Regulation of DNA end resection by cell stemness

Cintia Checa Rodríguez

Tesis doctoral realizada en el Departamento de Genética,
Universidad de Sevilla, y en el Centro Andaluz de Biología Molecular
y Medicina Regenerativa (CABIMER), para optar al título de Doctora
por la Universidad de Sevilla.

Sevilla, Octubre de 2019

Doctorando	Director y tutor de la tesis
Cintia Checa Rodríguez	Pablo Huertas Sánchez

Dicen que si caminas acompañado, llegarás más lejos…. y yo tengo la suerte de haber llegado muy lejos gracias a todos los que han estado conmigo de una forma u otra durante estos cinco años de tesis. Ahora, que por fin ha concluido este duro y largo camino, tengo mucho que agradecer:

En primer lugar a Pablo, por confiar en mí y darme la oportunidad de realizar esta tesis. Gracias por ese gran ambiente de trabajo que consigues que haya en el laboratorio, por ayudarnos y aconsejarnos siempre y por supuesto por todas las correcciones de la tesis.

A Dani, mi profe, mi compi, mi amigo y mi hermano mayor en el laboratorio. Gracias por enseñarme tanto desde que llegué, por guiarme, aconsejarme y estar siempre dispuesto a ayudarme incluso cuando ya no estabas en Sevilla. Gracias por las charlas sobre el futuro y la vida en general y por supuesto por todas las sugerencias y correcciones de la tesis. Infinitas gracias por ser una de las partes más importantes de esta tesis y de estos más de 5 años.

Gracias a todos mis compañeros del laboratorio PHS, los que aún siguen, los nuevos y los que ya no están. Gracias por todos los consejos, lo que me habéis enseñado y el buen ambiente de trabajo. Y en especial, a mi “gemela” Rosario y a Rosa, por estar siempre dispuestas a ayudar en cualquier cosa y por todas las risas y los grandes momentos compartidos dentro y fuera del laboratorio. Siempre llevaré un pedacito de vosotros conmigo.

Gracias a mis amigos de siempre y del baile por ayudarme a desconectar, por animarme y siempre interesarse por cómo iba la tesis. Esos momentos de desconexión ayudaban mucho.

Gracias a mis padres y mi hermano, mi familia, por enseñarme desde pequeña a ser constante con lo que hago y perseguir mis sueños. Por interesaros siempre en mi trabajo e intentar entender los experimentos y por apoyarme siempre en todo. Gracias por todo lo que me habéis enseñado y por todo lo que habéis hecho por mí durante toda mi vida para que yo pudiese llegar hasta aquí. Gracias porque sin vosotros no habría sido posible, os quiero.

Y por último, gracias a mi niño Ludo, por darme ánimos para seguir, por preguntarme siempre que tal en el laboratorio y por estar siempre ahí para levantarme cuando ya no podía más. Gracias por entender mis horarios de trabajo, muchas veces raros, y hacer ese tiempo más ameno, por acompañarme 3 meses a otro país y por estar siempre conmigo en las buenas y en las malas. Gracias mi amor, te quiero.

En largos caminos se conocen buenos amigos y en estos cinco años he encontrado muchos.

Gracias por todo.

Abstract

The ability to reprogram somatic cells into induced pluripotent stem cells (iPSCs) holds great potential for clinic applications, however, acquired genomic instability is one of the major concerns for its clinical use. The reprogramming process is accompanied by the induction of DNA damage, of which double-strand breaks (DSBs) are the most cytotoxic. To minimize the impact of these damages, cells have developed two main repair pathways: Homologous Recombination (HR) and Non-Homologous End-Joining (NHEJ). The choice between both mechanisms is a complex and highly regulated process and the right balance is critical to ensure the maintenance of genomic stability. One of the best-known decision points is DNA end resection which leads to HR activation, and the nuclear protein CtIP is its major regulator. In this thesis, we investigated the role of DNA end resection and CtIP during the reprogramming process and we revealed that reprogramming is associated with high CtIP protein levels and a hyper-activation of DNA end resection. Moreover, CtIP is essential for the maintenance of genomic stability and reprogramming in a resection-defective environment has long-term consequences on stem cell self-renewal and differentiation. Furthermore, we show that the pluripotency factor KLF4 plays an important role in regulating DNA damage repair processes and the balance between them acting upon the HR pathway, specifically promoting DNA-end resection.

Resumen

La inestabilidad genómica adquirida es una de las principales preocupaciones para el uso clínico de células madre pluripotentes inducidas (iPSCs). Todos los métodos de reprogramación están acompañados de la inducción de daño en el ADN, de los cuales los cortes de doble cadena (DSBs) son los más citotóxicos y mutagénicos. Para reparar los DSBs existen dos mecanismos principales y alternativos: la unión de extremos no homólogos (NHEJ) y la recombinación homóloga (HR). La elección entre uno u otro tipo de reparación es un proceso complejo y altamente regulado, y el uso de cualquiera de ellos en un momento inadecuado genera un aumento de la inestabilidad genómica, un fenómeno estrechamente asociado con el desarrollo del cáncer. Sin embargo, se considera que el procesamiento de los extremos del corte, o resección, es un paso clave en esta decisión ya que conduce a la reparación por HR y bloquea eficientemente la NHEJ. La resección está fuertemente regulada por muchas señales celulares diferentes, siendo CtIP el componente más conocido de esta red reguladora. Durante esta tesis, investigamos el papel de la resección y del factor CtIP durante el proceso de reprogramación celular y revelamos que durante el proceso los niveles de CtIP aumentan y la resección está hiper-activada. Además, demostramos que el factor de resección CtIP es esencial para el mantenimiento de la estabilidad genómica y la reprogramación en un entorno defectuoso de resección tiene consecuencias a largo plazo limitando el mantenimiento del estado de pluripotencia de las células reprogramadas o autorrenovación y su posterior diferenciación. Por otro lado, observamos que Krüppel-like factor 4 (KLF4), uno de los cuatro factores utilizados para reprogramar células diferenciadas adultas en células madre pluripotentes inducidas, parece tener un papel importante en la regulación de los procesos de reparación de cortes de doble cadena y en el balance entre ambas vías. De hecho, encontramos que KLF4 actúa sobre la vía de recombinación homóloga, promoviendo específicamente la resección.

Indexes

Table of contents

I. Introduction..21

1. DNA damage and genomic stability..23

1.1. DNA Double Strand Breaks (DSBs) 25

1.2. The DNA Damage Response (DDR)..26

1.3. DNA DSBs repair pathways..28

1.4. DNA end resection..31

1.5. DNA DSBs repair pathway choice..33

1.6. CtBP-interacting protein (CtIP)..35

1.7. Functions and regulation of CtIP 37

2. DNA Damage and genomic instability in Stem Cells................................39

2.1. Types of Stem Cell 39

2.2. DNA damage repair in Stem cells. 43

2.2.1 DNA damage repair in Adult Stem cells. 43

2.2.2. DNA damage repair in Embryonic Stem cells..45

2.2.3. DNA damage repair in iPSCs and the reprogramming process..................47

2.2.4. DNA damage repair in Cancer Stem Cells. 48

2.3. The reprogramming factor Krüppel like factor 4..49

2.4. Regulation of KLF4..50

2.5. Role of KLF4 in biological processes and cell homeostasis 51

2.6. Role of KLF4 in stem cells and reprogramming..52

2.7. Role of KLF4 in cancer 52

II. Objectives..55

III. Results 59

Chapter one 61

1. DNA end resection increases upon mouse cell reprogramming 61
2. CtIP levels increase during miPSCs formation 64
3. CtIP is required for efficient cell reprogramming 65
4. CtIP deficiency during mouse cell reprogramming causes genomic instability 68
5. CtIP deficiency triggers apoptosis during reprogramming. 69
6. Normal CtIP levels during reprogramming are required for maintenance of iPSCs 71
7. Normal CtIP levels during reprogramming are required for differentiation of iPSCs 75

Chapter two 78

1. KLF4 affects DNA break repair pathway choice..78
2. KLF4 is required for survival to DSB inducing agents 81
3. KLF4 is required for DNA end resection 84
4. KLF4 affects the recruitment of pro- and anti-resection proteins to DSBs sites....88
5. Arginine methylation of KLF4 is essential for its activity in DNA end resection.....91
6. KLF4 affects TIP60 protein level 94

IV. Discussion...97

V. Conclusions 113

VI. Materials and methods 117

1. Cell culture procedures 119

1.1. Cell lines, growth media and conditions 119

1.2. Mouse cell reprogramming into iPSC 120

1.3. Isolation of iPSCs colonies 120

1.4. Proliferation assay 121

1.5. GFP-shRNA stability analysis 121

1.6 Colony formation assay 121

1.7. Embryoid bodies formation...122

1.8. siRNA Transfection 122

1.9. Plasmid DNA Transfection 123

1.9.1. FuGENE transfection 123

1.9.2. Calcium phosphate transfection 123

1.10. Lentivirus production and transduction..124

2. Molecular biology procedures.. 124

2.1. Nucleic acids manipulations 124

2.1.1. Plasmid DNA amplification...124

2.1.2. DNA digestion with restriction enzymes..125

2.1.3. DNA electrophoresis in agarose gels 125

2.1.4. Site-directed mutagenesis 126

2.1.5. RNA extraction 126

2.1.6. Reverse transcription 126

2.1.7. Quantitative PCR (qPCR) 126

2.2. Protein analysis 127

2.2.1. Protein extraction under denaturing conditions 127

2.2.2. SDS-PAGE electrophoresis..127
2.2.3. Western blot analysis..128
3. Cell biology procedures..128
3.1. DSBs repair assays *in vivo*..128
3.1.1. Description of repair systems..128
3.1.2. DSBs repair assays in vivo..131
3.2. Flow cytometry..132
3.2.1. Cell cycle analysis..132
3.2.2. DNA-end resection analysis..133
3.2.3. Titration of lentiviral production..133
3.2.4. Fluorescent protein analysis..134
3.2.5. Cell sorting..134
3.3. Clonogenic survival assay..135
3.4. Microscopy..136
3.4.1. RPA/γH2AX foci immunofluorescence..136
3.4.2. BRCA1/γH2AX foci immunofluorescence..136
3.4.3. RIF1/γH2AX foci immunofluorescence..137
3.4.4. Single-molecule analysis of resection tracks (SMART)..137
3.5. Array Comparative Genomic Hybridization..138
4. Statistical analysis..139
5. Tables of materials..140
VII. Bibliography..147
VIII. Publications..179

Figures Index

Figure I1. Types, sources and repair of DNA lesions. 24

Figure I2. The DNA damage response (DDR) pathway. 27

Figure I3. DSB repair pathways. 30

Figure I4. DNA end resection mechanism. 32

Figure I5. Regulation of DSB repair pathway choice. 35

Figure I6. Structure and modifications of human CtIP protein. 37

Figure I7. Types of stem cells 40

Figure I8. General diagram of reprogramming process and differentiation 42

Figure I9. Structure and modifications of human KLF4 protein. 50

Figure R1. DNA end resection increases in reprogrammed cells 62

Figure R2. The extension of DNA end resection is promoted during reprogramming 63

Figure R3. CtIP levels increases in miPSCs 64

Figure R4. DDR proteins increases during reprogramming 66

Figure R5. CtIP loss impairs reprogramming and leads to DNA damage accumulation 67

Figure R6. CtIP deficiency causes genomic instability during reprogramming 68

Figure R7. CtIP deficiency during reprogramming activates apoptosis. 70

Figure R8. Stability of shRNA-eGFP cassette in reprogramming and miPSCs maintenance 72

Figure R9. The levels of CtIP during reprogramming affect iPSC maintenance 74

Figure R10. CtIP downregulation does not affect D3 embryonic stem cells colony formation 75

Figure R11. CtIP during reprogramming is essential for iPSCs differentiation. 76

Figure R12. CtIP downregulation does not affect D3 embryonic stem cells differentiation. 77

Figure R13. KLF4 controls the balance between NHEJ and HR. 79

Figure R14. Downregulation of KLF4 impairs single strand annealing. 80

Figure R15. Absence of KLF4 impairs gene conversion. 80

Figure R16. KLF4 depletion increases NHEJ. 81

Figure R17. KLF4 expression levels 82

Figure R18. KLF4 avoid cell sensitivity to DSB inducing agents 83

Figure R19. KLF4 does not affect cell cycle progression. 84

Figure R20. KLF4 affects DNA end resection. 86

Figure R21. KLF4 affects the speed of DNA end resection. 87

Figure R22. KLF4 affects RIF1 foci formation. 89

Figure R23. KLF4 affects BRCA1 foci formation. 90

Figure R24. PRMT5-mediated methylation of KLF4 is essential for its role in DNA end resection. 92

Figure R25. KLF4 overexpression does not rescue the defect on DNA end resection caused by PRMT5 depletion. 93

Figure R26. KLF4 does not affect the expression of some proteins involved in DSB repair. 95

Figure R27. KLF4 overexpression controls TIP60 levels. 95

Figure D1. DNA end resection during cell reprogramming. 105

Figure M1. DSBs repair reporter systems used. 130

Tables Index

Table M1. Cell lines used in this thesis. 140

Table M2. siRNAs used in this thesis.. 141

Table M3. Plasmids used in this thesis. 142

Table M4. Primers used in this thesis. 143

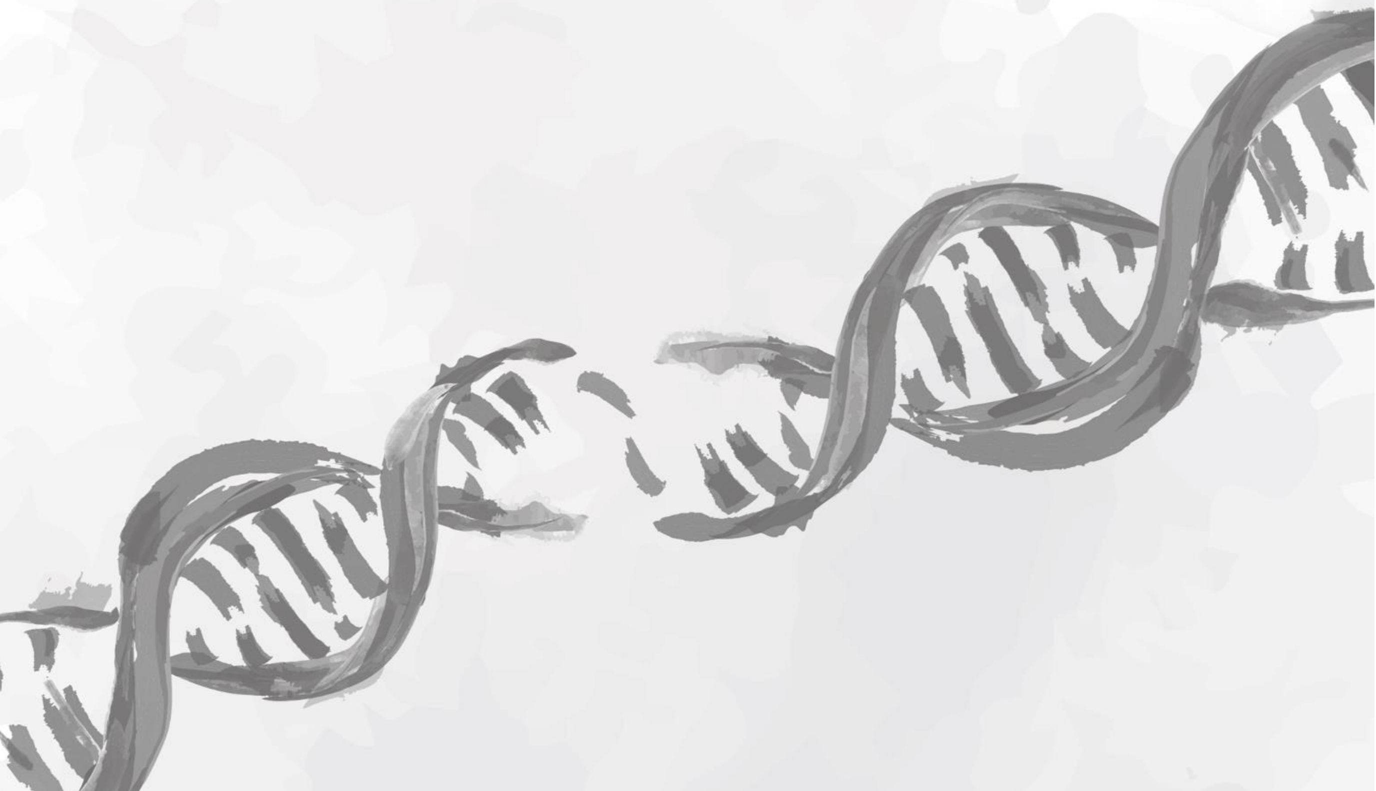
Table M5. Primary antibodies used in this thesis. 144

Table M6. Secondary antibodies used in this thesis. 145

Abbreviations

ASC	Adult stem cell
BFP	Blue fluorescent protein
BrdU	Bromodeoxyuridine
CNV	Copy number variation
CPT	Camptothecin
CSC	Cancer stem cell
DDR	DNA damage response
DOX	Doxycycline
DSB	Double strand break
EB	Embryoid body
eGFP	Enhanced green fluorescence protein
ESC	Embryonic stem cell
FACS	Fluorescence-activated cell sorting
GFP	Green fluorescence protein
HR	Homologous recombination
iPSC	Induced pluripotent stem cell
IR	Ionizing radiation
LIF	Leukaemia inhibitory factor
MMEJ	Microhomology-mediated NHEJ
NHEJ	Non-homologous end joining
NT	Non-target
OSKM	OCT4, SOX2, KLF4, c-MYC
ROS	Reactive oxygen species
SC	Stem cell
SEM	Standard error of the mean
SMART	Single molecule analysis of resection tracks
SSA	Single-strand annealing
ssDNA	Single-strand DNA
SSR	SeeSaw reporter system
VP-16	Etoposide
WB	Western-blot

I. Introduction



1. DNA damage and genomic stability

Life depends on the ability of cells to store, retrieve, and translate the information required to make and maintain a living organism. So, the high-fidelity transmission of the genetic information stored in DNA is critical for the maintenance of genome integrity and viability of the cells. However, DNA is constantly exposed to damage induced by environmental agents or generated spontaneously during cell metabolism. It has been estimated that every cell is subject to more than 10.000 DNA lesions per day (Hoeijmakers, 2009). To counteract DNA damage, cells have developed strategies that respond and repair DNA lesions to maintain genome integrity (Ciccia and Elledge, 2010). Thus, the absence or impairment of such mechanisms can lead to DNA mutations or chromosome aberrations that contribute to genomic instability which is associated with human premature ageing, predisposition to various types of cancer and with inherited diseases (Aguilera and Gómez-González, 2008).

Endogenous DNA damage occurs at a high frequency compared with exogenous damage and the types of damage produced are identical or very similar to those caused by some environmental agents (Jackson and Loeb, 2001) (Figure I1). Spontaneous DNA alterations may occur due to DNA mismatches, insertions and deletions caused during DNA replication (Jackson and Bartek, 2009). Moreover, by-products derived from cellular metabolism, including reactive oxygen species (ROS), can also attack DNA and produce DNA damage. These ROS include superoxide, hydrogen peroxide, hydroxyl radicals and singlet oxygen (De Bont and van Larebeke, 2004; Jackson and Bartek, 2009). Exogenous DNA damage can be produced by exposure to physical sources, including ultraviolet (UV) light and ionizing radiation (IR) (Ciccia and Elledge, 2010; Hoeijmakers, 2009). In addition, a variety of DNA lesions can be produced by chemical agents, such as the case of alkylating agents, crosslinking agents topoisomerase inhibitors or cigarette smoke (Cheung-Ong et al., 2013; Phillips et al., 1988) (Figure I1).

The wide diversity of types of DNA-lesion provokes the need of multiple and distinct DNA-repair mechanisms (Figure I1). Accordingly, cells have evolved specific repair

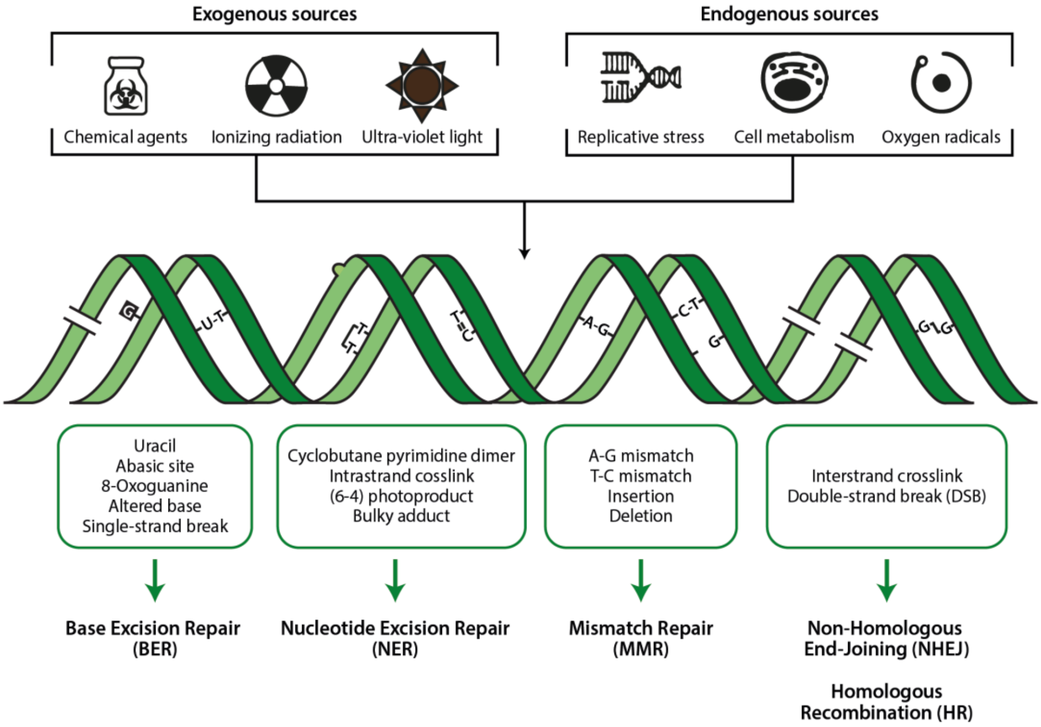

Exogenous sources
Chemical agents
Ionizing radiation
Ultra-violet light
Endogenous sources
Replicative stress
Cell metabolism
Oxygen radicals
Uracil
Abasic site
8-Oxoguanine
Altered base
Single-strand break
Cyclobutane pyrimidine dimer
Intrastrand cosslink
(6-4) photoproduct
Bulky adduct
A-G mismatch
T-C mismatch
Insertion
Deletion
Interstrand crosslink
Double-strand break (DSB)
Base Excision Repair (BER)
Nucleotide Excision Repair (NER)
Mismatch Repair (MMR)
Non-Homologous End-Joining (NHEJ)
Homologous Recombination (HR)

Figure I1. Types, sources and repair of DNA lesions.
DNA damage can be caused by various exogenous and/or endogenous sources (top). DNA lesions can occur in many different forms that can affect one or both DNA strands (middle) and each modification requires specific machinery for repair (bottom). (Modified from (Genois et al., 2014; Hoeijmakers, 2001).

mechanisms to counteract DNA damage and maintain genomic stability. DNA lesions can occur in many different forms. The most common lesions are chemical modifications that alter the structure of the nucleotides (Lindahl, 1993), including abasic sites, bulky adducts, covalent binding between nucleotides of both DNA strands (intrastrand and interstrand crosslinks), pyrimidine dimers and insertion/deletion mismatches (De Bont and van Larebeke, 2004). Another major type of DNA damage are single strand breaks (SSB) in which the phosphodiester backbone is cut in one strand of the double helix, interrupting the continuity of the DNA (Caldecott, 2008). Even though there are specific repair pathways for each type of modifications, which include base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR), they act basically by removing the damaged base of a nucleotide and inserting new bases to fill the gap using the complementary strand as a template (Hoeijmakers, 2009; Jackson and Bartek, 2009). On the other hand, if the damage affects both strands simultaneously, it produces a double strand break (DSB) which is the most cytotoxic form of damage and requires more complex repair mechanisms (Figure I1).

1.1. DNA Double Strand Breaks (DSBs)

DNA double-strand breaks (DSBs) arise regularly in cells and are highly toxic and potentially dangerous. In mammalian cells, DSBs can arise following exposure to exogenous agents, such as ionizing radiation (IR), topoisomerase poisons and radio-mimetic chemicals (Jackson, 2002), but can also be caused by endogenous sources. Naturally occurring DSBs are generated spontaneously during DNA synthesis when the replication fork encounters unrepaired DNA lesions, triggering fork collapse (Pfeiffer et al., 2000). Moreover, during certain specialized cellular processes such as meiosis or the diversification of immunoglobulins, DSBs are programmed by the cell. On the one hand, meiotic DSB repair is essential for correct chromosome segregation at the first meiotic division and generates gametes with allele combinations distinct from the parental germline. On the other hand, DSB-induced rearrangements at immunoglobulin genes are critical for the multiplicity of antigen receptor diversity over limited numbers of loci (Chapman et al., 2012).

Regardless of the source, accurate repair of DSBs is essential for the successful maintenance and propagation of genetic information. Failure or lack of repair may result in cell death or mutations and gross chromosomal rearrangements, including deletions and translocations (Aguilera and Gómez-González, 2008). Indeed, defective DSB repair is associated with various developmental, immunological, and neurological disorders, and is a major driver in cancer (Jackson and Bartek, 2009; McKinnon, 2009). The repair of DSBs is carried out by more complex repair mechanisms. Moreover, cells respond to DNA DSBs through the actions of systems that detect the DNA lesion and then trigger various downstream events. This network of cellular pathways is known as the DNA damage response (DDR) (Jackson and Bartek, 2009).

1.2. The DNA Damage Response (DDR)

The DNA damage response (DDR) is a complex network of signal transduction pathways that senses DNA damage, mainly DSBs, and trigger various downstream events to repair DNA lesions and maintain genome integrity. The DDR controls the coordination of important cellular processes including cell cycle checkpoints, transcription, translation, DNA repair, metabolism, and cell fate decisions, such as apoptosis or senescence, to prevent the replication of damaged DNA and the inheritance of damaged DNA by daughter cells (Ciccia and Elledge, 2010; Jackson and Bartek, 2009). Like all classic signal transduction cascades, the DDR uses sensors, mediators, transducers and effectors (Zhou and Elledge, 2000)(Figure I2). It requires proteins recruitment to damage, that normally occurs in a hierarchical manner, and also involves multiple posttranslational modifications (Lukas and Bartek, 2004; Polo and Jackson, 2011). DNA damage is directly recognized by sensor proteins that bind broken DNA in a sequence-independent manner. This event then triggers the recruitment and activation of transducer proteins, with the help of mediators. Then, the transducer proteins transmit the damage signal to various effector proteins that participate in a wide spectrum of cellular processes which leads to the repair of the lesion (Jackson, 2002; Maréchal and Zou, 2013; Polo and Jackson, 2011) (Figure I2).

In mammalian cells, the main sensors that recognize DNA DSBs are the KU70-KU80 heterodimer and the MRE11-RAD50-NBS1 (MRN) complex (Polo and Jackson, 2011) (Figure I2). They bind rapidly to the site of damage and lead to the activation of transducer kinases. Ataxia-telangiectasia mutated (ATM), ATM and Rad3-related (ATR), and DNA-dependent protein kinase catalytic subunit (DNA-PKcs) are the most upstream DDR kinases (Ciccia and Elledge, 2010; Maréchal and Zou, 2013) (Figure I2). Following the initial activation, the kinases phosphorylate many different substrates involved in DDR signalling, including histones and proteins involved in the regulation of chromatin structure. In response to DNA damage, those kinases phosphorylate over 700 different substrates in mammalian cells (Matsuoka et al., 2007). Chiefly among them, these kinases mediate the phosphorylation of the histone variant H2AX, considered one of the earliest events in the cellular response to DSBs. Phosphorylation of H2AX at the S139 residue, known as γH2AX, occurs within seconds after DNA damage and spreads over large chromatin domains

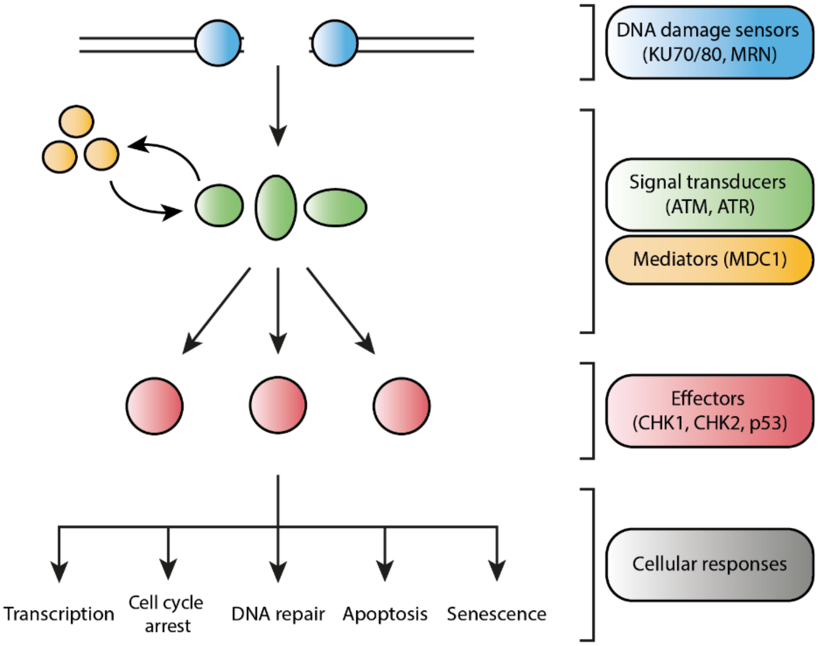

DNA damage sensors
(KU70/80, MRN)
Signal transducers
(ATM, ATR)
Mediators (MDC1)
Effectors
(CHK1, CHK2, p53)
Cellular responses
Transcription
Cell cycle
arrest
DNA repair
Apoptosis
Senescence

Figure I2. The DNA damage response (DDR) pathway.
Schematic representation of DDR molecular components that are recruited to DSBs in a hierarchical manner. DNA damage is recognized by sensor proteins (Ku70/80 or MRN) and triggers the recruitment and activation of transducer proteins (ATM, ATR) with the help of mediators (MDC1 and others). Transducer proteins amplify the signalling cascade by targeting many effector proteins (CHK1, CHK2, p53) that participate in a wide spectrum of cellular processes including transcriptional induction, cell cycle arrest, DNA repair, progression towards apoptosis or senescence, etc.

flanking the DNA breaks (Savic et al., 2009). Moreover, H2AX phosphorylation induces the binding of the mediator protein MDC1 which is required for γH2AX spreading (Stewart et al., 2003; Stucki et al., 2005) (Figure I2). Formation of γH2AX extensive regions amplifies the initial signal and initiates the recruitment and accumulation of many different DDR factors at DNA lesions which form structures known as ionizing radiation-induced foci (IRIF) that can be visualized by microscopy (Bekker-Jensen et al., 2006; Stucki et al., 2005).

One well-established feature of the DNA damage response is the activation of DNA-damage checkpoints which slow down or arrest the cell-cycle progression to avoid cell-cycle transitions until repair has taken place (Abraham, 2001; Kastan and Bartek, 2004). In response to DNA damage, three different and major cell cycle–checkpoints can be activated in the DDR: G1/S, intra-S phase, and G2/M checkpoints (Lukas et al., 2004). The main checkpoint effector kinases are CHK1 and CHK2 which are activated by ATR and ATM phosphorylation, respectively (Figure I2). Active effector kinases phosphorylate numerous downstream targets and leads to reduce cyclin-dependent kinase (CDK) activity which is essential to establish a cell cycle arrest (Kastan and Bartek, 2004). This arrest can be reversed once the damage has been repaired. However, if the damage cannot be repaired, DDR signalling triggers programmed cell death via induction of p53 (Jackson and Bartek, 2009).

1.3. DNA DSBs repair pathways

In mammals, there are different DSB repair pathways which can be categorized in two big families of mechanisms depending on whether sequence homology is required: non-homologous end joining (NHEJ) and homologous recombination (HR) (Figure I3).

