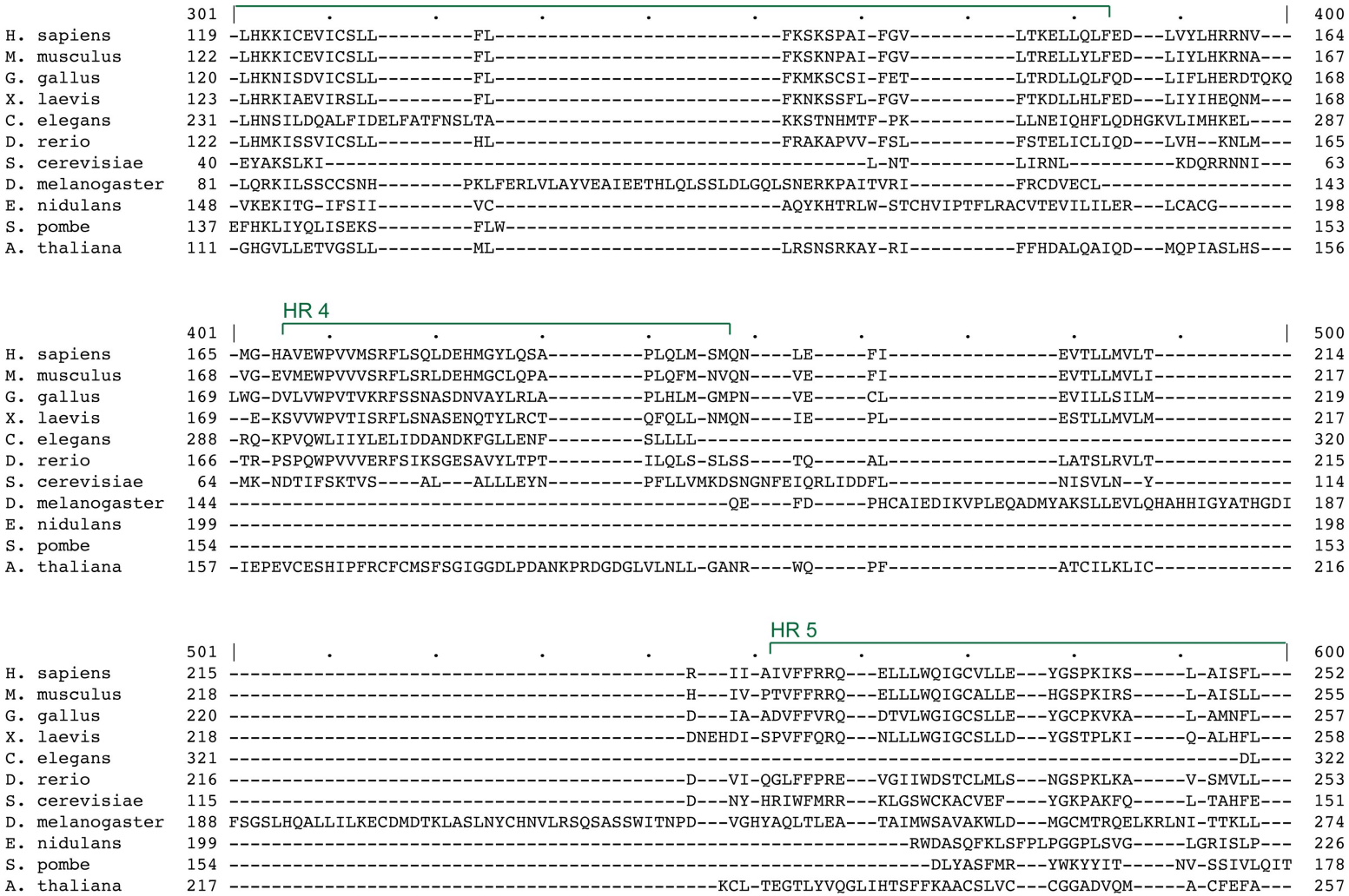
Alignment between *Homo sapiens* ATR protein and its orthologues from the following species: *Mus musculus* (mouse), *Gallus gallus*, (chicken), *Xenopus laevis* (frog), *Caenorhabditis elegans* (roundworm), *Danio rerio* (zebra fish), *Saccharomyces cerevisiae* (budding yeast), *Drosophila melanogaster* (fruit fly), *Aspergillus nidulans* (fungi), *Schizosaccharomyces pombe* (fission yeast), *Arabidopsis thaliana* (plants).

T-coffee online software was used to align these protein sequences and a script written by Dr. Andrew Flaus was used to obtain the desired output in landscape. The dots above the alignment are used as markers every 10 residues, while a line below the alignment is used to mark strongly conserved positions, as stated by the T-coffee software. "*" indicates positions which have a single, fully conserved residue. ":" indicates that one of the following 'strong' groups is fully conserved: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY or FYW. "." indicates that one of the following 'weaker' groups is fully conserved: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM or HFY. These are all the positively scoring groups that occur in the Gonnet Pam250 matrix. The strong and weak groups are defined as strong score > 0.5 and weak score =< 0.5 respectively.

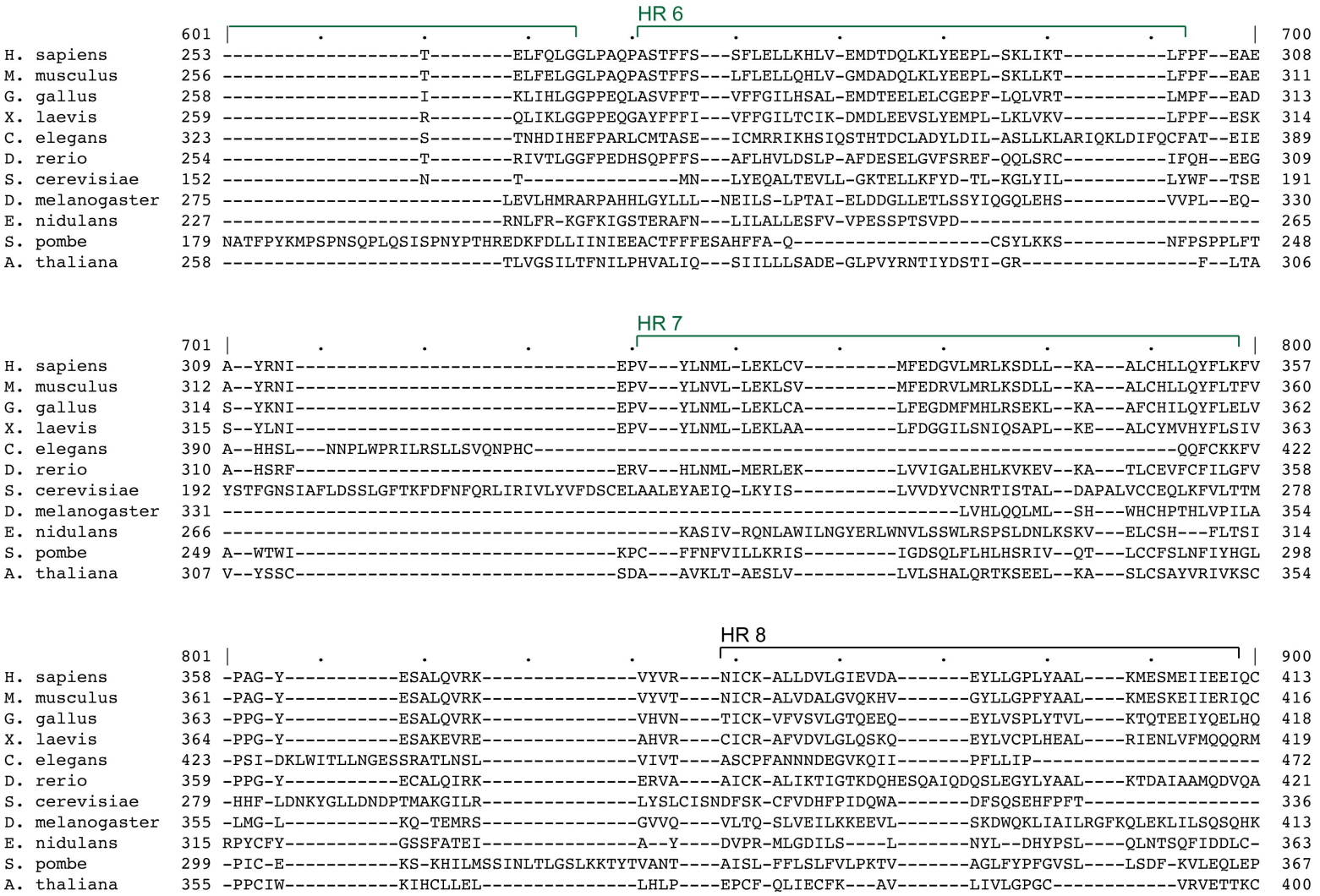
HEAT repeats, according to its definition within the human ATR sequence (Perry and Kleckner, 2003), are highlighted throughout the alignment. Important HEAT repeats, such as the ATR-specific units, the ATRIP binding domain and the FAT domain are marked in purple, green and orange, respectively. Other features, such as the kinase and FATC domains, are marked in yellow and red, respectively.



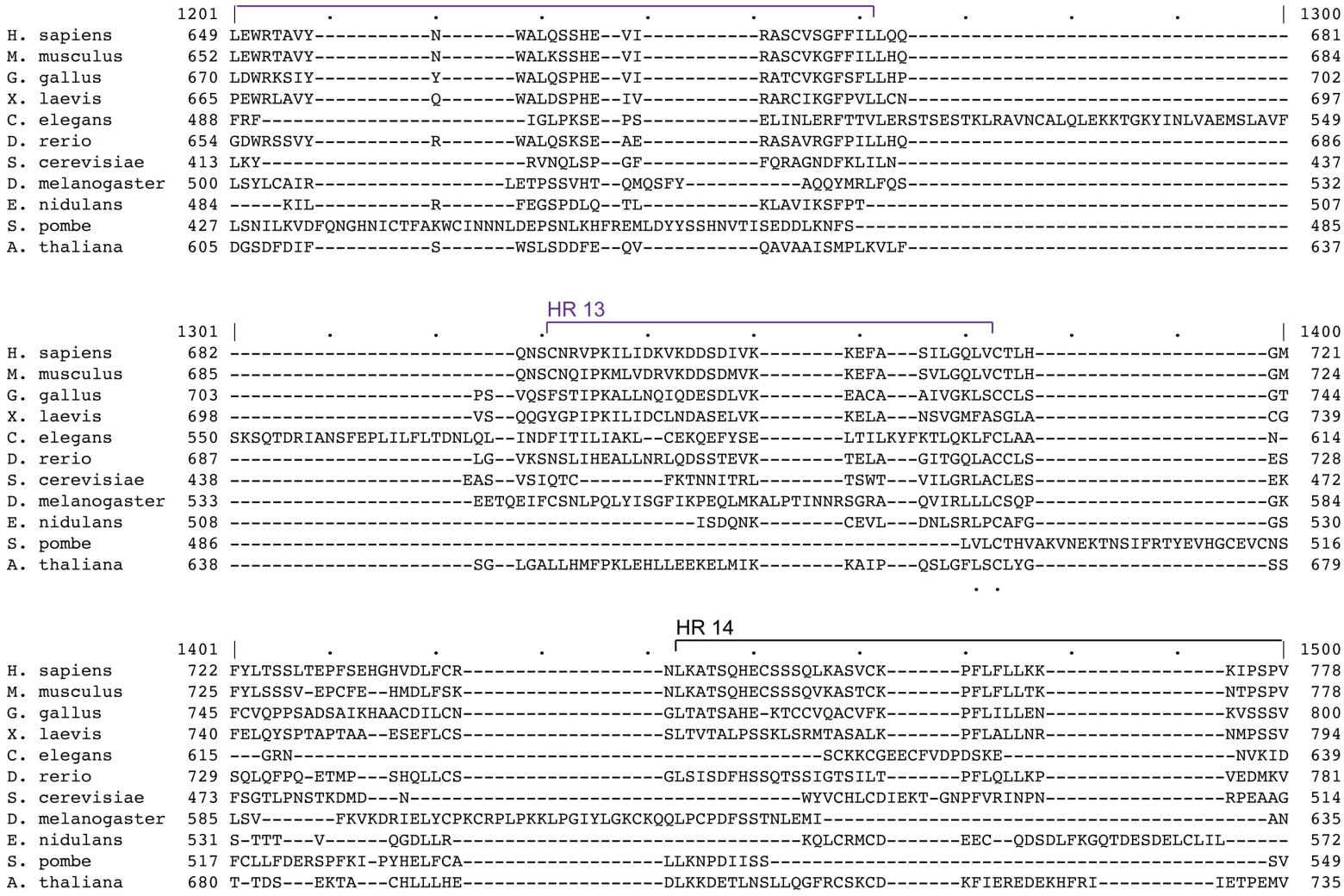

```
                    1 |    .         .         .         .         .         .         .         .         .         | 100
H. sapiens          1 MGEH-GLELASMIPALRELG-------SATP-EE-------------------------------------------------YNTV--  29
M. musculus         1 MGDH-GLELASMIPALRELA-------SATP-EE-------------------------------------------------YNTV--  29
G. gallus           1 MAEQ-GLEMASMIPALRELA-------SATP-EE-------------------------------------------------YNIV--  29
X. laevis           1 MATDPGLEMASMIPALRELA-------SAGA-EE-------------------------------------------------YNTT--  30
C. elegans          1 MNID-K-DVSALLCTAKKYVT-------Q---KEFNRQKMKFLLDDITKLLSKSFVSRIEQLDENEFMEKIMILQSITFTLHLITEKQELSSEEKIRL--  86
D. rerio            1 MEQ--GLEMSAMIPALQELA-------SASS-VE-------------------------------------------------YNQA--  28
S. cerevisiae       1 MESH-VKYLDELILAIKDLNS-------GVD-SKVQ-------------------------------------------------------  27
D. melanogaster     1 MSTQ-RKDMWKLLYN---HV-------NRNV-SN---------------------------------------------------FSGV--  26
E. nidulans         1 MGMS-DW--ASVEPVSSTLAAQLAPRLSSQGSQI---------------------------------------------------HHLTKE  37
S. pombe            1 MSQH-AKRKAGSL-D---LS-------PRGL-------------------------------------------------------  19
A. thaliana         1 MAKD-DNNLSSLVHELRERV-------AASASTP---------------------------------------------------ANNLRH  32
                      *

                  101 |    .         .         .         .         .         .         .         .         .         | 200
H. sapiens         30 -VQKPRQILCQ--F--------IDRILTDVN-------VVA---------------VELVKKTDSQPT-------------------------  64
M. musculus        30 -VQKPRQILCQ--F--------IDRILTDVN-------VVA---------------LELVKKTDAQPT-------------------------  64
G. gallus          30 -VQKPRQILCQ--F--------IDRILTDVD-------VVA---------------LELIKKSDSQPS-------------------------  64
X. laevis          31 -VQKPRQILCQ--F--------IDRILTDVD-------VVA---------------VELSKNTDSQPS-------------------------  65
C. elegans         87 -FQSVIASMCA--FEDEELAKCFDRILSVLM-------ISVTKQDPRETIYFLQGLLEVVKYCDVEPS-------------------------- 144
D. rerio           29 -VQKPRQILCQ--F--------IDRILTDVD-------VVA---------------LELCKKSSSEPA-------------------------  63
S. cerevisiae      28 ------------------------------------------------------IKKVPTDPS-------------------------  36
D. melanogaster    27 -YSVIEDILCQ--E--------PSLIS----------------------------CSLVRELHNKFQ-------------------------  54
E. nidulans        38 AFAQLRQELLEGKFSH----LSLDDSATDVSRLVCIVLKAG---------------LEPCM-NDEKPDE-----------------ADLNGQ-------  92
S. pombe           20 ---DDRQAFGQ--L--------LKEVLALDK---------E---------------HE-LGRSNSLPSMTSELVEVLIEVGLLAFKHDDSKSEFISPKML  81
A. thaliana        33 SS--GDEDAL--E--I----RFRAVIP----NLLNT----Y---------------VVPS--LGN------------------------GRE-------  65

                  201 |    .         .         .         .         .         .         .      .  .         .  | 300
H. sapiens         65 --SVMLLDFIQHIMKSSPL---MFVNVSG-------S---HEAKGSCIEFSNWIIT--------------------RLLRIAA----TP-SCHL------ 118
M. musculus        65 --SVMLLDFIQHIMKSSPL---MFVNVNG-------SQGQNEAKDSCIEFSHWIIT--------------------RLLRIAA----TP-SCHM------ 121
G. gallus          65 --SVMLLDFIQHIMKSSPL---MFVNVDG-------SQ--EETECNCIEFSNWIIT--------------------RLLRIAA----TP-SCIT------ 119
X. laevis          66 --SVMLLDFIQHIMKSTPL---MFLSANN-------GDQSAETNQNCVAFSNWIIS--------------------RLLRIGA----TP-SCKA------ 122
C. elegans        145 --SSLPQLITPEIASAISLGASGFFRYHLKWNIISTLF--GENAGECVHFDELVLFWPHCRSQLEDKLKNDCSMTENCLRAGL----TILKYKK------ 230
D. rerio           64 --CVMLLDFVQHIIKSSSL---MFINPAC------LSDHFKNSENSCTDFTKWIIN--------------------RLLRIAA----CP-ECED------ 121
S. cerevisiae      37 --SS-------------------------------------------------------------------------------Q------  39
D. melanogaster    55 ------------------------------------------DTFLLWLLN--------------------KLAKCLSESPDSS-ECIN------  80
E. nidulans        93 --LLDCLDIIKTAVSKAPH---TLTKVPD-------PE--ILGEKACAPLYVWLAV--------------------RLLHLSS----VR-CNEI------ 147
S. pombe           82 KEAHLSLQALMLILKRSPT---VLREI-----------------KSSVTLLDWILP--------------------RTISLFA----DI-RFIKLFDSLK 136
A. thaliana        66 --VTAVLKLVGHTARNIPG---VFYHGTP------------------SAILPVIA--------------------RIIPFFA----EP-EFVP------ 110
```