NHEJ involves the direct ligation of DNA ends through a process largely independent of homology. It is highly efficient but it is also considered to be an error-prone process because nucleotides can be lost or gained at the ends previous to ligation (Chiruvella et al., 2013). This is the major pathway that repair DSBs in vertebrate cells and it can function throughout the cell cycle, but is of particular importance during G0 and G1 phases when HR is highly suppressed (Rothkamm et al., 2003) (Figure I3). NHEJ is initiated by the high-affinity binding of KU70/KU80 heterodimer to both ends of the broken DNA molecule, which protects them from degradation (Dynan and Yoo, 1998) (Figure I3). DSB-bound KU then recruits and activates DNA-PKcs that initiates an extensive signalling cascade and facilitate the recruitment of XRCC4/XLF/DNA Ligase IV complex which carries out the ligation of the DNA ends to complete repair (Cottarel et al., 2013; Grawunder et al., 1997) (Figure I3). HR, by contrast, is commonly considered to be error-free and requires the use of an undamaged homologous sequence as a template to copy and accurately restore the DNA sequence (San Filippo et al., 2008). Therefore, this pathway is restricted to the S and G2 phases of the cell cycle when a sister chromatid is available (Rothkamm et al., 2003)

(Figure I3). In HR, the search for a homologous sequence requires the processing of DNA ends to generate single-strand DNA (ssDNA) containing a 3′-hydroxyl overhang. This intermediate is generated by nucleolytic degradation in a process termed DNA end resection, which is mediated by the MRN complex and CtBP-interacting protein (CtIP) (Huertas, 2010) (Figure I3). The ssDNA overhangs generated by DNA resection are rapidly coated by the replication protein A (RPA) complex, which prevents the formation of secondary structures. Afterwards, RPA is displaced by RAD51 forming RAD51 filaments needed for strand invasion (San Filippo et al., 2008) (Figure I3).

Alternatively, end resection can also provide an intermediate for other minor repair pathways: single-strand annealing (SSA) and one form of alternative NHEJ (alt-NHEJ) known as microhomology-mediated NHEJ (MMEJ) (Figure I3). On the one hand, MMEJ seal the break in a mechanism that shares elements with both NHEJ and HR. It is a KU-independent NHEJ pathway that requires limited resection and involves alignment and subsequent ligation of microhomologous sequences of 5-25 bp close to the break, thereby resulting in deletions flanking the original break (McVey and Lee, 2008; Sfeir and Symington, 2015) (Figure I3). On the other hand, SSA require more extensive end resection and it is used only when two long homologous regions flank the DSB site, leading to the annealing of complementary strands from each repeated sequence (Figure I3). This results in the deletion of one of the repeats and the intervening sequence and, therefore, loss of genetic information (Morrical, 2016). Thus, both SSA and MMEJ are always mutagenic pathways associated with deletions and chromosomal rearrangements (Symington and Gautier, 2011).

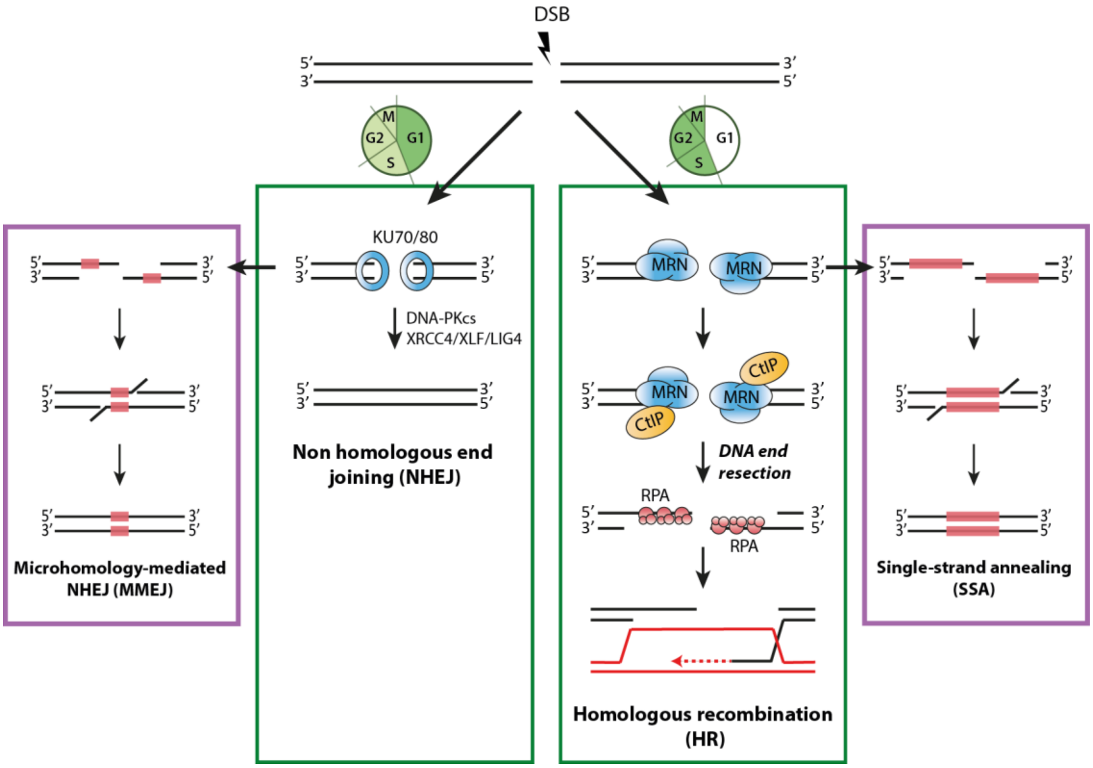

DSB
5′
3′
3′
5′
M
G2
G1
S
KU70/80
DNA-PKcs
XRCC4/XLF/LIG4
Non homologous end joining (NHEJ)
MRN
CtIP
DNA end resection
RPA
Homologous recombination (HR)
Microhomology-mediated NHEJ (MMEJ)
Single-strand annealing (SSA)

Figure I3. DSB repair pathways.

The two main repair pathways to repair DSBs are non-homologous end joining (NHEJ) and homologous recombination (HR). NHEJ (left green panel) starts with the recognition of the DNA ends by the KU70/80 heterodimer and then, others factors (DNA-PKcs, XRCC4/XLF/LIG4) are subsequently recruited to carry out the direct ligation of both DNA ends. This mechanism can function throughout the cell cycle but is the predominant repair pathway used in G1. During HR (right green panel), the repair of DSB is initiated by the resection of the DNA ends which requires the action of the MRN complex and CtIP to generate single-stranded DNA that is protected by RPA complex. The single-strand DNA is required to search for a homologous template for repair. Alternatively, there are minor repair pathways that require end resection: microhomology-mediated NHEJ (MMEJ) and single-strand annealing (SSA). SSA (right purple panel) require extensive end resection that exposes single strand regions with long homologous sequences (red boxes) of greater than 25 nucleotides flanking the DSB site. Then, the complementary single strands from each repeated sequence anneal and are processed for ligation resulting in the deletion of the sequence between the repeats. MMEJ (left purple panel) share molecular components and intermediates with both NHEJ and SSA and it requires limited resection and subsequent ligation of microhomologous sequences of 5-25 bp (red boxes) close to the break, resulting in genomic deletions.

1.4. DNA end resection

The decision between HR and NHEJ relies, mostly, on the activation or not of extensive processing of the DNA ends known as DNA end resection (Huertas, 2010; Symington, 2014). This consists in the degradation of one strand of the DNA at each side of the break in a 5’-3’ polarity that uncovers long tails of 3’OH protruding ssDNA that is immediately coated by the protecting complex RPA (Huertas, 2010; Symington, 2014) (Figure I4). Resected DNA is required for HR, but efficiently blocks canonical NHEJ.

In eukaryotes, DNA end resection occurs in two distinct steps (Gravel et al., 2008; Mimitou and Symington, 2008; Zhu et al., 2008) (Figure I4). First, there is an initial short-range resection that is slow and consists of limited 5′ DNA end degradation (Mimitou and Symington, 2008). This step is mainly carried out by the MRE11-RAD50-NBS1 (MRN) complex and CtIP in human cells (Paull and Gellert, 1998; Sartori et al., 2007). MRE11 is the catalytic component of the MRN complex and appears to be the critical nuclease responsible for short-range resection. It has a variety of enzymatic activities, including 3’ to 5’ exonuclease activity on double stranded DNA (dsDNA) and endonuclease activity on 5′-terminated DNA strands and on other DNA structures (Paull and Gellert, 1998; Trujillo et al., 1998). Since resection generates 3’ overhang, a bidirectional model has been proposed (Garcia et al., 2011). According to this, DNA is first nicked in the 5’-terminated strand away from the break by MRE11 endonuclease activity (Figure I4). Then, starting from the nick, the 3’ to 5’ exonuclease activity can degrade the DNA toward the DSB end leaving 3’ ssDNA tails (Garcia et al., 2011; Mimitou and Symington, 2008; Shibata et al., 2014) (Figure I4). Besides the importance of MRN in resection, the complex by itself is not enough to initiate the process. Recruitment of CtIP and its physical interaction with MRN complex is also required; in fact, it has been shown that phosphorylation of CtIP specifically promotes the MRE11 endonuclease activity (Anand et al., 2016; Huertas et al., 2008; Mimitou and Symington, 2008; Peterson et al., 2013; Sartori et al., 2007). Moreover, the tumour suppressor protein Breast Cancer 1 (BRCA1) binds to CtIP-MRN complex to add processivity to DNA end resection (Cruz-García et al., 2014). CtIP possesses a 5’-flap endonuclease

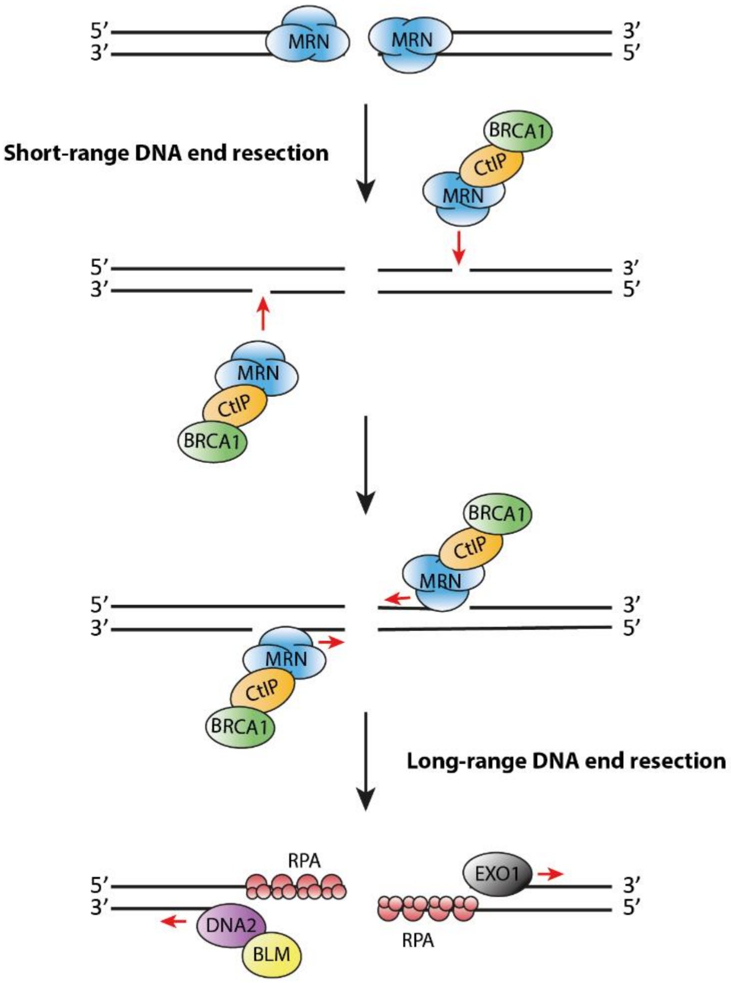

5′
3′
MRN
MRN
3′
5′
Short-range DNA end resection
BRCA1
CtIP
MRN
5′
3′
3′
5′
MRN
CtIP
BRCA1
BRCA1
CtIP
MRN
5′
3′
3′
5′
MRN
CtIP
BRCA1
Long-range DNA end resection
RPA
5′
3′
DNA2
BLM
EXO1
3′
5′
RPA

Figure I4. DNA end resection mechanism.

Homologous recombination (HR) is initiated by DNA-end resection which occurs in two distinct steps. First, the MRN complex recruits CtIP and BRCA1 and starts the short-range DNA end resection with an endonucleolytic cleavage of the 5′-terminated DNA strand mediated by the endonuclease activity of MRN. Then, the 3′ to 5′ exonuclease activity of the MRN complex degrade the DNA toward the DSB end leaving short 3′ ssDNA tails that can be extended by a second phase known as long-range resection. The formation of extensive 3′ ssDNA overhangs by the long-range resection is catalysed by either the exonuclease activity of EXO1 (right) or the combined helicase/nuclease activities of BLM/DNA2 (left), and extensive ssDNA tails are rapidly coated by RPA complex for protection.

activity but there is controversy whether this activity is necessary (Makharashvili et al., 2014; Wang et al., 2014a) or not (Anand et al., 2016) for resection. Upon initiation of resection by MRN complex and CtIP, DNA end resection tracks are extended by a second phase known as long-range resection (Figure I4). Two separate pathways have been described for long-range resection in human cells, dependent on either the exonuclease activity of EXO1 or the combined helicase/nuclease activities of BLM/DNA2 (Gravel et al., 2008; Nimonkar et al., 2011; Zhu et al., 2008). These pathways lead to the formation of extensive 3′ ssDNA overhangs that facilitate successful homology search and repair.

Short-range resection is essential for the repair of DSBs with protein or chemical adducts at the 5′ ends because the nucleases activities of EXO1 and DNA2 cannot directly process these ends. However, it has been shown that resection at a clean DSB with free DNA ends can be initiated by DNA2 via its endonuclease activity. In this model, the MRN complex promotes the recruitment of DNA2 to the DNA end which cleaves the 5′ strand DNA 10–20 nucleotides away from the end approximately (Paudyal et al., 2017).

1.5. DNA DSBs repair pathway choice

The right balance between different DNA DSB repair pathways is critical to ensure the maintenance of genomic stability, as failure to accurate repair is associated with the acceleration of tumorigenesis and several human genetic syndromes (Huertas, 2010). Consequently, it is a complex and highly regulated process, and although the exact mechanisms that control the repair pathway choice are not completely understood yet, multiple layers of control have been involved in this regulation (Chapman et al., 2012).

One of the key events that has a major impact in regulating the balance between HR and NHEJ is DNA end resection which leads to HR activation and impairs NHEJ (Huertas, 2010). DNA end resection is highly regulated and responds to many different cellular signals (Hartlerode and Scully, 2009; Huertas, 2010; Sartori et al., 2007). Regarding regulation, cell cycle has a major influence on the choice of the repair pathway through the cyclin-dependent kinases (CDKs) activity, which mediates phosphorylation of multiple substrates to promote or not end resection (Huertas, 2010; Wohlbold and Fisher, 2009) (Figure I5). In addition, there are accessory factors that also contribute to repair pathway choice via modulation of end resection. For instance, the balance between the pro-resection factor BRCA1 and the anti-resection factors 53BP1, RIF1 and the shieldin complex (SHLD1, SHLD2 and SHLD3) modulate pathway choice by either promoting or preventing end resection (Daley and Sung, 2014; Dev et al., 2018; Escribano-Díaz et al., 2013; Gupta et al., 2018; Noordermeer et al., 2018) (Figure I5). Resection is reduced in G1 phase of the cell cycle and is activated in S/G2, when the sister chromatid is available to perform a precise repair by homologous recombination (Branzei and Foiani, 2008). In G1 cells, 53BP1 is localized at the break and is phosphorylated by ATM which leads to the recruitment of RIF1, an effector that antagonize with BRCA1 (Chapman et al., 2013; Escribano-Díaz et al., 2013; Zimmermann et al., 2013) (Figure I5). Moreover, shieldin complex is recruited to double-strand-break sites in a 53BP1- and RIF1-dependent manner and interacts with RIF1 by SHLD3 subunit (Gupta et al., 2018; Noordermeer et al., 2018). Then, SHLD2 interacts with single-stranded DNA leading to DNA end resection inhibition and directing the repair towards NHEJ, possibly by occluding access to resection nucleases (Noordermeer et al., 2018) or by antagonising the replacement of RPA with BRCA2/RAD51 on resected ssDNA (Dev et al., 2018) (Figure I5). As cells progress to S and G2 phases, CDK phosphorylates CtIP at T847 which is required for DNA end resection initiation (Huertas and Jackson, 2009). Moreover, CDK-mediated phosphorylation of CtIP at S327 promotes its interaction with BRCA1 which in turn antagonize RIF1 (Chen et al., 2008; Escribano-Díaz et al., 2013) which is required for SHLD3 assembly and shieldin function (Gupta et al., 2018) and increase (Gupta et al., 2018) and accelerates the speed of resection (Cruz-García et al., 2014) (Figure I5).

Other factors involved in DSB pathway choice are the regulation of chromatin status, in which histone post-translational modifications (PTM) play a key role (Price and D’Andrea, 2013), the transcriptional status of loci (Aymard et al., 2014), the structure of ends (Liao et al., 2016) or nuclear position (Lemaître et al., 2014).

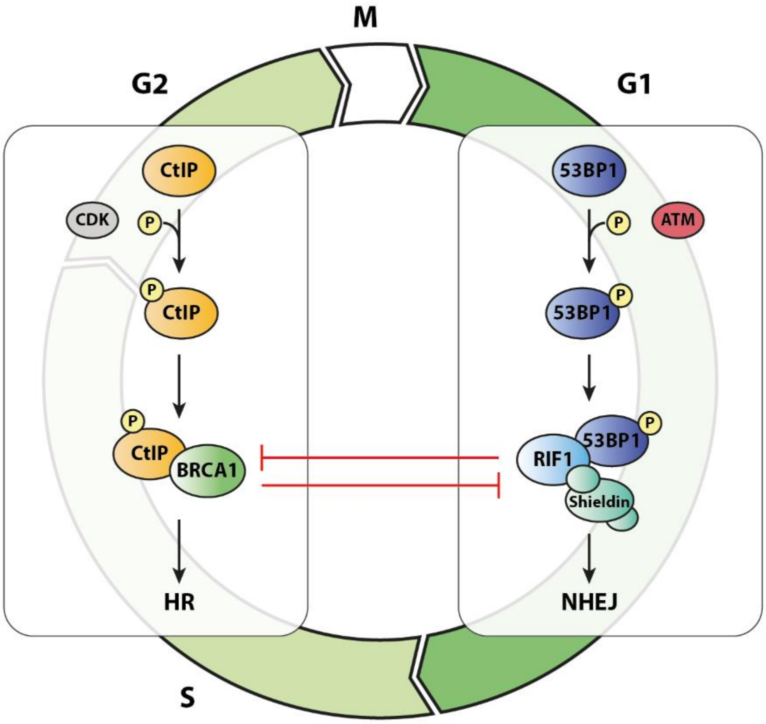


Figure I5. Regulation of DSB repair pathway choice.
In G1 cells, ATM-phosphorylated 53BP1 recruits RIF1 to the break and the shieldin complex interacts with RIF1, which in turn inhibits DNA end resection by antagonising BRCA1 and directs the repair towards canonical NHEJ. During S/G2 phase of the cell cycle, CDK-phosphorylated CtIP both directly increases its resection activity and facilitates interaction with BRCA1 which negatively regulates RIF1 and promotes DNA end resection and therefore, HR.

1.6. CtBP-interacting protein (CtIP)

The human *CtIP* gene, that is located at chromosome 18q11.2, encodes for an 897-amino acid nuclear protein (Fusco et al., 1998) (Figure I6). CtIP was initially identified as a cofactor for the transcriptional repressor CtBP (carboxy-terminal binding protein) (Schaeper et al., 1998). Subsequently, in a yeast two-hybrid screen, CtIP was also identified as a direct interaction partner for other proteins including breast cancer 1 protein (BRCA1) (Wong et al., 1998; Yu et al., 1998) and retinoblastoma protein (Rb) (Fusco et al., 1998), which gives the CtIP alternative name: Retinoblastoma-binding protein 8 (RBBP8). Since its identification, CtIP has been shown to modulate transcription, DNA replication, DNA repair, cell cycle progression and checkpoint activation (You and Bailis, 2010). However, it is now better known for its role in DNA double-strand break (DSB) repair by promoting end resection (Huertas and Jackson, 2009; Sartori et al., 2007).

CtIP functional homologs have been identified in many eukaryotic species: *Saccharomyces cerevisiae* (Sae2), *Schizosaccharomyces pombe* (Ctp1), *Caenorhabditis elegans* (COM-1), *Xenopus* (xCtIP), chicken, mice, and *Arabidopsis thaliana* (AtGR1) (You and Bailis, 2010). Despite low sequence homology, there are conserved sequences limited to small regions in its amino and carboxyl terminal ends (Andres and Williams, 2017) (Figure I6). The N-terminal region of CtIP contains a coiled-coil domain (amino acid residues 45 to 160) that is responsible for protein homodimerization (Dubin et al., 2004) (Figure I6). Recent evidence suggests that CtIP exists as constitutive tetramers. Tetramerization is mediated by short sequence motifs present at the start of parallel coiled-coil segments located in the amino-terminal region. The ends of 2 coiled-coil dimers come together generating a homotetramer, also known as the tetrameric helical dimer of dimers or THDD (Andres et al., 2015; Forment et al., 2015). Moreover, the amino-terminus of CtIP (especially residues 22–45), as well as the carboxy-terminal region (the last 108 residues), are independently capable of binding to MRN, playing a critical role in targeting CtIP to sites of DNA breaks, DNA-end resection and G2/M checkpoint activation (Sartori et al., 2007; Yuan and Chen, 2009) (Figure I6). The C-terminal region also contains a CDK phosphorylation site (T847), essential for CtIP function in DNA end resection, which is functionally analogous to S267 in *S. cerevisiae* Sae2 (Huertas and Jackson, 2009). As well, CCAR2 interacts with CtIP thought the C-terminal part of CtIP (from amino acid 650 to the end of the protein) to block DNA end resection (López-Saavedra et al., 2016) (Figure I6). Additionally, a DNA binding domain is present in the 509-557 region of CtIP, and two key lysine residues conserved in vertebrates within this region (K513 and K515) are important for CtIP recruitment to DNA damage (You et al., 2009) (Figure I6).

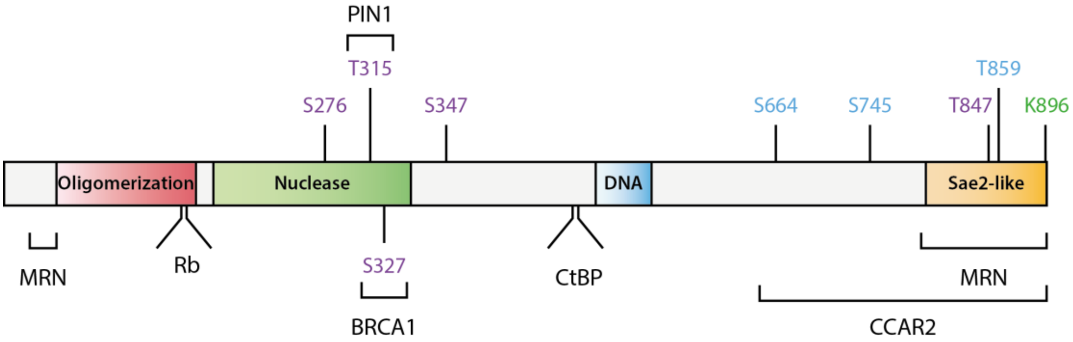


Figure I6. Structure and modifications of human CtIP protein.
The structure of human CtIP presents the oligomerization domain (in red) which contains a coiled-coil domain, the nuclease domain (in green), the DNA binding domain (in blue) and the Sae2-like region which is conserved from yeast to humans (in orange). CtIP is subjected to many different post-translational modifications. Residues subject to different modifications are marked in purple (CDK-mediated phosphorylation), blue (ATM or ATR-mediated phosphorylation) and green (sumoylation). The sites of interaction with other proteins are marked with brackets.

1.7. Functions and regulation of CtIP

A large body of evidence suggests that CtIP coordinates transcription, DNA replication and DNA damage repair through its interactions with diverse families of proteins (Figure I6). CtIP was originally identified as a transcription regulator, and is known to interact with several proteins to act either as a correpressor or a coactivator for the expression of different subsets of genes (Li et al., 1999; Liu and Lee, 2006; Schaeper et al., 1998; Yu et al., 1998). Moreover, CtIP is able to regulate cell cycle progression by its association with Rb. In the G1 phase of the cell cycle, CtIP binds Rb and releases Rb-mediated transcriptional repression of E2F, thus promoting expression of genes required for S phase entry, such as Cyclin D1 (Liu and Lee, 2006). CtIP is also required for DNA damage-dependent cell cycle arrest. BRCA1 interacts at G2 phase with the transcriptional co-repressor complex of CtIP and CtBP through its BRCT domains (Yu et al., 1998) in a manner dependent on the CDK-dependent phosphorylation of CtIP S327 (Figure I6). This phosphorylation of CtIP and its association with BRCA1 facilitate the ATM/ATR-dependent phosphorylation of CtIP after DNA damage (Foray et al., 2003) at S664 and S745, leading to dissociation of the CtIP–CtBP repressor complex from BRCA1, which in turn, activate transcription of cell cycle inhibitor genes required for G2/M checkpoint activation, such as p21 and Gadd45A (Li et al., 2000). Other studies indicate that CtIP participates in the surveillance of ongoing DNA replication and in the processing of DNA intermediates that arise during replication. Accordingly, CtIP directly localizes at DNA replication foci by direct interaction with PCNA through a Replication Foci Targeting Sequence (RFTS). The CtIP-PCNA interaction is implied in the stabilization of stalled replication forks (Gu and Chen, 2009). Additionally, CtIP has been found to participate in the detection and resolution of replication stress caused by interstrand cross-link (ICL) by directly binding with components of the FANC complex (Duquette et al., 2012; Murina et al., 2014; Unno et al., 2014). Despite its contribution in transcription, replication and cell cycle, the best known role of CtIP is to promote initiation of DNA end resection and consequently, activate the repair by HR (Huertas and Jackson, 2009; Sartori et al., 2007; You et al., 2009). Interaction between CtIP and the MRN complex is necessary for this function (Sartori et al., 2007), but the exact mechanism is not fully understood. At 'clean' broken ends, the MRN-CtIP complex seems to promote the recruitment of DNA2 which initiates DNA end resection (Paudyal et al., 2017). In contrast, at DNA ends with an altered structure like topoisomerase adducts, there is controversy about the role of CtIP. On the one hand, CtIP has been proposed to stimulate the endonuclease activity of MRE11 (Anand et al., 2016). On the other hand, it has been suggested that CtIP possess an endonuclease activity that is required for the processing of these 'dirty' ends (Makharashvili et al., 2014).

CtIP is subjected to many different post-translational modifications which control the interaction with other proteins and its multiple roles in different branches of DNA metabolism (Makharashvili and Paull, 2015) (Figure I6). Among these modifications, CtIP is known to be a target of phosphorylation induced by CDK, as well as ATM and ATR kinases (Huertas and Jackson, 2009; Peterson et al., 2013; Sartori et al., 2007; Wang et al., 2013a; Yu and Chen, 2004) (Figure I6). Importantly, CDK-mediated phosphorylation of T847 is essential for DNA end resection initiation (Huertas and Jackson, 2009), whereas CDK-dependent phosphorylation of S327 is required for its interaction with BRCA1 (Yu and Chen, 2004) which accelerates the speed of resection (Cruz-García et al., 2014). CDK2-mediated phosphorylation of T315 and S276 phosphorylation are involved in PIN1-CtIP interaction and PIN1-mediated isomerization of CtIP respectively (Steger et al., 2013). This modification leads to CtIP polyubiquitylation and trigger its degradation. Moreover, a cluster of CDK-mediated phosphorylation sites situated in the middle of CtIP is involved in DNA end resection. Phosphorylation of the middle cluster promotes the interaction of CtIP with NBS1, which is a prerequisite for DNA damage-induced CtIP phosphorylation by ATM (Wang et al., 2013a). The ATM phosphorylation sites on CtIP, S664 and S745 (Figure I6), were reported to block interaction with BRCA1 after DNA damage (Li et al., 2000) and ATR-dependant phosphorylation of T859 (Figure I6) in response to damage is essential for CtIP recruitment to DSB and, consequently, for DNA end resection and HR (Peterson et al., 2013). Nevertheless, recruitment of CtIP to break sites also depends on two ubiquitin-mediated post-translational modifications (Schmidt et al., 2015; Yu et al., 2006). CtIP has also been shown to be constitutively sumoylated by CBX4 at K896 (Figure I6) which is essential for its recruitment to damaged DNA (Soria-Bretones et al., 2017)

2. DNA Damage and genomic instability in Stem Cells

2.1. Types of Stem Cell

Stem cells (SCs) are non-specialized cells found in all multicellular organisms from the early stages of development to the end of life. They are characterized by the ability to unlimited or prolonged self-renew and the capacity to differentiate into a diverse range of specialized cell types, thus to form tissues and organs (Hui, Hongxiang; Tang, Yongming; Hu, Min; Zhao, 2011). There are many types of stem cells that can be categorized according to their differentiation potential and sources. While some of them are physiological stem cells (embryonic and adult stem cells), others are artificial (induced pluripotent stem cells) or pathological (cancer stem cells) (Figure I7).

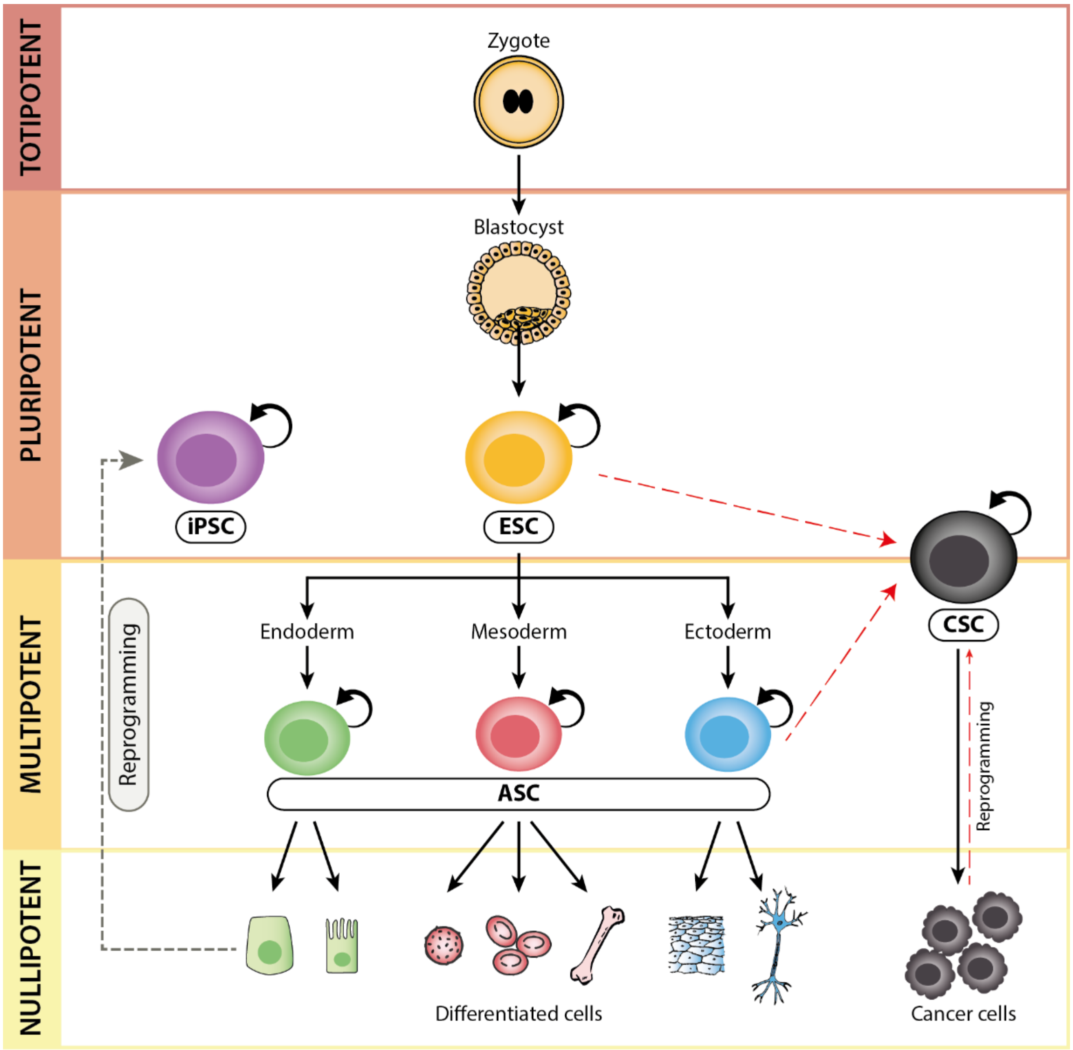

TOTIPOTENT
PLURIPOTENT
MULTIPOTENT
NULLIPOTENT
Zygote
Blastocyst
iPSC
ESC
CSC
Reprogramming
Endoderm
Mesoderm
Ectoderm
ASC
Reprogramming
Differentiated cells
Cancer cells

Figure I7. Types of stem cells

A totipotent cell, such as the zygote, has the potential to create an entire organism. The totipotent zygote undergoes several mitotic divisions to form the blastocyst. The blastocyst has the inner cell mass from which derive the embryonic stem cells (ESC). ESCs are pluripotent and have the capacity to differentiate into cells of all three dermal layers (endoderm, mesoderm and ectoderm). Adult stem cells (ASC) are more specialized cells that are produced from pluripotent stem cells and are present in all tissues. They are multipotent and have the capacity to differentiate into multiple specialised cell types of their tissue of origin. Finally, differentiated cells are nullipotent and are not capable to give rise to other cell types. There are other types of stem cells that are artificial (induced pluripotent stem cells) or pathological (cancer stem cells). The induced pluripotent stem cells (iPSC) are pluripotent cells generated by reprogramming from differentiated cells. Cancer stem cells (CSC) can be created from stem cells (ESC, ASC) which obtain the ability to generate tumours due to a genetic mutation or can be reprogrammed from cancer cells (red arrows). The turning arrow indicates the self-renewal capacity of stem cells.