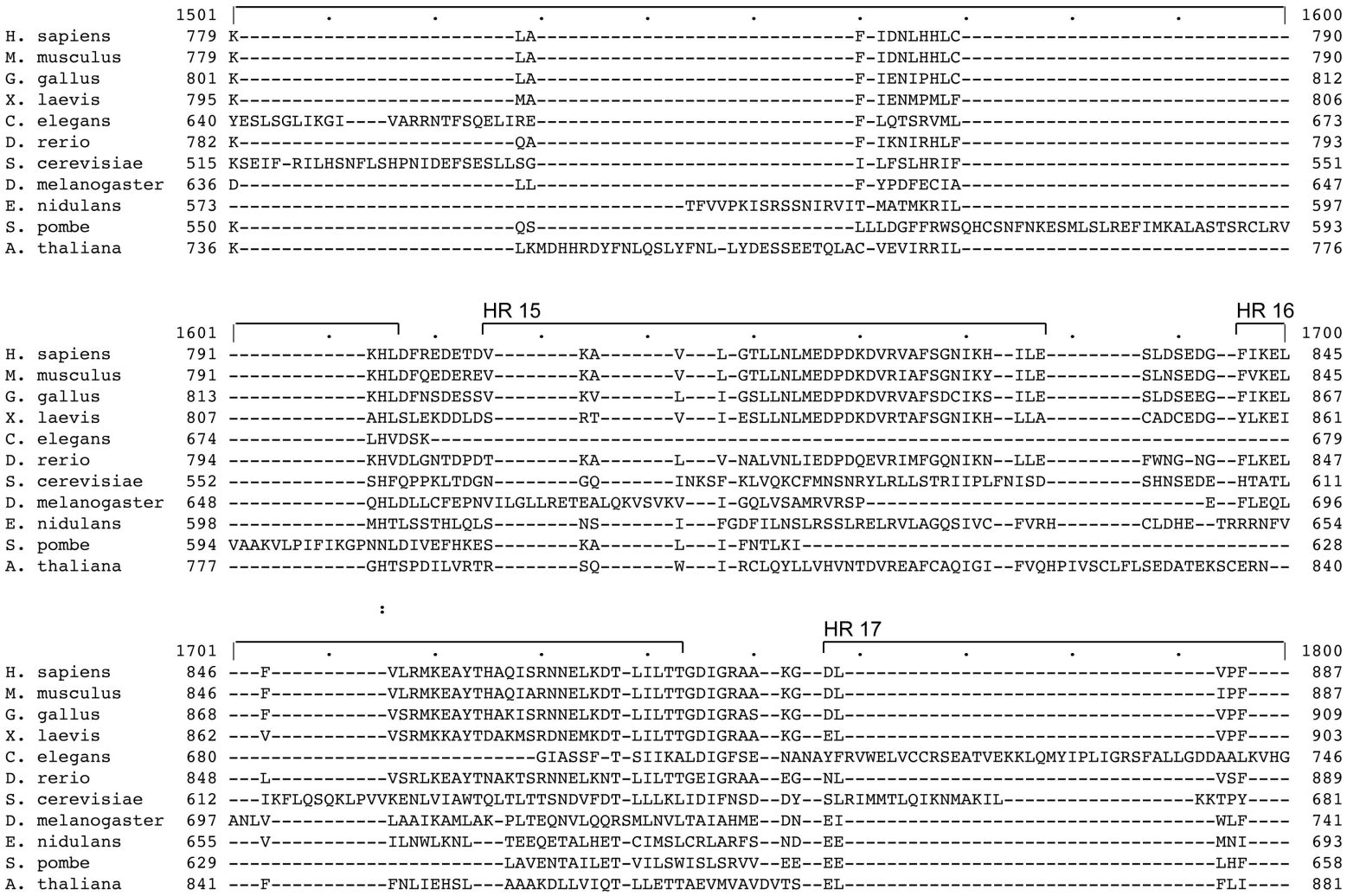


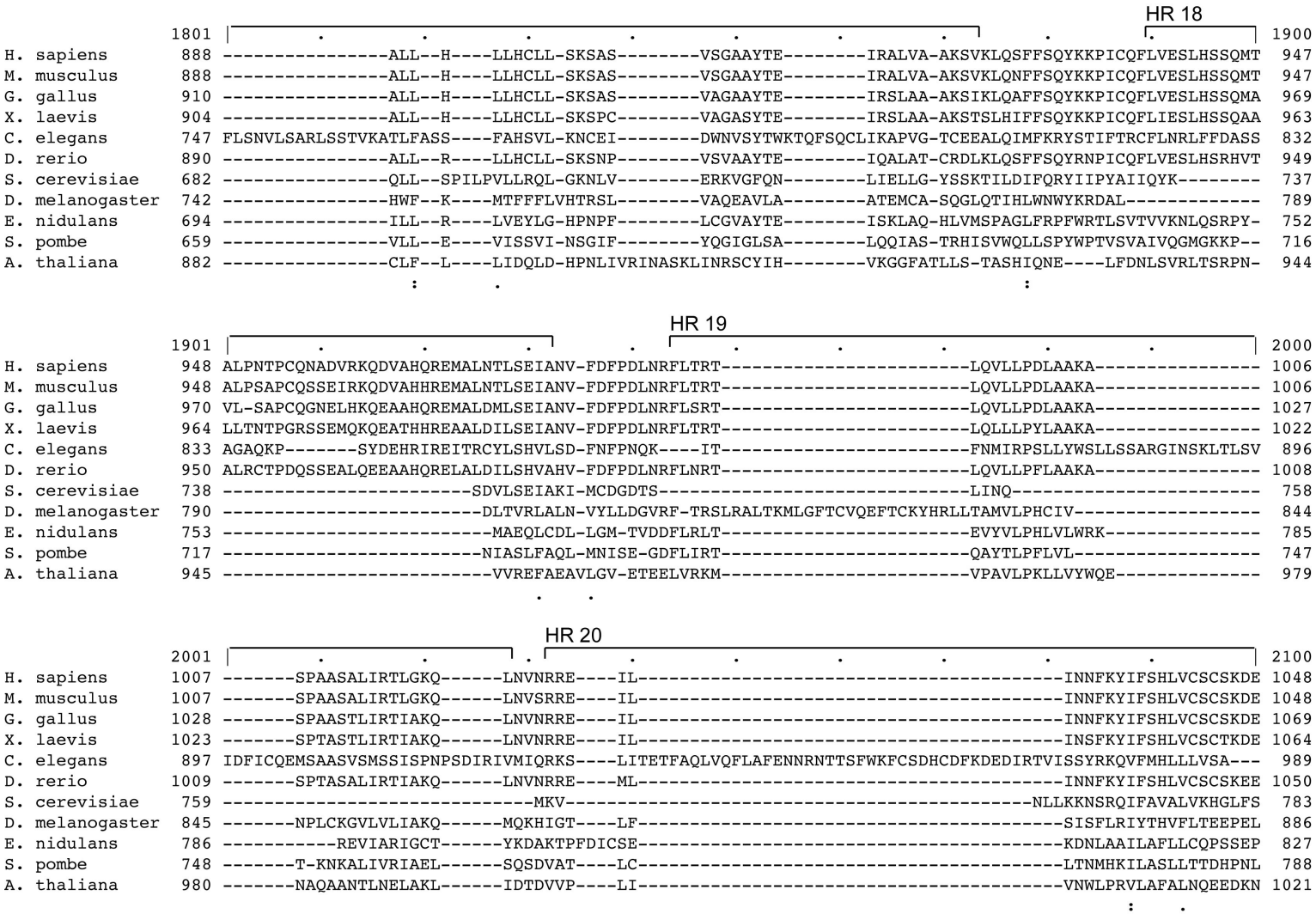

```
                 301 |       .         .         .         .         .         .         .         .         | 400
H. sapiens       119 -LHKKICEVICSLL---------FL-------------------------FKSKSPAI-FGV----------LTKELLQLFED---LVYLHRRNV---  164
M. musculus      122 -LHKKICEVICSLL---------FL-------------------------FKSKNPAI-FGV----------LTRELLYLFED---LIYLHKRNA---  167
G. gallus        120 -LHKNISDVICSLL---------FL-------------------------FKMKSCSI-FET----------LTRDLLQLFQD---LIFLHERDTQKQ  168
X. laevis        123 -LHRKIAEVIRSLL---------FL-------------------------FKNKSSFL-FGV----------FTKDLLHLFED---LIYIHEQNM---  168
C. elegans       231 -LHNSILDQALFIDELFATFNSLTA-------------------------KKSTNHMTF-PK----------LLNEIQHFLQDHGKVLIMHKEL----  287
D. rerio         122 -LHMKISSVICSLL---------HL-------------------------FRAKAPVV-FSL----------FSTELICLIQD---LVH--KNLM---  165
S. cerevisiae     40 -EYAKSLKI-------------------------------------------L-NT----------LIRNL----------KDQRRNNI---   63
D. melanogaster   81 -LQRKILSSCCSNH---------PKLFERLVLAYVEAIEETHLQLSSLDLGQLSNERKPAITVRI----------FRCDVECL-----------  143
E. nidulans      148 -VKEKITG-IFSII---------VC-------------------------AQYKHTRLW-STCHVIPTFLRACVTEVILILER---LCACG-------  198
S. pombe         137 EFHKLIYQLISEKS---------FLW-------------------------------------------------------------------  153
A. thaliana      111 -GHGVLLETVGSLL---------ML-------------------------LRSNSRKAY-RI----------FFHDALQAIQD---MQPIASLHS---  156

                 401 |       .         .         .         .         .         .         .         .         | 500
H. sapiens       165 -MG-HAVEWPVVMSRFLSQLDEHMGYLQSA---------PLQLM-SMQN----LE-----FI----------EVTLLMVLT-------------  214
M. musculus      168 -VG-EVMEWPVVVSRFLSRLDEHMGCLQPA---------PLQFM-NVQN----VE-----FI----------EVTLLMVLI-------------  217
G. gallus        169 LWG-DVLVWPVTVKRFSSNASDNVAYLRLA---------PLHLM-GMPN----VE-----CL----------EVILLSILM-------------  219
X. laevis        169 --E-KSVVWPVTISRFLSNASENQTYLRCT---------QFQLL-NMQN----IE-----PL----------ESTLLMVLM-------------  217
C. elegans       288 -RQ-KPVQWLIIYLELIDDANDKFGLLENF---------SLLLL-------------------------------------------------  320
D. rerio         166 -TR-PSPQWPVVVERFSIKSGESAVYLTPT---------ILQLS-SLSS----TQ-----AL----------LATSLRVLT-------------  215
S. cerevisiae     64 -MK-NDTIFSKTVS----AL---ALLLEYN---------PFLLVMKDSNGNFEIQRLIDDFL----------NISVLN--Y-------------  114
D. melanogaster  144 ------------------------------------------QE----FD-----PHCAIEDIKVPLEQADMYAKSLLEVLQHAHHIGYATHGDI  187
E. nidulans      199 -------------------------------------------------------------------------------------------  198
S. pombe         154 -------------------------------------------------------------------------------------------  153
A. thaliana      157 -IEPEVCESHIPFRCFCMSFSGIGGDLPDANKPRDGDGLVLNLL-GANR----WQ-----PF----------ATCILKLIC-------------  216

                 501 |       .         .         .         .         .         .         .         .         | 600
H. sapiens       215 ------------------------------------------R---II-AIVFFRRQ---ELLLWQIGCVLLE---YGSPKIKS-----L-AISFL---  252
M. musculus      218 ------------------------------------------H---IV-PTVFFRRQ---ELLLWQIGCALLE---HGSPKIRS-----L-AISLL---  255
G. gallus        220 ------------------------------------------D---IA-ADVFFVRQ---DTVLWGIGCSLLE---YGCPKVKA-----L-AMNFL---  257
X. laevis        218 ------------------------------------------DNEHDI-SPVFFQRQ---NLLLWGIGCSLLD---YGSTPLKI-----Q-ALHFL---  258
C. elegans       321 -------------------------------------------------------------------------------------------DL---  322
D. rerio         216 ------------------------------------------D---VI-QGLFFPRE---VGIIWDSTCLMLS---NGSPKLKA-----V-SMVLL---  253
S. cerevisiae    115 ------------------------------------------D---NY-HRIWFMRR---KLGSWCKACVEF----YGKPAKFQ-----L-TAHFE---  151
D. melanogaster  188 FSGSLHQALLILKECDMDTKLASLNYCHNVLRSQSASSWITNPD---VGHYAQLTLEA---TAIMWSAVAKWLD---MGCMTRQELKRLNI-TTKLL---  274
E. nidulans      199 ---------------------------------------------------------RWDASQFKLSFPLPGGPLSVG-----LGRISLP---  226
S. pombe         154 ------------------------------------------------------DLYASFMR---YWKYYIT-----NV-SSIVLQIT  178
A. thaliana      217 ------------------------------------------KCL-TEGTLYVQGLIHTSFFKAACSLVC---CGGADVQM-----A-CFEFA---  257
```