On the one hand, embryonic stem cells (ESCs) are derived from the inner cell mass of the blastocyst after approximately five days of development (Evans and Kaufman, 1981; Thomson, 1998). ESCs are considered to be pluripotent as they are able to differentiate into cells derived from the three germ layers (ectoderm, endoderm, and mesoderm) and they can also be maintained in an undifferentiated state for a prolonged period in culture under defined conditions (Yao et al., 2006) (Figure I7). Several factors have been associated with the stemness state and are considered as pluripotency markers. Among them, NANOG and OCT4 are essential transcription factors that regulate self-renewal and pluripotency of ES cells (Liang et al., 2008; Shi and Jin, 2010).

On the other hand, adult or somatic stem cells (ASCs) are found among differentiated cells in postnatal tissues. They have the ability to renew themselves and they are multipotent, therefore they can differentiate into specialized cell types of their tissue of origin (Figure I7). They play important roles contributing to tissue homeostasis and repairing adult tissues in which they reside (Alison and Islam, 2009). Adult stem cells can be found in a number of various tissues of the adult organism including bone marrow, peripheral blood, brain, spinal cord, dental pulp, blood vessels, skeletal muscle, epithelia of the skin and digestive system, cornea, retina, liver, and pancreas (Alison and Islam, 2009).

Induced pluripotent stem cells (iPSC) are the third type of stem cells. They are an artificial type of pluripotent stem cells that have similar properties to embryonic stem cells but they are derived from a non-pluripotent cell, typically an adult somatic cell, by a process called reprogramming (Takahashi and Yamanaka, 2006) (Figure I7). The first evidence that differentiation of cells is reversible was reported in 1962 when Gurdon could develop normal tadpoles by transplanting the nucleus from a mature intestinal cell of *Xenopus laevis* tadpoles into enucleated eggs (Gurdon, 1962). These experiments were essential for following studies and later, in 2006, mouse somatic cells were reprogrammed into induced pluripotent stem cells by simple expression of four defined transcription factors: octamer-binding transcription factor 3/4 (OCT3/4 or OCT4), SRY-related high-mobility group box protein-2 (SOX2), the oncoprotein c-MYC, and Krüppel-like factor 4 (KLF4) (Takahashi and Yamanaka, 2006) (Figure I8). iPSCs have also been derived from a

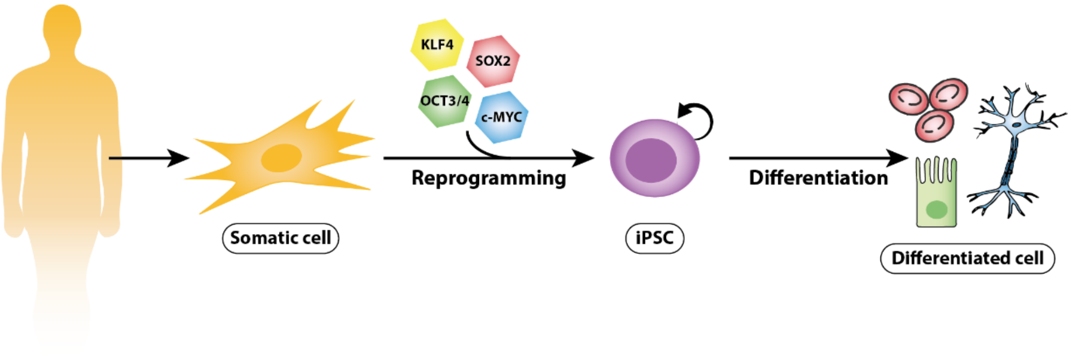


Figure I8. General diagram of reprogramming process and differentiation
Somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs) by using the four Yamanaka transcription factors: OCT3/4, SOX2, KLF4 and c-MYC. The resulting iPSCs can be expanded (represented as a turning arrow) and differentiated into the desired cell type from any of the three germ layers of the human body.

number of different species, including humans (Takahashi et al., 2007; Yu et al., 2007), rats (Li et al., 2009), rhesus monkeys (Liu et al., 2008), pigs (Ezashi et al., 2009) and dogs (Shimada et al., 2010) among others, by expression of the so-called four Yamanaka factors. iPSCs are currently useful tools for drug development, modelling of diseases, and regenerative medicine, and potentially have therapeutic uses without the controversial use of embryos and avoiding immune rejection. Nevertheless, regardless of the methods used to reprogram adult cells, iPSCs still present significant risks that limit its clinical application.

Finally, emerging evidence has pointed out a sub-group of stem-like cells within tumours, known as cancer stem cells (CSC), which exhibit characteristics of both cancer cells and stem cells such as self-renewal ability and multi-lineage differentiation to drive tumour growth and heterogeneity. Most CSCs are believed to exhibit resistance mechanisms against conventional anti-cancer therapies such as radiation and chemotherapy (Bao et al., 2006; Jeon et al., 2011) eventually resulting in tumour recurrence. The first evidence of CSCs existence came in 1994 with a study of human acute myeloid leukaemia where a subpopulation of cells CD34+/CD38- capable of initiating tumours were isolated (Lapidot et al., 1994). Two different theories about the origin of

CSCs have been suggested: one theory believes that CSCs arise from stem cells which obtain the ability to generate tumours due to a genetic mutation or environmental alteration (Rubio et al., 2005; Shiras et al., 2007); the other suggests that cancer cells can be reprogrammed to become CSCs (Eun et al., 2017) (Figure I7).

Stem cells, including ESCs and iPSCs, harbour significant potential for regenerative medicine, but prior to its use in clinic, ensuring genomic stability of stem cells is required for safe clinical application (Vitale et al., 2017).

2.2. DNA damage repair in Stem cells.

Stem cells are vital for the generation and maintenance of intercellular heterogeneity and tissue homeostasis. Nevertheless, SCs are particularly exposed to DNA damage of endogenous or exogenous origin (Blanpain et al., 2011; Vitale et al., 2017), which can have catastrophic consequences for the tissues and lead to embryonic lethality, developmental defects, aging-related degenerative disorders, and oncogenesis (Behrens et al., 2014). Nevertheless, both pluripotent and multipotent stem cells have a large capacity to repair and prevent the accumulation of DNA lesions or avoid their propagation to daughter cells by their DNA damage response and apoptosis induction capacity which vary in the different types of stem cells (Vitale et al., 2017). Thus, SCs rely on a very robust DNA damage response (DDR), which is an important mechanism for stem cell safety and efficacy for regenerative purposes.

2.2.1 DNA damage repair in Adult Stem cells.

Adult stem cells (ASCs) resident in adult tissues possess high proliferative capacity allowing to carry out regenerative activities in response to tissue damage throughout the life of the organism. ASCs are susceptible to DNA damage due to constant genotoxic stress from endogenous processes, which leads to the activation of the DNA damage response and initiates DNA repair, cell cycle arrest, and ultimately apoptosis or cellular senescence.

Nevertheless, the mechanisms of response to DNA damage vary between tissues (Blanpain et al., 2011).

Hematopoietic stem cells (HSCs), which regenerate the blood system, are one of the best-studied adult tissues in terms of its hierarchical development and the properties of HSCs in each site differ (Orkin and Zon, 2008). HSCs exist in two dynamic pools: one of active cells that proliferate and progressively differentiate; and quiescent cells which persist for prolonged periods in a reversible non-proliferating state that ensure the lifelong production of all the diverse hematopoietic cell types (Orkin and Zon, 2008). HSCs are very sensitive to the acquisition of mutations upon exposure to endogenous and exogenous sources of DNA damage, including replication stress and microenvironmental genotoxins. Such pronounced sensitivity is, at least in part, due to the lack of an efficient G1-S DNA damage checkpoint (Moehrle et al., 2015). Stem cell quiescence has been shown to be one of the mechanisms that protect adult HSCs function, since there is evidence that with each round of division the HSC function is deteriorated, so older HSCs are less capable to regenerate the blood system (Beerman et al., 2013; Chambers et al., 2007). This state of quiescence results in low metabolic activity, low levels of reactive oxygen species (ROS) and less replication to protect HSCs from the accumulation of endogenous DNA damage and telomere shortening (Pietras et al., 2011; Walter et al., 2015; Wilson et al., 2008). However, there is also evidence that quiescent stem cells are also vulnerable to the accumulation of DNA damage and mutations because they employ the error-prone NHEJ pathway to repair DSBs, while proliferating HSCs use the high-fidelity HR (Mohrin et al., 2010). To ensure that damage accrued during quiescence is repaired prior to differentiation or self-renewal divisions, cells exit from G0 and entry to cell cycle which promotes the induction of DNA repair and the switch of the repair mechanism from NHEJ toward HR (Beerman et al., 2014; Walter et al., 2015) or the combination of both pathways.

On the other hand, non-hematopoietic adult stem cells prevent the propagation of DNA damage differently depending on the tissue of origin. For instance, mammary stem cells (MaSCs) have been shown to possess more resistance to apoptosis and they efficiency repair DSBs by increasing NHEJ activity. This is possibly due to the upregulation of 53BP1, a pro-NHEJ protein, and downregulation of pro-apoptotic proteins like BAX and Caspase-3 (Chang et al., 2015). Hair follicle stem cells (HFSCs) are also more resistant to cell death due to the upregulation of BCL-2 and they use NHEJ preferentially, at least during quiescence (Sotiropoulou et al., 2010). Upon activation of SCs, BRCA1 becomes essential, so it is possible that HR and NHEJ are differentially used at distinct stages of development and activation (Sotiropoulou et al., 2013). DNA damage in mesenchymal stem cells (MSCs) activates key repair pathways and cell cycle checkpoints (Prendergast et al., 2011). This high resistance to DNA damage seems to depend on NHEJ since there is an increase in the level of KU70 and in the phosphorylation of DNA-PKcs as well as ATM after damage by radiation (Oliver et al., 2013). Skeletal muscle stem cells or satellite cells, accurately repair DSBs by NHEJ which depends on DNA-PKcs (Vahidi Ferdousi et al., 2014). Unlike hematopoietic stem cells, where substantial genomic rearrangements and mutations are observed, skeletal muscle stem cells have the capacity to perform accurate and high efficient repair (Mohrin et al., 2010; Vahidi Ferdousi et al., 2014). Neural stem cells (NSCs) were also shown to be more resistant to DNA DSBs than differentiated neurons. During nervous system development, each DNA DSB repair pathway is required at different stages, being HR particularly important for proliferating cells and NHEJ for differentiating cells (Lee and McKinnon, 2007). In fact, NSCs carry out HR to repair DNA damage occurring during S and G2 phases (Rousseau et al., 2012) and high levels of DNA-PKcs expression and activity were found in NSPCs and neurons indicating that NHEJ activity also occurs (Kashiwagi et al., 2018). In the same manner, ASCs resident in the rest of tissues of the organism activate different DNA damage responses to repair DNA lesions. Despite such control and activation of DNA damage repair pathways, ASCs tend to accumulate mutations and chromosomal abnormalities (Ben-David et al., 2011).

2.2.2. DNA damage repair in Embryonic Stem cells.

Embryonic stem cells show a high proliferation rate and the time required to complete a full cell cycle is shorter than differentiated cells. The cell cycle in ESCs displays abbreviated G1 and G2 phases and so, S phase occupies a large proportion of time (50-

70%) (White and Dalton, 2005). ESCs are susceptible to generation of endogenous damage from replication errors due to the shorter G1 phase which causes replication stress in the following S phase (Ahuja et al., 2016). In addition, it has been reported that mouse ESCs lack a G1 checkpoint following DNA damage, which can be explained by two theories. One supports that p53 is not activated after DNA damage which prevents the transcription of target genes like the cell cycle inhibitor p21 (Suvorova et al., 2016). The other is that Chk2 does not phosphorylate Cdc25a which promotes an increase in Cdk2 activity and thus the cell cycle continues (Hong and Stambrook, 2004). Human ESCs has also been shown to fail in G1 checkpoint activation in response to ionizing radiation, but can temporarily arrest in the G2 phase upon ATM activation (Filion et al., 2009; Momčilović et al., 2009).

In spite of this, the accumulation rate of mutations in ESCs is low (Cervantes et al., 2002), suggesting that ESCs may use mechanisms to ensure genomic stability. ESCs employ two main strategies to diminish DNA damage and genomic instability. On one hand, more efficient DNA repair activity is presented in ESCs compared with differentiated cells (Maynard et al., 2008) and, on the other hand, apoptotic or differentiation programs are activated to eliminate cells with damaged DNA from the stem cell population (Qin et al., 2007). In response to DSBs, ESCs preferentially employ HR over the error-prone NHEJ (Adams et al., 2010a; Serrano et al., 2010; Tichy et al., 2010). In fact, the levels of proteins involved in HR have been shown to be highly elevated in ESCs (Adams et al., 2010a; Serrano et al., 2010; Tichy et al., 2010). In addition to HR-mediated DSB repair, it has been shown that human ESCs can also use high fidelity NHEJ through a DNA-PK-independent mechanism to repair DSBs induced by I-SceI endonuclease (Adams et al., 2010b) or DNA-PK-dependent NHEJ mechanism for repair of radiation-induced DSBs during the late G2 stage of the cell cycle (Bogomazova et al., 2011). The other mechanisms that ESCs use to avoid the propagation of genetic lesions are to initiate apoptosis or to lose pluripotency by differentiation (Qin et al., 2007). In response to DNA damage, human ESCs undergo cell death via Caspase-related mitochondrial apoptosis dependent of p53 (Filion et al., 2009; Grandela et al., 2007; Qin et al., 2007) while mouse ESCs mainly undergo apoptosis via p53-independent mechanisms (Aladjem et al., 1998). Moreover, p53 activation leads to differentiation by suppressing the expression of pluripotency genes, such as *NANOG* (Lin et al., 2005; Qin et al., 2007).

2.2.3. DNA damage repair in iPSCs and the reprogramming process.

One of the major concerns about the use of induced pluripotent stem cells (iPSCs) in clinic applications is the accumulation of diverse genetic abnormalities (Mayshar et al., 2010; Peterson and Loring, 2014). These genetic variations, such as aneuploidies, chromosomal copy number variations (CNVs) and single nucleotide variations (SNVs), can come from different sources during iPSC generation and maintenance. For instance, they can be inherited from reprogrammed somatic cells, induced by the reprogramming process or accumulated during subsequent expansion of iPSCs in culture (Liang and Zhang, 2013). iPSCs were initially derived integrating the four factors OCT4, SOX2, KLF4, and c-MYC (OSKM) (Figure I8) by retroviral transduction but since then, many methods have been developed and modified to increase efficiency and minimize or remove the integration of vector sequences into the iPSC genome (Malik and Rao, 2013). However, genomic instability is intrinsically associated with the reprogramming process itself and does not depend on the method (Ben-David and Benvenisty, 2012; Gore et al., 2011; Mayshar et al., 2010). Reprogramming factors accelerate growth rate, so in iPSCs, the metabolisms change from oxidative respiration to oxidative glycolysis (Folmes et al., 2011; Varum et al., 2011). This can increases the oxidative stress by accumulation of reactive oxygen species (ROS) and replication stress by collapsed replication forks, leading to DNA damage (Banito et al., 2009; Esteban et al., 2010; Ruiz et al., 2015). Moreover, the reprogramming process also induces DNA DSBs independently of the method used (González et al., 2013).

iPSCs share numerous resemblances with ESCs, such as the susceptibility to accumulate DNA lesions, a shortened G1 phase (Ghule et al., 2011) and the lack of G1 checkpoint (Momcilovic et al., 2010). Furthermore, iPSCs also share numerous similarities in their DDR with ESCs, including G2/M cell cycle arrest, efficient DNA repair and high expression of genes involved in DNA damage signalling and repair (Momcilovic et al., 2010). It has been found higher expression of several HR (*RAD52, MRE11, NBS1*) and NHEJ (*XRCC4 and LIGIV*) genes (Momcilovic et al., 2010). Different DDR pathways have also been found to be involved during the reprogramming process. In fact, ATM deficiency impairs reprogramming efficiency and increases genomic instability in iPSCs (Kinoshita et al., 2011). Furthermore, reprogramming process has also been observed to require both HR and NHEJ genes and a decrease in reprogramming efficiency has been reported due to the lack of HR genes, including *BRCA1, BRCA2*, and *RAD51* (González et al., 2013), and NHEJ genes, including *LIGIV* and *DNA-PKc* (Felgentreff et al., 2014; Tilgner et al., 2013). Reprogramming is limited to ensure iPSC genomic integrity by a p53-mediated DNA damage response. The activation of p53 acts as a barrier for cell reprogramming by inducing apoptosis to eliminate suboptimal cells carrying DNA damage (Marión et al., 2009). Defects in p53 have been shown to improve reprogramming efficiency but also increase genetic instability (Sarig et al., 2010).

2.2.4. DNA damage repair in Cancer Stem Cells.

Similar to ASCs, cancer stem cells (CSCs) can exist in a cycling and quiescent state (Kreso and Dick, 2014). The dormant cells were shown to survive to conventional chemotherapy and contribute to tumour recurrence (Kreso and Dick, 2014). CSCs also present a robust DNA damage response compared to differentiated cancer cells. They possess a highly efficient ability to protect cellular DNA by activating DNA repair pathways, including BER, NER, MMR and DSBs repair, to process DNA damage. Different types of CSCs have shown an increased efficiency of DSBs repair by HR (Lim et al., 2012), as well as NHEJ (Yuan et al., 2014), and they display increased levels of multiple repair components (Ahmed et al., 2015; Bao et al., 2006; Desai et al., 2014). This efficient DDR appears to confer radioresistance (Ahmed et al., 2015; Bao et al., 2006; Desai et al., 2014; Yuan et al., 2014). Moreover, CSCs activate intracellular systems to eliminate ROS in order to safeguard themselves from damage or cell death caused by ROS (Diehn et al., 2009; Nagano et al., 2013). In this way, they maintain low ROS levels and redox homeostasis which are required for CSCs self-renewal (Wang et al., 2013b).

Such robust activation of these pathways in CSCs permits to put up with high levels of replication stress and survive to DNA damaging agents. Therefore, specific inhibition of these repair pathways together with the treatment with radiotherapy or chemotherapy could improve therapy response and clinical outcome.

2.3. The reprogramming factor Krüppel like factor 4

The Krüppel-like factor 4 was initially identified by two groups who differently named it gut enriched Krüppel like factor (GKLF) because of its high expression in the gastrointestinal tract (Shields et al., 1996), and epithelial zinc finger (EZF) as it was highly expressed in the skin epithelium (Garrett-Sinha et al., 1996). Later, GKLF/EZF was renamed Krüppel like factor 4 (KLF4) to avoid confusion since it has been also found in other tissues including lung (Jean et al., 2013), male and female reproductive tracts (Behr and Kaestner, 2002; Godmann et al., 2010), thymus (Panigada et al., 1999), lymphocytes (Fruman et al., 2002), cornea (Chiambaretta et al., 2004) and cardiac myocytes (Cullingford et al., 2008). It was named Krüppel like due to its high homology with the *Drosophila* Krüppel gene involved in segmentation of the developing embryo (Preiss et al., 1985).

The human *KLF4* gene that is mapped on chromosome 9q31, encodes for a 513-amino acid protein. The gene is conserved among vertebrates species, having 91% sequence similarity between human and mouse *KLF4* (Yet et al., 1998). KLF4 belong to the Sp/XKLF family, a group of transcription factors containing three C2H2-type zinc fingers, which play an important role in many fundamental biologic processes (Philipsen and Suske, 1999). Several domains have been characterized in KLF4, which include a transcriptional activation domain within its amino terminus, a repression domain next to it (Geiman, 2002; Yet et al., 1998), followed by a NLS (nuclear localization signal) sequence having PKRGRR repeats (Shields and Yang, 1997) and the carboxyl DNA binding domain which has the three C2H2 zinc finger motifs (Philipsen and Suske, 1999) (Figure I9). Each C2H2 zinc finger consists of two antiparallel β-strands followed by an α-helix. Although one report suggests that KLF4 prefers to bind a RRGGYGY sequence (R=A/G and Y=C/T) (Shields and Yang, 1998), in general, KLF4 interacts with CACCC elements on target genes (Yet et al., 1998). In addition to these domains, a PEST sequence, that may act as a signal peptide for protein degradation, is reported to be located between the activation and inhibitory domains (Chen et al., 2005b) (Figure I9).

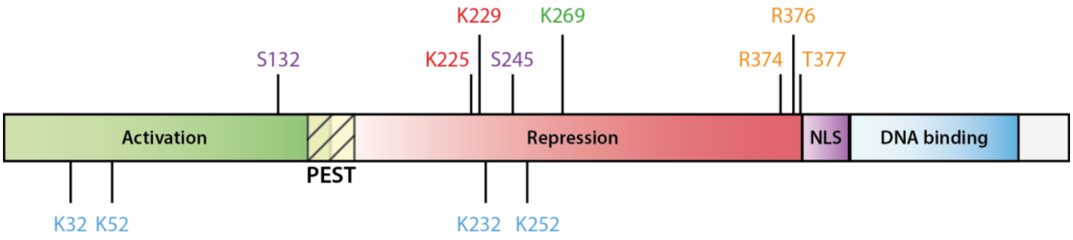


Figure I9. Structure and modifications of human KLF4 protein.
The structure of human KLF4 presents an activation domain (in green), a repression domain (in red), a PEST sequence that may act as a signal peptide for protein degradation (yellow box with diagonal lines), a nuclear localization signal (in purple) and a DNA binding domain which contains three C2H2 zinc finger (in blue). KLF4 is subjected to many different post-translational modifications. Residues subject to different modifications are marked in purple (phosphorylation), blue (ubiquitination), red (p300/CBP-mediated acetylation), green (PIAS1-mediated sumoylation) and orange (PRMT5-mediated methylation).

2.4. Regulation of KLF4

Expression of KLF4 is highly regulated at both transcriptional and post-transcriptional levels. Previous studies have shown that hypermethylation in the KLF4 promoter downregulates its expression in many types of cancer (Cho et al., 2007; Nakahara et al., 2015; Yang and Zheng, 2014). Another important mechanism in modulating the expression of KLF4 are micro-RNA (Tian et al., 2010; Xu et al., 2009). Furthermore, many others stimuli, like serum starvation (Chen et al., 2005b), oxidative stress (Cullingford et al., 2008), or DNA damage (Zhou et al., 2009), can take part in KLF4 regulation.

At post-translational levels, KLF4 is also known to be subjected to multiple modifications including phosphorylation, sumoylation, acetylation and methylation (Figure I9). Phosphorylation at S132 negatively regulates KLF4 resulting in induction of ESC differentiation and triggering the ubiquitination and degradation of Klf4 (Kim et al., 2012). Phosphorylation at S245 results in suppression of its transcriptional activities (Gunasekharan et al., 2016). KLF4 is acetylated by p300/CBP at K225 and K229 which regulates the ability of KLF4 to transactivate target genes and to inhibit proliferation (Evans et al., 2007). Mouse KLF4 has been shown to be both sumoylated at a single lysine residue (K275) and interacts physically with SUMO-1 via SUMO-interacting motif (SIM) which is required for KLF4- transactivation of target genes (Du et al., 2010). In human, K269 sumoylation inhibits KLF4-dependent transcription (Tahmasebi et al., 2013). Furthermore, sumoylation of human KLF4 mediated by PIAS1 promotes its degradation (Kawai-Kowase et al., 2009). Ubiquitination of multiple lysine residues, including K32, K52, K232 and K252, has been shown to be responsible for proteasomal degradation of KLF4 (Lim et al., 2014). Recently, it has been published that KLF4 is methylated by the protein arginine methyltransferase PRMT5 at R374, R376 and R377, which stabilizes KLF4 (Hu et al., 2015).

2.5. Role of KLF4 in biological processes and cell homeostasis

KLF4 is a transcription factor that can both activate or repress transcription depending on the target gene. Initially, KLF4 was associated with growth arrest (Shields et al., 1996) and later, it was shown that KLF4 inhibits cell proliferation by up- or down-regulating some genes involved in cell cycle control (Chen et al., 2003) including, for example, downregulation of cyclin D1 expression (Chen et al., 2003; Shie, 2000) and upregulation of cyclin-dependent kinase inhibitor 1 (CDKN1A/p21) (Chen et al., 2003; Zhang et al., 2000). Moreover, KLF4 has been shown to block G1/S progression of the cell cycle (Chen et al., 2001) and it is also required for p53-dependent G1/S (Yoon et al., 2003) and G2/M (Yoon and Yang, 2004) checkpoints in response to DNA damage. KLF4 is also involved in the decision of cells to undergo apoptosis or not. In fact, KLF4 has been found to suppress p53-dependent apoptotic pathway (Talmasov et al., 2015; Zhou et al., 2009), but under certain conditions and contexts, KLF4 may switch to a pro-apoptotic role (Li et al., 2010b; Wang et al., 2015). After DNA damage induced by γ-radiation, KLF4 was found to prevent centrosome amplification by suppressing cyclin E, which control centrosome duplication together with Cdk2 (El-Karim et al., 2013; Yoon et al., 2005). Finally, KLF4 appears to be important to regulate inflammation (Feinberg et al., 2005).

2.6. Role of KLF4 in stem cells and reprogramming

KLF4 is one of the four factors used to induce pluripotency in somatic cells since 2006, when Takahashi and Yamanaka demonstrated for the first time that mouse fibroblasts could be reprogrammed into induced pluripotent stem cells (Takahashi and Yamanaka, 2006). Nevertheless, the precise role of KLF4 in cell reprogramming is not clear. Indeed, in 2007 human iPSC were produced using NANOG and LIN28 instead of c-MYC and KLF4 (Yu et al., 2007). Therefore, it has even been suggested that KLF4 is not absolutely required but it might increase the reprogramming efficiency.

Recently, it has been proposed that KLF4 plays a dual role by repressing factors associated with differentiation at early time points and activating pluripotency targets at late time points of reprogramming process (Polo et al., 2012). Moreover, overexpression of KLF4 inhibits differentiation of stem cells, suggesting a role in self-renewal (Li et al., 2005). The self-renewal ability and pluripotency characterize stem cells and it has been shown that in culture it is required the leukaemia inhibitory factor (LIF) to maintain stem cells in an undifferentiated state (Smith et al., 1988), through activation of STAT3 (Matsuda et al., 1999). In response to LIF, KLF4 expression is activated rapidly and KLF4 forms a complex with OCT4 and SOX2 to regulate NANOG expression (Zhang et al., 2010). NANOG is well established as a key factor in maintenance of stemness (Mitsui et al., 2003). Therefore, KLF4 acts as a mediator between LIF-STAT3 and NANOG to regulate self-renewal and pluripotency (Zhang et al., 2010).

2.7. Role of KLF4 in cancer

Numerous evidences have implicated KLF4 in the development and progression of many cancers, acting as both an oncogene or a tumour suppressor depending on the cellular, tissue, and genetic contexts (Rowland et al., 2005).

KLF4 expression is frequently decreased in various human cancer types, such as colorectal (Zhao et al., 2004), gastric (Wei et al., 2005), lung (Hu et al., 2009), intestinal

(Ton-That et al., 1997) or prostate cancer (Jerónimo et al., 2011). In colorectal and gastric cancers, KLF4 was shown to undergo promoter hypermethylation, mutations in the open reading frames and loss of heterozygosity (Wei et al., 2005; Zhao et al., 2004). This downregulation could contribute to cellular hyperproliferation and malignant transformation, supporting the role of KLF4 as a tumour suppressor and consistent with its role in cell cycle arrest and growth inhibition. Furthermore, KLF4 overexpression inhibits colony formation, migration, and invasion in colon cancer cells (Dang et al., 2003).

Although KLF4 can acts as tumour suppressor in multiple tissues, KLF4 has also been identified as an oncogene and high levels of KLF4 expression have been found in primary breast ductal carcinoma (Foster et al., 2000; Pandya et al., 2004) and squamous cell carcinoma (Foster et al., 2005; Huang et al., 2005). It has been shown that elevated KLF4 expression is detected in up to 70% of breast carcinomas and nuclear localization of KLF4 is associated with a more aggressive phenotype and poor prognosis (Pandya et al., 2004). Moreover, KLF4 is essential for the maintenance of breast cancer stem cells and cell migration and invasion (Yu et al., 2011).

II. Objectives

Objectives

1. Analysis of the requirement for DNA end resection and its major regulator, CtIP, during the reprogramming process.

2. Investigation of the consequences of CtIP downregulation during cell reprogramming for self-renewal and differentiation.

3. Study the role of the pluripotency factor KLF4 in DNA double strand breaks repair.

III. Results

CHAPTER 1

1. DNA end resection increases upon mouse cell reprogramming

Maintenance of genomic stability is a key element for the proper self-renewal and differentiation capacities of stem cells, being crucial for the suitable use of induced pluripotent stem cells (iPSCs) in clinic applications (Peterson and Loring, 2014). Given that cellular reprogramming increases DNA damage and genetic instability, mainly by replication stress (Ruiz et al., 2015) after the forced expression of Yamanaka factors (OCT4, SOX2, KLF4, and c-MYC), we wonder if DNA end resection, an essential step in Homologous Recombination, could be required for the reprogramming to iPSC.

To address this objective, we used a previously reported mouse embryonic fibroblast (MEF) cell line bearing a doxycycline-inducible system with the four murine reprogramming factors *Klf4, Oct4, Sox2*, and *c-Myc* to carry out the reprogramming process (Abad et al., 2013). This cell line will be abbreviated as MEFs-i4F hereinafter. Thus, we generated mouse induced pluripotent stem cells (miPSCs) by doxycycline treatment and we analysed cellular levels of DNA end resection in MEFs-i4F and their corresponding iPSCs. For this, we developed a novel strategy as a readout of DNA end resection based on a long pulse of bromodeoxyuridine (BrdU) to allow one strand of every DNA molecule to be labelled and its subsequent detection by fluorescence-activated cell sorting (FACS)

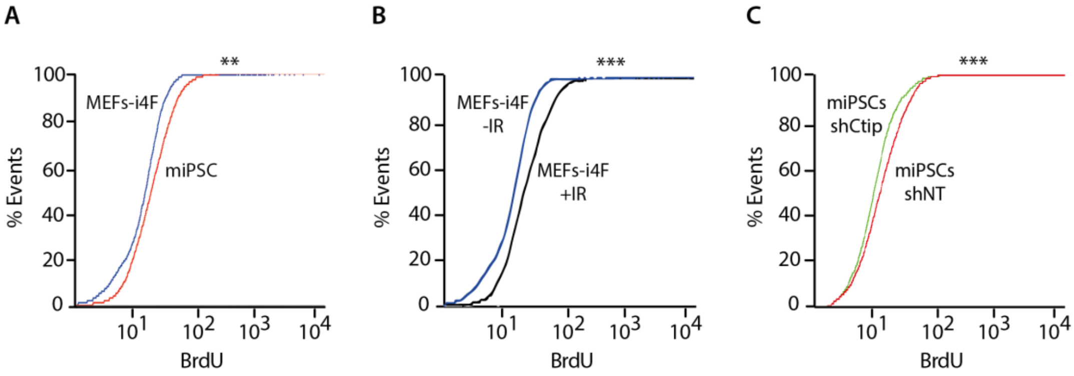


Figure R1. DNA end resection increases in reprogrammed cells
(A) Fluorescence-activated cell sorting (FACS) analysis of BrdU exposed under native conditions after the formation of ssDNA in MEFs-i4F (blue line) and their respective reprogrammed miPSCs (red line). p values were calculated using the Kolmogorov-Smirnov test by CellQuest Pro software (* $p<0.05$; ** $p<0.01$; *** $p<0.001$). A representative graph out of at least three independent experiments is shown. **(B)** Same as (A) but in MEFs-i4F either untreated (-IR; blue line) or 1 h after irradiation with 10 Gy (+IR; black line). **(C)** Same as (A) but in miPSCs reprogrammed upon depletion of *Ctip* (shCtip; green line) or expressing shRNA control (shNT; red line).

analysis using native conditions. BrdU epitope is hidden in double-stranded DNA and thereby it is not accessible for an anti-BrdU antibody under native conditions unless the epitope is exposed after the formation of ssDNA by resection. With this new method we could observe more exposed BrdU in miPSCs than in primary MEFs-i4F, showing that a higher amount of endogenously arising ssDNA appeared in reprogrammed cells (Figure R1A). In fact, the increased BrdU signal intensity in miPSCs following reprogramming was comparable to that seen after treating primary cells with high doses of ionizing radiation (IR) (Figure R1B). We confirmed that such excess ssDNA that appear during reprograming was due to canonical DNA end resection, as it disappeared when the key resection factor CtIP was depleted (Figure R1C).