HR 6
HR 7
HR 8
601 700
H. sapiens 253 -------------------T--------ELFQLGGLPAQPASTFFS---SFLELLKHLV-EMDTDQLKLYEEPL-SKLIKT----------LFPF--EAE 308
M. musculus 256 -------------------T--------ELFELGGLPAQPASTFFS---LFLELLQHLV-GMDADQLKLYEEPL-SKLLKT----------LFPF--EAE 311
G. gallus 258 -------------------I--------KLIHLGGPPEQLASVFFT---VFFGILHSAL-EMDTEELELCGEPF-LQLVRT----------LMPF--EAD 313
X. laevis 259 -------------------R--------QLIKLGGPPEQGAYFFFI---VFFGILTCIK-DMDLEEVSLYEMPL-LKLVKV----------LFPF--ESK 314
C. elegans 323 -------------------S--------TNHDIHEFPARLCMTASE---ICMRRIKHSIQSTHTDCLADYLDIL-ASLLKLARIQKLDIFQCFAT--EIE 389
D. rerio 254 -------------------T--------RIVTLGGFPEDHSQPFFS---AFLHVLDSLP-AFDESELGVFSREF-QQLSRC----------IFQH--EEG 309
S. cerevisiae 152 -------------------N--------T---------------MN---LYEQALTEVLL-GKTELLKFYD-TL-KGLYIL----------LYWF--TSE 191
D. melanogaster 275 --------------------------LEVLHMRARPAHHLGYLLL--NEILS-LPTAI-ELDDGLLETLSSYIQGQLEHS----------VVPL--EQ- 330
E. nidulans 227 ----------------------RNLFR-KGFKIGSTERAFN---LILALLESFV-VPESSPTSVPD---------------------------------- 265
S. pombe 179 NATFPYKMPSPNSQPLQSISPNYPTHREDKFDLLIINIEEACTFFFESAHFFA-Q-------------------CSYLKKS----------NFPSPPLFT 248
A. thaliana 258 --------------------------TLVGSILTFNILPHVALIQ---SIILLLSADE-GLPVYRNTIYDSTI-GR----------------F--LTA 306
701 800
H. sapiens 309 A--YRNI--------------------------------EPV---YLNML-LEKLCV---------MFEDGVLMRLKSDLL--KA---ALCHLLQYFLKFV 357
M. musculus 312 A--YRNI--------------------------------EPV---YLNVL-LEKLSV---------MFEDRVLMRLKSDLL--KA---ALCHLLQYFLTFV 360
G. gallus 314 S--YKNI--------------------------------EPV---YLNML-LEKLCA---------LFEGDMFMHLRSEKL--KA---AFCHILQYFLELV 362
X. laevis 315 S--YLNI--------------------------------EPV---YLNML-LEKLAA---------LFDGGILSNIQSAPL--KE---ALCYMVHYFLSIV 363
C. elegans 390 A--HHSL---NNPLWPRILRSLLSVQNPHC---QQFCKKFV 422
D. rerio 310 A--HSRF--------------------------------ERV---HLNML-MERLEK---------LVVIGALEHLKVKEV--KA---TLCEVFCFILGFV 358
S. cerevisiae 192 YSTFGNSIAFLDSSLGFTKFDFNFQRLIRIVLYVFDSCELAALEYAEIQ-LKYIS----------LVVDYVCNRTISTAL--DAPALVCCEQLKFVLTTM 278
D. melanogaster 331 --LVHLQQLML--SH---WHCHPTHLVPILA 354
E. nidulans 266 --KASIV-RQNLAWILNGYERLWNVLSSWLRSPSLDNLKSKV---ELCSH---FLTSI 314
S. pombe 249 A--WTWI--------------------------------KPC---FFNFVILLKRIS---------IGDSQLFLHLHSRIV--QT---LCCFSLNFIYHGL 298
A. thaliana 307 V--YSSC--------------------------------SDA---AVKLT-AESLV----------LVLSHALQRTKSEEL--KA---SLCSAYVRIVKSC 354
801 900
H. sapiens 358 -PAG-Y-----------ESALQVRK---------------VYVR----NICK-ALLDVLGIEVDA-------EYLLGPLYAAL----KMESMEIIEEIQC 413
M. musculus 361 -PAG-Y-----------ESALQVRK---------------VYVT----NICR-ALVDALGVQKHV-------GYLLGPFYAAL----KMESKEIIERIQC 416
G. gallus 363 -PPG-Y-----------ESALQVRK---------------VHVN----TICK-VFVSVLGTQEEQ-------EYLVSPLYTVL----KTQTEEIYQELHQ 418
X. laevis 364 -PPG-Y-----------ESAKEVRE---------------AHVR----CICR-AFVDVLGLQSKQ-------EYLVCPLHEAL----RIENLVFMQQQRM 419
C. elegans 423 -PSI-DKLWITLLNGESSRATLNSL---------------VIVT----ASCPFANNNDEGVKQII-------PFLLIP---------------------- 472
D. rerio 359 -PPG-Y-----------ECALQIRK---------------ERVA----AICK-ALIKTIGTKDQHESQAIQDQSLEGYLYAAL----KTDAIAAMQDVQA 421
S. cerevisiae 279 -HHF-LDNKYGLLDNDPTMAKGILR---------------LYSLCISNDFSK-CFVDHFPIDQWA-------DFSQSEHFPFT----------------- 336
D. melanogaster 355 -LMG-L-----------KQ-TEMRS---------------GVVQ----VLTQ-SLVEILKKEEVL-------SKDWQKLIAILRGFKQLEKLILSQSQHK 413
E. nidulans 315 RPYCFY-----------GSSFATEI---------------A--Y----DVPR-MLGDILS----L--------NYL--DHYPSL----QLNTSQFIDDLC- 363
S. pombe 299 -PIC-E-----------KS-KHILMSSINLTLGSLKKTYTVANT----AISL-FFLSLFVLPKTV-------AGLFYPFGVSL----LSDF-KVLEQLEP 367
A. thaliana 355 -PPCIW-----------KIHCLLEL---------------LHLP----EPCF-QLIECFK---AV-------LIVLGPGC------------VRVETTKC 400

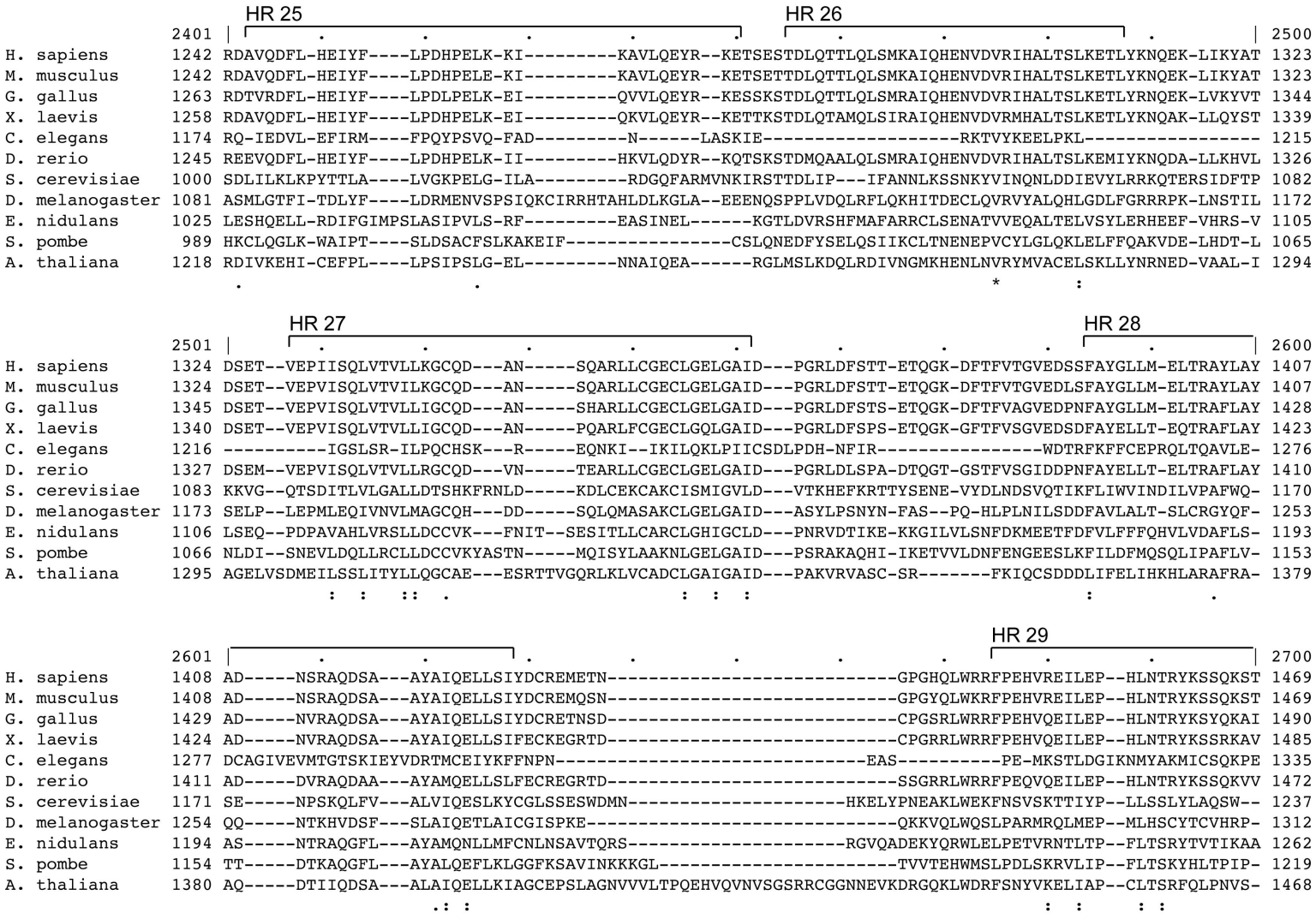
```
                                                                            HR 9
                    901 |        .         .         .         .         .         .         |.         .         .         | 1000
H. sapiens          414 -QTQ-----QEN-------------LSSNSDGISPKRRRLSSSLN-PSKRAP-KQTEEIKH-VDMNQKSILWSALKQKAESLQISLEY---------SGL  482
M. musculus         417 -QAQ-----QEN-------------LSGNNDEVSPKRRKLSSSLS-SYKKPS-RQPEEIIH-VDMDKKSILWNVLKQKAESLQISLEC---------GTL  485
G. gallus           419 SQLD--TSDTSG-------------WSSSSDEVPQKRRRLSLPAK-HPKK-S-PAPDLLSP-VNMKLKSTLWNAITKKAASLISVLEK---------ESL  490
X. laevis           420 -QPLSTDSEGGG-------------SSSSSDEVQEKRPRLSLTAK-PLRR---NTPSVPAP-VDMKTKSILWKAVSAKFSSILCKLEG---------DEV  491
C. elegans          473 ----------------------------------------------------------------------------------------------------  472
D. rerio            422 -SVS----DPAD-------------PRMYEEELPAKRPNLSLPTQ--LRS---RDKTQPMQ-VDMKSRSEVWAAVDCRLEELLTQMRN---------HTV  488
S. cerevisiae       337 -----------------------------------------------QLTNK---ALSIVY-FDLKRRSLPVEAL-------------------------  360
D. melanogaster     414 --IA----EHEG-------------H-------------IDSSVLAM----------LPLQCEIIKVADTNWNNLSMQLVELESKC--------------  457
E. nidulans         364 ------------LPDN-----CTEG------RVKCLRGLLSPLSD-TLKKATFESLE-------PHLQAAVRNAFTK-AMRVVSTVESATLLTNHAVGTM  431
S. pombe            368 ------------------------------------------------------------D-SDLKKAIILFKCRYQSSEIDQTTLRAFGEICT---GKL  403
A. thaliana         401 -GSH-----TSATSDRPVQGINAGKKRHIEDESTYKRKRQKVGDD-I-RRGVYFAPEFADE-TDGKDAASLREMLISTVESLKPPPAG---------PSL  482

                                                                                      HR 10
                   1001 |         .         .     |   .         .         .         .         |.         .         | 1100
H. sapiens          483 -------KN-PVIEMLEGIAVVLQLTALCTVHCSHQNMNCR---TFKDCQHKS---------KKKP---SV-VITWMSLDFYTKVLKSCRSLLESVQ-KL  557
M. musculus         486 -------KN-SVAEALEGITVVLQLTALCTVHCSHQDMDGH---NVKDHQHKY---------KKKP---PV-VVTWMSLDFYTKVLKSCRSLLESVQ-KL  560
G. gallus           491 -------TA-NVLQTAEGIAIILQLAALCVVHCSPPTNSSSTDHKVTHLCNTDTCKNTPCSLDSVE---NI-QFTWISSDFITRVVKACGNLLEAATKMV  578
X. laevis           492 -------TDEEMVSLLEGLNTTVRVAALNTVHIFTNNST-DTDQLVSDLSNTS-------GIQSVE---IVPHVFWLSPEDILKILKICRKVLDSAHQRA  573
C. elegans          473 ----------------------------------------------------------------------------------------------------  472
D. rerio            489 -------SQ--CVSAVQGLAVIFHLAALCSEHSLKGPQ------------RGE---------KSSESSVVA-QLIWLKPQMLAQVVESCRTVLANNYNEE  557
S. cerevisiae       361 -----------------------------------------------------------------------------------------------KYDNK  365
D. melanogaster     458 SADRRHIHL-EICSLLMQITFIRHFLKTQTQH--------------------------------------------------------------------  488
E. nidulans         432 --------P-DSLSCSQGLLAINQ--------------------------------------EVE---GD----EFM----RASKRYCRPQPRNGQDGS  472
S. pombe            404 E-------N-TLFSNSELNLFLLHYLSL------------------------------------------------------------------------  423
A. thaliana         483 -------------SQTESSIVALSMLTNAFCFCPWTDMT-----------HRL---------FNQM---YA-WIPWIA----GQVEETN-------PIMF  534

                                                    HR 11                                                      HR 12
                   1101 |         .         |.         |.         .         .         .         .|          .         . |-| 1200
H. sapiens          558 DL-EATIDKVVKIYDALIYMQVNS-----SFEDHILEDLCGMLSLPWIYSHSDDGCLKLTTFAANLLTLSCRISDSYSPQAQSRCVFLLTLFPR---RIF  648
M. musculus         561 EL-ELVIDSMVRICDALMYMQVKS-----SFKDHVLEELCGMLSLPWIYSYSDDNSLKMTTFATNLLPLSQRVWDSYSPQAQSKCVFLLTLFPR---RIF  651
G. gallus           579 NL-EQVIDRTVKIFDALLYAQIKI-----PLNHQILDNMCGLLSLPWICSHPDVSPFKTSAFGTDPLVLSQKTAQSFSPRTQSRCVFLLSLFPR---NIH  669
X. laevis           574 NI-NDILMKIIKIFDAILYIHAGN-----RLNDQTLKDLCSMISLPWLQNHSNHASFKVASFDPTLMTISERIGQHYSPEIQSQLVFLLCLFPK---MLC  664
C. elegans          473 ----------------------------------------------CFQDFRDRVGSKWDN---------------------------------------  487
D. rerio            558 HL-EHLVEGIVRILDAVLYLSNNLFAANSESDKSFHRNICALLSVPWVLEHSSQSVYQTASFPASLIALSQKLAPVYNAQTRAHCVFLLAPLSK---SVC  653
S. cerevisiae       366 FNI-------------------------------------------WVYQSEPDSSLKNVT---------SPFDDRYKQLEKLRL-LVLKKFNKTERGTL  412
D. melanogaster     489 ------------------------------------------------------------------------------------QLLAILQR---HLK  499
E. nidulans         473 NTTQQLISKSL-----------------------------------------------------------------------------------------  483
S. pombe            424 -------------------------------------------------------------------------------------------------DND  426
A. thaliana         535 D-ISIYLEGIHNLLLVGVDPQYEY-----TSKGNDLVAIQFLLKLPWTHYM---------------------LFKTPSSLVKSKCLSVGIWTKL---GLQ  604
```