As BrdU detection by FACS is a low-resolution technique to study DNA end resection and is unable to measure how fast resection is taking place, we performed Single Molecule Analysis of Resection Tracks (SMART) assay (Cruz-García et al., 2014) in miPSCs and MEFs-i4F. SMART is a high-resolution technique that measures the length of resected DNA in

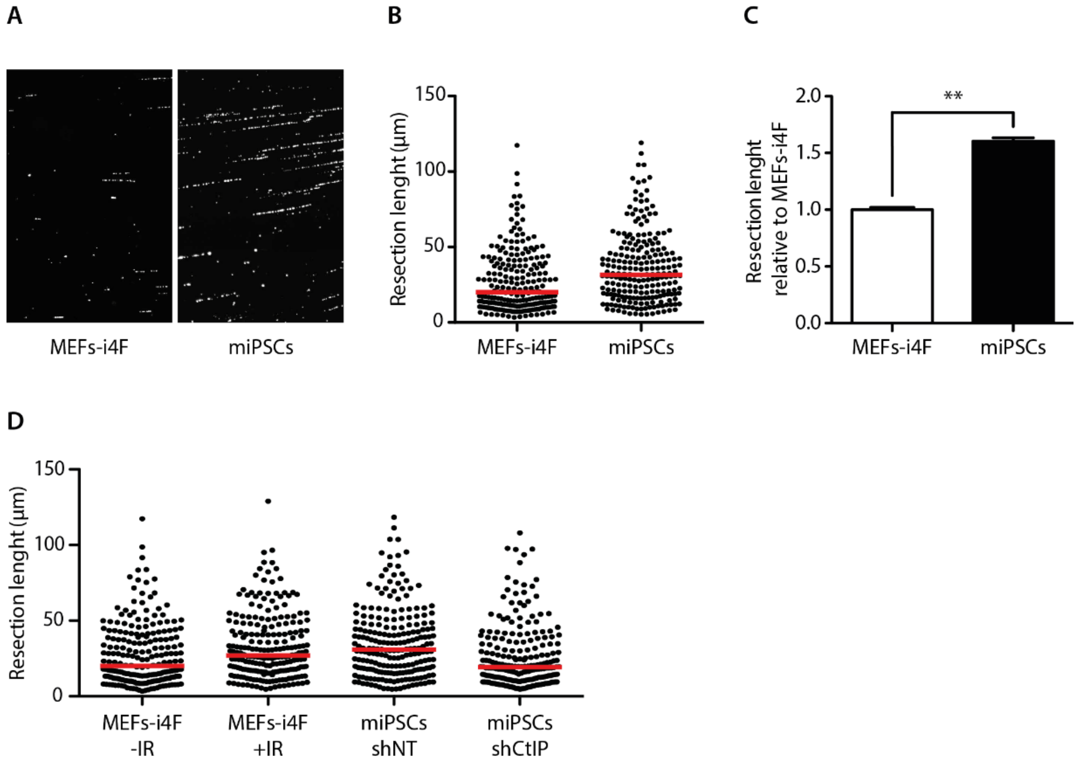


Figure R2. The extension of DNA end resection is promoted during reprogramming
(A) Representative fluorescence microscopy images of individual DNA fibres obtained by SMART in MEFs-i4F and miPSCs obtained from them. **(B)** Representative scatter plot of SMART in MEFs-i4F and miPSCs. Each dot corresponds to the length of an individual ssDNA fibre and the red line shows the median length of the population. **(C)** Median of resected DNA length obtained by SMART technique relative to MEFs-i4F. For each replica, at least 300 individual ssDNA fibres were measured. The average and standard error of the mean (SEM) of three independent experiments are shown. A Student's t-test was performed to compare both samples (* $p<0.05$; ** $p<0.01$; *** $p<0.001$). **(D)** Same as (B) but in MEFs-i4F untreated (-IR) or 1 h after irradiation with 10 Gy (+IR) and in miPSCs reprogrammed in the presence of shCtIP or the control shNT.

individual DNA fibres (see Figure R2A for representative images). Interestingly, miPSCs greatly increased the extension of resection, showing significantly longer tracks of ssDNA compared with the primary differentiated parent cells (Figure R2B and R2C). Again, we also confirmed that the higher processivity of the resection machinery in miPSCs was similar to that observed in MEFs-i4F exposed to IR and depends on CtIP activity (Figure

R2D).

All of these results demonstrated a hyper-activation of DNA end resection due to cell reprogramming in mouse cells, which likely reflects an attempt of repair by HR in order to minimize the impact of the DNA damage caused endogenously during such process.

2. CtIP levels increase during miPSCs formation

The role of CtIP as a major regulator of DNA end resection is clearly established (Huertas and Jackson, 2009; Sartori et al., 2007; You et al., 2009). As just mentioned above, we observed that CtIP was required for an hyperactivation of resection in miPSCs (Figure R1C and R2D), so we wondered if CtIP was essential for reprogramming and maintenance of miPSC. In agreement, studying CtIP expression either by mRNA or protein levels, we observed an increase in miPSCs respect to primary MEFs-i4F (Figure R3A and R3B). Moreover, this increase was concurrent with the expression of the pluripotency marker

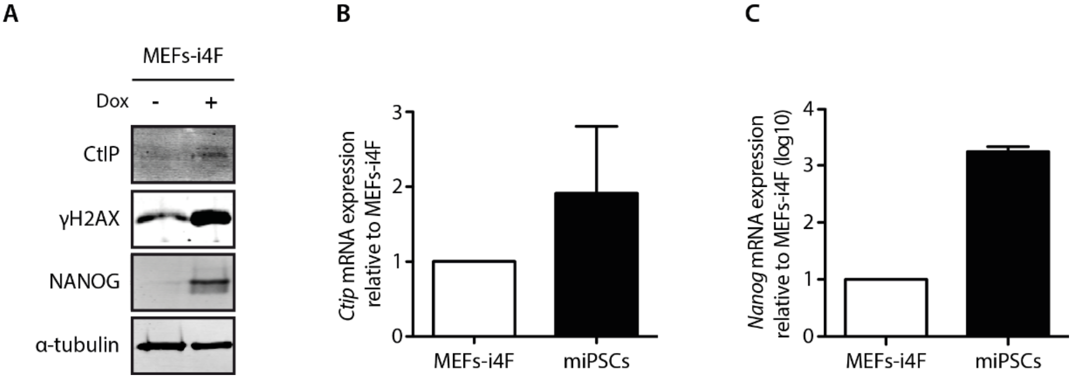


Figure R3. CtIP levels increases in miPSCs
(A) Representative western blot image of the indicated proteins and α-tubulin as loading control from whole protein extract of MEFs-i4F not induced with doxycycline (-) and miPSC formed after induction with 1 µg/ml doxycycline (+). At least three experiments were performed. **(B)** *Ctip* mRNA expression analysed by RT-qPCR in MEFs-i4F and their respective miPSC. Error bars represent SEM of three independent experiments performed with technical triplicates. Values were normalized to the housekeeping gene actin and relativized to MEFs-i4F. **(C)** Same as (B) but for *Nanog* mRNA.

NANOG (Figure R3A and R3C). Furthermore, miPSCs also showed higher levels of γH2AX (Figure R3A), a marker of DSB induction, which is consistent with the raised incidence of DSBs already known in iPSC due to oxidative and replicative stress caused by the accelerating growth rate (Banito et al., 2009; Esteban et al., 2010; Ruiz et al., 2015).

These results clearly suggested that CtIP upregulation caused an hyper-activation of DNA end resection in iPSCs, most likely to manage an increased load of DSBs that have to be repaired by homologous recombination.

3. CtIP is required for efficient cell reprogramming

Since our data supported a CtIP-dependent hyperactivation of DNA end resection in reprogrammed cells, and given that the reprogramming process induces DSBs in iPSCs (González et al., 2013) and efficient reprogramming requires functional repair pathways to favour error-free dedifferentiation (Felgentreff et al., 2014; González et al., 2013; Tilgner et al., 2013), we hypothesized that CtIP and its upregulation might be imperative during cell reprogramming to avoid genomic instability. In agreement with this, we used MEFs-i4F to study the expression at the protein level of some components involved in the DDR at different time points during reprogramming and we observed a progressive increase in CHK1, γH2AX, and CtIP protein levels that follows the progression of the reprogramming process (Figures R4), supporting the idea that replication stress causes DNA damage and CHK1 activation during reprogramming (Ruiz et al., 2015), and CtIP is expressed to counteract such effect.

Then, we decided to test the functional requirement of CtIP for cell reprogramming. Therefore, we firstly transduced MEFs-i4F with lentivirus harbouring short hairpin RNA (shRNA) specifically targeting *Ctip* (shCtip) or a non-target sequence as a control (shNT), and we forced the reprogramming process by doxycycline treatment. Then, three weeks after induction, we analysed reprogramming efficiency measured as the number of miPSCs colonies formed from the same number of MEFs-i4F cells. We observed a dramatic and

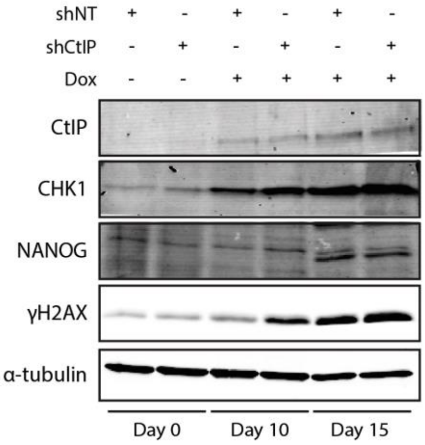


Figure R4. DDR proteins increases during reprogramming
Representative Western blot showing CtIP, CHK1, γH2AX, NANOG and α-tubulin protein levels. Samples were obtained from MEFs-i4F at day 0, 10 and 15 after reprogramming induction with 1 μg/ml doxycycline. At least three independent experiments were performed.

significant reduction of the number of colonies with stem cell-like morphology in CtIP depleted cells compared with shNT control cells (Figure R5A). Moreover, CtIP-depleted colonies were smaller than control ones (Figure R5B). As an additional readout of CtIP depletion in primary MEFs-i4F, and in agreement with results published previously (Chen et al., 2005a), we observed less proliferation in CtIP-downregulated MEFs-i4F(Figure R5C).

CtIP loss leads to DNA damage accumulation associated with defective homologous recombination (Polato et al., 2014) so, as expected, we also detected that the amount of the DNA damage marker γH2AX increased slightly but significantly in cells containing shCtip versus those with the control shNT, as shown by both immunoblot and FACS analysis (Figure R5D and R5E). Even though this difference was also observed in MEFs-i4F (Figure R5E -Dox), it was more intense in reprogrammed cells (Figure R5E +Dox), in agreement with CtIP playing a more critical role in repairing DNA damage during cell reprogramming. Pluripotent status was confirmed by the observation of NANOG protein levels (Figure R4 and R5D). Hence, we conclude that during reprogramming the response to DNA damage is enhanced, including CtIP expression. As a consequence, the lack of CtIP hampers considerably the efficiency of cell reprogramming, most likely by the accumulation of damage due to replication stress.

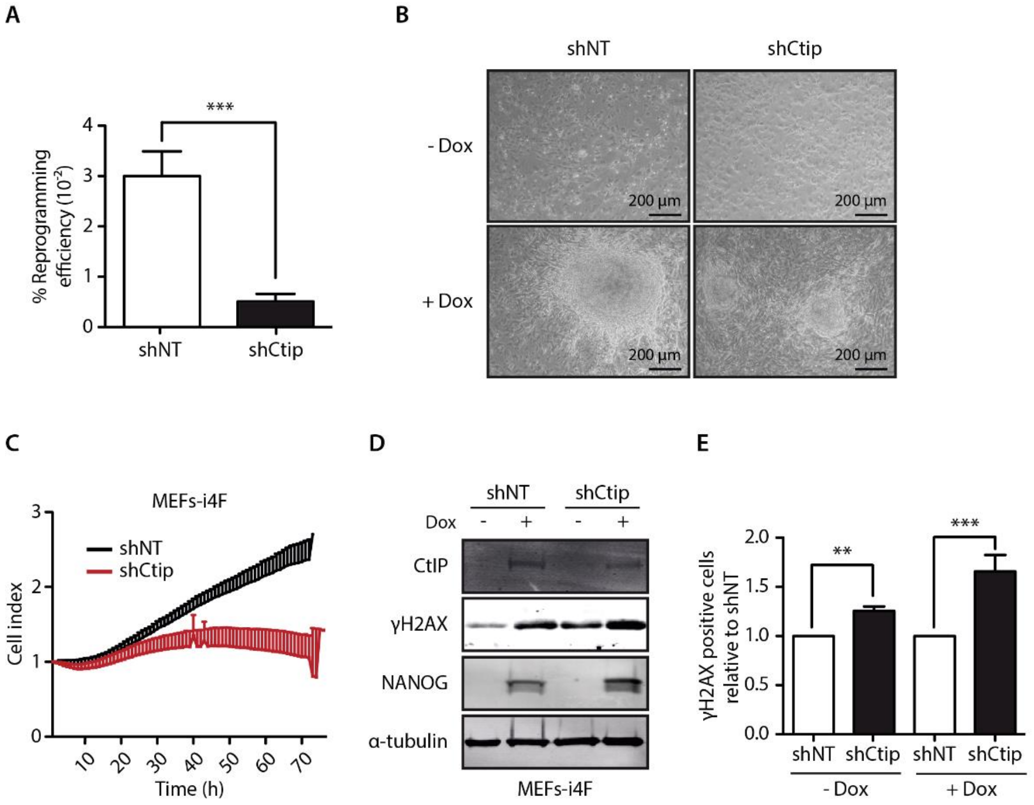


Figure R5. CtIP loss impairs reprogramming and leads to DNA damage accumulation

(A) The efficiency of reprogramming in miPSCs transduced with shCtip and shNT lentivirus prior to induction of reprogramming with doxycycline. Reprogramming efficiency was measured as the number of formed colonies normalized to the number of cells seeded. Error bars represent SEM of three independent experiments performed with technical triplicates. **(B)** Representative optical microscope pictures of morphology and size of miPSC colonies (+Dox) reprogrammed in the presence of shCtip or shNT. Non-reprogrammed MEFs-i4F (-Dox) were used as colony absence control. **(C)** Cell growth in MEFs-i4F depleted for CtIP (red) or expressing a control shRNA (black). Cell proliferation was monitored each hour for 72 hours in the xCELLigence Real Time Cell Analyzer. At least three independent experiments using technical duplicate were performed and a representative growth curve is shown. **(D)** A representative western blot of the indicated proteins in samples from MEFs-i4F (-Dox) and their respective miPSCs formed after 10 days of doxycycline treatment (+Dox). shCtip and shNT lentivirus were transduced prior to reprogramming. **(E)** FACS quantification of cells positive for γH2AX in MEFs-i4F (-Dox) or reprogrammed miPSC at day 10 (+Dox) downregulated for CtIP (black bars) or expressing a control shRNA (white bars) previous to reprogramming. Error bars indicate SEM of three independent experiments. A Student's t-test statistical analysis was performed in (A) and (E) (* p<0.05; ** p<0.01; *** p<0.001).

4. CtIP deficiency during mouse cell reprogramming causes genomic instability

To strengthen our hypothesis that CtIP is necessary for cell reprogramming to avoid genomic instability, we performed a Comparative Genomic Hybridization array (aCGH) for the detection of genomic copy number variations (CNV). CNVs involve alterations in the number of copies of a region of DNA which can either be amplifications and deletions ranging from 1 kb to many megabases. We analysed CNVs between two clones of miPSC reprogrammed in the presence of shCtip or shNT at early passes after reprogramming. We found a large difference in the number of CNV between miPSCs reprogrammed in CtIP depletion condition and their respective control cells (Figure R6). Thus, in agreement with our hypothesis, genomic instability was found to be enhanced during cell reprogramming under CtIP deficiency.

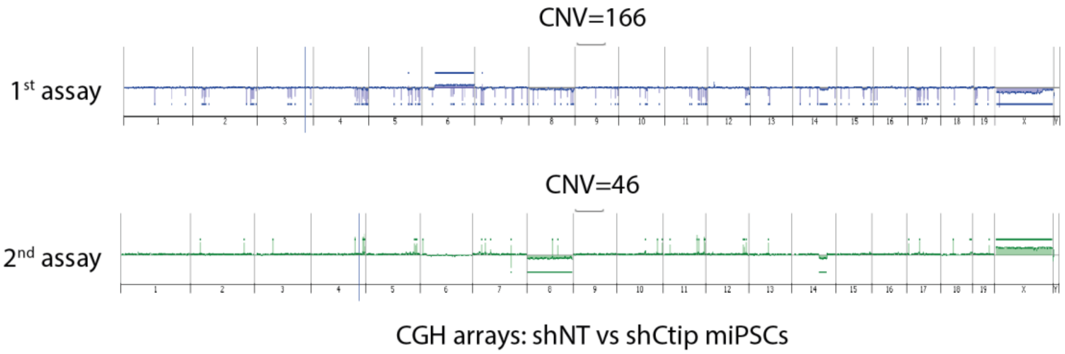


Figure R6. CtIP deficiency causes genomic instability during reprogramming
Copy number variation (CNV) of genomic DNA between miPSCs obtained in the presence of shCtip or shNT during cell reprogramming. CNVs were detected using SurePrint G3 Mouse CGH 4x180K Microarray. Two independent experiments are shown.

5. CtIP deficiency triggers apoptosis during reprogramming.

Since we observed a noteworthy drop in the number of miPSC colonies formed by reprogramming in a CtIP-downregulated situation, we set out to further investigate the causes of this reduced viability of miPSCs. Taking into account that pluripotent cells control genome integrity and avoid the propagation of genetic lesions by the elimination of damaged cells by undergoing regulated cell death (RCD) or losing pluripotency when DNA damage is unresolvable (Vitale et al., 2017), we considered that the induction of apoptosis could explain the low number of colonies obtained. The cell death by apoptosis results in extensive fragmentation of DNA, so apoptotic cells can be identified on DNA content frequency histograms as cells with a sub-G1 DNA content (Nicoletti et al., 1991). Thus, to test if unrepaired DNA damage during cell reprogramming in the absence of CtIP triggers apoptosis, we performed DNA fluorescence histograms of propidium iodide-stained reprogrammed mouse cells expressing the shCtip or the control one and we quantify the percentage of apoptotic cells in the sub-G1 peak. Interestingly, we observed that reprogrammed mouse cells expressing an shRNA against *Ctip* showed a strong and significative increase in sub-G1 cells 10 days after doxycycline induction (Figures R7A and R7B), compared with reprogrammed cells expressing a control shRNA or MEFs-i4F depleted for CtIP. This large rise was exclusive for reprogramming cells given that it was not noticed between CtIP-downregulated MEFs-i4F and shNT expressing MEFs-i4F (Figure R7A -Dox). Therefore, the reduced proliferation observed in CtIP-downregulated MEFs-i4F (Figure R5C) was not associated with an increase in sub-G1 phase cells while the reduced number of miPSC colonies formed in such circumstances did.

Caspases are protease enzymes that play essential roles in programmed cell death. Among them, caspase-3 is a central death protease in the process of apoptosis which initiates DNA fragmentation and consequently, is an important marker of the apoptotic signalling pathway (Porter and Jänicke, 1999; Wolf et al., 1999). Indeed, we observed an increase of proteolytic cleavage and activation of caspase-3 measured at protein level in cells reprogrammed in the presence of short hairpin Ctip (shCtip) at day 10 after reprogramming (Figure R7C), confirming that the increased levels of sub-G1 were due to induced apoptosis.

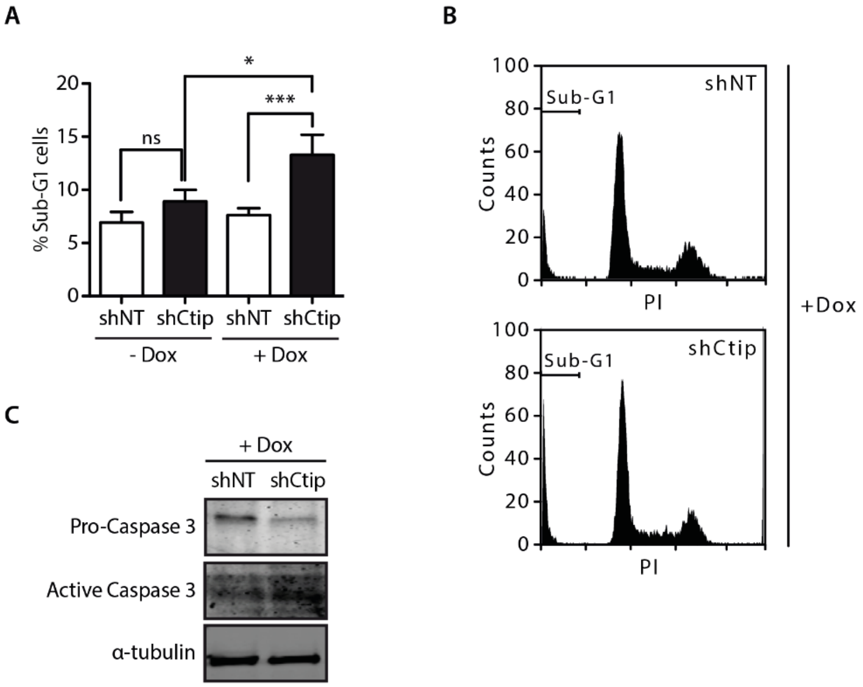


Figure R7. CtIP deficiency during reprogramming activates apoptosis.
(A) The percentage of cells with a sub-G1 DNA content was quantified by FACS analysis in MEFs-i4F (-Dox) and their respective miPSCs (+Dox) downregulated for CtIP (black bars) or expressing a control shRNA (white bars). For each analysis 10.000 single cells were analysed and the average and SEM of three independent experiments are shown. A Student's t-test statistical analysis was performed (ns: not significant; * p<0.05; ** p<0.01; *** p<0.001). **(B)** Representative flow cytometry histograms of cell cycle distribution of miPSCs reprogrammed in the presence of shCtip or shNT. Marker indicates the sub-G1 population of cells. **(C)** Immunoblot of proactive and active/cleavage caspase-3 in reprogrammed MEFs-i4F (+Dox) bearing the indicated shRNAs, at day 10. α-tubulin was used as a loading control. A representative western blot is shown out of three independent experiments.

These data altogether provide strong evidence that CtIP plays a role avoiding the genomic instability generated specifically during cell reprogramming. Nevertheless, when CtIP is absent, it is suggested that excessive damage causes the observed induction of apoptosis to avoid spreading the damage.

6. Normal CtIP levels during reprogramming are required for maintenance of iPSCs

Our previous results showed that fewer cells reprogrammed in the absence of CtIP resulted in viable miPSCs-forming colonies due to an increased apoptosis. Nonetheless, some cells did managed to survive and become miPSCs, but at the cost of accumulating genomic variations as we have described previously. Such increased genomic instability might alter the ability of iPSCs to self-renew and to differentiate. Thus, and to deepen our understanding about the role of CtIP in iPSCs, we wondered whether the downregulation of CtIP during reprogramming could affect the subsequent self-renewal and differentiation capacity of the resultant miPSCs. For this, we first selected several miPSC clones produced in the presence of shCtip or the control shRNA. Colonies from each clone were analysed for the presence of the shRNA taking advantage of the presence of an enhanced green fluorescence protein (*eGFP*) gene on the plasmid encoding the shRNAs. The eGFP gives rise to brighter fluorescence than normal GFP from *Aequoria victoria*. Interestingly, we determined that all control shNT-harbouring miPSC clones maintained higher numbers of eGFP-expressing cells than shCtip-harbouring miPSCs, suggesting that the cells that continued growing and maintaining cell pluripotency had a tendency to lose the cassette containing *eGFP* and the shRNA against *Ctip* (Figure R8A-C). Thus, collectively, our data suggested that miPSC colonies reprogrammed in the presence of CtIP depletion suffer a natural selection that favours cells with higher CtIP expression. This could explain the still high CtIP levels observed in cells reprogramed in the presence of shCtip (Figure R4).

In order to study the kinetics of loss of shCtip expression, we separated the population of eGFP-positive miPSCs generated from cells containing either shNT or shCtip by fluorescence-activated cell sorting (FACs). Then, starting from a population enriched for eGFP positive miPSCs, we investigated the eGFP disappearance during growth as a proxy for the maintenance of the shRNA. We found that eGFP cells were rapidly purified from the cell population containing shCtip compared with control cells, with a reduction in the number of eGFP-positive cells of up to 50% after only three passages (Figure R8D).

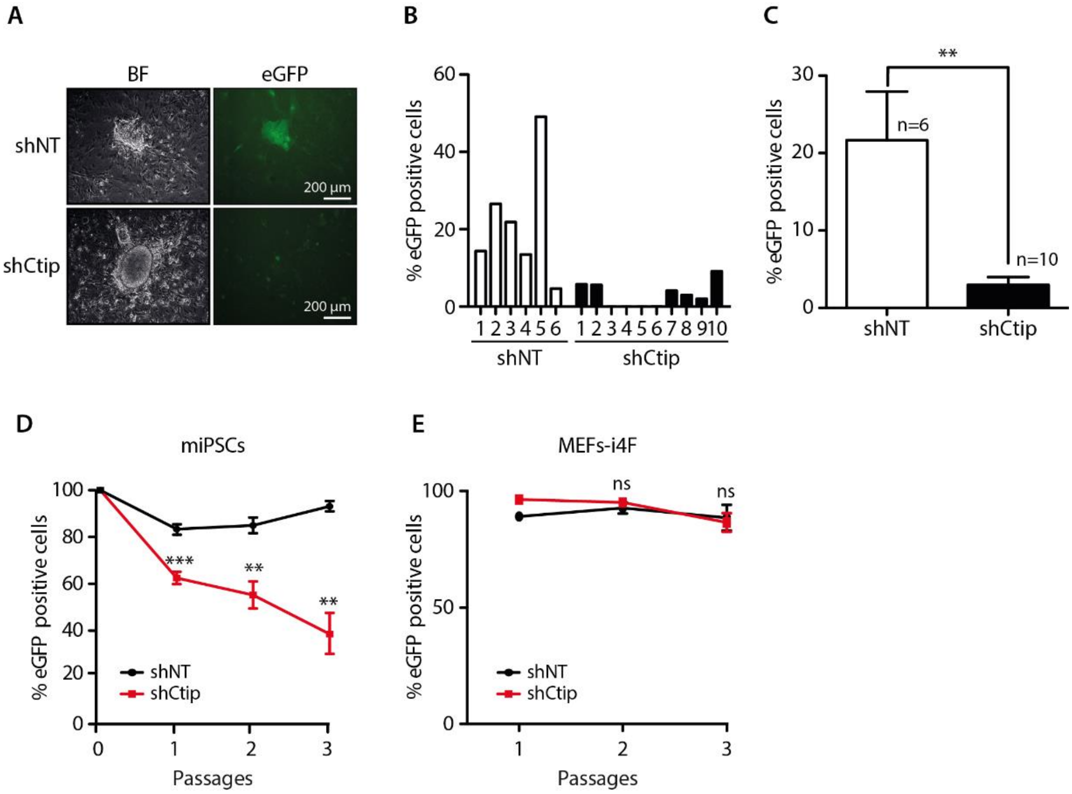


Figure R8. Stability of shRNA-eGFP cassette in reprogramming and miPSCs maintenance.

(A) Representative bright field (BF) and eGFP fluorescence images of miPSC colonies obtained after reprogramming in the presence of shCtip or shNT. **(B)** Percentage of eGFP-positive cells in miPSC colonies isolated after reprogramming in the presence (white bars) or absence (black bars) of CtIP measured by flow cytometry. 6 independent clones of shNT expressing iPSC (1-6) and 10 of shCtIP bearing (1-10) were analysed. **(C)** Average percentage of eGFP-positive cells of the clones of shNT and shCtip miPSCs shown in (B). Error bars indicate SEM of different clones (n=6 and n=10). **(D)** miPSCs bearing a cassette containing eGFP and shRNA against Ctip (red line) or a control sequence (black line) were enriched for eGFP expression by fluorescence-activated cell sorting (FACs). From the eGFP-enriched population, the percentage of eGFP-positive miPSCs was measured during subsequence passages of the cells by FACS. The average and SEM of five independent experiments are plotted. **(E)** Same as (D) but in primary MEFs-i4 bearing the shCtip-eGFP or the shNT-eGFP. At least three independent experiments were performed. Statistical analysis was performed in (C), (D) and (E) using a Student's t-test (ns: not significant; * $p<0.05$; ** $p<0.01$; *** $p<0.001$).

Strikingly, this effect was specific for iPSC cells that had been reprogrammed in the presence of shCtip and was not observed in primary MEFs-i4F with shCtip (Figure R8E), suggesting that iPSC are particularly sensitive to the loss of CtIP.

Next, we performed a self-renewal assay after enriching the populations of eGFP-positive cells. For this, we measured the number of colonies formed from the same amount of miPSCs reprogrammed either in the presence or absence of CtIP. Unsurprisingly, reprogrammed cells depleted for CtIP formed less-viable colonies than control cells (Figure R9A) and were also smaller (Figure R9B). CtIP knockdown efficiency and pluripotent status were checked at protein level (Figure R9C). To differentiate whether this effect of CtIP deficiency was due to defects gained during cell reprogramming or acquired during the pluripotent state of the reprogrammed cells, we transduced already-reprogrammed miPSCs with either the shRNA against *Ctip* or against a control sequence and analysed them for self-renewal by colony formation. Interestingly, in this scenario, CtIP downregulation did not modify self-renewal capacity compared with control cells, as a similar number of colonies of the same size were formed in both conditions (Figure R9D and R9E). Again, CtIP knockdown efficiency and pluripotent status were checked at protein level (Figure R9F). Along the same lines, we performed a self-renewal assay in a physiological stem cell line, the murine D3 embryonic stem cell line (ES-D3), downregulated or not for CtIP. Likewise, ES-D3 was not affected by CtIP depletion either in the number of colonies (Figure R10A) or their size (Figure R10B). CtIP knockdown efficiency and pluripotent status were checked at protein level (Figure R10C).