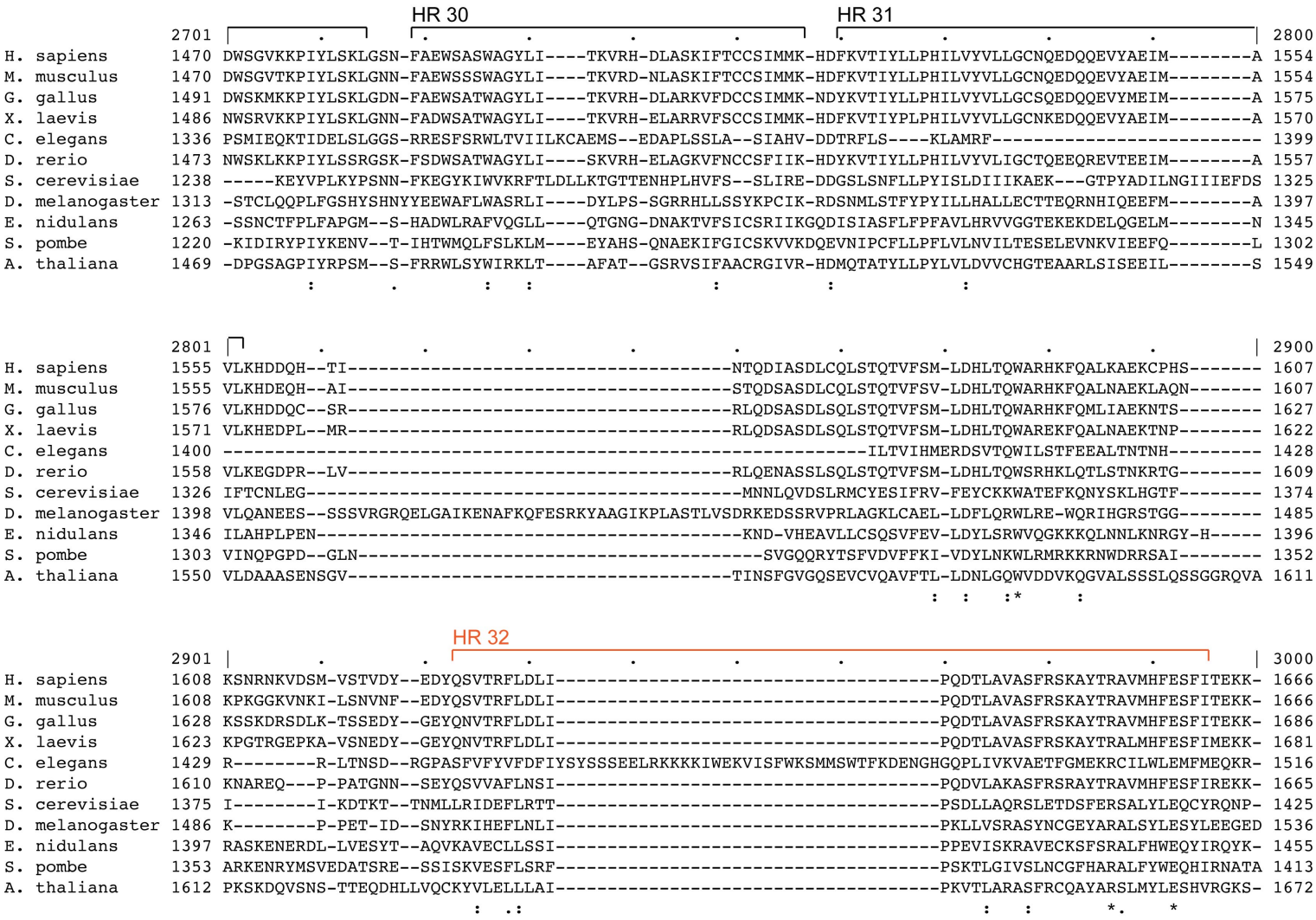

HR 13
HR 14


HR 15
HR 16
HR 17
H. sapiens
M. musculus
G. gallus
X. laevis
C. elegans
D. rerio
S. cerevisiae
D. melanogaster
E. nidulans
S. pombe
A. thaliana

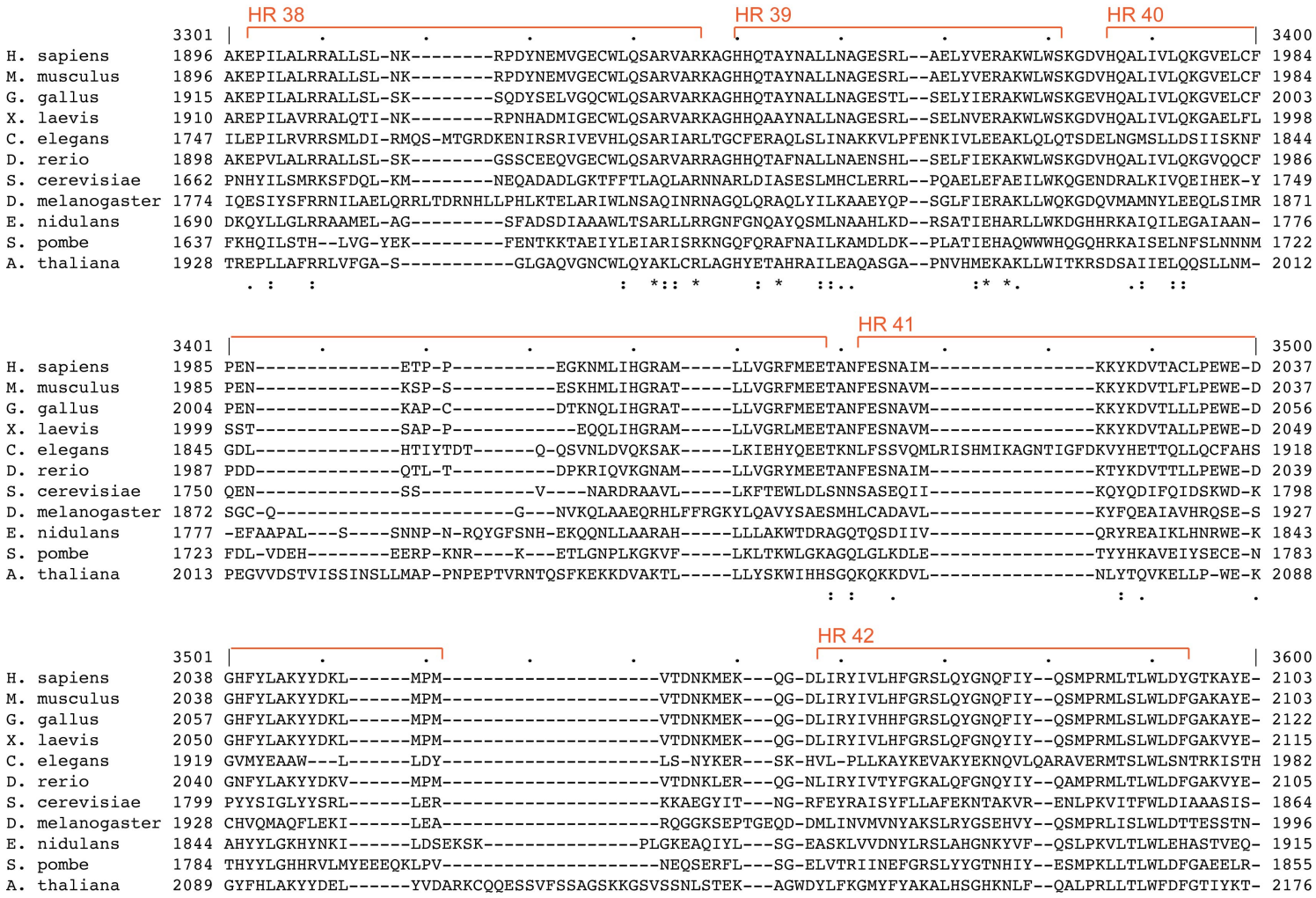

HR 18
HR 19
HR 20
H. sapiens
M. musculus
G. gallus
X. laevis
C. elegans
D. rerio
S. cerevisiae
D. melanogaster
E. nidulans
S. pombe
A. thaliana

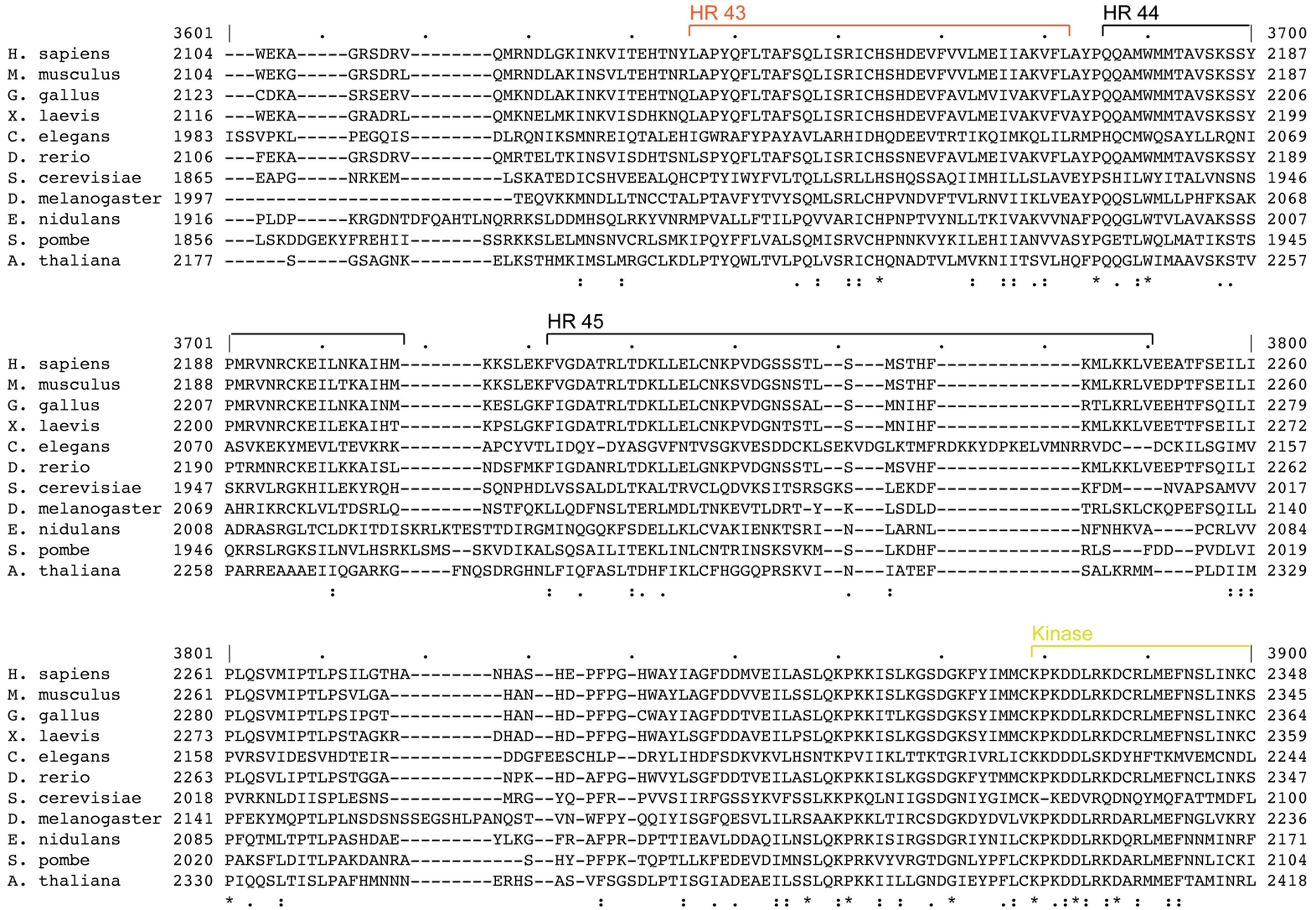
```
                  HR 21
                  2101 |                                                                                  | 2200
H. sapiens        1049 LERA-LHYLK----NETEIE------------------------------LGSLLRQD--FQGLHNELLLRIG-E-H--------YQQVFNGLSILASF- 1100
M. musculus       1049 LERA-LHYLK----NETEIE------------------------------LGSLLRQD--FQGLHNELLLRIG-E-H--------YQQVFNGLSILASF- 1100
G. gallus         1070 LERA-LHYLK----NETEIE------------------------------LGSLLRQD--YQGLHNELLLRIG-E-H--------YQQVFNGLSILASF- 1121
X. laevis         1065 LEKS-LHYLK----NETEIE------------------------------LGSLLRQD--YQGLHNELLLRLG-E-H--------YQQVFSGLSILATY- 1116
C. elegans         990 ------------------------------REFL---------NPRNPSMIVEQLRAT--YENIKKGEKFMIS-T-A---------------------- 1023
D. rerio          1051 LERA-FHYLK----NETEIE------------------------------LGSLLRQD--FQGLHNELLLRLG-E-H--------YQQVFNGLAILASF- 1102
S. cerevisiae      784 LDILETLFLNR-APTFDKGYITAYLPDYKTLAEITKLYKNSVTKDASDSENANMILCSLRFLITNFEKDKRHG-S-K--------YKNIN----------  862
D. melanogaster    887 ANSC-IELVV----SCTQSS------------------------------LQQLMNAD--VKQTVAELLIYFN-R-N--------PTFVMRSFQSLLQLS  939
E. nidulans        828 QKMI-MSTLSAVDPAFNGRT------------------------------LAELVRIE--PILIACDLLKGLG-D-S----EDEKGAKFQQALRILASL-  887
S. pombe           789 EESV-MLLLSLATSDFEKVD------------------------------LTSLLRSD--PISITVELLQLYQNDVP--------HEKIENALRKVAMI-  846
A. thaliana       1022 LLSV-LQLYHSQI----GSD------------------------------NQEIFAAA--LPALLDELVCFVD-I-ADTPETDRRLQRLPDAIKKI----  1078
                                                                            . .

                                 HR 22                                                    HR 23
                  2201 |                                                                                  | 2300
H. sapiens        1101 ASSDDPYQGPRDIIS--PELMADYLQPKLLGILAF-FNMQLLS--SSV--GI-EDKKMALNSLMSLMKLMGPKHVSSVR-VKMMTTLRTG--LRFKDDFP 1189
M. musculus       1101 ASSDDPYQGPRDITS--PELMADYLQPKLLGILAF-FNMQLLS--SSV--GI-EDKKMALTSLMSLMKLMGPKHVSSVR-VKMMTTLRTG--LRFKDDFP 1189
G. gallus         1122 ASHDDPYQGPREITS--SEQMADYLQPKLLGILAF-FNMQLLS--SSV--GI-EDKKMALNSLMSLMKLMGPKHISSVR-VKMMTTLRTA--LRYKDDFP 1210
X. laevis         1117 ASNDDPYQGPRNFAK--PEIMADYLQPKLLGILAF-FNMHLLS--SSI--GI-EDKKMALNSLVSLMKLMGPKHISSVR-VKMMTTLRTG--LRYKEEFP 1205
C. elegans        1024 ------------------------IETRNEGIEFLFIFRTYFT--NDQ--YF-LTREATAESFRILLQNMKQDFFDTNWYIILMTLRQIP--LNLDS-TQ 1091
D. rerio          1103 ASSDDPYQGPREITT--PGCMADYLQPKLLGILAF-FNMQLLS--SSA--GEKEKKKMALNSLMSLMKLMGPKHISSVR-VKMMTTLRTG--LRYKDDFP 1192
S. cerevisiae      863 -----------NWTDDQEQAFQKKLQDNILGIFQV-FSSDIHDVEGRT--TY-YEKLRVINGISFLIIYAPKKSIISAL-AQISICLQTG--LGLKE-VR  943
D. melanogaster    940 IGS---LEELSSQTA--NAEFANFIAERFLGVITY-FESCLSE--PSF--EK-PLKEETLYSLGQIMRFVGSQHVTQFR-FKIIAMLSFVHTLQE-PRLQ 1026
E. nidulans        888 VPRKS----AYASKK--SNLVGHFIEEHVLGIITQ-FAHAIND--FQIRQPL-VEKKRNIMAIGAMIKVAP-GHVSSAL-PQICACLRSA--LEI-KELC  972
S. pombe           847 VSQ---VVNDEDLSN--KELLYDFFNNHILGILAE-FSNILNDLKGKT--SI-NEKIKTIVGIEKMLSLCGG-AVKLGL-PQILSNLQSA--FQN-EHLR  932
A. thaliana       1079 ---------SKVLTN--AEDLPGFLQNHFVGLLNS-IDRKMLH--AD---DI-FLQKQALKRIKLLIEMMG-HYLSTYV-PKLMVLLMHA--IEK-DALQ 1155
                                            :  .  *:    :                   :      :  ::       .       :        :

                                   HR 24
                  2301 |                                                                                  | 2400
H. sapiens        1190 ELCCRAWDCFVRCLDH---ACLGSLLSHVIVALLP-----------------------------------LIH-------IQP--K-ETAAIFHYLIIEN- 1241
M. musculus       1190 ELCCRAWDCFVRCLDH---AYLGPLLSHVIVALLP-----------------------------------LIH-------MQP--K-ETAAIFHYLIIEN- 1241
G. gallus         1211 ELCCRAWDCFVRSLDH---SYLGSLLSHVIVALLP-----------------------------------LIH-------IQP--K-ETTAVFHFLIVEN- 1262
X. laevis         1206 GLCCSAWDLFVRCLDQ---AYLGPLLSHVIVALLP-----------------------------------LLH-------IQP--K-ETVAVFRYLIVEN- 1257
C. elegans        1092 ----SSWISFIEQINY---SILRSNIWRILTDISRIENNDILIEKVWNRLSYDMVIDKISTDAMIKRVFWLIP------LEVEK----RIDLKFEISKS- 1173
D. rerio          1193 ELCCRTWDCFVRCLDP---SYLGPLLSHVIVALLP-----------------------------------LIS-------IQP--K-ETAVIMHYLIVEN- 1244
S. cerevisiae      944 YEAFRCWHLLVRHLND---EELSTVIDSLIAFIL-----------------------------------QKWSE------FNGKLRNIVYSILDTLIKEK-  999
D. melanogaster   1027 RICLKIWHIFLHVVNV---QELGPSLGRIVATLQP-----------------------------------LLA-------DNESVK-QVNDLYEFIILRN- 1080
E. nidulans        973 NNAFRVWGVLVSSLHK---EEVEPLLDQTLAIVIK-----------------------------------HWVTFT-----------EDTRKFAYELVEHI 1024
S. pombe           933 FYAIKAWFSLILATKE---PEYSSIAGLSLVILPP-----------------------------------LFP------YLEPQEAELVIQIFDFISSDT-  988
A. thaliana       1156 SEGLLVLHFFTRKLADVSPSSIKYVISQIFAALIP-----------------------------------FLEKEKEGPHVYL--D-EVVKILEELVLKN- 1217
                                :                     .. :
```