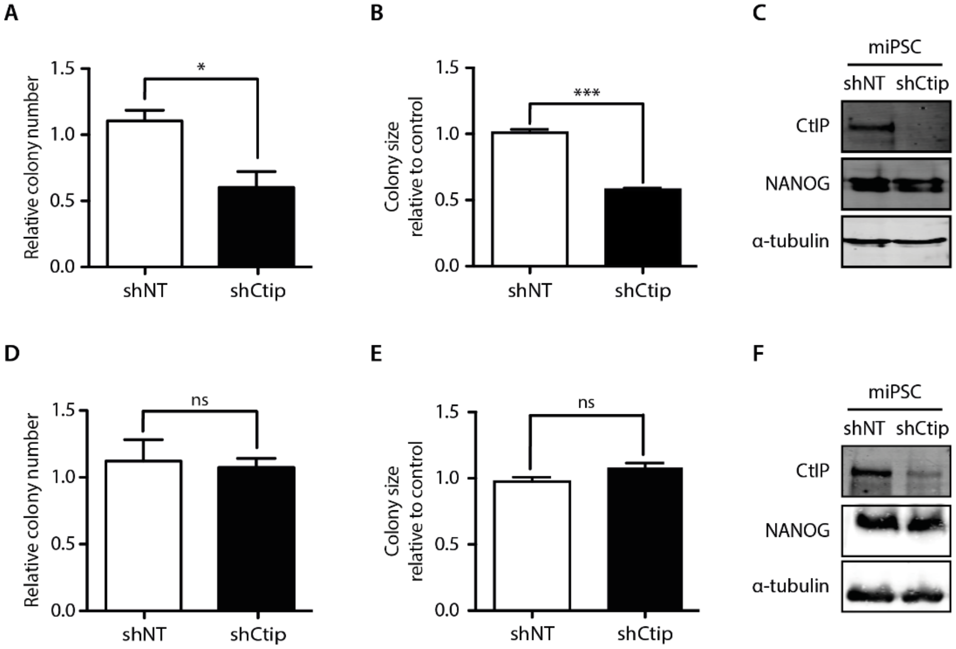


Figure R9. The levels of CtIP during reprogramming affect iPSC maintenance

(A) miPSCs reprogrammed in the presence of shCtip-eGFP (black bars) or the shNT-eGFP (white bars) were enriched for eGFP expression by FACs. The same amount of eGFP-positive miPSCs were seeded at low density and the number of colonies formed was measured and relativized to control. Error bars represent SEM of three independent experiments performed with technical triplicates. **(B)** The size of each colony formed from miPSCs reprogrammed in the presence of shCtip-eGFP (black bars) or shNT-eGFP (white bars) was measured using Adobe Photoshop CS6 and relativized to control. Error bars represent SEM of three independent experiments performed with technical triplicates. **(C)** miPSCs reprogrammed in the presence of shCtIP or control shNT were immunoblotted to analyse CtIP, NANOG and α-tubulin as loading control. A representative western blot is shown out of three independent experiments. **(D)** Same as (A) but with miPSCs reprogrammed in the presence of CtIP, and then transduced with the shRNA against CtIP (black bars) or control shNT (white bars). **(E)** Same as (B) but with miPSCs transduced with the shRNA against CtIP after reprogramming. **(F)** Same as in (C) but with cells transduced with the shCtip after reprogramming. Statistical analysis was performed in (A), (B), (D) and (E) using a Student’s t-test (ns: not significant; * $p<0.05$; ** $p<0.01$; *** $p<0.001$).

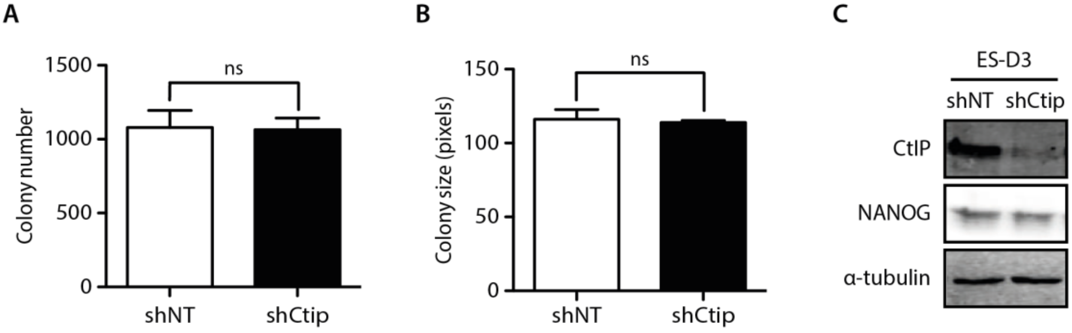


Figure R10. CtIP downregulation does not affect D3 embryonic stem cells colony formation.
(A) The number of colonies formed from the same amount of D3 embryonic stem cells (ES-D3) transduced with shNT (white bars) or shCtip (black bars) seeded at low density was measured. Error bars represent SEM of three independent experiments. **(B)** The size of colonies formed from D3 embryonic stem cells depleted (black bars) or not (white bars) of CtIP. At least 300 colonies for each condition were measured. The average and SEM of three independent experiments are plotted. **(C)** A representative western blot of the indicated proteins in samples from ES-D3 cells expressing the indicated shRNAs. A Student's t-test was performed in (A) and (B) (ns: not significant; * $p<0.05$; ** $p<0.01$; *** $p<0.001$).

7. Normal CtIP levels during reprogramming are required for differentiation of iPSCs

As mentioned in the introduction section, the main characteristics of stem cells are the ability of long-term self-renewal and the capacity to differentiate into specialized cells. Therefore, as self-renewal was compromised by inherited defects aroused upon cell reprogramming in the absence of CtIP, we wondered whether the differentiation process could also be jeopardized in a similar way. To study differentiation, we took advantage of the formation of embryoid bodies (EBs). EBs are complex three-dimensional cell aggregates comprising the three embryonic germ layers (Itskovitz-Eldor et al., 2000) and whose formation is an intermediate step during the differentiation of pluripotent stem cells into specialized cell types. Pluripotent stem cells are induced to differentiate into EBs by seeding them in suspension on non-adhesive plates and removing factors that support pluripotency. In this way, we forced mouse embryoid body differentiation from several clones, by removing the growth factor, leukaemia inhibitory factor (LIF) for 6 days. CtIP

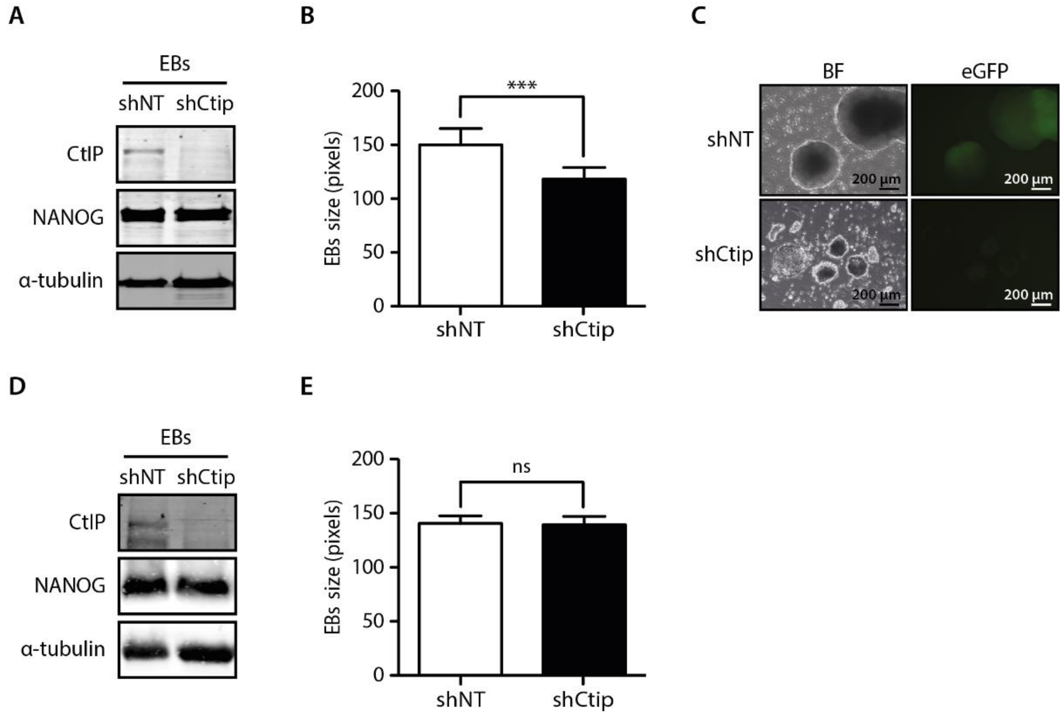


Figure R11. CtIP during reprogramming is essential for iPSCs differentiation.
(A) Western blot analysis of indicated proteins in embryoid bodies (EBs) generated from miPSCs reprogrammed in the presence of shCtip or shNT. A representative western blot is shown of three independent experiments. **(B)** Size of EBs generated at 6 days after spontaneous differentiation of miPSCs reprogrammed in the presence of shCtIP or control shNT and enriched for eGFP expression by FACs. At least 300 EBs of each condition were measured. The average and SEM of three independent experiments are plotted. **(C)** Representative bright field (BF) and eGFP fluorescence images of EBs formed from miPSC downregulated or not for CtIP. **(D)** Same as (A), but with cells reprogrammed in the presence of CtIP and then transduced with shCtip or shNT. **(E)** Same as (B) but with cells transduced with shCtip or shNT after reprogramming. Statistical significance in (B) and (E) was obtained by a Student's t-test (ns: not significant; * $p<0.05$; ** $p<0.01$; *** $p<0.001$).

and NANOG protein level, to follow differentiation, were checked by western blot (Figure R11A). We observed that miPSCs reprogrammed under CtIP depletion also generated embryoid bodies smaller than their respective control cells, hence showing a clear deficiency in differentiation (Figure R11B and R11C).

Then, we assessed if this defect was a consequence of problems inherent to

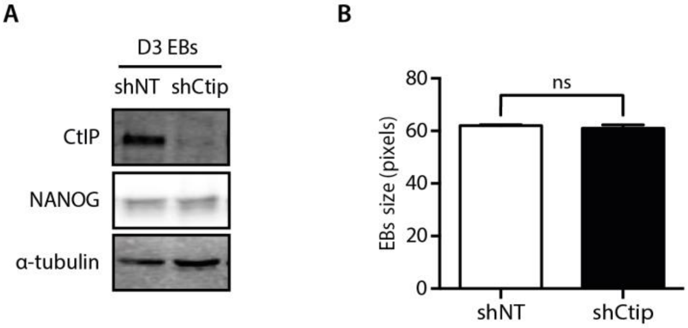


Figure R12. CtIP downregulation does not affect D3 embryonic stem cells differentiation.
(A) A representative western blot of the indicated proteins. Samples were obtained from embryoid bodies derived from ES-D3 cells expressing the indicated shRNAs. **(B)** Size of EBs derived from ES-D3 cells expressing shCtip or a control shRNA. At least 300 EBs of each condition were measured. The average and SEM of three independent experiments are plotted. A Student's t-test statistical analysis was performed (ns: not significant; * $p<0.05$; ** $p<0.01$; *** $p<0.001$).

reprogramming in a CtIP-defective environment rather than a consequence of the loss of an active role of CtIP in stem cell differentiation. So, we analysed differentiation by inducing EBs formation in miPSC transduced with shCtip or shNT after reprogramming as well as in murine ES-D3 embryonic stem cells. Again, we checked CtIP and NANOG protein level in EBs formed from miPSCs (Figure R11D) and from ES-D3 (Figure R12A). Similarly to self-renewal, CtIP downregulation in already-reprogrammed miPSCs (Figure R11E) or ES-D3 cells (Figure R12B) did not affect their differentiation capacity.

Collectively, our data confirmed that genomic instability created during cell reprogramming under CtIP deficiency is the main cause of impairment of iPSCs self-renewal and differentiation.

CHAPTER 2

1. KLF4 affects DNA break repair pathway choice

Since our data in chapter 1 support that cells hyper-activate DNA end resection during cell reprogramming in order to deal with the damage induced by the process, we decided to analyse if some of the master regulators of cell stemness might regulate resection, and therefore the balance between homologous recombination (HR) and non-homologous end joining (NHEJ). As mentioned in the introduction (section 2.1), cell stemness is controlled by a series of master regulators, including the aforementioned Yamanaka factors (NANOG, OCT3/4, SOX2 and KLF4), among others. Interestingly, in a general screening searching for regulators of the balance between DNA double strand break repair pathways (López-Saavedra et al., 2016) using the SeeSaw Reporter system (SSR) (Gomez-Cabello et al., 2013), KLF4 was detected as a candidate gene that favour HR. On the contrary, neither NANOG nor the other factors used by Yamanaka for reprogramming did significantly alter the HR/NHEJ balance in the screening.

First, to validate the screening result we analysed again the balance between NHEJ and HR using the SSR system (Figure R13A) in the human osteosarcoma cell line (U2OS) downregulated for KLF4 using an additional siRNA. Moreover, siRNA targeting the canonical recombination and resection factor *CtIP* was used as a control. As expected, we obtained an imbalance towards NHEJ when either KLF4 or CtIP are depleted (Figure R13B), as predicted by the screening. Next, in order to confirm the results of KLF4 with the SSR, we used different reporter systems in U2OS to analyse the ability to perform specific DSB

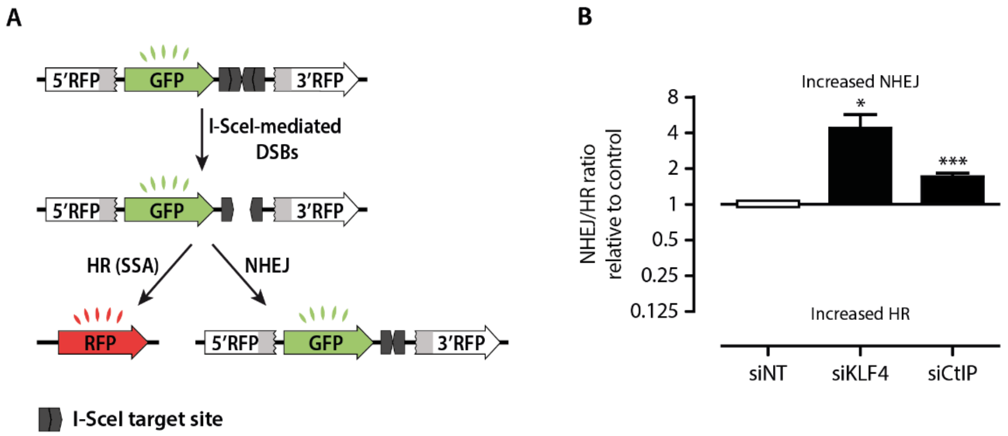


Figure R13. KLF4 controls the balance between NHEJ and HR.
(A) Schematic representation of the SeeSaw reporter system (SSR). A GFP gene is flanked by two truncated parts of RFP gene that share 302 bp of homology (marked in grey) and two I-SceI-target sites were cloned downstream of GFP. After generation of a DSB by I-SceI expression the damage can be repaired by NHEJ, restoring the GFP gene or by HR using homologous sequence, thus creating a functional RFP gene. **(B)** U2OS cells harbouring a single copy of SSR were transfected with the indicated siRNAs. The siNT and siCtIP were used as negative and positive controls, respectively. The NHEJ:HR ratio was calculated and normalized to siNT control. The average and SEM of four independent experiments are shown. A Student's t-test statistical analysis was performed (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

repair pathways. As the SSR system compares non-homologous end joining with single-strand annealing (SSA), we first used U2OS cells harbouring a single copy of the SA-GFP reporter system, which measures such recombination subpathway, by the formation of an active GFP gene upon I-SceI-induced DSB (Figure R14A). SSA is a specific type of HR that is dependent on extent of resection but Rad51-independent (Stark et al., 2004). As expected, we observed that KLF4 downregulation impairs, slightly but statistically significant, this particular repair pathway, albeit to a lesser extent than the canonical resection factor CtIP (Figure R14B). Next, to check classical recombination, we used the DR-GFP reporter (Figure R15A) in which an active GFP gene is formed when cells repair an I-SceI-induced DSB by Rad51-dependent gene conversion (Pierce et al., 1999). Importantly, we also observed an impairment of this pathway when KLF4 was downregulated (Figure R15B). So, we conclude that the two different subtypes of HR were compromised when KLF4 was depleted. Finally,

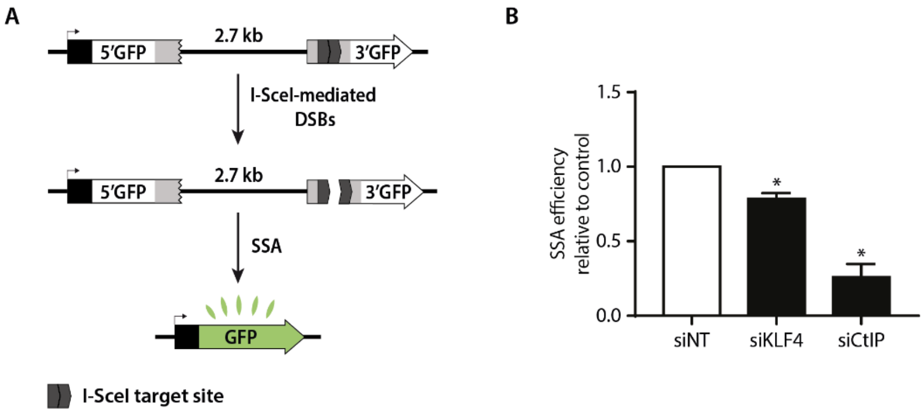


Figure R14. Downregulation of KLF4 impairs single strand annealing.
(A) Schematic representation of the SA-GFP reporter. Such reporter consists of two truncated copies of the GFP which have 266 bp of homology (marked in grey) and are separated by 2.7 Kb. Repair of the I-SceI-mediated DSB by single strand annealing (SSA) restores an active GFP gene with the deletion of one of the repeats and the intervening region. **(B)** The efficiency of SSA was calculated as the percentage of GFP positive cells using U2OS cells bearing a single copy of the SA-GFP reporter and depleted for the indicated genes by siRNAs. This percentage was normalized with the control siNT and plotted. Bars represent the average and SEM of at least three independent experiments and a Student's t-test statistical analysis was performed (* p<0.05; ** p<0.01; *** p<0.001).

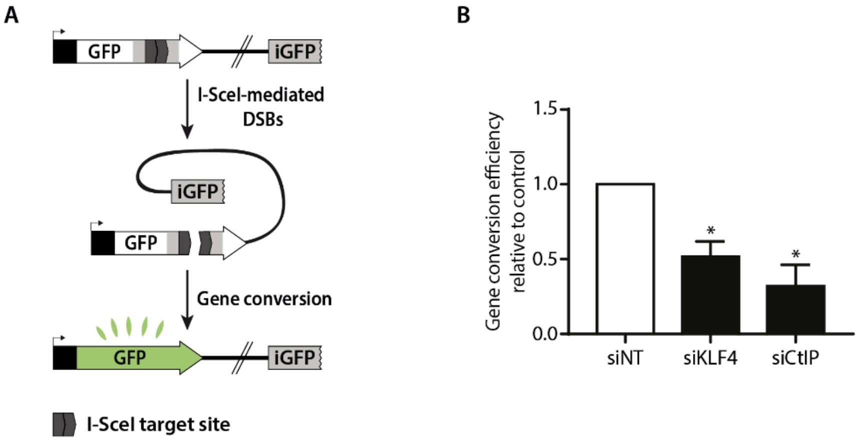


Figure R15. Absence of KLF4 impairs gene conversion.
(A) Schematic representation of the DR-GFP reporter which is formed by two non-functional copies of the GFP (homologous sequences are marked in grey). The upstream repeat contains the recognition site for the I-SceI endonuclease and an 812-bp internal GFP fragment (iGFP) can repair the I-SceI-mediated DSB by gene conversion and results in a functional GFP gene. **(B)** The efficiency of gene conversion was calculated as the percentage of GFP positive cells using U2OS cells bearing a single copy of the DR-GFP reporter and transfected with siNT, siKLF4 or siCtIP. This percentage was normalized with the cells transfected with siNT. Bars represent the average and SEM of at least three independent experiments and statistical significance was calculated with a Student's t-test (* p<0.05; ** p<0.01; *** p<0.001).

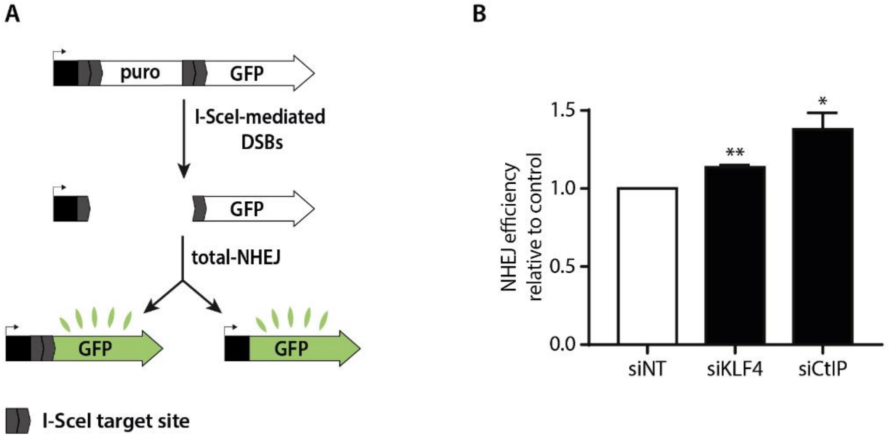


Figure R16. KLF4 depletion increases NHEJ.
(A) Schematic representation of the EJ5-GFP reporter which contains GFP separated from a promoter by a puromycin gene that is flanked by two I-SceI sites. I-SceI induced DSB can be repaired by NHEJ, both conservative (right) or mutagenic (left), recreating an active GFP gene, containing or not a functional I-SceI target site. **(B)** The efficiency of NHEJ was calculated as the percentage of GFP positive cells normalized with the siNT using U2OS cells bearing a single copy of EJ5-GFP reporter and transfected with the indicated siRNAs. Bars represent the average and SEM of at least three independent experiments and statistical significance was calculated with a Student's t-test (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

we used EJ5-GFP reporter (Figure R16A) to measures NHEJ (Bennardo et al., 2008). This reporter leads to GFP expression when a DSB induced by I-SceI is repaired through multiple classes of NHEJ events, including classical and alternative NHEJ. In agreement with a role of KLF4 in controlling DSB repair pathway choice, its absence mildly increased NHEJ (Figure R16B). Thus, we confirmed that KLF4 depletion unbalance the ratio between DSB repair factors both hampering recombination and increasing NHEJ.

2. KLF4 is required for survival to DSB inducing agents

Several studies have shown that cells defective in DSB repair are sensitive toward many DNA damaging agents. Thus, as KLF4 affects DSB repair pathway choice, we wonder if KLF4 could also have an effect in cell sensitivity to various genotoxic agents. To test this

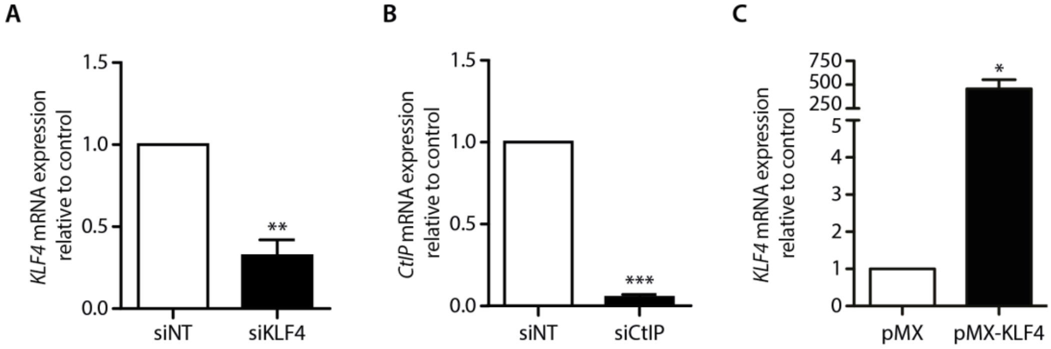


Figure R17. KLF4 expression levels

(A) Depletion efficiency of *KLF4* at the level of mRNA, measured by RT-qPCR, in U2OS cells transfected either with siNT control sequence or siKLF4. Error bars represent SEM of three independent experiments performed with technical triplicates. Values were normalized to the housekeeping gene actin and relativized to the control siNT. A Student's t-test statistical analysis was performed (* p<0.05; ** p<0.01; *** p<0.001). **(B)** Depletion efficiency of *CtIP* at the level of mRNA, measured by RT-qPCR, in U2OS cells transfected either with siNT control sequence or siCtIP. Other details as in (A). **(C)** Same as (A) but in cells transfected with the pMX-KLF4 vector, to overexpress the protein, or the control pMX plasmid.

hypothesis, we performed clonogenic assays and measured colony formation in cells either downregulated or overexpressing KLF4 upon treatment with three different DSB induced agents: camptothecin (CPT), etoposide (VP-16) and ionizing radiation (IR). CPT is a topoisomerase I inhibitor which induces replication-associated DSBs in the S phase of the cell cycle while VP-16, an inhibitor of topoisomerase II, and IR cause DSBs directly in all phases of the cell cycle. Hence, U2OS cells downregulated or overexpressed for KLF4 were treated with different doses of IR or different concentrations of CPT or VP-16. Due to the low levels of KLF4 endogenous protein in U2OS cells that makes complicated its detection by western blot, we analysed protein downregulation and overexpression efficiency by mRNA levels using qRT-PCR (Figure R17A-C). As expected, impaired homologous recombination by KLF4 downregulation rendered cells hypersensitive to IR (Figure R18A), or the topoisomerase poisons CPT (Figure R18B) and VP-16 (Figure R18C). Interestingly, overexpression of KLF4 promotes hyper-resistance to CPT (Figure R18E) but displayed no significant difference to control in cells treated with IR (Figure R18D) or etoposide (Figure R18F). CPT creates single-strand breaks (SSB) that are converted into DSBs during

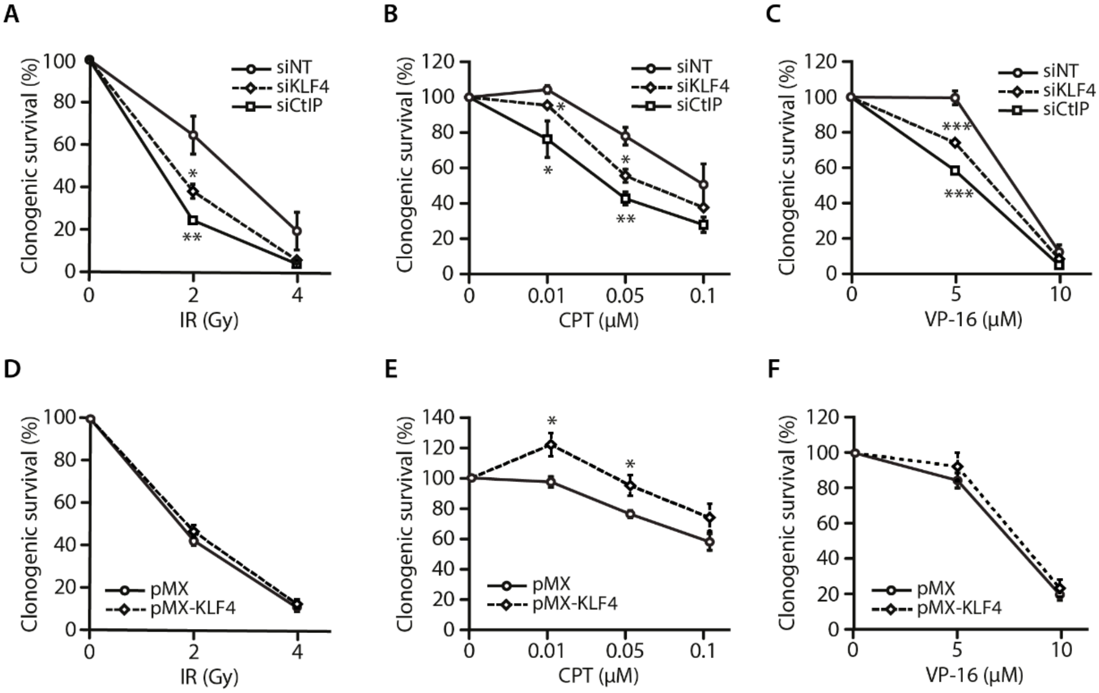


Figure R18. KLF4 avoid cell sensitivity to DSB inducing agents

(A) Percentage of survival population, compared with the untreated control, in U2Os cells transfected with siRNAs against the indicated genes upon exposure to the mentioned doses of ionizing radiation. The average and SEM of three independent experiments are plotted. Colony number at each dose were normalized with the corresponding untreated sample. **(B)** Same as (A), but in cells treated for 1 hour with the indicated concentrations of camptothecin (CPT). **(C)** Same as (A), but in cells treated for 1 hour with the indicated concentrations of etoposide (VP-16). **(D)** Same as (A), but in cells transfected with a plasmid expressing KLF4 (pMX-KLF4) or an empty vector pMX. **(E)** Same as (B), but in cells transfected with a plasmid expressing KLF4 (pMX-KLF4) or an empty vector pMX. **(F)** Same as (C), but in cells transfected with a plasmid expressing KLF4 (pMX-KLF4) or an empty vector pMX. Statistical significance was calculated with a Student's t-test (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

replication and, therefore, could not be repaired by NHEJ and are strictly dependent on the action of HR. Thus, this hyper-resistance to CPT in cells overexpressing KLF4 could be explained by a more active HR. On the contrary, most breaks caused by IR or etoposide will be readily repaired by either NHEJ or HR, thus a small increase in HR will have little or no effect in the sensitivity to those agents.

However, considering that the usage of DSB repair pathway is regulated by the cell

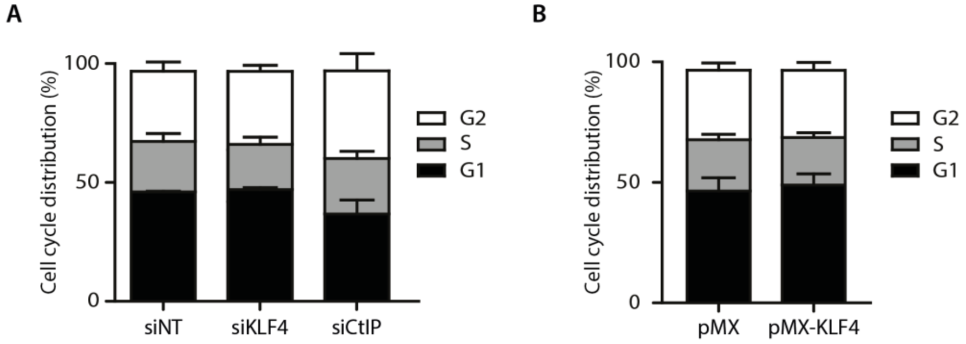


Figure R19. KLF4 does not affect cell cycle progression.
(A) Cell cycle distribution of U2OS cells transfected either with siNT control sequence or with siRNAs against KLF4 and CtIP as indicated. The average and SEM of at last three independent experiments are plotted. **(B)** Same as (A) but in cells transfected with a plasmid expressing KLF4 (pMX-KLF4) or the control pMX vector.

cycle (Rothkamm et al., 2003), and that KLF4 has been proposed to affect cell cycle progression (Ghaleb and Yang, 2017), we reasoned that was critical to exclude that the phenotypes observed in both reporter systems and cell sensitivity were due to a cell cycle arrest in our cellular system. Indeed, in our experimental setup, cell cycle progression was not affected by altering KLF4 levels (Figure R19A and R19B). Thus, these findings strongly suggest that KLF4 is involved in DSB repair pathway choice and it is necessary to avoid cell sensitivity to treatments that induce DSB.

3. KLF4 is required for DNA end resection

As mentioned in the introduction, the balance between HR and NHEJ is mostly regulated by controlling the licensing of DNA end resection. Resection gives rise to long single-stranded DNA (ssDNA) tails that are immediately coated by RPA protein complex for protection. The accumulation of RPA at the sites of the break produces foci that can be visualized by immunofluorescence. In order to test the impact of KLF4 to DNA end resection, we depleted KLF4 by siRNA or overexpressed the protein in U2OS and analysed the formation of RPA foci in response to exposure to high dose of IR (10Gy), as a readout of resection (Figure R20). The damage-induced phosphorylation of H2AX (γH2AX) was used as a sensitive marker of DSB (Figure R20A and R20C). In agreement with a role favouring HR over NHEJ, KLF4 depletion using two different siRNAs partially impaired RPA foci formation, albeit to a lesser extent than downregulation of the core resection factor CtIP (Figure R20A and R20B). Conversely, KLF4 overexpression showed the opposite phenotype and a mild increase of cells positive for RPA foci was readily observed (Figure R20C and R20D).

RPA foci formation is a low-resolution technique so, in order to confirm the phenotype observed we used the high-resolution method SMART, that gives information about the length of DNA resected in individual DNA fibres (Cruz-García et al., 2014). Interestingly, not only the number of breaks that were resected responded to KLF4 levels but the length of ssDNA formed during resection was reduced in cells depleted of KLF4 (Figure R21A-C) and increased in cells overexpressing KLF4 (Figure R21D-F).

Hence, these results lead us to conclude that KLF4 plays an active role in DSB repair by promoting both initiation and speed of DNA end resection.

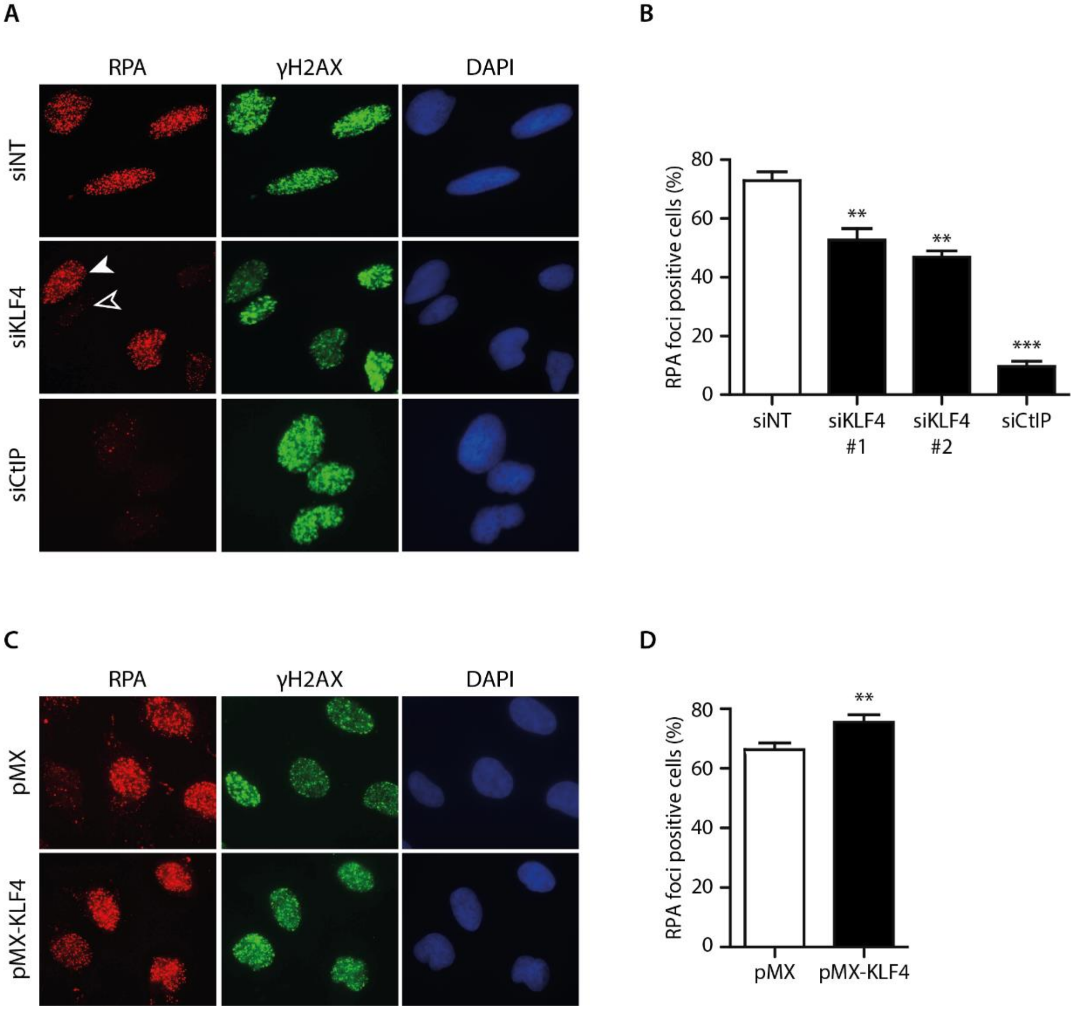

A
RPA
γH2AX
DAPI
siNT
siKLF4
siCtIP
B
RPA foci positive cells (%)
80
60
40
20
0
**
**
***
siNT
siKLF4 #1
siKLF4 #2
siCtIP
C
RPA
γH2AX
DAPI
pMX
pMX-KLF4
D
RPA foci positive cells (%)
80
60
40
20
0
**
pMX
pMX-KLF4

Figure R20. KLF4 affects DNA end resection.

(A) Representative fluorescence microscopy images of RPA foci (red) and γH2AX foci (green). Cells transfected with the indicated siRNAs were irradiated (10 Gy) and collected one hour later for immunofluorescence. An example of cell defined as RPA foci positive cells is indicated by a full white arrow, while an example of RPA foci negative cell is marked with an empty arrow. Cell nuclei are stained with DAPI (blue). **(B)** The percentage of cells positive for RPA foci upon DSB induction by ionizing radiation is plotted. Graph represents the average and SEM of three independent experiments. At least 200 cells were measured in each experiment. **(C)** Same as (A) but in cells bearing the pMX-KLF4 or the pMX empty vector, as indicated. **(D)** Same as (B) but in cells bearing the pMX-KLF4 or the pMX empty vector, as indicated. Statistical significance in (B) and (D) was calculated by Student's t-test comparing each condition to siNT or pMX cells, respectively (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

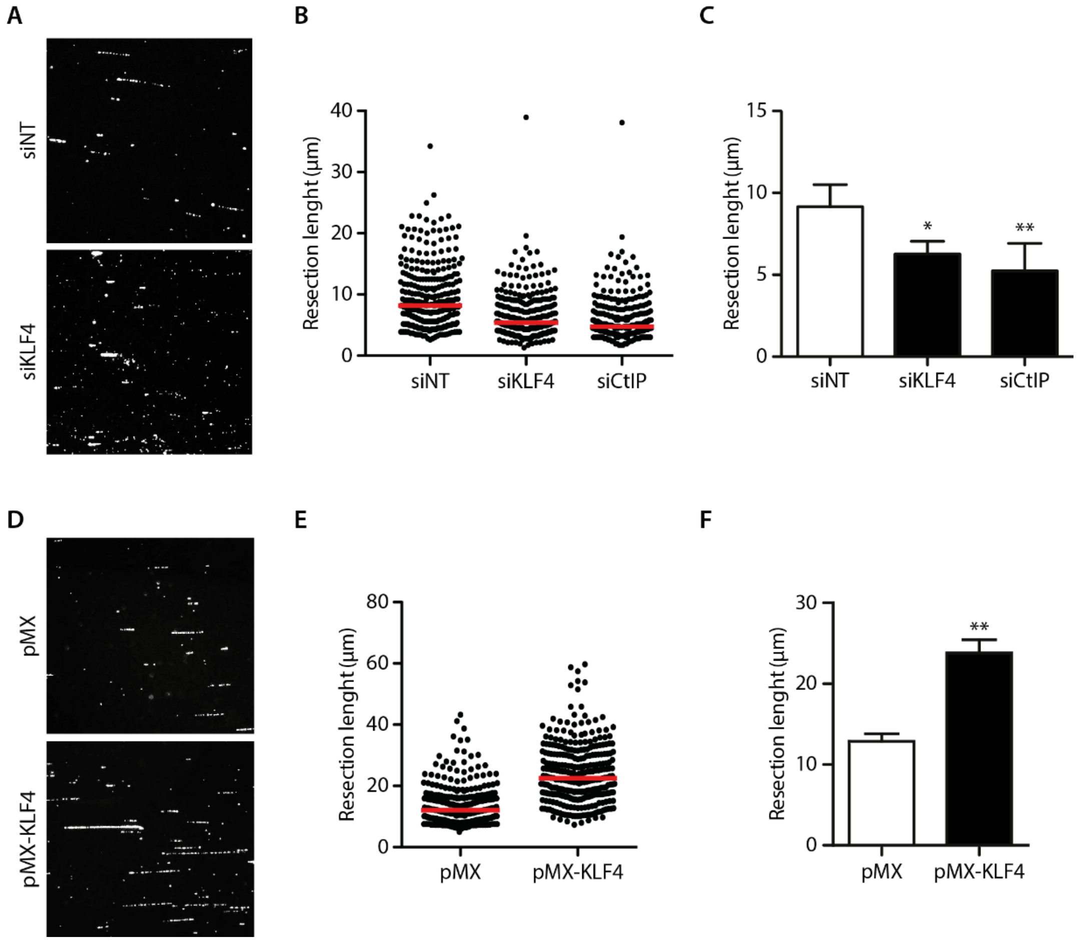


Figure R21. KLF4 affects the speed of DNA end resection.

(A) Representative fluorescence microscopy images of individual DNA fibres obtained by SMART in U2OS cells transfected with siNT or siKLF4. **(B)** Representative scatter plot of SMART of one representative experiment in U2OS cells transfected with the indicated siRNAs. Each dot corresponds to the length of an individual ssDNA fibre and the red line shows the median length of the population. **(C)** Median of resected DNA length obtained by SMART technique. The graph represents the average and SEM of three independent experiments. For each replica, at least 300 individual ssDNA fibres were measured. Statistical significance was calculated with a Student's t-test (* $p<0.05$; ** $p<0.01$; *** $p<0.001$). **(D)** Same as (A), but in cells overexpressing KLF4 (pMX-KLF4) and control cells (pMX). **(E)** Same as (B), but in cells overexpressing KLF4 (pMX-KLF4) and control cells (pMX). **(F)** Same as (C), but in cells overexpressing KLF4 (pMX-KLF4) and control cells (pMX).

4. KLF4 affects the recruitment of pro- and anti-resection proteins to DSBs sites

It has been shown previously that the license of resection is strongly regulated by many different aspects of cellular metabolism, being the mutually exclusive localization of pro- and anti-resection factors at sites of DSBs one of the best defined regulatory events (Chapman et al., 2013; Dev et al., 2018; Escribano-Díaz et al., 2013; Gupta et al., 2018; Zimmermann et al., 2013). Thus, we decided to investigate the implication of KLF4 in the foci formation of the pro-resection factor BRCA1 and the anti-resection protein RIF1. For this, we depleted or overexpressed KLF4 in U2OS and analysed the foci formation of both proteins by immunofluorescence after exposure to 10 Gy of ionizing radiation (Figure R22A, R22C, R23A and R23C). Indeed, we observed that KLF4 downregulation, in agreement with a hampered resection, caused a mild but consistent increased accumulation of RIF1 measured as the average number of foci per cell (Figure R22A and R22B) while KLF4 overexpression impaired RIF1 localization (Figure R22C and R22D).

On the contrary, also in concordance with the previous data, we observed a reduction of BRCA1 recruitment in KLF4 downregulated cells (Figure R23A and R23B) as well as an increase in BRCA1 foci formation when KLF4 is overexpressed (Figure R23C and R23D).

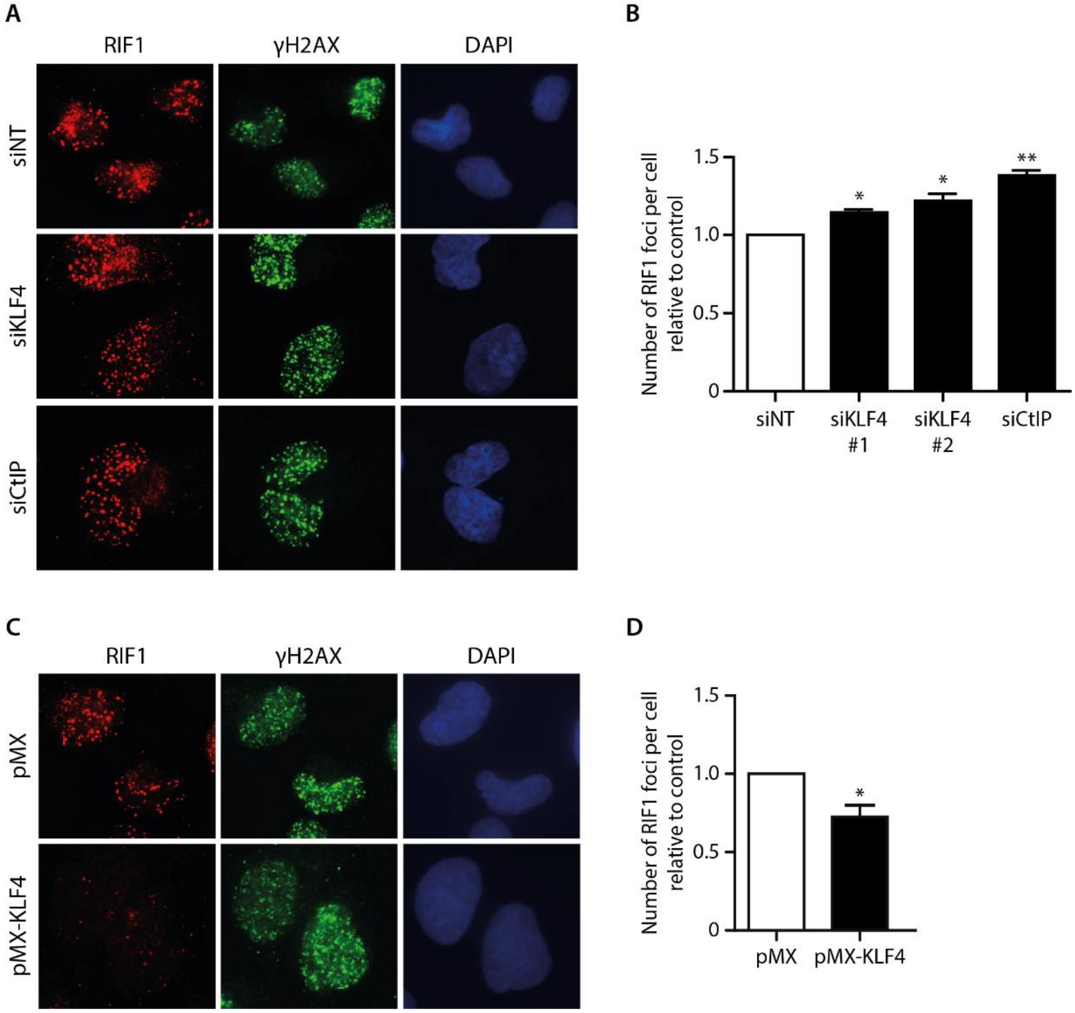


Figure R22. KLF4 affects RIF1 foci formation.

(A) Representative fluorescence microscopy images of RIF1 (red), γH2AX (green) and DAPI staining. Cells transfected with the indicated siRNAs were irradiated (10 Gy) and collected one hour later for immunofluorescence using an anti-RIF1 antibody. **(B)** The number of RIF1 foci per cell relative to control is plotted. Graph represents the average and SEM of three independent experiments. (C) Same as (A) but in cells overexpressing wild type KLF4 (pMX-KLF4) or not (pMX). **(D)** Same as (B) but in cells bearing the pMX-KLF4 or the pMX empty vector, as indicated. Statistical significance in (B) and (D) was calculated by Student's t-test (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

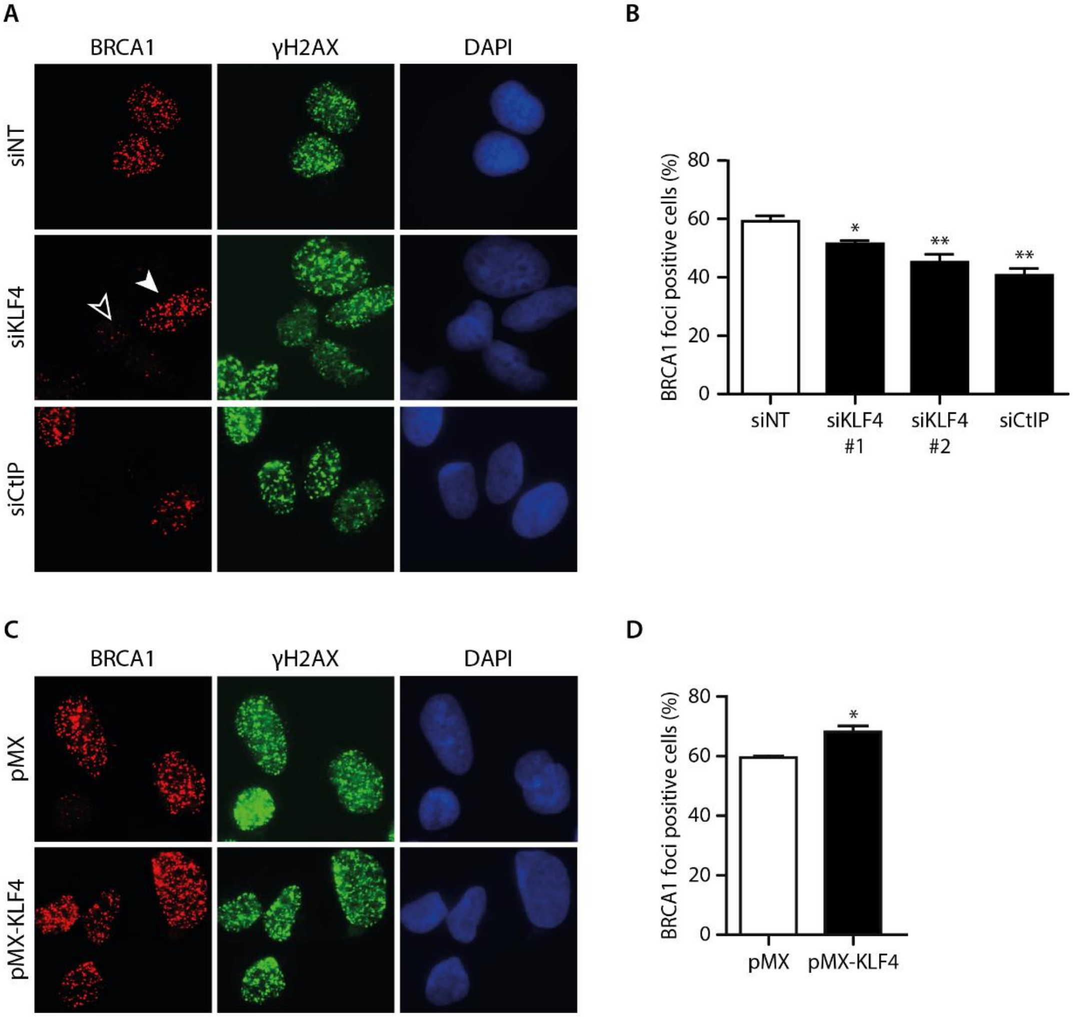

A
BRCA1
γH2AX
DAPI
siNT
siKLF4
siCtIP
B
BRCA1 foci positive cells (%)
80
60
40
20
0
*
**
**
siNT
siKLF4 #1
siKLF4 #2
siCtIP
C
BRCA1
γH2AX
DAPI
pMX
pMX-KLF4
D
BRCA1 foci positive cells (%)
80
60
40
20
0
*
pMX
pMX-KLF4

Figure R23. KLF4 affects BRCA1 foci formation.

(A) Representative fluorescence microscopy images of BRCA1 (red), γH2AX (green) and DAPI staining. Cells transfected with the indicated siRNAs were irradiated (10 Gy) and collected one hour later for immunofluorescence using an anti-BRCA1 antibody. An example of cell defined as BRCA1 foci positive cells is indicated by a full white arrow, while an example of BRCA1 foci negative cell is marked with an empty arrow. **(B)** The percentage of cells positive for BRCA1 foci upon DSB induction by ionizing radiation is plotted. Graph represents the average and SEM of three independent experiments. At least 200 cells were measured in each experiment. **(C)** Same as (A) but in cells bearing the pMX-KLF4 or the pMX empty vector, as indicated. **(D)** Same as (B) but in cells overexpressing or not wild type KLF4. Statistical significance in (B) and (D) was calculated by Student's t-test (* $p<0.05$; ** $p<0.01$; *** $p<0.001$).

5. Arginine methylation of KLF4 is essential for its activity in DNA end resection

As mentioned in the introduction section, KLF4 is subjected to multiple post-translational modifications. Recently, it has been shown that KLF4 directly interacts with the protein arginine N-methyltransferase 5 (PRMT5) which catalyses the methylation of three specific arginine residues R374, R376 and R377 on KLF4 (Hu et al., 2015). This post-translational modification is proposed to affect KLF4 roles in maintaining genomic stability and carcinogenesis (Hu et al., 2015). Additionally, preliminary data from our laboratory showed that PRMT5 also interact with the key resection factor CtIP and its downregulation hampers both Rad51-dependent and Rad51-independent recombination, as well as DNA end resection, mimicking KLF4 and CtIP depletion (López-Saavedra and Cepeda-García, unpublished results). Thus, we wondered if PRMT5 effect in DNA end resection might be mediated, at least partially, through the methylation of KLF4.

In order to establish a genetic relationship between PRMT5 and KLF4 for DNA end resection, we first produced a mutant form of KLF4 that cannot be methylated by PRMT5 (KLF4-3RK). For this, we mutated the KLF4 methylation sites replacing the three arginines by lysines (R374K, R376K and R377K). Then, we analysed the formation of RPA foci in U2OS cells depleted or not of endogenous KLF4 and complemented by transient expression with either the control empty vector pMX, the siRNA resistant wild type KLF4 or the methylation deficient mutant KLF4-3RK (Figure R24A and R24B). In agreement with our previous results shown in Figure R20D, expression of ectopic KLF4 wild type protein in cells that retain endogenous KLF4 expression rendered cells more prone to resect DNA breaks (Figure R24B). Moreover, such wild type form of the protein complemented the resection defect observed upon depletion of endogenous KLF4 (Figure R24B). On the contrary, overexpression of the non-methylable KLF4, neither caused hyper-resection in control cells that maintain endogenous KLF4 nor complemented the defect in DNA end resection caused by endogenous KLF4 downregulation (Figure R24B). Levels of ectopic KLF4 are shown in Figure R24C.

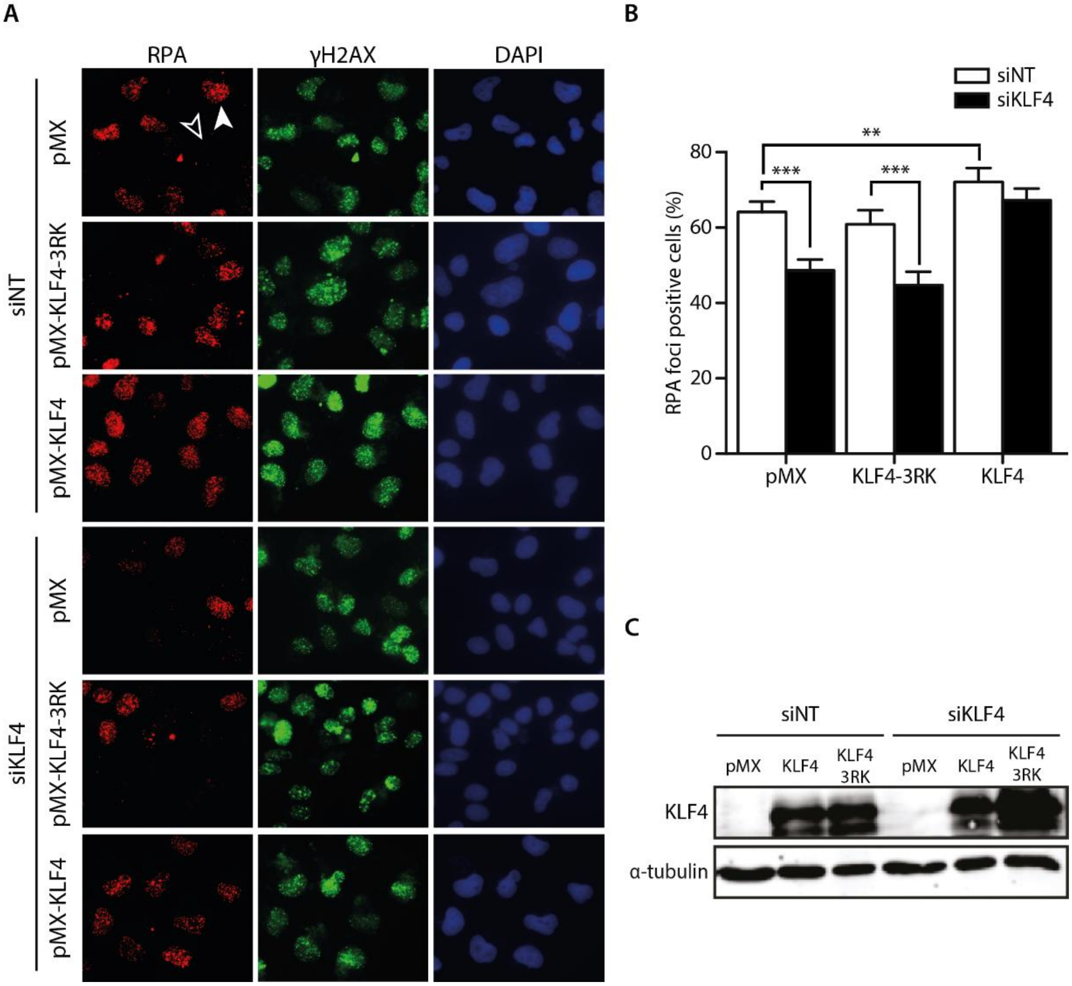

A
RPA
γH2AX
DAPI
siNT
pMX
pMX-KLF4-3RK
pMX-KLF4
siKLF4
pMX
pMX-KLF4-3RK
pMX-KLF4
B
siNT
siKLF4
**
***
***
RPA foci positive cells (%)
0
20
40
60
80
pMX
KLF4-3RK
KLF4
C
siNT
siKLF4
pMX
KLF4
KLF4 3RK
pMX
KLF4
KLF4 3RK
KLF4
α-tubulin

Figure R24. PRMT5-mediated methylation of KLF4 is essential for its role in DNA end resection.

(A) Representative fluorescence microscopy images of RPA (red), γH2AX (green) and DAPI staining in cells transfected with a siRNA against KLF4 or a control sequence, as indicated, and bearing a plasmid to express the wild type or non-methylable version of KLF4 (KLF4 and KLF4-3RK, respectively) or an empty vector (pMX). Cells were irradiated (10 Gy) and collected one hour later for immunofluorescence using an anti-RPA antibody. An example of cell defined as RPA foci positive cells is indicated by a full white arrow, while an example of RPA foci negative cell is marked with an empty arrow. **(B)** The percentage of cells positive for RPA foci upon DSB induction by ionizing radiation as described in (A) is plotted. Graph represents the average and SEM of three independent experiments. At least 200 cells were measured in each experiment. Statistical significance was calculated using a two-way ANOVA, but only biologically relevant statistical differences are shown for simplicity (* $p<0.05$; ** $p<0.01$; *** $p<0.001$). **(C)** Cells overexpressing wild type or non-methylable KLF4 were depleted for endogenous KLF4 or transfected with a control siRNA. Protein samples from those cells were resolved in an SDS-PAGE gel and blotted with the indicated antibodies.

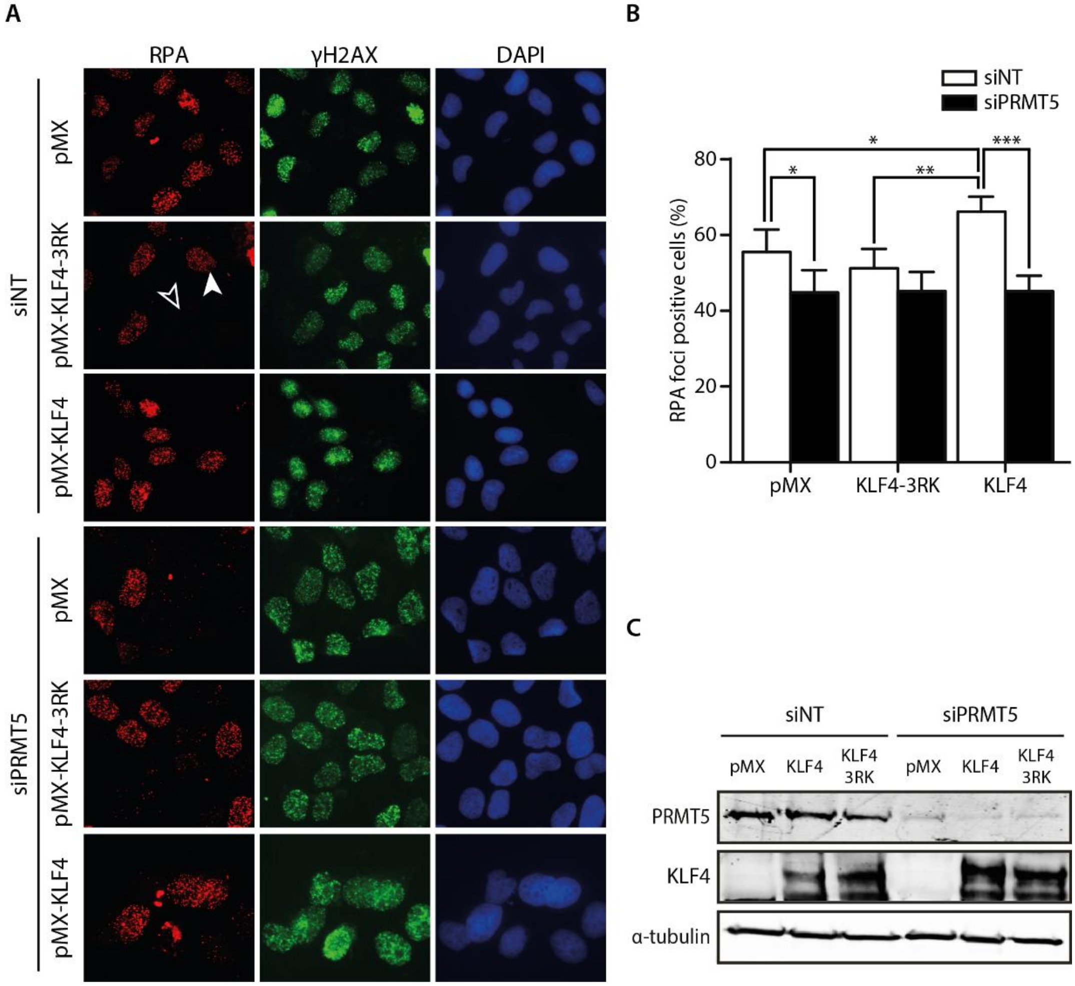


Figure R25. KLF4 overexpression does not rescue the defect on DNA end resection caused by PRMT5 depletion.

(A) Representative fluorescence microscopy images of RPA (red), γH2AX (green) and DAPI staining in cells transfected with a siRNA against PRMT5 or a control sequence, as indicated, and bearing a plasmid to express the wild type or non-methylable version of KLF4 (KLF4 and KLF4-3RK, respectively) or an empty vector (pMX). Cells were irradiated (10 Gy) and collected one hour later for immunofluorescence using an anti-RPA antibody. An example of cell defined as RPA foci positive cells is indicated by a full white arrow, while an example of RPA foci negative cell is marked with an empty arrow. **(B)** The percentage of cells positive for RPA foci upon DSB induction by ionizing radiation as described in (A) is plotted. Graph represents the average and SEM of three independent experiments. At least 200 cells were measured in each experiment. Statistical significance was calculated using a two-way ANOVA, but only biologically relevant statistical differences are shown for simplicity (* $p<0.05$; ** $p<0.01$; *** $p<0.001$). **(C)** Cells overexpressing wild type or non-methylable KLF4 were depleted for endogenous PRMT5 or transfected with a control siRNA. Protein samples from those cells were resolved in an SDS-PAGE gel and blotted with the indicated antibodies.

Given that PRMT5 seems to be upstream and in the same pathway as KLF4 for DNA end resection, RPA foci formation was also examined after overexpression of both the wild type KLF4 and the non-methylable mutant in cells downregulated for PRMT5 (Figure R25A and R25B). Strikingly, PRMT5 depletion was epistatic over KLF4 overexpression, both in cells expressing wild type or non-methylable variant of the protein (Figure R25B), in agreement with a genetic connection. Depletion of PRMT5 even eliminates the KLF4-overexpression induction of resection. Levels of ectopic KLF4 and endogenous PRMT5 are shown in Figure R25C.

Thus, genetically we can conclude that KLF4 role in resection is controlled by PRMT5-mediated methylation.

6. KLF4 affects TIP60 protein level

KLF4 is a transcription factor involved in a wide range of cellular processes. Thus, we reasoned that its role in resection might imply a control in the expression levels of proteins involved in DNA resection. However, we failed to see any changes in the amounts of core pro- or anti- resection proteins such as CtIP, etc (Figure R26). Thus, we hypothesized that it might affect other accessory resection proteins. Interestingly, PRMT5 has been already involved in regulating homologous recombination by modulating the action of the histone acetyltransferase (HAT) TIP60 at DNA breaks (Clarke et al., 2017). It has been described that some histone modifications regulate the recruitment and retention of 53BP1 to DSBs. In fact, the affinity of 53BP1 for H4K20me2 has been shown to be reduced by acetylation of histone H4 on lysine 16 by TIP60 and therefore allowing DNA end resection and HR (Jacquet et al., 2016). Accordingly, we wondered if KLF4 role in HR might be related also to TIP60. For this, we measured TIP60 expression at the protein level after overexpression of the wild type and non-methylable mutant version of KLF4. Indeed, KLF4 overexpression increased TIP60 protein levels (Figure R27A and R27B). Therefore, we conclude that there is a regulatory axis encompassing KLF4-PRMT5-TIP60 that affects DNA end resection and homologous recombination.