HR 25
HR 26
HR 27
HR 28
HR 29
2401 2500
2501 2600
2601 2700
H. sapiens 1242 RDAVQDFL-HEIYF----LPDHPELK-KI---------KAVLQEYR--KETSESTDLQTTLQLSMKAIQHENVDVRIHALTSLKETLYKNQEK-LIKYAT 1323
M. musculus 1242 RDAVQDFL-HEIYF----LPDHPELE-KI---------KAVLQEYR--KETSETTDLQTTLQLSMKAIQHENVDVRIHALTSLKETLYKNQEK-LIKYAT 1323
G. gallus 1263 RDTVRDFL-HEIYF----LPDLPELK-EI---------QVVLQEYR--KESSKSTDLQTTLQLSMRAIQHENVDVRIHALTSLKETLYRNQEK-LVKYVT 1344
X. laevis 1258 RDAVQDFL-HEIYF----LPDHPELK-EI---------QKVLQEYR--KETTKSTDLQTAMQLSIRAIQHENVDVRMHALTSLKETLYKNQAK-LLQYST 1339
C. elegans 1174 RQ-IEDVL-EFIRM----FPQYPSVQ-FAD---------N------LASKIE-----------------RKTVYKEELPKL----------------- 1215
D. rerio 1245 REEVQDFL-HEIYF----LPDHPELK-II---------HKVLQDYR--KQTSKSTDMQAALQLSMRAIQHENVDVRIHALTSLKEMIYKNQDA-LLKHVL 1326
S. cerevisiae 1000 SDLILKLKPYTTLA----LVGKPELG-ILA---------RDGQFARMVNKIRSTTDLIP---IFANNLKSSNKYVINQNLDDIEVYLRRKQTERSIDFTP 1082
D. melanogaster 1081 ASMLGTFI-TDLYF----LDRMENVSPSIQKCIRRHTAHLDLKGLA--EEENQSPPLVDQLRFLQKHITDECLQVRVYALQHLGDLFGRRRPK-LNSTIL 1172
E. nidulans 1025 LESHQELL-RDIFGIMPSLASIPVLS-RF---------EASINEL------KGTLDVRSHFMAFARRCLSENATVVEQALTELVSYLERHEEF-VHRS-V 1105
S. pombe 989 HKCLQGLK-WAIPT----SLDSACFSLKAKEIF----------------CSLQNEDFYSELQSIIKCLTNENEPVCYLGLQKLELFFQAKVDE-LHDT-L 1065
A. thaliana 1218 RDIVKEHI-CEFPL----LPSIPSLG-EL---------NNAIQEA------RGLMSLKDQLRDIVNGMKHENLNVRYMVACELSKLLYNRNED-VAAL-I 1294
H. sapiens 1324 DSET--VEPIISQLVTVLLKGCQD---AN-----SQARLLCGECLGELGAID---PGRLDFSTT-ETQGK-DFTFVTGVEDSSFAYGLLM-ELTRAYLAY 1407
M. musculus 1324 DSET--VEPVISQLVTVILKGCQD---AN-----SQARLLCGECLGELGAID---PGRLDFSTT-ETQGK-DFTFVTGVEDLSFAYGLLM-ELTRAYLAY 1407
G. gallus 1345 DSET--VEPVISQLVTVLLIGCQD---AN-----SHARLLCGECLGELGAID---PGRLDFSTS-ETQGK-DFTFVAGVEDPNFAYGLLM-ELTRAFLAY 1428
X. laevis 1340 DSET--VEPVISQLVTVLLIGCQD---AN-----PQARLFCGECLGQLGAID---PGRLDFSPS-ETQGK-GFTFVSGVEDSDFAYELLT-EQTRAFLAY 1423
C. elegans 1216 ----------IGSLSR-ILPQCHSK---R-----EQNKI--IKILQKLPIICSDLPDH-NFIR----------------WDTRFKFFCEPRQLTQAVLE- 1276
D. rerio 1327 DSEM--VEPVISQLVTVLLRGCQD---VN-----TEARLLCGECLGELGAID---PGRLDLSPA-DTQGT-GSTFVSGIDDPNFAYELLT-ELTRAFLAY 1410
S. cerevisiae 1083 KKVG--QTSDITLVLGALLDTSHKFRNLD-----KDLCEKCAKCISMIGVLD---VTKHEFKRTTYSENE-VYDLNDSVQTIKFLIWVINDILVPAFWQ- 1170
D. melanogaster 1173 SELP--LEPMLEQIVNVLMAGCQH---DD-----SQLQMASAKCLGELGAID---ASYLPSNYN-FAS--PQ-HLPLNILSDDFAVLALT-SLCRGYQF- 1253
E. nidulans 1106 LSEQ--PDPAVAHLVRSLLDCCVK---FNIT--SESITLLCARCLGHIGCLD---PNRVDTIKE-KKGILVLSNFDKMEETFDFVLFFFQHVLVDAFLS- 1193
S. pombe 1066 NLDI--SNEVLDQLLRCLLDCCVKYASTN-----MQISYLAAKNLGELGAID---PSRAKAQHI-IKETVVLDNFENGEESLKFILDFMQSQLIPAFLV- 1153
A. thaliana 1295 AGELVSDMEILSSLITYLLQGCAE---ESRTTVGQRLKLVCADCLGAIGAID---PAKVRVASC-SR-------FKIQCSDDDLIFELIHKHLARAFRA- 1379
H. sapiens 1408 AD-----NSRAQDSA---AYAIQELLSIYDCREMETN----------GPGHQLWRRFPEHVREILEP--HLNTRYKSSQKST 1469
M. musculus 1408 AD-----NSRAQDSA---AYAIQELLSIYDCREMQSN----------GPGYQLWKRFPEHVREILEP--HLNTRYKSSQKST 1469
G. gallus 1429 AD-----NVRAQDSA---AYAIQELLSIYDCRETNSD----------CPGSRLWRRFPEHVQEILEP--HLNTRYKSYQKAI 1490
X. laevis 1424 AD-----NVRAQDSA---AYAIQELLSIFECKEGRTD----------CPGRRLWRRFPEHVQEILEP--HLNTRYKSSRKAV 1485
C. elegans 1277 DCAGIVEVMTGTSKIEYVDRTMCEIYKFFNPN----------EAS----------PE-MKSTLDGIKNMYAKMICSQKPE 1335
D. rerio 1411 AD-----DVRAQDAA---AYAMQELLSLFECREGRTD----------SSGRRLWRRFPEQVQEILEP--HLNTRYKSSQKVV 1472
S. cerevisiae 1171 SE-----NPSKQLFV---ALVIQESLKYCGLSSESWDMN----------HKELYPNEAKLWEKFNSVSKTTIYP--LLSSLYLAQSW-- 1237
D. melanogaster 1254 QQ-----NTKHVDSF---SLAIQETLAICGISPKE----------QKKVQLWQSLPARMRQLMEP--MLHSCYTCVHRP- 1312
E. nidulans 1194 AS-----NTRAQGFL---AYAMQNLLMFCNLNSAVTQRS----------RGVQADEKYQRWLELPETVRNTLTP--FLTSRYTVTIKAA 1262
S. pombe 1154 TT-----DTKAQGFL---AYALQEFLKLGGFKSAVINKKKGL----------TVVTEHWMSLPDLSKRVLIP--FLTSKYHLTPIP- 1219
A. thaliana 1380 AQ-----DTIIQDSA---ALAIQELLKIAGCEPSLAGNVVVLTPQEHVQVNVSGSRRCGGNNEVKDRGQKLWDRFSNYVKELIAP--CLTSRFQLPNVS- 1468