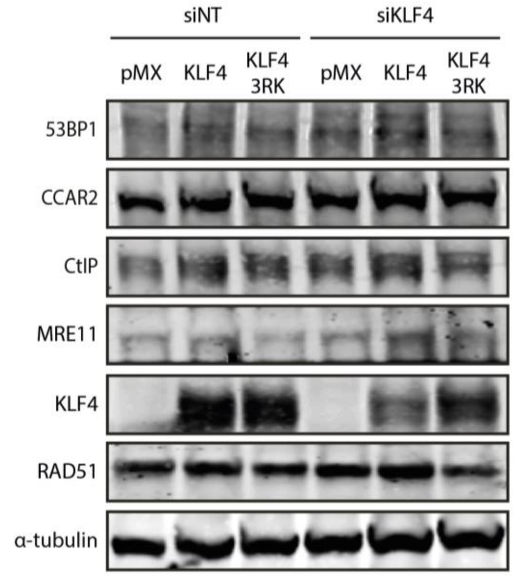


Figure R26. KLF4 does not affect the expression of some proteins involved in DSB repair.
A representative western blot of protein samples from cells downregulated or not for KLF4 and bearing a plasmid to express the wild type KLF4, a non-methylable form of the protein (KLF4-3RK) or an empty vector (pMX) resolved in SDS-PAGE and blotted with the indicated antibodies. α-tubulin is used as a loading control.

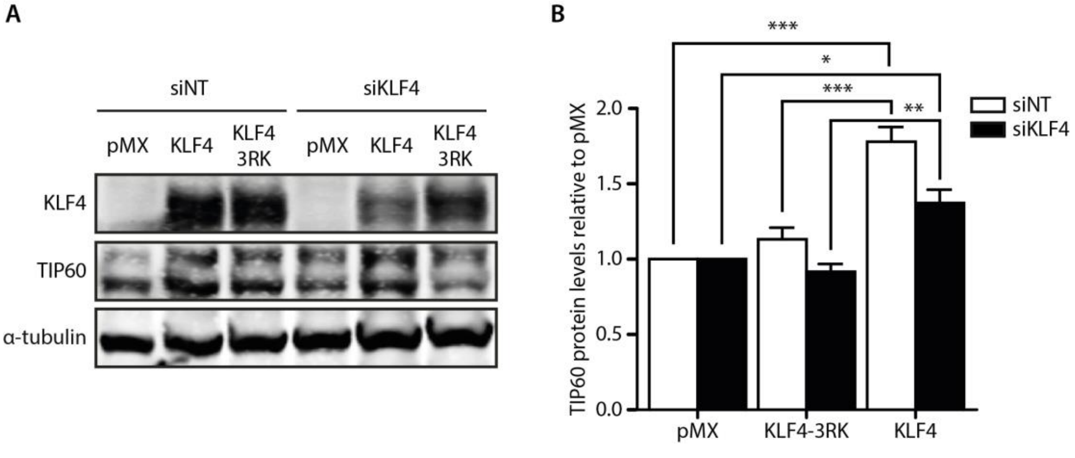


Figure R27. KLF4 overexpression controls TIP60 levels.
(A) A representative western blot of protein samples from cells downregulated or not for KLF4 and bearing a plasmid to express the wild type KLF4, a non-methylable form of the protein (KLF4-3RK) or an empty vector (pMX) resolved in SDS-PAGE and blotted with the indicated antibodies. **(B)** TIP60 protein levels were quantified from western blot images under the same conditions as (A) using a Li-Cor Odyssey system. Data are shown normalized to α-tubulin (loaging control) and relative to pMX control. The average and SEM of at least three independent experiments are shown. Statistical significance was calculated using a two-way ANOVA (* $p<0.05$; ** $p<0.01$; *** $p<0.001$)

IV. Discussion

The ability to reprogram somatic cells into induced pluripotent stem cells (iPSCs), which possess similar properties to physiological embryonic stem cells, holds great potential for biomedical research and clinic applications. However, evidence of DNA damage and genomic instability in iPSCs is one of the biggest concerns for their proper clinical use (Mayshar et al., 2010; Peterson and Loring, 2014) since genomic instability is associated with cancer and compromises the safety of iPSCs. Thus, an in-depth understanding of the causes and consequences of genetic abnormalities that arise during iPSCs formation is required. These genetic variations can come from different sources during iPSC generation and maintenance, however, they are highly linked with the reprogramming process itself, mainly by replication stress (Ruiz et al., 2015). The reprogramming process induces DNA double strand breaks (DSBs), the most cytotoxic type of damage, most likely due to replication problems that increase the risk of fork collapse (Zeman and Cimprich, 2014). Therefore, we aimed to study the role of DNA end resection as a relevant inherent mechanism that minimizes genomic instability during this process. Here, we demonstrate that DNA end resection is, indeed, hyper-activated during iPSCs formation and it is required for efficient reprogramming. In fact, depletion of CtIP, a key protein in resection (Huertas and Jackson, 2009; Sartori et al., 2007), impairs reprogramming. Several studies have shown that embryonic stem cells (ESCs) preferentially employ homologous recombination (HR) over the error-prone non-homologous end joining (NHEJ) in response to DSBs (Adams et al., 2010a; Serrano et al., 2010; Tichy et al., 2010), although it is suggested that NHEJ also have a role in repair of DSBs in humans ESCs (Adams et al., 2010b; Bogomazova et al., 2011). Likewise, iPSCs, which share numerous similarities with ESCs including enhanced DNA repair and high expression of genes involved in DNA damage signalling and repair, have been shown to require an intact HR pathway for an efficient reprogramming (González et al., 2013). Accordingly, it is not surprising that DNA end resection is essential for the process since it is the key decision point involved in the decision between DSBs repair pathways (Huertas, 2010). We have revealed that cells undergoing reprogramming not only have an increase in the number of breaks that are resected, but also an increase in resection processivity, measured as the length of resected DNA, compared with the parental somatic cells or already-differentiated cells. Moreover, we found a progressive upregulation of CtIP and checkpoint kinase 1 (CHK1) that follows the progression of the reprogramming. In agreement, increasing levels of CHK1 has been suggested to increase the efficiency of reprogramming limiting the replication stress (Ruiz et al., 2015). These findings are consistent with the requirement of an efficient repair mechanism and DDR machinery to allow successful cell reprogramming and to control genetic stability during the process (Kinoshita et al., 2011). Besides, a higher expression of several HR (RAD52, MRE11, NBS1) and NHEJ (XRCC4 and LIGIV) genes in pluripotent cells have already been reported (Momcilovic et al., 2010).

A probable explanation for this upregulation of resection and its bona fide regulator CtIP is the high incidence of DNA damage during reprogramming, mostly DSBs. CtIP is known to be a key player in maintaining genomic stability (Huertas and Jackson, 2009), and we reasoned that CtIP could be appropriately activated to repair DSBs generated by the presence of reprogramming factors. This idea is consistent with the increased γH2AX level, a robust marker for DSBs, that we found during reprogramming and that was already reported in previous works (González et al., 2013; Ruiz et al., 2015). Phosphorylation of histone H2AX has been implicated in the maintenance of genomic stability in response to DSBs (Rogakou et al., 1998), however, also occurs in response to replication stress at the sites of stalled replication forks, being also a hallmark of cells enduring replication stress (Ward and Chen, 2001). Moreover, CHK1 plays a central role in DNA replication checkpoint induced by replication stress. Helicase and polymerase uncoupling leads to accumulation of single-stranded DNA (ssDNA) bound by RPA at stalled forks. RPA-coated ssDNA recruits the complex ATR-ATRIP and activates ATR which subsequently phosphorylates and activates the main effector kinase CHK1, turning on the intra-S checkpoint to stabilize or repair stalled forks and to prevent late origin firing (Jossen and Bermejo, 2013; Zou and Elledge, 2003). Reprogramming in human and mouse cells has been associated with increased levels of replication stress (Ruiz et al., 2015), so it could be hypothesized that the detected ssDNA is formed by helicase-polymerase uncoupling during replication instead of resection. Moreover, the increase in CHK1 and CtIP during reprogramming could be in concordance with this hypothesis, since CHK1 is essential in replication stress response and CtIP has been reported to have a role in the response to replication fork stalling (Przetocka et al., 2018). However, while the role of CtIP initiating DNA end resection generates ssDNA overhangs, in stalled replication forks CtIP limits the accumulation of ssDNA by protecting stalled forks from nucleolytic degradation (Przetocka et al., 2018). Thus, we reason that the increase in ssDNA observed in this thesis is mainly due to an increase in the resection process since CtIP depletion led to a decrease in BrdU exposure and fibres length, contrary to an expected RPA-ssDNA increase if CtIP was protecting replication forks (Przetocka et al., 2018). Similarly to our observations with CtIP depletion, reprogramming efficiency was also reported to be impaired by downregulation of BRCA1, which accelerates DNA resection through its interaction with CtIP (Cruz-García et al., 2014). Moreover, other HR factors, like BRCA2 and RAD51, are required for successful reprogramming (González et al., 2013).

The low efficiency of reprogramming might be attributed to apoptosis since our results show an increase in the expression of apoptotic marker cleaved caspase 3 during reprogramming in the absence of CtIP, suggesting activation of the mitochondrial-dependent apoptosis pathway. Thus, impairment of HR-mediated DSB repair by selective CtIP depletion during mouse cell reprogramming interferes with iPSC generation by triggering apoptosis despite the existence of a functional NHEJ in these cells. This observation is in agreement with the two main strategies that pluripotent cells employ to diminish DNA damage and genomic instability, as activation of apoptosis or differentiation programs to eliminate cells with damaged DNA from the stem cell pool (Qin et al., 2007). These data are consistent with an increase of apoptotic cells observed during reprogramming of MEFs deficient in HR genes *Brca1, Brca2* or *Rad51* (González et al., 2013). Conversely, it has been previously reported that caspases 3 and 8 are activated by OCT4 and play key roles in mediating the human fibroblasts reprogramming through inactivation of retinoblastoma (Rb) protein, which regulates cell cycle progression (Li et al., 2010a). This could explain the basal level of activation of caspase 3 also detected in cells containing shNT, however, this activation was significantly higher in cells reprogrammed in CtIP deficiency, probably to activate apoptosis.

We also observed that CtIP deficiency during mouse cell reprogramming not only drastically impairs the reprogramming process but also endangers the future of the formed iPSCs by critically hampering their growth, maintenance of pluripotent state and their further differentiation to embryoid bodies (EBs). However, curiously and unexpectedly, the drastic effects of CtIP deficiency are highly specific to the cell reprogramming process, since the depletion of CtIP in already established iPSCs or ESCs does not reduce the ability of cells to self-renew or differentiate in EBs. It is suggested that CtIP contributes to the regulation of cell cycle progression at the G1/S transition by its association with Rb. Unphosphorylated Rb represses S-phase transition by directly binding to E2F transcription factor family and, therefore, cannot activate the transcription of the genes encoding the key components of the G1/S transition and the replication of the DNA. CDK-mediated phosphorylation of RB and direct interaction between CtIP and Rb are required for disrupting its association with E2F factors, thus activating S-phase genes expression, as well as for increasing CtIP expression (Liu and Lee, 2006). Indeed, inactivation of CtIP leads to early embryonic lethality by arresting cells in G1, being CtIP essential for the S-phase entry in inner stem cells of blastocysts (Chen et al., 2005a). In pluripotent stem cells, including ESCs and iPSCs, Rb is mostly phosphorylated, being unable to inhibit E2F and therefore, allowing the passage of cells into S phase leading a fast G1/S phase transition (Hindley and Philpott, 2013; Nevins, 2001). However, further investigations are necessary to see if there is also a connection between CtIP and Rb during reprogramming to regulate the cell cycle progression. Moreover, the influence of cell cycle in the reprogramming process is not completely known but, it has been observed that more slowly dividing or older cells reprogram worst and the activity of cyclins and CDKs correlates with efficiency of reprogramming (Mahmoudi and Brunet, 2012; Ruiz et al., 2011), like cyclin D1 overexpression and localization controls reprogramming (Edel et al., 2010). Thus, it is possible that reprogramming requires a different cell cycle structure, being CtIP more important, to that required for the maintenance of pluripotency, which would be supported by the observation that the treatment of ESCs with chemical inhibitors to slow the cell cycle has no impact on self-renewal or differentiation potential (Stead et al., 2002). Although this could explain the importance of CtIP essentially during reprogramming, our data support the idea that the requirement of CtIP during reprogramming is mostly associated with its role in resection and the maintenance of genomic stability. Thus, most of all of the effects due to CtIP deficiency during reprogramming are likely caused by the genomic aberrations acquired by cells during reprogramming. This is supported by the fact that cells that survive during reprogramming in the absence of CtIP and become iPSCs have increased genomic variations with respect to iPSCs generated under normal CtIP level. Thus, in the absence of CtIP, DSBs created during reprogramming would be repaired through more mutagenic repair pathways, thereby increasing mutagenesis and chromosomal rearrangements (Ceccaldi et al., 2016). This is consistent with an increase in DNA damage and genetic instability that would ultimately lead to apoptosis during reprogramming. These severe genomic consequences correlate with the well-established roles of CtIP in DNA end resection, HR and DSB repair pathway choice (Gomez-Cabello et al., 2013; Huertas and Jackson, 2009; Sartori et al., 2007), rather than reflecting a novel role of CtIP in the reprogramming process. Moreover, previous studies have shown that defects in Fanconi anemia (FA) complementation group decreases the efficiency of reprogramming (Müller et al., 2012), suggesting a potential link between HR and reprogramming since several FA pathway components promote HR repair (Kee and D’Andrea, 2010). Indeed, it has been shown that in response to interstrand crosslinks (ICL) DNA damage, FANCD2 physically interacts with CtIP which is essential to promote DNA end resection at the DSB generated during ICL processing to facilitate HR (Murina et al., 2014; Unno et al., 2014; Yeo et al., 2014). Likewise, MEFs deficient in BRCA1, a central component in HR, have problems for cell reprogramming and the derived iPSCs are unable to establish colonies (Gonzalez et al., 2013).

It is quite well established that ESCs and iPSCs exhibit a cell cycle with an abbreviated G1 phase and a longer S phase. This would be in concordance with the requirement of HR in pluripotent cells, as during S and G2 phases cells preferentially repair DNA by HR when sister chromatids are available to act as homologous templates (Rothkamm et al., 2003; Yang et al., 2016). Thus, we suggest that although both pluripotent and differentiated cells can likely repair their DNA by both NHEJ and HR, the latter would be the main mechanism able to deal with the endogenous damage during cell reprogramming. We hypothesize that this is due to the amount of DNA damage and the nature of the DNA lesion in these different situations. Most of the DSBs that can be attributed to endogenous processes are caused by stress during DNA replication. Replication stress, due to the presence of a stalled replication fork, readily causes the appearance of one-ended DNA DSBs and are commonly resolved mainly by HR or microhomology-mediated end joining to restore replication

(Aguilera and Gómez-González, 2008; Petermann and Helleday, 2010; Shibata, 2017). However, this would not fully explain why CtIP is more important during reprogramming than in cells already reprogrammed because replicative stress levels are also high in already established iPSCs. It is well-known that other endogenous sources, like reactive oxygen species (ROS), can produce DSBs (Cannan and Pederson, 2016). Moreover, ROS can alter replication through different mechanisms, leading to replication forks arrest. The metabolic change in iPSCs from oxidative respiration to oxidative glycolysis (Folmes et al., 2011; Varum et al., 2011) maintains low levels of intracellular ROS limiting the risk of cellular and genomic damage during self-renewal (Zhou et al., 2016). Nevertheless, ROS levels have been reported to early increase during reprogramming and gradually decreased over time, being required increased ROS levels as well as ROS signalling during the initial stage of nuclear reprogramming but not during proliferation and maintenance of iPSC (Zhou et al., 2016). Hence, increased ROS levels during reprogramming could lead to more collapsed forks and the generation of DSBs, which could explain why CtIP is essential during iPSC formation for effective repair but not in maintenance of iPSCs already established. In the absence of CtIP, those breaks would be erroneously repaired, inducing the observed CNV differences and other chromosomal abnormalities. Indeed, CtIP is essential for cell viability in cells that have been exposed to mutagens that result in high levels of DNA damage (Huertas and Jackson, 2009; Sartori et al., 2007) or chromosomal aberrations (Huertas and Jackson, 2009). Strikingly, and in agreement with this idea, our data suggest that cell reprogramming in wild type mouse fibroblasts causes an increase in DNA end resection that is comparable with high doses of ionizing radiation.

In summary, we propose a novel role of DNA end resection during cell reprogramming. Our data reveal that CtIP and resection are essential during reprogramming to minimize the impact of replication-induced DNA damage and maintain genomic stability. Moreover, reprogramming in a CtIP-defective environment has long-term consequences on self-renewal and differentiation (Figure D1).

Collectively, the available data from us and previous reports suggest that cells from patients with deficient DDR and DNA repair pathways could not be efficiently reprogrammed and used in regenerative medicine. However, to yield this powerful tool a future as a clinical standard procedure, more investigations are needed to elucidate which others genes are specifically required to avoid genomic instability during the reprogramming process.

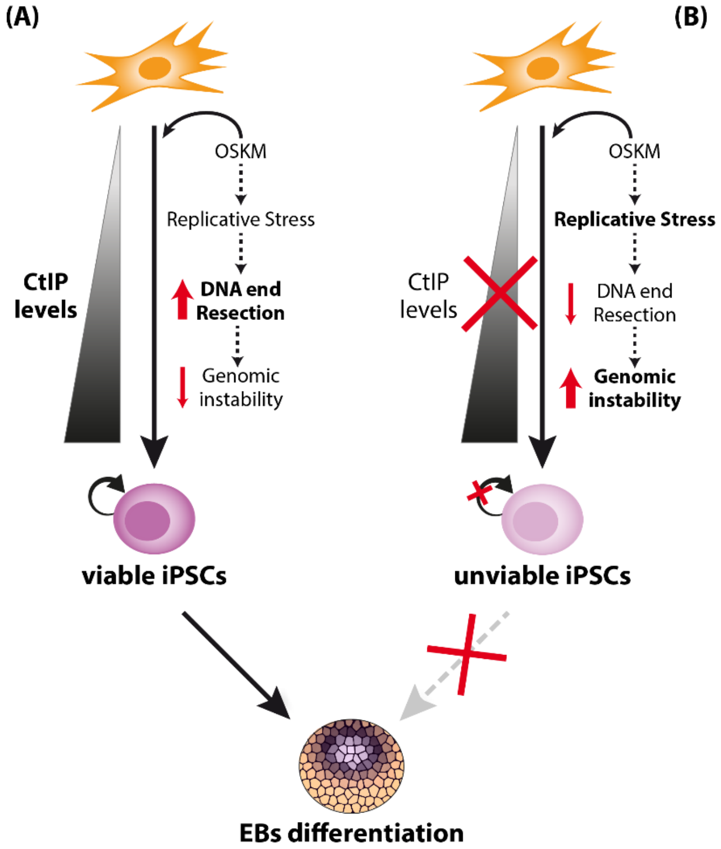

(A)
(B)
CtIP levels
OSKM
Replicative Stress
DNA end Resection
Genomic instability
CtIP levels
OSKM
Replicative Stress
DNA end Resection
Genomic instability
viable iPSCs
unviable iPSCs
EBs differentiation

Figure D1. DNA end resection during cell reprogramming.

(A) During cell reprogramming, the CtIP levels increase progressively leading to an increase in DNA end resection to repair DSBs by the error-free pathway homologous recombination. CtIP and resection are essential to minimize the genomic instability caused by the large amount of DNA damage caused due to replicative stress, thus obtaining viable iPSCs that maintain their capacity for self-renewal and differentiation. **(B)** During **r**eprogramming in a CtIP-defective environment DNA end resection, and therefore, HR, cannot be activated to repair the replication-induced damage, accumulating a large amount of genomic instability. CtIP deficiency during reprogramming has long-term consequences and leads to the generation of unviable or defective iPSCs that present hampered self-renewal and differentiation.

Interestingly, in a general screening searching for regulators of the NHEJ versus HR balance (López-Saavedra et al., 2016), the master regulator of cell stemness KLF4 was found as a positive candidate. KLF4 plays important roles in many fundamental biologic processes, being its role in cell reprogramming well established, as one of the four factors sufficient to induce pluripotency (Takahashi and Yamanaka, 2006). Moreover, KLF4 was found to play an essential role in modulating the DNA damage response and repair mechanisms to maintain genetic stability (El-Karim et al., 2013; Hagos et al., 2009; Yoon et al., 2005). According to these data, KLF4 might be involved in regulating DNA damage repair during mouse reprogramming, so we aimed to investigate the role of KLF4 in DSBs repair more in-depth.

Here, we have demonstrated that KLF4 levels control the repair pathway choice for DSBs, mainly by controlling the licensing of DNA end resection. KLF4 promotes both initiation and processivity of DNA end resection, leading to repair by HR. Interestingly, this means that during induced cell reprogramming with Yamanaka factor, whereas the process itself induces the accumulation of DNA damage (Blasco et al., 2011), the overexpression of KLF4 facilitates the repair of those breaks preferentially by HR instead of the more frequent NHEJ of somatic cells. Strikingly, this induced preference for HR could also explain, at least partially, why cells become so dependent of DNA end resection and HR during reprogramming (González et al., 2013).

KLF4 downregulation clearly impairs resection even in cells with naturally low endogenous levels of KLF4, as U2Os cells. Therefore, our data support a model in which KLF4 levels are intrinsically relevant for the DNA damage response. In ESCs, HR is enhanced (Adams et al., 2010a; Serrano et al., 2010; Tichy et al., 2010) and this could be mediated by KLF4. By controlling DNA end resection, KLF4 might also modulate the activation of DNA damage checkpoint, as created RPA-coated ssDNA stimulates the binding of ATRIP which enables the recruitment of the ATR-ATRIP complex to the sites of DNA damage and initiates the checkpoint signalling (Zou and Elledge, 2003). Indeed, both master checkpoint kinases, ATM and ATR, are oppositely regulated by DNA end resection in a length-dependent manner. Single-stranded overhangs reduce ATM activation and potentiate ATR activation, thereby DNA end resection is an essential element in promoting the ATM-ATR switch (Shiotani and Zou, 2009). This might explain why ESCs are largely dependent on ATR for efficient DSBs repair by HR (Adams et al., 2010a), which is consistent with the early embryonic lethality observed in mice due to ATR gene inactivation (de Klein et al., 2000). Otherwise, INO80, which increases expression of key pluripotency factors including OCT4, NANOG, SOX2, KLF4 and ESRRB in ESCs (Wang et al., 2014b), has been shown to function together with YY1 to regulate HR (Wu et al., 2007). Moreover, it has been recently reported that YY1 regulates the expression of KLF4 at transcriptional levels (Morales-Martinez et al., 2019) and loss of YY1 is related to chromatid and chromosome aberrations (Wu et al., 2007). Thus, we can speculate that YY1 and INO80 would increase KLF4 levels leading to an increment in the DSBs repair by the HR pathway, in concordance with enhanced HR in ESCs. We show that KLF4 does not affect cell cycle progression, however, it has been previously reported that overexpression of KLF4 induces cell cycle arrest through the transcriptional activation of p21 and the inhibition of cyclin D1 (Chen et al., 2001; Shie, 2000; Song et al., 2013) among others. More importantly, upon ionising radiation and in response to DNA damage, KLF4 expression is induced and it is required to activate G1/S (Yoon et al., 2003) and G2/M (Yoon and Yang, 2004) checkpoints. In spite of the effect of KLF4 over cell cycle that could modify the balance between NHEJ and HR which are regulated by the cell cycle (Rothkamm et al., 2003), we suggest that the role of KLF4 in resection is independent of its role in cell cycle progression as cell cycle was not affected by altering KLF4 levels in our hands. Considering the role of KLF4 as checkpoint activator and now, based on our results, as inductor of the error-free HR repair pathway, is not surprising that cells knocked out for KLF4 are genetically unstable and accumulate DNA breaks (El-Karim et al., 2013; Hagos et al., 2009). KLF4 has also been found to suppress p53-dependent apoptotic pathway by directly inhibiting p53 and suppressing BAX expression (Choi et al., 2018; Ghaleb et al., 2007; Talmasov et al., 2015; Zhou et al., 2009) but under certain conditions and contexts, KLF4 may switch its role from anti-apoptotic to pro-apoptotic role (Li et al., 2010b; Wang et al., 2015). Indeed, although KLF4 was first identified as a potential tumour suppressor agreeing with its role in cell cycle arrest, it has also been identified as an oncogene that could be explained by its role in suppressing p53-dependent apoptosis. Consistent with this, KLF4 expression has been found either increased (Foster et al., 2000; Pandya et al., 2004) or decreased (Wei et al., 2005; Zhao et al., 2004) in various human cancer types. Thus, KLF4 might possess a context-dependent activity in checkpoint activation and cell cycle progression as well as in apoptosis.

KLF4 is highly regulated by many stimuli and modifications (Chen et al., 2005b; Cullingford et al., 2008; Yang and Zheng, 2014). Post-translational modifications (PTMs), such as phosphorylation, acetylation and ubiquitylation among others, can regulate protein stability, subcellular localization, interactions with other proteins and activities of target proteins. KLF4 is known to be subjected to multiple modifications including phosphorylation (Gunasekharan et al., 2016; Kim et al., 2012), sumoylation (Du et al., 2010; Tahmasebi et al., 2013), acetylation (Evans et al., 2007), ubiquitylation (Lim et al., 2014), glutamylation (Ye et al., 2018) and methylation (Hu et al., 2015). Therefore, understanding KLF4 regulation could enhance our comprehension about the double function of KLF4 as both tumour suppressor and oncogene. In fact, in human papillomaviruses-positive keratinocytes, it has been observed that KLF4 expression and activities are disturbed by post-transcriptional and post-translational modifications (Gunasekharan et al., 2016).

Based on the importance of modifications in KLF4, we investigated whether any post-translational modification could be important for its role in resection. Interestingly, PRMT5-mediated arginine methylation of KLF4 has an important effect on tumorigenesis and DNA damage response (Hu et al., 2015). Furthermore, since an increase of KLF4 methylation as well as interaction between KLF4 and PRMT5 was observed after ionizing radiation (Hu et al., 2015), we anticipated that this post-translational modification might be involved in DNA end resection role of KLF4. As hypothesized, we found that this methylation is, indeed, required for proficient DNA end resection since either the ectopic expression of a methylation-deficient mutant (KLF4-3RK) or downregulation of PRMT5 impaired the resection. KLF4 methylation by PRMT5 is proposed to stabilize the protein by preventing its degradation, so it would be reasonable to believe that the requirement for resection is probably due to the fact that such post-translational modification controls KLF4 levels (Hu et al., 2015). In fact, we can clearly show a direct correlation of KLF4 expression and DNA resection efficiency. Despite the aforementioned role in controlling

KLF4 level, we propose a more direct role of this modification since DNA end resection is also impaired even when the protein level of non-methylable KLF4 (KLF4-3RK) has not yet decreased by degradation.

Given that KLF4 is a transcription factor, it is possible that KLF4 affects resection and recombination through the control of other factors, either directly or indirectly by regulating genes important for repair at the level of gene expression. Surprisingly, we found that the expression levels of histone acetyltransferase KAT5, also known as TIP60, were affected by KLF4. Indeed, the overexpression of KLF4 increased TIP60 protein levels while the KLF4-3RK mutant was unable to do it. Interestingly, TIP60 promotes HR repair over NHEJ by acetylating the histone H4K16, which disrupts 53BP1 binding and thereby releases DNA end resection inhibition (Jacquet et al., 2016). Moreover, PRMT5 has been shown to control homologous recombination-mediated repair by regulating the methylation of RUVBL1, which is critically required for the acetyltransferase activity of TIP60 at DNA breaks to facilitate recombination (Clarke et al., 2017). Here we show that all these three proteins are related and form a regulatory axis in promoting homologous recombination repair. Thus, we can speculate that the role of KLF4 in DNA end resection is through TIP60 acetyltransferase activity, by regulating the expression of TIP60 or by the physical interaction between TIP60 and KLF4 that has been previously demonstrated (Ai et al., 2007). Indeed, KLF4 has also been involved in the modulation of chromatin structure controlling histone H4 acetylation and deacetylation at the promoters of target genes through recruitment of p300/CBP histone acetyltransferases or histone deacetylases respectively (Evans et al., 2007). Moreover, the histone acetyltransferases CBP and p300 have been shown to promote homologous recombination at different levels, both by contributing to the transcriptional activation of recombination genes, such as BRCA1 and RAD51 (Ogiwara and Kohno, 2012) but also by acetylating recombination proteins at DSB sites (Qi et al., 2016; Yasuda et al., 2018). However, if this relies on KLF4 and/or PRMT5 too or is an independent event must be discovered in future studies.

Altered expression of both, PRMT5 and KLF4, has been related with many types of cancers and more specifically breast cancer (Fletcher et al., 2011; Li et al., 2013; Wang et al., 2017; Yang and Bedford, 2013). Some breast cancers are strongly associated with altered homologous recombination and mutations in BRCA1 are linked with triple-negative breast cancer (Afghahi et al., 2016). However, the influence of KLF4 as tumour suppressor or oncogene remains controversial in breast cancer. Indeed, although some reports have shown that 70% of breast carcinomas have elevated KLF4 expression that is associated with poor prognosis as it is required for cell migration and invasion (Foster et al., 2000; Pandya et al., 2004; Yu et al., 2011), others suggest that KLF4 inhibits tumour progression and metastasis (Yori et al., 2010, 2011). Recently, it has been discovered that PRMT5 controls TIP60 RNA splicing (Hamard et al., 2018). In hematopoietic cells, PRMT5 depletion caused aberrant splicing of TIP60 which cannot carry out its acetyltransferase activity leading to impaired HR (Hamard et al., 2018). Interestingly, KLF4 is also subjected to alternative splicing producing the cytoplasmic version KLF4α (Ferralli et al., 2014). Moreover, KLF4α has been shown to antagonize with the KLF4 full-version by retaining it in the cytoplasm and thereby impeding its function as tumour-suppressor in the nucleus (Ferralli et al., 2014). Thereby, the alternative splicing of KLF4, and maybe also TIP60, might explain the different role of KLF4 as a tumour suppressor or oncogene in different types of cancers by allowing or impeding the already known KLF4 functions including cell cycle arrest, apoptotic suppression and the new function discovered in this thesis, homologous recombination induction.

Actually, numerous studies have implicated KLF4 in the regulation of cancer stem cells (CSCs). In fact, both PRMT5 and KLF4 expression are essential for breast CSCs survival and propagation (Chiang et al., 2017; Qi et al., 2019; Yu et al., 2011). It is known that CSCs have, in general, an increased repair efficiency (Lim et al., 2012; Yuan et al., 2014) which confers resistance to conventional chemotherapy and radiotherapy (Ahmed et al., 2015; Bao et al., 2006; Desai et al., 2014; Kreso and Dick, 2014; Yuan et al., 2014), which might be explained by higher levels of KLF4 driving DSBs repair through HR. Moreover, deregulation of KLF4 PRMT5-mediated methylation induces carcinogenesis (Hu et al., 2015). Thus, it is possible that the role of KLF4 and PRMT5 in cancer might be caused by an altered recombination pattern. Indeed, our data would suggest that KLF4 inhibition might sensitize cancer cell towards ionising radiation in tumours with high rates of HR, while in those tumours in which HR is impaired, such as triple-negative breast cancer, an increase in KLF4 can lead to repair DNA lesions through the hampered recombination, causing cell death in the process. Hence, although additional studies are required, our observations might have an impact on the understanding and potential treatment of cancer in the future.

V. Conclusions

Conclusions

1. CtIP protein levels increase upon mouse cell reprogramming leading to a hyper-activation of DNA end resection, most likely to deal with the increase in DNA damage caused endogenously during the process in order to maintain genomic stability.

2. CtIP deficiency during reprogramming causes genomic instability and hence, hinders the process efficiency by inducing apoptosis to avoid spreading the damage. Moreover, iPSC formed in the absence of CtIP are compromised in self-renewal and their ability to differentiate.

3. KLF4 is involved in DSB repair pathway choice and its absence impairs HR and stimulates NHEJ. As a consequence, KLF4 is required to avoid cell sensitivity to DSB inducing agents such as camptothecin, etoposide and ionizing radiation.

4. KLF4 promotes both initiation and processivity of DNA end resection, leading to repair by HR. Mechanistically, KLF4 affects the recruitment of pro- and anti-resection proteins to DSBs sites such as RIF1 foci and BRCA1.