HR 30
HR 31
HR 32
2701
2800
H. sapiens 1470 DWSGVKKPIYLSKLGSN-FAEWSASWAGYLI----TKVRH-DLASKIFTCCSIMMK-HDFKVTIYLLPHILVYVLLGCNQEDQQEVYAEIM--------A 1554
M. musculus 1470 DWSGVTKPIYLSKLGNN-FAEWSSSWAGYLI----TKVRD-NLASKIFTCCSIMMK-HDFKVTIYLLPHILVYVLLGCNQEDQQEVYAEIM--------A 1554
G. gallus 1491 DWSKMKKPIYLSKLGDN-FAEWSATWAGYLI----TKVRH-DLARKVFDCCSIMMK-NDYKVTIYLLPHILVYVLLGCSQEDQQEVYMEIM--------A 1575
X. laevis 1486 NWSRVKKPIYLSKLGNN-FADWSATWAGYLI----TKVRH-ELARRVFSCCSIMMK-HDFKVTIYPLPHILVYVLLGCNKEDQQEVYAEIM--------A 1570
C. elegans 1336 PSMIEQKTIDELSLGGS-RRESFSRWLTVIILKCAEMS--EDAPLSSLA--SIAHV-DDTRFLS----KLAMRF-------------------------- 1399
D. rerio 1473 NWSKLKKPIYLSSRGSK-FSDWSATWAGYLI----SKVRH-ELAGKVFNCCSFIIK-HDYKVTIYLLPHILVYVLIGCTQEEQREVTEEIM--------A 1557
S. cerevisiae 1238 -----KEYVPLKYPSNN-FKEGYKIWVKRFTLDLLKTGTTENHPLHVFS--SLIRE-DDGSLSNFLLPYISLDIIIKAEK---GTPYADILNGIIIEFDS 1325
D. melanogaster 1313 -STCLQQPLFGSHYSHNYYEEWAFLWASRLI----DYLPS-SGRRHLLSSYKPCIK-RDSNMLSTFYPYILLHALLECTTEQRNHIQEEFM--------A 1397
E. nidulans 1263 -SSNCTFPLFAPGM--S-HADWLRAFVQGLL----QTGNG-DNAKTVFSICSRIIKGQDISIASFLFPFAVLHRVVGGTEKEKDELQGELM--------N 1345
S. pombe 1220 -KIDIRYPIYKENV--T-IHTWMQLFSLKLM----EYAHS-QNAEKIFGICSKVVKDQEVNIPCFLLPFLVLNVILTESELEVNKVIEEFQ--------L 1302
A. thaliana 1469 -DPGSAGPIYRPSM--S-FRRWLSYWIRKLT----AFAT--GSRVSIFAACRGIVR-HDMQTATYLLPYLVLDVVCHGTEAARLSISEEIL--------S 1549
2801
2900
H. sapiens 1555 VLKHDDQH--TI-----------------------------------NTQDIASDLCQLSTQTVFSM-LDHLTQWARHKFQALKAEKCPHS------- 1607
M. musculus 1555 VLKHDEQH--AI-----------------------------------STQDSASDLCQLSTQTVFSV-LDHLTQWARHKFQALNAEKLAQN------- 1607
G. gallus 1576 VLKHDDQC--SR-----------------------------------RLQDSASDLSQLSTQTVFSM-LDHLTQWARHKFQMLIAEKNTS-------- 1627
X. laevis 1571 VLKHEDPL--MR-----------------------------------RLQDSASDLSQLSTQTVFSM-LDHLTQWAREKFQALNAEKTNP-------- 1622
C. elegans 1400 ---ILTVIHMERDSVTQWILSTFEEALTNTNH---------- 1428
D. rerio 1558 VLKEGDPR--LV-----------------------------------RLQENASSLSQLSTQTVFSM-LDHLTQWSRHKLQTLSTNKRTG-------- 1609
S. cerevisiae 1326 IFTCNLEG-------------------------------------MNNLQVDSLRMCYESIFRV-FEYCKKWATEFKQNYSKLHGTF-------- 1374
D. melanogaster 1398 VLQANEES--SSSVRGRQELGAIKENAFKQFESRKYAAGIKPLASTLVSDRKEDSSRVPRLAGKLCAEL-LDFLQRWLRE-WQRIHGRSTGG-------- 1485
E. nidulans 1346 ILAHPLPEN-------------------------------------KND-VHEAVLLCSQSVFEV-LDYLSRWVQGKKKQLNNLKNRGY-H----- 1396
S. pombe 1303 VINQPGPD--GLN-------------------------------------SVGQQRYTSFVDVFFKI-VDYLNKWLRMRKKRNWDRRSAI-------- 1352
A. thaliana 1550 VLDAAASENSGV-------------------------------------TINSFGVGQSEVCVQAVFTL-LDNLGQWVDDVKQGVALSSSLQSSGGRQVA 1611
2901
3000
H. sapiens 1608 KSNRNKVDSM-VSTVDY--EDYQSVTRFLDLI-------------------------------------PQDTLAVASFRSKAYTRAVMHFESFITEKK- 1666
M. musculus 1608 KPKGGKVNKI-LSNVNF--EDYQSVTRFLDLI-------------------------------------PQDTLAVASFRSKAYTRAVMHFESFITEKK- 1666
G. gallus 1628 KSSKDRSDLK-TSSEDY--GEYQNVTRFLDLI-------------------------------------PQDTLAVASFRSKAYTRAVMHFESFITEKK- 1686
X. laevis 1623 KPGTRGEPKA-VSNEDY--GEYQNVTRFLDLI-------------------------------------PQDTLAVASFRSKAYTRALMHFESFIMEKK- 1681
C. elegans 1429 R--------R-LTNSD--RGPASFVFYVFDFIYSYSSSEELRKKKKIWEKVISFWKSMMSWTFKDENGHGQPLIVKVAETFGMEKRCILWLEMFMEQKR- 1516
D. rerio 1610 KNAREQ---P-PATGNN--SEYQSVVAFLNSI-------------------------------------PQDVLAKASFRSRAYTRAVMHFESFIREKK- 1665
S. cerevisiae 1375 I--------I-KDTKT--TNMLLRIDEFLRTT-------------------------------------PSDLLAQRSLETDSFERSALYLEQCYRQNP- 1425
D. melanogaster 1486 K--------P-PET-ID--SNYRKIHEFLNLI-------------------------------------PKLLVSRASYNCGEYARALSYLESYLEEGED 1536
E. nidulans 1397 RASKENERDL-LVESYT--AQVKAVECLLSSI-------------------------------------PPEVISKRAVECKSFSRALFHWEQYIRQYK- 1455
S. pombe 1353 ARKENRYMSVEDATSRE--SSISKVESFLSRF-------------------------------------PSKTLGIVSLNCGFHARALFYWEQHIRNATA 1413
A. thaliana 1612 PKSKDQVSNS-TTEQDHLLVQCKYVLELLLAI-------------------------------------PKVTLARASFRCQAYARSLMYLESHVRGKS- 1672

```
                        HR 33                                                       HR 34
                   3001 |                                                                                             | 3100
H. sapiens         1667 --QNI----------Q-EHLGFLQKLYAAMHEPDGVAGVSAIRKAEPS---LKEQILEHESLGLLRDATACYDRA-I----QLEPDQIIHYHGVVKSMLG 1745
M. musculus        1667 --QNI----------Q-KHLGFLQKLYAAMHEPDGVAGVSAIRKAEPS---LKEQILEHESIGLLRDATACYDRA-I----QLEPDQIIHYHGVVKSMLG 1745
G. gallus          1687 --QNI----------Q-EHLGFLQKLYAAMHEPDGVAGVSAIRKAEPS---LKEQILEHESIGLLRDATACYDRA-I----QLEPDQIIHYHGVVKSMLG 1765
X. laevis          1682 --QEI----------Q-EHLGFLQKLYAAMHEPDGVAGVSAIRKKEAS---LKEQILEHESIGLLRDATACYDRA-I----QLKPEEIIHYHGVVKSMLG 1760
C. elegans         1517 --KTST---EIESETEAAYYFTLMNLYGRIHELNGVRGAYARL-SRIQIDHVYGKISMREAFGDFNSA-ACFARMTG----KGKPFNS-T---------- 1594
D. rerio           1666 --QNI----------Q-DHLSFLQTLYAAMHEPDGVRGVNALRKEEPS---LREQILEHESIGLLRESTACYDRA-I----QLEPDQLAHYHGVMNSMLG 1744
S. cerevisiae      1426 --HDKN---QN----G-QLLKNLQITYEEIGDIDSLDGVLRTF-ATGN---LVSKIEELQYSENWKLAQDCFNVL-G----KFSDDPK-TTTRMLKSMYD 1505
D. melanogaster    1537 KSQRL----------L-EQFTFLVEVYGSLRDPDSVEGAVQVRSYDMS---VTQDILVNRLVERQQDMITSYEQL-LSSTDQMQPDHVR---AMIDAYLR 1618
E. nidulans        1456 --TRPK--TQQYTSLE-SLYEHLQGIYSQIDEPDGIEGISTHL-HVLN---IDQQVLEHRKAGRWATAQSWYELQ-L----EREPDNLDAQWNLFTCLKE 1541
S. pombe           1414 PYAAL----------E-SDYRVLQEIYAGIDDPDEIEAVSL-NFHDYS---FDQQLLLHENSGTWDSALSCYEII-I----QKDPENKKAKIGLLNSMLQ 1493
A. thaliana        1673 --GSLNPAAEKTGIFENADVSSLMGIYSCLDEPDGLSGFASLS-KSLN---LQDQLLINKKSGNWADVFTACEQA-L----QMEPTSVQRHSDVLNCLLN 1761
                                              *   *  : : : : .         .   .  .:   .                    : .

                                    HR 35                                                          HR 36
                   3101 |                                                                                             | 3200
H. sapiens         1746 LGQLSTVITQVNGVH------ANRSEWTDEL--NTYRVEAAWKLSQWDLVENYLA-------AD-G---------KSTTWSVRLGQLLLSAKKR--DITA 1818
M. musculus        1746 LGQLSTVITQVNGVH------ANRSEWTDEL--NTYRVEAAWKLSQWDLVENYLA-------AD-G---------KSTTWSVRLGQLLLSAKKR--DTTT 1818
G. gallus          1766 LGQLSTVITQVNGVL------ANRSEWISEL--NTYRVEAAWKLSQWDLLENYLA-------SD-V---------KSSAWSVRLGHLLLSAKKK--NEVA 1838
X. laevis          1761 LGQLSTVITQVNGIL------NSRSEWTAEL--NTYRVEAAWKLSQWDLVEEYLS-------AD-R---------KSTTWSIRLGQLLLSAKKG--ERDM 1833
C. elegans         1595 -EAIQKLIDELNCLEYS---QIERNEQEDYL--NSLKT-----LSQWVNIDNDIG--PSPHIFS-R---------NIEYWATEST-ILKMIRND--ERDE 1668
D. rerio           1745 LGQLSTVITQVNGVL------ASRPRWKSDL--NAYRVEAAWKLSKWDLVEDYLA-------SD-C---------KSSTWGVRLGQMLLAAKKQ--DSEA 1817
S. cerevisiae      1506 HQLYSQIISN-SSFHSSDGKISL-SPDVKEW--YSIGLEAANLEGNVQTLKNWVEQIESLRNID-D---------REVLLQYNIAKALIAISNE--DPLR 1589
D. melanogaster    1619 DT-PKTAQLIADGLW------QRLSDRYSDQCFAECKSELLWRLGSYDEMEELQS----------------------NWPAQCSQGCLKLRRPLTTRIE 1688
E. nidulans        1542 SGQQDAILTRFEIL-------QNTSS-VPRF--LPFAVEASWMTGKWEKMHNYLE-------LC--------PQQATADFNIGIGLALDAFRRG--EPQQ 1614
S. pombe           1494 SGHYESLVLSLDSFI------INDNHEYSKM--LNLGIEASWRSLSIDSLKKCLS-------KS-----------NLESFEAKLGSIFYQYLRK--DS-- 1563
A. thaliana        1762 MCHHQTMVTHVDGLI------SRVPEYKKTW--CTQGVQAAWRLGKWDLMDEYLD-------GADAEGLLFSSSDSNASFDRDVAKILHAMMKK--DQYS 1844
                            .      . .                             .  :..                                     .

                                                      HR 37
                   3201 |                                                                                             | 3300
H. sapiens         1819 ----FYDSLKLVRAEQIVPLS-AA-SFERGSYQRGYEYIVRLHMLCELEHSIKPLFQHSPG-D------S--SQE------DSLNWVARLEMTQN--SYR 1895
M. musculus        1819 ----FYDTLKLVRAEQIVPLS-AA-SFERGSYQRGYEFIVRLHMLCELEHSLKPLFRKSPG-D------S--CNE------DSLNWGARLEMTQN--SYR 1895
G. gallus          1839 ----FYETLKVVRAEQIVPLS-AA-SFERGSYQRGYEHIIRLHMLCELEHSIGPMFQQLDG-E---------HGR------DSLNWCARIEMTQN--SYR 1914
X. laevis          1834 ----FYETLKVVRAEQIVPLS-AA-SFERGSYQRGYEYIVRLHMLCELEHSVKMFLQKPSV-E---------PAV------DSLNLPARLEMTQN--SYR 1909
C. elegans         1669 I---VNNAIENAKSKVIERLS-ECAIGGSCSYEIATPFLVELQKLNEIVELKNVS-NDELS-A------F--NSD------FWKNIQKRTDDSEQ--KIS 1746
D. rerio           1818 ----FYEKLKIARKEQVVPLS-AA-SFECGTYQRGYEYIVRLHMLCELEHVFTELQKESSEAG------RSKKPE------PKLNWDAHLLMTQN--SFR 1897
S. cerevisiae      1590 TQKYIHNSFRLIGTN-------FITSSKETTLLKKQNLLMKLHSLYDLSFLSSA--KDKF---------E--YKS------NTTILDYRMERIGA--DFV 1661
D. melanogaster    1689 ----FDSLLDGMRESVLEELRSCS-AVQQHSYANAYDAVLKLHLVHELHCSQ-ELVEKLEQ-D------RDEDNQEKLMKNYFDDWQYRLQIVQP--QVR 1773
E. nidulans        1615 ----FREIVDKLRLSVARSLT-AN-S--VTSLQSCHDSMLKLHALTEIESVVLAGGADGS---------QGSRSC------LRDALDRRLDVLGG--YIS 1689
S. pombe           1564 ----FAELTERLQPLYVDAATAIA-NTGAHSAYDCYDILSKLHAINDFSRIA-------ET-D------GI--VSD------NLDIVLRRRLSQVAPYGK 1636
A. thaliana        1845 ----VAEGIAISKQALIAPLA-AA-G--MDSYTRAYPFVVKLHLLRELEDFQAVLNGDSYL-EKSFSTSDQVFSK------AVDNWENRLRFTQS--SLW 1927
                            . .                       :       :  .*: :  ::                              :
```


HR 38
HR 39
HR 40
HR 41
HR 42
H. sapiens
M. musculus
G. gallus
X. laevis
C. elegans
D. rerio
S. cerevisiae
D. melanogaster
E. nidulans
S. pombe
A. thaliana

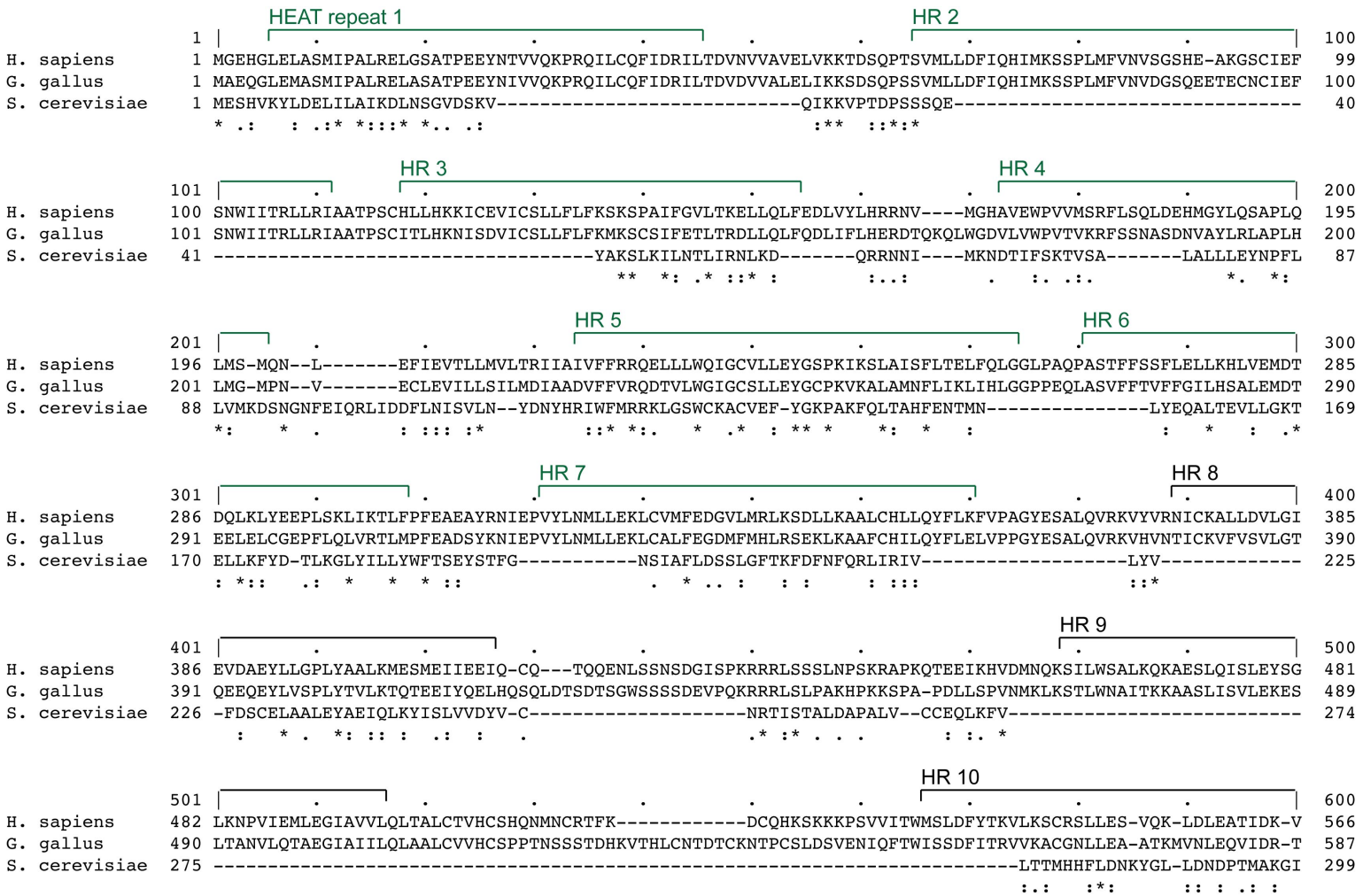

HR 43
HR 44
HR 45
Kinase
H. sapiens
M. musculus
G. gallus
X. laevis
C. elegans
D. rerio
S. cerevisiae
D. melanogaster
E. nidulans
S. pombe
A. thaliana

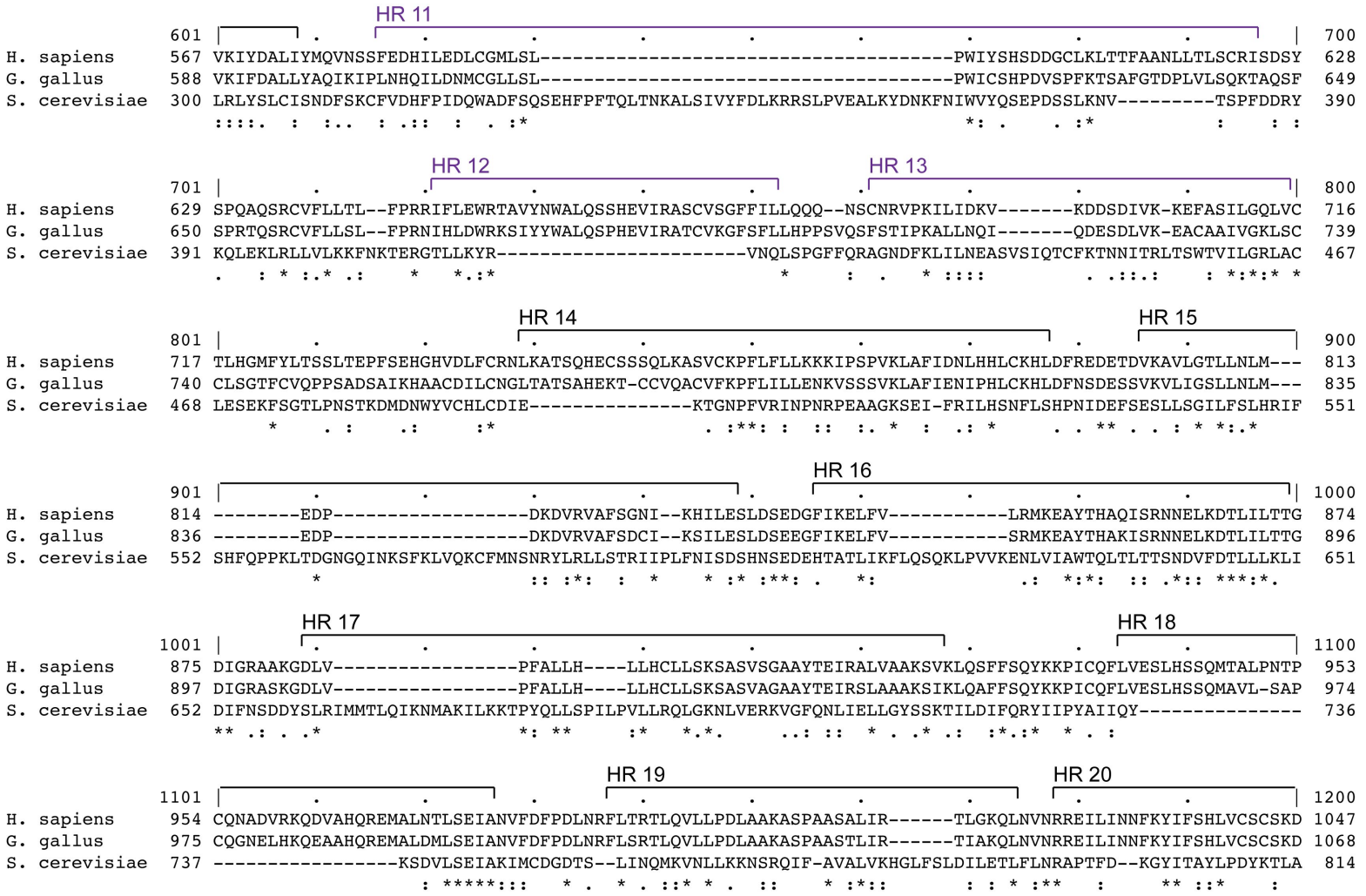

FATC

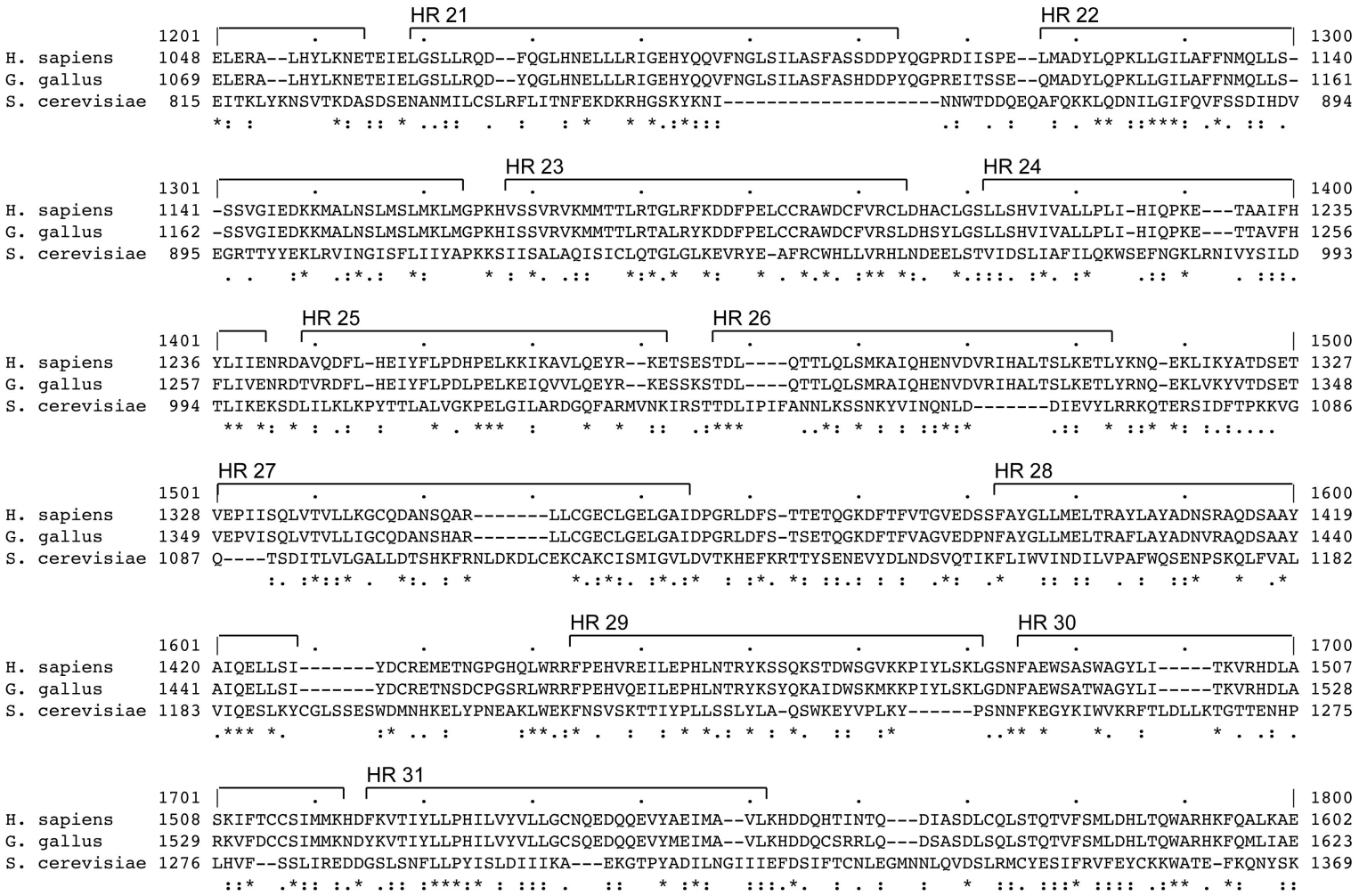
8.5 Alignment of ATR, Atr and Mec1

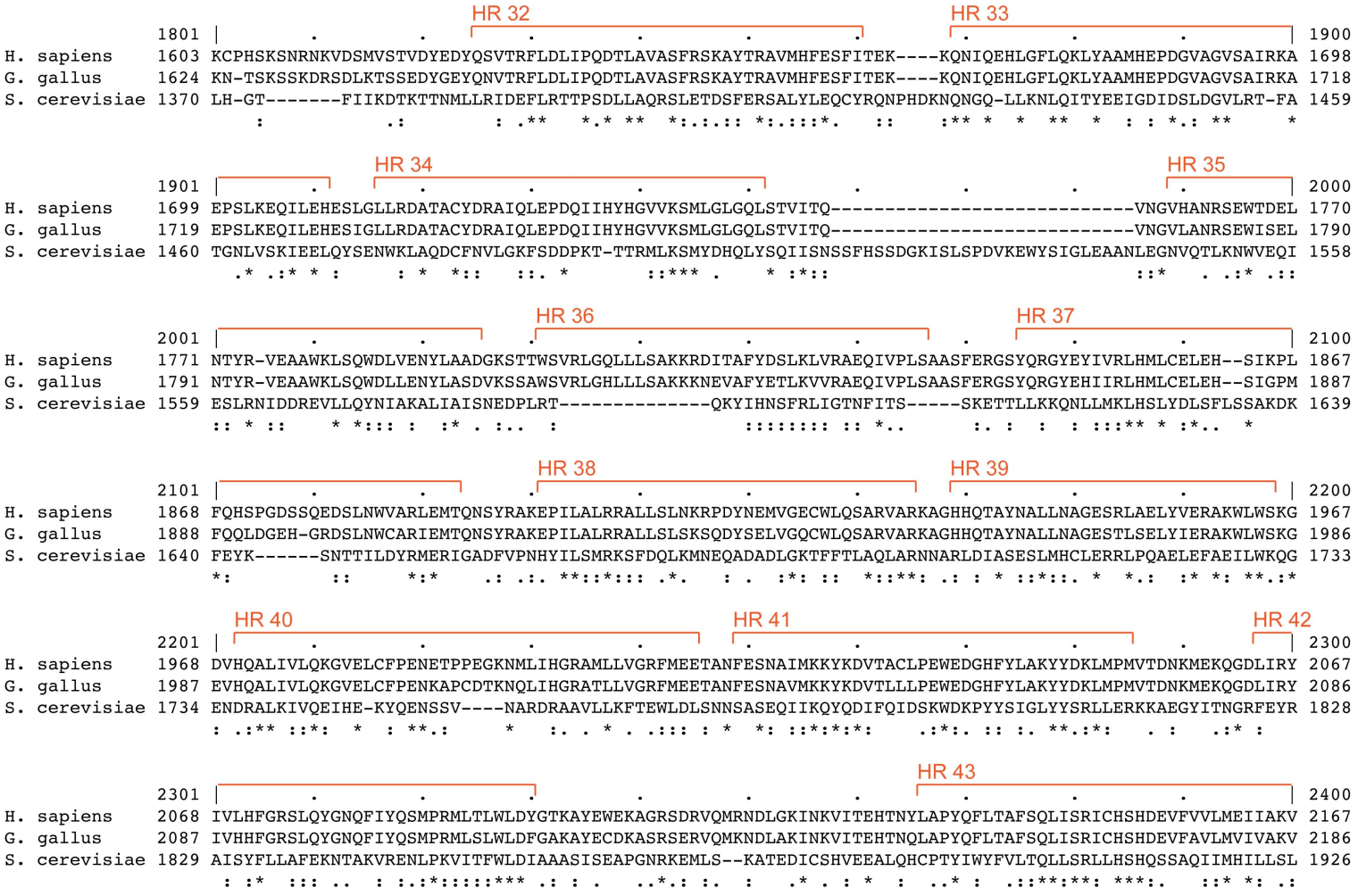
Alignment of human ATR, chicken Atr and yeast Mec1 proteins.

T-coffee online software was used to align these protein sequences and a script written by Dr. Andrew Flaus was used to obtain the desired output in landscape. Details to interpret the alignment results can be found in section 8.4.


HEAT repeat 1
HR 2
HR 3
HR 4
HR 5
HR 6
HR 7
HR 8
HR 9
HR 10
H. sapiens
G. gallus
S. cerevisiae


HR 11
HR 12
HR 13
HR 14
HR 15
HR 16
HR 17
HR 18
HR 19
HR 20
601 700
H. sapiens 567 VKIYDALIYMQVNSSFEDHILEDLCGMLSL----------PWIYSHSDDGCLKLTTFAANLLTLSCRISDSY 628
G. gallus 588 VKIFDALLYAQIKIPLNHQILDNMCGLLSL----------PWICSHPDVSPFKTSAFGTDPLVLSQKTAQSF 649
S. cerevisiae 300 LRLYSLCISNDFSKCFVDHFPIDQWADFSQSEHFPFTQLTNKALSIVYFDLKRRSLPVEALKYDNKFNIWVYQSEPDSSLKNV---------TSPFDDRY 390
701 800
H. sapiens 629 SPQAQSRCVFLLTL--FPRRIFLEWRTAVYNWALQSSHEVIRASCVSGFFILLQQQ--NSCNRVPKILIDKV-------KDDSDIVK-KEFASILGQLVC 716
G. gallus 650 SPRTQSRCVFLLSL--FPRNIHLDWRKSIYYWALQSPHEVIRATCVKGFSFLLHPPSVQSFSTIPKALLNQI-------QDESDLVK-EACAAIVGKLSC 739
S. cerevisiae 391 KQLEKLRLLVLKKFNKTERGTLLKYR----------VNQLSPGFFQRAGNDFKLILNEASVSIQTCFKTNNITRLTSWTVILGRLAC 467
801 900
H. sapiens 717 TLHGMFYLTSSLTEPFSEHGHVDLFCRNLKATSQHECSSSQLKASVCKPFLFLLKKKIPSPVKLAFIDNLHHLCKHLDFREDETDVKAVLGTLLNLM--- 813
G. gallus 740 CLSGTFCVQPPSADSAIKHAACDILCNGLTATSAHEKT-CCVQACVFKPFLILLENKVSSSVKLAFIENIPHLCKHLDFNSDESSVKVLIGSLLNLM--- 835
S. cerevisiae 468 LESEKFSGTLPNSTKDMDNWYVCHLCDIE----------KTGNPFVRINPNRPEAAGKSEI-FRILHSNFLSHPNIDEFSESLLSGILFSLHRIF 551
901 1000
H. sapiens 814 --------EDP----------DKDVRVAFSGNI--KHILESLDSEDGFIKELFV----------LRMKEAYTHAQISRNNELKDTLILTTG 874
G. gallus 836 --------EDP----------DKDVRVAFSDCI--KSILESLDSEEGFIKELFV----------SRMKEAYTHAKISRNNELKDTLILTTG 896
S. cerevisiae 552 SHFQPPKLTDGNGQINKSFKLVQKCFMNSNRYLRLLSTRIIPLFNISDSHNSEDEHTATLIKFLQSQKLPVVKENLVIAWTQLTLTTSNDVFDTLLLKLI 651
1001 1100
H. sapiens 875 DIGRAAKGDLV----------PFALLH----LLHCLLSKSASVSGAAYTEIRALVAAKSVKLQSFFSQYKKPICQFLVESLHSSQMTALPNTP 953
G. gallus 897 DIGRASKGDLV----------PFALLH----LLHCLLSKSASVAGAAYTEIRSLAAAKSIKLQAFFSQYKKPICQFLVESLHSSQMAVL-SAP 974
S. cerevisiae 652 DIFNSDDYSLRIMMTLQIKNMAKILKKTPYQLLSPILPVLLRQLGKNLVERKVGFQNLIELLGYSSKTILDIFQRYIIPYAIIQY---------- 736
1101 1200
H. sapiens 954 CQNADVRKQDVAHQREMALNTLSEIANVFDFPDLNRFLTRTLQVLLPDLAAKASPAASALIR------TLGKQLNVNRREILINNFKYIFSHLVCSCSKD 1047
G. gallus 975 CQGNELHKQEAAHQREMALDMLSEIANVFDFPDLNRFLSRTLQVLLPDLAAKASPAASTLIR------TIAKQLNVNRREILINNFKYIFSHLVCSCSKD 1068
S. cerevisiae 737 ----------KSDVLSEIAKIMCDGDTS--LINQMKVNLLKKNSRQIF-AVALVKHGLFSLDILETLFLNRAPTFD--KGYITAYLPDYKTLA 814

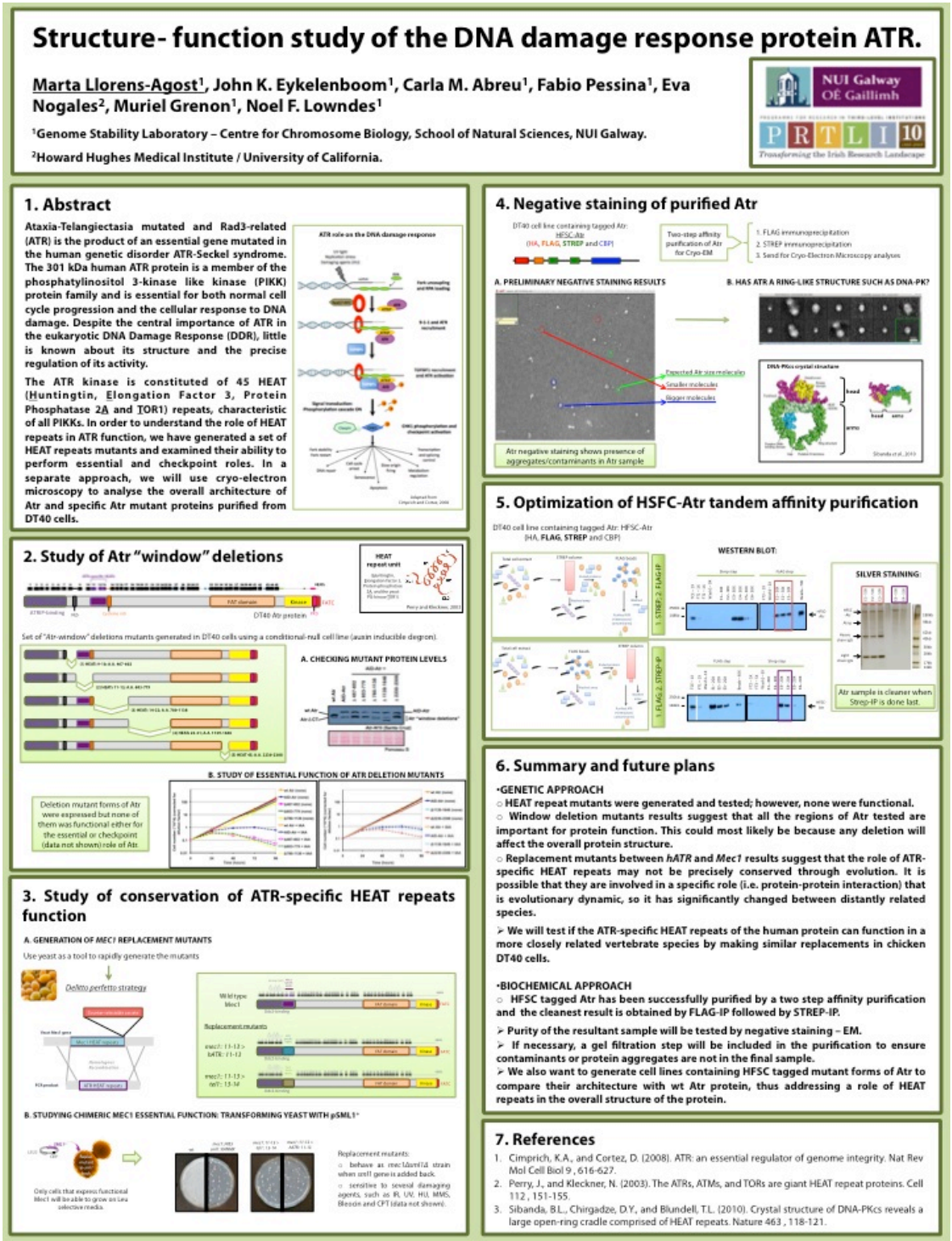

HR 21
HR 22
HR 23
HR 24
HR 25
HR 26
HR 27
HR 28
HR 29
HR 30
HR 31
1201 1300
H. sapiens 1048 ELERA--LHYLKNETEIELGSLLRQD--FQGLHNELLLRIGEHYQQVFNGLSILASFASSDDPYQGPRDIISPE--LMADYLQPKLLGILAFFNMQLLS- 1140
G. gallus 1069 ELERA--LHYLKNETEIELGSLLRQD--YQGLHNELLLRIGEHYQQVFNGLSILASFASHDDPYQGPREITSSE--QMADYLQPKLLGILAFFNMQLLS- 1161
S. cerevisiae 815 EITKLYKNSVTKDASDSENANMILCSLRFLITNFEKDKRHGSKYKNI----------NNWTDDQEQAFQKKLQDNILGIFQVFSSDIHDV 894
1301 1400
H. sapiens 1141 -SSVGIEDKKMALNSLMSLMKLMGPKHVSSVRVKMMTTLRTGLRFKDDFPELCCRAWDCFVRCLDHACLGSLLSHVIVALLPLI-HIQPKE---TAAIFH 1235
G. gallus 1162 -SSVGIEDKKMALNSLMSLMKLMGPKHISSVRVKMMTTLRTALRYKDDFPELCCRAWDCFVRSLDHSYLGSLLSHVIVALLPLI-HIQPKE---TTAVFH 1256
S. cerevisiae 895 EGRTTYYEKLRVINGISFLIIYAPKKSIISALAQISICLQTGLGLKEVRYE-AFRCWHLLVRHLNDEELSTVIDSLIAFILQKWSEFNGKLRNIVYSILD 993
1401 1500
H. sapiens 1236 YLIIENRDAVQDFL-HEIYFLPDHPELKKIKAVLQEYR--KETSESTDL----QTTLQLSMKAIQHENVDVRIHALTSLKETLYKNQ-EKLIKYATDSET 1327
G. gallus 1257 FLIVENRDTVRDFL-HEIYFLPDLPELKEIQVVLQEYR--KESSKSTDL----QTTLQLSMRAIQHENVDVRIHALTSLKETLYRNQ-EKLVKYVTDSET 1348
S. cerevisiae 994 TLIKEKSDLILKLKPYTTLALVGKPELGILARDGQFARMVNKIRSTTDLIPIFANNLKSSNKYVINQNLD-------DIEVYLRRKQTERSIDFTPKKVG 1086
1501 1600
H. sapiens 1328 VEPIISQLVTVLLKGCQDANSQAR-------LLCGECLGELGAIDPGRLDFS-TTETQGKDFTFVTGVEDSSFAYGLLMELTRAYLAYADNSRAQDSAAY 1419
G. gallus 1349 VEPVISQLVTVLLIGCQDANSHAR-------LLCGECLGELGAIDPGRLDFS-TSETQGKDFTFVAGVEDPNFAYGLLMELTRAFLAYADNVRAQDSAAY 1440
S. cerevisiae 1087 Q----TSDITLVLGALLDTSHKFRNLDKDLCEKCAKCISMIGVLDVTKHEFKRTTYSENEVYDLNDSVQTIKFLIWVINDILVPAFWQSENPSKQLFVAL 1182
1601 1700
H. sapiens 1420 AIQELLSI-------YDCREMETNGPGHQLWRRFPEHVREILEPHLNTRYKSSQKSTDWSGVKKPIYLSKLGSNFAEWSASWAGYLI-----TKVRHDLA 1507
G. gallus 1441 AIQELLSI-------YDCRETNSDCPGSRLWRRFPEHVQEILEPHLNTRYKSYQKAIDWSKMKKPIYLSKLGDNFAEWSATWAGYLI-----TKVRHDLA 1528
S. cerevisiae 1183 VIQESLKYCGLSSESWDMNHKELYPNEAKLWEKFNSVSKTTIYPLLSSLYLA-QSWKEYVPLKY------PSNNFKEGYKIWVKRFTLDLLKTGTTENHP 1275
1701 1800
H. sapiens 1508 SKIFTCCSIMMKHDFKVTIYLLPHILVYVLLGCNQEDQQEVYAEIMA--VLKHDDQHTINTQ---DIASDLCQLSTQTVFSMLDHLTQWARHKFQALKAE 1602
G. gallus 1529 RKVFDCCSIMMKNDYKVTIYLLPHILVYVLLGCSQEDQQEVYMEIMA--VLKHDDQCSRRLQ---DSASDLSQLSTQTVFSMLDHLTQWARHKFQMLIAE 1623
S. cerevisiae 1276 LHVF--SSLIREDDGSLSNFLLPYISLDIIIKA---EKGTPYADILNGIIIEFDSIFTCNLEGMNNLQVDSLRMCYESIFRVFEYCKKWATE-FKQNYSK 1369


HR 32
HR 33
HR 34
HR 35
HR 36
HR 37
HR 38
HR 39
HR 40
HR 41
HR 42
HR 43
1801 | . . . . . | 1900
H. sapiens 1603 KCPHSKSNRNKVDSMVSTVDYEDYQSVTRFLDLIPQDTLAVASFRSKAYTRAVMHFESFITEK----KQNIQEHLGFLQKLYAAMHEPDGVAGVSAIRKA 1698
G. gallus 1624 KN-TSKSSKDRSDLKTSSEDYGEYQNVTRFLDLIPQDTLAVASFRSKAYTRAVMHFESFITEK----KQNIQEHLGFLQKLYAAMHEPDGVAGVSAIRKA 1718
S. cerevisiae 1370 LH-GT-------FIIKDTKTTNMLLRIDEFLRTTPSDLLAQRSLETDSFERSALYLEQCYRQNPHDKNQNGQ-LLKNLQITYEEIGDIDSLDGVLRT-FA 1459
1901 | . . . . . | 2000
H. sapiens 1699 EPSLKEQILEHESLGLLRDATACYDRAIQLEPDQIIHYHGVVKSMLGLGQLSTVITQ----------VNGVHANRSEWTDEL 1770
G. gallus 1719 EPSLKEQILEHESIGLLRDATACYDRAIQLEPDQIIHYHGVVKSMLGLGQLSTVITQ----------VNGVLANRSEWISEL 1790
S. cerevisiae 1460 TGNLVSKIEELQYSENWKLAQDCFNVLGKFSDDPKT-TTRMLKSMYDHQLYSQIISNSSFHSSDGKISLSPDVKEWYSIGLEAANLEGNVQTLKNWVEQI 1558
2001 | . . . . . | 2100
H. sapiens 1771 NTYR-VEAAWKLSQWDLVENYLAADGKSTTWSVRLGQLLLSAKKRDITAFYDSLKLVRAEQIVPLSAASFERGSYQRGYEYIVRLHMLCELEH--SIKPL 1867
G. gallus 1791 NTYR-VEAAWKLSQWDLLENYLASDVKSSAWSVRLGHLLLSAKKKNEVAFYETLKVVRAEQIVPLSAASFERGSYQRGYEHIIRLHMLCELEH--SIGPM 1887
S. cerevisiae 1559 ESLRNIDDREVLLQYNIAKALIAISNEDPLRT----------QKYIHNSFRLIGTNFITS-----SKETTLLKKQNLLMKLHSLYDLSFLSSAKDK 1639
2101 | . . . . . | 2200
H. sapiens 1868 FQHSPGDSSQEDSLNWVARLEMTQNSYRAKEPILALRRALLSLNKRPDYNEMVGECWLQSARVARKAGHHQTAYNALLNAGESRLAELYVERAKWLWSKG 1967
G. gallus 1888 FQQLDGEH-GRDSLNWCARIEMTQNSYRAKEPILALRRALLSLSKSQDYSELVGQCWLQSARVARKAGHHQTAYNALLNAGESTLSELYIERAKWLWSKG 1986
S. cerevisiae 1640 FEYK------SNTTILDYRMERIGADFVPNHYILSMRKSFDQLKMNEQADADLGKTFFTLAQLARNNARLDIASESLMHCLERRLPQAELEFAEILWKQG 1733
2201 | . . . . . | 2300
H. sapiens 1968 DVHQALIVLQKGVELCFPENETPPEGKNMLIHGRAMLLVGRFMEETANFESNAIMKKYKDVTACLPEWEDGHFYLAKYYDKLMPMVTDNKMEKQGDLIRY 2067
G. gallus 1987 EVHQALIVLQKGVELCFPENKAPCDTKNQLIHGRATLLVGRFMEETANFESNAVMKKYKDVTLLLPEWEDGHFYLAKYYDKLMPMVTDNKMEKQGDLIRY 2086
S. cerevisiae 1734 ENDRALKIVQEIHE-KYQENSSV----NARDRAAVLLKFTEWLDLSNNSASEQIIKQYQDIFQIDSKWDKPYYSIGLYYSRLLERKKAEGYITNGRFEYR 1828
2301 | . . . . . | 2400
H. sapiens 2068 IVLHFGRSLQYGNQFIYQSMPRMLTLWLDYGTKAYEWEKAGRSDRVQMRNDLGKINKVITEHTNYLAPYQFLTAFSQLISRICHSHDEVFVVLMEIIAKV 2167
G. gallus 2087 IVHHFGRSLQYGNQFIYQSMPRMLSLWLDFGAKAYECDKASRSERVQMKNDLAKINKVITEHTNQLAPYQFLTAFSQLISRICHSHDEVFAVLMVIVAKV 2186
S. cerevisiae 1829 AISYFLLAFEKNTAKVRENLPKVITFWLDIAAASISEAPGNRKEMLS--KATEDICSHVEEALQHCPTYIWYFVLTQLLSRLLHSHQSSAQIIMHILLSL 1926

HR 44
HR 45
Kinase
FATC

8.6 Poster presentations

RAMI Biomedical Sciences Section Meeting, NUIG (2012-06-14)



30th Ernst Klenk Symposium: DNA Damage Response and Repair Mechanisms in Aging and Disease, University of Cologne, Germany (2014-09 21st to 23rd).

Modelling *ATR* disease - related mutations using chicken DT40 cells

Marta Llorens-Agost, John K. Eykelenboom, Noel F. Lowndes

Genome Stability Laboratory - Centre for Chromosome Biology, Biosciences Building, NUI Galway, Ireland

INVESTING IN YOUR FUTURE

1. Abstract

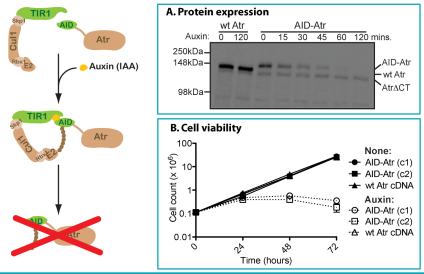
The ATR (ATM and RAD3 related) protein functions as a master regulator of the DNA Damage Response (DDR). *ATR* is an essential gene in mammalian cells and mutations in this gene have been linked to rare cases of Seckel syndrome.

To date, all *ATR* Seckel patients belong to one of two related families, which harbour the identical mutation in *ATR*. This is a limiting factor for the characterisation of the clinical phenotype conferred by ATR deficiency.

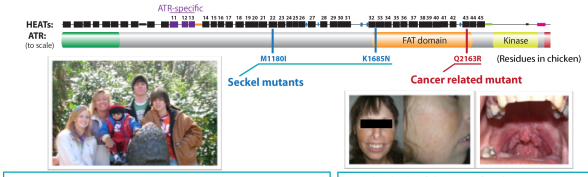
In the past year, two novel mutations in *ATR* have been found in unrelated Seckel patients by our collaborators in the UK. Besides, mutations in the *ATR* gene have also been linked recently to cancer predisposition in humans.

The goal of our study is to use a conditional-null DT40 cell line to generate a model system that will facilitate mechanistic studies of these mutations. This will also help elucidating the role of ATR in the maintenance of genome stability.

2. Conditional null Atr cell line



3. Atr structure: Seckel and cancer related mutations



Seckel patients suffer from mycrocephaly and developmental defects.

Classic Seckel syndrome: exonic mutation (2101A>G) that affects splicing and results in low levels of expression of functional ATR.

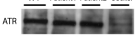
New Seckel mutations:
- One allele of ATR contains the point mutation (M1159I or K1665N in human).
- The other allele carries a new splicing mutation.

ATR protein levels are very low in patient cells.

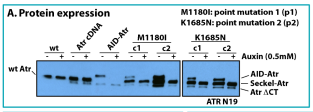
24 individuals of the same family contain a point mutation in ATR (6431A>G) that causes oropharyngeal cancer and skin telangiectases.

First ATR mutation related to cancer predisposition => NOT Seckel.

The mutation:
- is located within the FAT domain.
- potential S/Q phosphorylation site.
- does not affect ATR protein levels.

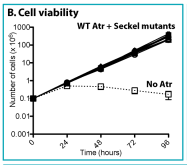


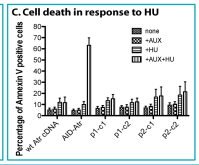
4. Modelling ATR Seckel mutations



Seckel point mutants M1180I-Atr and K1685N-Atr behave as wild type:

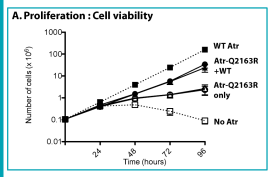
-The proteins are expressed at wt level.
-They support cell viability.
-They are chekpoint proficient in response to HU.

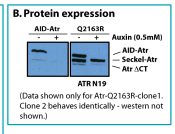




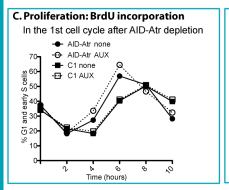
The point mutations could be a false lead.
Our preliminary data using patient-derived cell lines suggests that these mutations (M1159I and K1665N in humans) might affect splicing in the same fashion as the classical Seckel syndrome (exon skipping).

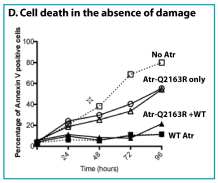
5. Modelling an ATR cancer - related mutation

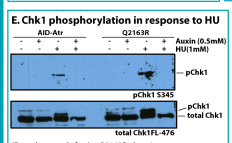




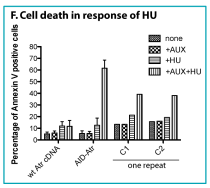
(Data shown only for Atr-Q2163R-clone1. Clone 2 behaves identically - western not shown.)







(Data shown only for Atr-Q2163R-clone1. Clone 2 behaves identically - western not shown.)



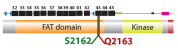
Cancer-related Q2163R mutants have a defect in growth (dominant negative effect)

Due to a defect in signalling?

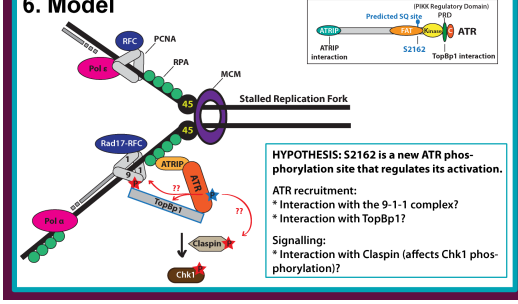
Chk1 phosphorylation is completely abrogated!

The serin adjacent to the mutated residue is a putative SQ phosphorylation site:

* Docking site for an interacting protein?
* Conformational activation of ATR?



6. Model



HYPOTHESIS: S2162 is a new ATR phosphorylation site that regulates its activation.

ATR recruitment:
* Interaction with the 9-1-1 complex?
* Interaction with TopBp1?

Signalling:
* Interaction with Claspin (affects Chk1 phosphorylation)?

7. Future Work:

M1180I and K1685N: These mutations do not seem to affect the function in DT40 cells => *Confirm if they affect splicing in patient cells.*

Q2163R: Atr-Q2163R mutants have a signalling defect.

In DT40 cells: Study the phosphorylation of S2162
- Purify tagged Atr and perform mass spectrometry to confirm phosphorylation.
- Generate an antibody to study the importance of this post-translational modification.

In human cells: Confirm Q2144R mutation affects signalling:
- CRISPR-Cas9 technology: Generate a stable cell line containing the Q2144R mutation.
- Co-IPs: Pull down ATR and check if any interaction in the signalling pathway is lost because of the Q2144R mutation (9-1-1, TopBp1...).

Acknowledgements We would like to thank Dr. Grant Stewart for kindly providing human LCLs derived from Seckel patients to study the splicing defects caused by K1665N and M1159I mutations. We would also like to thank PRLTI, SFI and the National University of Ireland, in Galway, for funding this research project.