5. PRMT5-mediated methylation of arginine residues R374, R376 and R377 on KLF4 is required for its role in DNA end resection.

VI. Materials and methods

1. Cell culture procedures

1.1. Cell lines, growth media and conditions

Cell lines used in this thesis were C57BL/6 mouse embryonic fibroblast carrying a doxycycline-inducible tetracistronic cassette encoding the four murine reprogramming factors Oct4, Sox2, Klf4, and c-Myc (MEFs-i4F), mouse induced-pluripotent stem cells (miPSCs), D3 mouse embryonic stem cells (ES-D3), human embryonic kidney (HEK293T), human osteosarcoma (U2OS) or cells lines derived from them harbouring different constructs (Table M1).

MEFs-i4F, HEK293T and U2OS cells were cultured in high-glucose Dulbecco's Modified Eagle Medium (DMEM; Sigma, D6546), supplemented with 10% fetal bovine serum (FBS; Sigma, F7524), 2 mM L-glutamine (Gibco, 25030024), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco, 15140122). miPSCs were cultured on plates coated with 0.1% gelatin (Sigma, G1890) using Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 with GlutaMAX Supplement (DMEM/F12 + GlutaMAX; Gibco, 10565018) supplemented with 20% knockout serum replacement (Gibco, 10828028), 1.000 U/mL Leukaemia Inhibitory Factor (LIF; Millipore, ESG1107), 1% non-essential amino acids (Gibco, 11140035), 0.1 mM b-mercaptoethanol (Gibco, 31350010), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco, 15140122); and ES-D3 which were cultured in DMEM supplemented with 15% FBS, 1% non-essential amino acids, 1 mM sodium pyruvate, 100 U/ml penicillin, 100 µg/ml streptomycin, 0.1 mM 2-mercaptoethanol, and 1,000 U/ml mouse LIF. Stable cell lines derived from the above-mentioned ones were cultured in the same conditions but adding either 0.5 mg/ml G418 (Sigma, A1720) or 1 µg/ml puromycin (Sigma, P8833) to the medium, according to the selection marker of the inserted plasmid.

HEPA class 100 incubators (Thermo) were used to maintained cell lines at 37ºC and 5% CO2.

Trypsin-EDTA solution (Sigma, T4049) or accutase (Invitrogen, 00-4555-56) were used to detach the cells when required. To seed specific numbers of cells, an automatic cell counter was used (Z2 Coulter Counter, Beckman Coulter).

For long-term preservation of the cells, they were harvested and pelleted by centrifugation at 500 g for 3 minutes. Cells were then resuspended in freezing solution (10% dimethyl sulfoxide [DMSO] in FBS; or 10% DMSO in KSR for iPSCs), aliquoted in tubes and maintained at -80ºC for at least 24 hours. Then, they were transferred to liquid nitrogen-containing tanks for long-term storage.

1.2. Mouse cell reprogramming into iPSC

1.5 x 10^5 C57BL/6 primary MEFs carrying a doxycycline-inducible tetracistronic cassette encoding the four murine reprogramming factors *Oct4, Sox2, Klf4*, and *c-Myc* (MEFs-i4F) were seeded per well on a 12-well plate. MEFs-i4F were cultured in miPSC medium supplemented with 1 mg/mL doxycycline (Sigma, D9891) to induce the expression of OSKM and to promote reprogramming. Medium was changed every 24 hours for 21 days or until iPSCs colonies appeared. iPSCs were then expanded in 6-well gelatin-coated plates and in iPSC medium without doxycycline.

1.3. Isolation of iPSCs colonies

After reprogramming, iPSCs were seeded in 6-well plates to isolate colonies grown from single cells. When iPSCs had formed isolated colonies of a suitable size, culture medium was removed and PBS was added to the plate. Each colony was identified under an inverted optical microscope and manually picked up by aspirating with the aid of an automatic pipette in sterile conditions. Each colony was placed in a 96-well plate and trypsin was added for 5 min at 37ºC to detach the cells. Then, separated cells from a colony were seeded in a well of a 24-well plate coated with 0.1% gelatin and containing 1 ml of iPSC media.

1.4. Proliferation assay

After transduction with shCtip and shNT particles, 2x10^3 MEFs-i4F per well were seeded in 96-well plates to measure real-time proliferation in the xCELLigence RTCA DP (ACEA Biosciences), which uses non-invasive electrical impedance monitoring to quantify cell proliferation. The impedance of electron flow caused by adherent cells is reported using a parameter called Cell Index (CI), where CI = (impedance at time point n – impedance in the absence of cells)/nominal impedance value. Cell proliferation was monitored each hour for 72 hours and all experiments were performed using technical duplicate for each sample in three independent experiments.

1.5. GFP-shRNA stability analysis

The populations of miPSCs reprogrammed either in the presence or absence of CtIP were enriched by cell sorting (see paragraph 3.2.5.) taking advantage of the presence of a GFP gene on the plasmids shNT and shCtip. After the enrichment, 1.5 x 10^5 miPSCs were seeded in 6-well plates and maintained for 3 weeks changing the media every day and passed in a 1:3 ratio once a week. The percentage of GFP was measured by flow cytometry during each passage of the cells using a BD FACSCalibur Flow Cytometer (BD Biosciences, 342975).

1.6 Colony formation assay

miPSCs reprogrammed either in the presence or absence of CtIP were enriched by cell sorting (see paragraph 3.2.5.) and 1 x 10^5 miPSCs were seeded in 6 well plates in triplicates. The media was changed every day until colony formation and then, several pictures were taken with an optical microscope. The number of colonies was counted and relativized to the control siNT. Colonies were also analysed for size using Adobe Photoshop CS6 (Adobe Systems Incorporated). At least 300 colonies for condition of three independent experiments were measured.

For ES-D3, 2000 cells depleted or not of CtIP were seeded in 6-well plate by triplicate. 5 days after, the number of colonies was measured as mentioned above.

1.7. Embryoid bodies formation

ES-D3 and miPSCs were detached with accutase, counted, and replated onto ultra-low attachment 6-well plates with iPSC regular medium without LIF for 3–4 days to induce differentiation. Embryoid bodies were analysed for size and number through microscopy images using Adobe Photoshop CS6 (Adobe Systems Incorporated).

1.8. siRNA Transfection

For transient protein knockdown, siRNA transfection was performed using Lipofectamine RNAiMAX transfection reagent (Thermo Fisher, 13778150) following manufacturer's instructions.

Briefly, cells were seeded in different plate formats, accordingly to the requirements of each specific experiment, and grown for 24 hours. The day of transfection, medium was replaced by fresh DMEM supplemented with L-glutamine but without antibiotics and cells were incubated with a mix of siRNA and RNAiMAX diluted in Opti-MEM (Gibco, 11058-021). Cells were then incubated at 37ºC for 6 hours before replacing the media with fresh complete DMEM to minimise cell death. All siRNA-mediated knockdowns were validated 48 hours after transfection by quantitative RT-PCR or western blot. The list of all siRNAs used in this thesis can be found in Table M2.

1.9. Plasmid DNA Transfection

Plasmids used in this thesis are listed in Table M3. Two alternative strategies were used to introduce plasmid DNA into human cells depending on the experiment: Transfection with FuGene or transfection with calcium phosphate.

1.9.1. FuGENE transfection

FuGENE 6 Transfection Reagent (Promega, E2691) was used to transfect expression vectors following manufacturer's instructions. Different plate formats were used for transfection depending on the requirements of each specific experiment. For 60-mm plates, 300.000 cells were seeded the day before transfection. Then, 3 µg of plasmid DNA was mixed with FuGENE in a 3:1 FuGENE:DNA ratio, and diluted in Opti-MEM. Solution was incubated 15 min at room temperature and finally was added dropwise to the plate with gentle rocking.

1.9.2. Calcium phosphate transfection

Calcium phosphate transfection protocol was used to introduce the plasmids required for lentivirus production in HEK293T cells. For 100-mm plates, 2.7×10^6 cells were seeded and grown for 24 hours. The media was changed with fresh one at least 30 minutes before transfection. The three plasmids required were mixed in a 3:2:1 ratio, i.e. 15 µg of the vector containing DNA of interest, 10 µg of the vector required to produce the viral capsid (p8.91), and 5µg of the vector encoding the viral envelope (pVSVG) (Table M3).

The DNA mix was prepared in a volume of 0.5 ml containing 250 mM $CaCl_2$ and was added dropwise while bubbling into 0.5 ml of 2x HEPES buffered saline (HBS; Sigma, 51558). The mix was incubated for 30 min at room temperature and added dropwise to the plates with gentle rocking. The day after transfection, the medium was removed and cells were washed once with PBS before adding fresh medium to minimise cell death. Some experiments required larger or smaller plates, so volumes were scaled accordingly in those cases.

1.10. Lentivirus production and transduction

Lentiviral particles harbouring shRNA specifically targeting *Ctip* or a non-target sequence as a control, or bearing the pRRL_sEF1a_HA.NLS.SceOPT.T2A.TagBFP vector, to express the restriction enzyme I-SceI and the blue fluorescent protein (BFP), were produced in HEK293T cells by transfecting with the specific plasmids together with p8.91 and pVSVG as described in the previous section. From that point onward, all procedures were carried out in a P2 biological safety room. 48 h after transfection, medium was collected and filtered using 0.45 µm Surfactant-free Cellulose Acetate (SFCA) Sterile filters (Minisart NML; Sartorius, 16555) to retrieve the viral particles while removing cellular debris. Then, lentiviral particles were pelleted by centrifugation at 22.000 rpm for 1 hour and 30 min at 4ºC. Supernatant was removed and viruses were resuspended in DMEM and stored in aliquots at -80ºC until use. Before infection, viral production was titrated by BFP (for particles containing the I-SceI enzyme) or GFP (for viruses bearing the shRNA constructs) expression using flow cytometer (see paragraph 3.2.3).

For transduction, lentiviral particles were diluted in DMEM supplemented with 8 µg/ml hexadimethrine bromide (Polybrene; Sigma, H9268) with a multiplicity of infection (MOI) of 10, i.e. 10 virus particles are added per cell during infection, and this solution was added to the plates. Cells were washed the next day with fresh medium to remove viral residues and hexadimethrine bromide. CtIP knockdown was validated by western blot.

2. Molecular biology procedures

2.1. Nucleic acids manipulations

2.1.1. Plasmid DNA amplification

Plasmid DNA was amplified by transforming competent cells of the DH5α strain of Escherichia coli (Model Organism Service, CABIMER) using the heat shock protocol. Briefly, 0.1 ml of competent bacteria were mixed with plasmid DNA and incubated on ice for 30 min. Cells were then placed at 42ºC for 85 seconds and returned to the ice for 5 min for a heat shock. Then, 1 ml of LB (Formedium, LB-Broth Lennox) broth was added and the transformed cells were incubated at 37ºC for 30 min. Cultures were harvested by centrifugation at 6000 rpm for 2 min and plated on LB agar plates supplemented with 100 µg/ml ampicillin (Sigma, A9518) or 25 µg/ml kanamycin (Sigma, K4000) depending on the antibiotic resistance cassette of the transformed plasmid.

Plasmid DNA was purified from one single colony using PureYield Plasmid Maxiprep System (Promega, A2393), PureYield Plasmid Midiprep System (Promega, A2492) or NucleoSpin Plasmid EasyPure Miniprep System (Macherey-Nagel, 740727), following the manufacturer's instructions. DNA concentration was quantified by measuring 260 nm absorbance using a NanoDrop ND-1000 spectrophotometer.

2.1.2. DNA digestion with restriction enzymes

For DNA cloning and plasmid checking, restriction endonucleases from Takara or New England Biolabs were used according to manufacturer's instructions. DNA fragments were resolved in agarose gels as described in the following section.

2.1.3. DNA electrophoresis in agarose gels

DNA electrophoresis was performed on gels containing a variable percentage of agarose (Pronadisa, 8010.22) depending on the size of the bands to differentiate, and RedSafe (Intron Biotechnology, 21141) for DNA staining, both diluted in 1x TAE buffer (40 mM Tris-HCl pH7.6, 20 mM acetic acid, and 1 mM EDTA). Loading buffer (Takara) was added to DNA samples prior to loading in the gel, and 1 kb DNA ladder (gTPbio, GTPBM0002) was used for size estimation of the bands. Stained DNA fragments were visualized using an ultraviolet transilluminator (Bio-Rad) and analysed by Quantity One software.

2.1.4. Site-directed mutagenesis

Directed changes in the DNA sequence of plasmids were performed using QuickChange Lightning Site-Directed Mutagenesis kits (Agilent Technologies, 210518) according to manufacturer's instructions. Briefly, mutagenesis was performed by PCR using designed mutagenic primers (Table M4). After the PCR, parental DNA was degraded by DpnI nuclease digestion. Finally, mutation-containing synthesized DNA was transformed into competent bacteria and amplified. Candidates were sent to CNIO Genome Unit (Madrid, Spain) for DNA sequencing.

2.1.5. RNA extraction

RNA extracts were obtained from cells using NZY Total RNA Isolation kit (NZYtech, MB13402) according to manufacturer's instructions. RNA concentration was quantified by measuring 260 nm absorbance using a NanoDrop ND-1000 spectrophotometer, and the RNA quality was checked by visualization on a 1% agarose gel.

To remove trace amounts of DNA, 1 µg RNA was treated with RQ1 RNase-Free DNase (Promega, M6101) according to manufacturer's instructions

2.1.6. Reverse transcription

To synthesize complementary DNA (cDNA), 1 µg RNA was subjected to reverse transcription reaction using Maxima H Minus First Strand cDNA Synthesis kit (Thermo Scientific, K1652) according to manufacturer's instructions.

2.1.7. Quantitative PCR (qPCR)

Quantitative PCR from cDNA was performed in triplicate with iTaq Universal SYBR Green Supermix (Bio-Rad, 172-5124) following manufacturer's instructions. DNA primers used for qPCR are listed in Table M4. qPCR was performed in an Applied Biosystem 7500 FAST Real-Time PCR system. The comparative threshold cycle (Ct) method was used to determine relative transcripts levels (Bulletin 5279, Real-Time PCR Applications Guide, Bio-Rad), using an endogenous housekeeping gene, β-actin, as an internal control. Expression levels relative to β-actin were determined with the formula 2-ΔΔCt (Livak and Schmittgen, 2001).

2.2. Protein analysis

2.2.1. Protein extraction under denaturing conditions

Protein extraction was carried out by lysing the cells in Laemmli buffer 2x (125 mM Tris-HCl pH 6.8, 4% SDS, and 20% glycerol) using a plastic cell scraper (Sarstedt, 83.1830). Alternatively, cells were harvested with trypsin, rinsed with PBS and resuspended in Laemmli 2X buffer. To reduce sample viscosity, protein extracts were passed through a syringe with a 0.5x16 mm needle (BD Plastipak, 303175) at least 10 times.

Protein concentration was determined by the measurement of absorbance at 280 nm using a NanoDrop ND-1000 spectrophotometer.

2.2.2. SDS-PAGE electrophoresis

Proteins were separated by size in 29:1 acrylamide:bis-acrylamide gels prepared at different concentrations according to the molecular weight of the proteins to analyse.

SDS-PAGE was performed according to previously described method (Laemmli, 1970). Samples were diluted to obtain similar protein concentration and then SDS-PAGE loading sample buffer 4x (250 mM Tris-HCl pH6.8, 8% SDS, 40% glycerol, 20% β-mercaptoethanol (Sigma, M6250) and bromophenol blue) was added to a final concentration of 1x. Samples were incubated at 100ºC for 5 min and then loaded into the gels. Prestained protein ladder (gTPbio, GTPBM003) was also loaded as a molecular weight marker. Electrophoresis was performed in a Mini-PROTEAN Tetra Cell (Bio-Rad) with running buffer (25 mM Tris-HCl pH 8.3, 190 mM glycine, and 0.1% SDS) at 100-150 V.

2.2.3. Western blot analysis

After electrophoresis, proteins were wet-transferred using Mini Trans-Blot system (Bio-Rad) for 1-2 h at 400 mA in transfer buffer (25 mM Tris-HCl pH 8.3, 190 mM glycine, 20% methanol and 0.1% SDS) into PVDF membranes (Immobilon-FL; Millipore, IPFL00010) previously activated in methanol for 1 min and equilibrated in transfer buffer. Commercial Odyssey Blocking Buffer (LI-COR Biosciences, 927-40000) was used for blocking the membrane for at least 1 hour at room temperature to reduce background signal with the antibodies, and then, membranes were incubated overnight at 4ºC with the appropriate primary antibodies (Table M5) diluted in Odyssey Blocking Buffer containing 0.1% Tween-20. Then, three washes of 5 minutes were performed with 0.1% Tween-20 in TBS, followed by 1 hour incubation at room temperature protected from ambient light with the corresponding IRDye secondary antibodies (Table M6) diluted in blocking buffer containing 0.1% Tween-20. Afterwards, membranes were washed again three times with 0.1% Tween-20 in TBS and dried before scanning. Image acquisition was performed in Odyssey CLx Imaging System (LI-COR Biosciences) at two infrared wavelengths (700 and 800 nm) for differential imaging of both anti-rabbit and anti-mouse secondary antibodies at the same time. ImageStudio v.2.1 software was used for scanning and analysis of the images.

3. Cell biology procedures

3.1. DSBs repair assays *in vivo*

3.1.1. Description of repair systems

To study the effect of KLF4 in DSB repair pathways, we used different repair systems in which a DSB is generated by the rare-cutting I-SceI endonuclease, whose 18-bp recognition sequence has been integrated into the sequence. The repair of the damage results in the expression of a fluorescent protein (GFP or RFP) that could be quantified by flow cytometry.

To analyse DSB repair by homologous recombination, SA-GFP and DR-GFP reporters, which each one measures a different HR pathway, (Figure M1A and M1B) were used.

SA-GFP reporter (Stark et al., 2004) (Figure M1A) consists of two GFP gene fragments (5’GFP and 3’GFP), which have 266 bp of homology (marked in grey) and are separated by 2.7 Kb. I-SceI site is present in the downstream GFP fragment inside the region of homology. Repair of the I-SceI-generated DSB by single-strand annealing (SSA) results in a functional GFP gene after annealing of complementary strands of both GFP fragments, followed by appropriate DNA-processing steps that produce a 2.7 kb deletion in the chromosome. The reporter can also be repaired by other homology-directed repair (HDR), but this repair does not restore a functional GFP gene.

DR-GFP system (Pierce et al., 1999) (Figure M1B) was used to measure homology-directed repair efficiency, specifically short tract gene conversion. This reporter is composed of two differentially mutated GFP genes oriented as direct repeats and separated by the puromycin N-acetyltransferase gene. The upstream repeat is a full-length GFP gene mutated at a BcgI restriction site to contain the recognition site for the I-SceI endonuclease. This mutation consists of a substitution of 11 bp of wild-type GFP sequence with those of the I-SceI site that supplies two in-frame stop-codons, which terminate translation and inactivate the protein. The downstream repeat is an 812-bp internal GFP fragment (iGFP). Homologous sequences in the two mutated GFP genes (marked in grey) are separated by 3.7 Kb. Repair of the I-SceI-induced DSB by a non-crossover gene conversion, using iGFP sequence as a donor of wild-type sequence information to the broken gene for repairing, reconstructs a functional GFP gene, expression of which can be scored by cellular fluorescence. There are other possible homologous recombination outcomes, including crossover recombination, long tract gene conversion or SSA pathway. These events retain only the 5’ fragment of the GFP gene that would encode a carboxy-terminal truncation, so there will be no accumulation of fluorescence in the cells.

Otherwise, NHEJ was measured by EJ5-GFP (Bennardo et al., 2008) (Figure M1C) which detects multiple classes of NHEJ events and thus can be considered an assay for total-NHEJ. This reporter contains a promoter that is separated from a GFP coding cassette

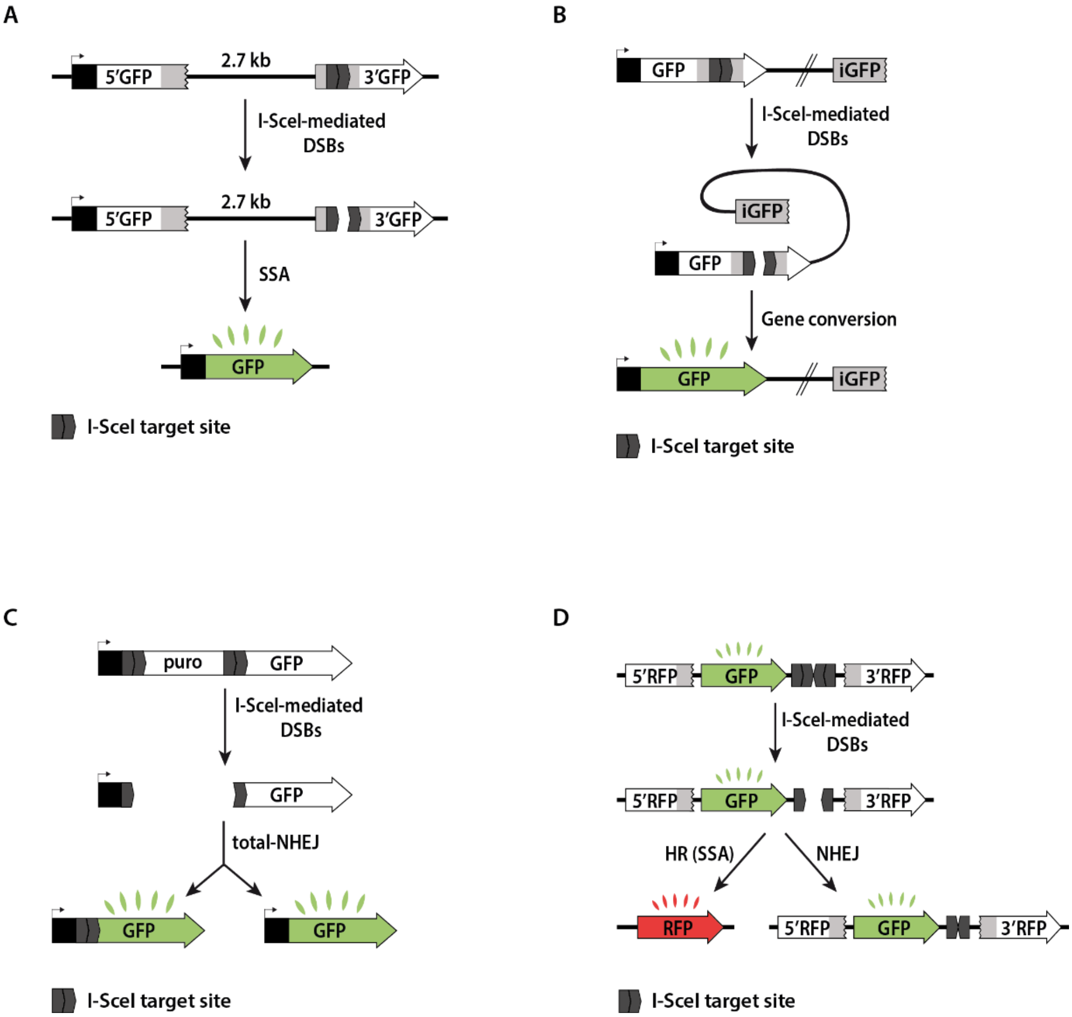

A
5'GFP
2.7 kb
3'GFP
I-SceI-mediated DSBs
SSA
GFP
I-SceI target site
B
GFP
iGFP
I-SceI-mediated DSBs
Gene conversion
I-SceI target site
C
puro
GFP
I-SceI-mediated DSBs
total-NHEJ
I-SceI target site
D
5'RFP
GFP
3'RFP
I-SceI-mediated DSBs
HR (SSA)
NHEJ
RFP
I-SceI target site

Figure M1. DSBs repair reporter systems used.
(A) Schematic representation of the SA-GFP reporter which measures DSB repair by single strand annealing (SSA). **(B)** Schematic representation of the DR-GFP reporter which measures DSB repair by gene conversion. **(C)** Schematic representation of the EJ5-GFP reporter which measures DSB repair by total-NHEJ, both conservative (right) or mutagenic (left). **(D)** Schematic representation of the SeeSaw reporter system (SSR) that measures the balance between NHEJ and HR.

by a puromycin gene that is flanked by two I-SceI sites in the same orientation. Once the two DSBs are produced after I-SceI expression, the puromycin gene is excised by NHEJ repair and the promoter is joined to the rest of the expression cassette, leading to restoration of GFP gene. The repair of the breaks results in reconstitution of the I-SceI site as the two I-SceI-induced DSBs have complementary 3’ overhangs. Alternatively, NHEJ could fail to restore the I-SceI site, generating an I-SceI resistant site that shows evidence of microhomology, suggesting that this repair product is one measure of alternative-NHEJ.

Finally, SeeSaw Reporter (SSR) system (Gómez-Cabello et al., 2013) (Figure M1D) was used to analyse the influence in the repair pathway choice by measuring the balance between NHEJ and HR. This reporter consists of the GFP gene flanked by a 5’- and a 3’-end truncated portions of the RFP gene that share 302 bp of homology with each other (marked in grey). Two sequences recognized by the meganuclease I-SceI were inserted downstream of the GFP gene, close to each other in an inverted orientation. Since the I-SceI target site is not palindromic, the repair of the inverted I-SceI-mediated breaks by NHEJ destroys the target sequence. Once DSBs are induced by the expression of the I-SceI meganuclease, cells could repair the damage through a classical NHEJ-type of repair, leading to GFP gene restoration and hence fluoresce in green. Alternatively, when resection takes place, thereby inhibiting classical NHEJ, the homologous regions of the RFP fragments are exposed and used to repair the damage by SSA. In this case, the repair excises the GFP gene and creates a functional RFP gene, giving rise to cells that fluoresce in red.

3.1.2. DSBs repair assays in vivo

Cells bearing a single copy integration of SA-GFP, DR-GFP, EJ5-GFP or SSR systems (Figure M1A-D) were used to analyse the role of KLF4 in the different DSB repair pathways.

To carry out the assays, 60.000 cells were plated per well in 6-well plates. One day after seeding, the indicated proteins were downregulated by using siRNAs against them. The medium was changed after 6-8 hours. The following day, cells were infected with lentivirus harbouring I-SceI and BFP with a MOI of 10. DMEM containing lentiviral particles was supplemented with 8 µg/ml hexadimethrine bromide to enhance transduction, as indicated in section 1.10. DMEM including hexadimethrine bromide but without lentivirus was added to another set of depleted cells as control of basal fluorescence. After 24 hours, cells were washed with fresh medium and maintained in culture during an additional day. Cells were then harvested and fixed with 4% paraformaldehyde followed by the analysis by flow cytometry of the blue, green and, in the case of SSR, red fluorescence, as described in paragraph 3.2.4.

The repair frequency for SA-GFP, DR-GFP and EJ5-GFP reporters was calculated as the percentage of green cells from 10.000 events positive for blue fluorescence (i.e. transduced with the I-SceI construct). The same cells not infected with pBFP-ISceI plasmid were considered as the background of green fluorescence (Bennardo et al., 2008; Pierce et al., 1999; Stark et al., 2004). Otherwise, the balance between NHEJ and HR with the SSR was calculated by dividing the number of cells expressing GFP by the number of cells expressing RFP from the total 10.000 blue-positive events analysed, again normalizing with the basal fluorescence observed in the samples without I-SceI infection (Gomez-Cabello et al., 2013). To facilitate the comparison between experiments, the percentage or the ratio calculated were normalized with the control siNT.

3.2. Flow cytometry

3.2.1. Cell cycle analysis

Cells were harvested by trypsinization, centrifuged at 400 g for 5 min, and washed once with PBS. Cells were fixed by adding 2 ml of cold 70% ethanol dropwise while vortexing at low speed and incubated at 4ºC overnight. Cells were then centrifuged, washed once with PBS to remove excess of ethanol, and incubated with PBS containing 250 µg/ml RNaseA (Sigma, R6148) and 10 µg/ml propidium iodide (Fluka, 81845) for 30 min at 37ºC in the dark. Finally, samples were analysed in a BD FACSCalibur Flow Cytometer (BD Biosciences, 342975) and CellQuest Pro software. At least 10,000 events were recorded for each sample.

Cell cycle profiles were analysed using ModFit LT 3.0 (Verity Software House, Inc; Topsham, ME, USA).

3.2.2. DNA-end resection analysis

3.5 x 10^5 MEFs-i4F and miPSCs were grown on 60-mm plates in the presence of 10 mM BrdU (Sigma, B5002) for 16–18 hours, harvested using accutase, centrifuged at 400 g for 5 min and washed once with PBS. Cells grown in absence of BrdU were used as FACS-negative control. Cells were fixed with 4% paraformaldehyde for 10 min at 4ºC, centrifuged, washed once with PBS and permeabilized with 0.1% Triton X-100 in PBS. Cells were then centrifuged, washed with PBS, and blocked with 5% FBS in PBS for 45 min at room temperature. After blocking, cells were centrifuged and incubated with an anti-BrdU mouse monoclonal antibody (Table M5) for 1-2 hours at room temperature, and then centrifuged and incubated with the appropriate secondary antibody (Table M6) for 30 min at room temperature. Additional control cells without primary antibody were used to set up FACS conditions. Cells were then washed with PBS, centrifuged and resuspended in PBS. Finally, samples were analysed in a BD FACSCalibur Flow Cytometer (BD Biosciences, 342975) and CellQuest Pro software. At least 10,000 events were recorded for each sample.

3.2.3. Titration of lentiviral production

As mentioned before (section 1.10.), production of lentiviruses harbouring pBFP-ISceI plasmid was titrated by measuring BFP expression using a flow cytometer and shRNA bearing particles by measuring GFP. To do so, 25.000 cells per well were seeded in 12-well plates. The following day, culture medium was removed and 300 µl fresh medium supplemented with 8 µg/ml hexadimethrine bromide was added. Then, viruses were diluted 1/10 in medium, and 0, 5, 10 or 20 µl of this solution was added to each well for infection. After 6 hours of incubation at 37ºC in a P2 biological safety room, the volume was completed with 700 µl fresh DMEM. Next day, cells were washed with DMEM and incubated for another 24 hours upon addition of fresh medium. Finally, cells were harvested by trypsinization and pelleted by centrifugation at 800 g for 5 minutes in FACs tubes. Samples were washed with PBS and resuspended in 200 µl PBS. The percentage of BFP or GFP positive cells was measured by cytometry using BD FACSAria (BD Biosciences) and FACSDiva v5.0.3 software. The highest dilution factor that resulted in 25-50% of the cells positive for BFP or GFP signal was used to calculate viral titer.

3.2.4. Fluorescent protein analysis

DNA double-strand breaks repair systems restore the expression of a fluorescent protein when a specific pathway repairs the damage (see paragraph 3.1.1.). Cells bearing a single copy integration of the reporters SA-GFP, DR-GFP, EJ5-GFP or SSR were downregulated for the indicated genes and infected with lentiviruses harbouring pBFP-ISceI plasmid to generate the break (see paragraph 3.1.2.). Cells were then harvested with trypsin, spun down at 500 g for 5 minutes and washed with PBS. Cells were fixed with 4% paraformaldehyde for 20 minutes at 4ºC in the dark and later rinsed and resuspended in 150 µl of fresh PBS. Samples were analysed by flow cytometry using BD FACSAria and FACSDiva v5.0.3 software. The percentage of green cells, and red cells in the case of SSR system, was calculated from a total of 10.000 events that were BFP positive, i.e. harbouring the I-SceI nuclease.

Furthermore, as mentioned before (section 1.5.), taking advantage of the presence of a GFP gene on the plasmids shNT and shCtip, miPSC self-renewal was analysed by flow cytometry. miPSCs transducted with shNT or shCtip were seeded in 6-well plates and maintained changing the media every day. Cells were passed once a week for three weeks and in each pass, part of the cells were collected in FACs tubes. Cells harvested were centrifuged at 300 g for 5 min and washed with PBS once. Samples were then centrifuged, resuspended in PBS and GFP expression in each cell was then measured by cytometry using BD FACSCalibur Flow Cytometer (BD Biosciences, 342975) and CellQuest Pro software. At least 10,000 events were scored for each sample.

3.2.5. Cell sorting

Cells expressing GFP were harvested using accutase, centrifuged at 300 g for 5 min, and washed once with PBS. Cells were then centrifuged, $5x10^6$ cells were resuspended in 500 µl sorting solution (1X PBS and 5 mM EDTA) and filtered using 70 µm Sterile Cell

Strainers (Corning, 352350). Cells expressing GFP were separate from the total population by cytometry using BD FACSAria (BD Biosciences), and they were collected in FBS (Sigma, F7524). Sorted cells were centrifuged, resuspended in DMEM and cultured.

3.3. Clonogenic survival assay

To study cell survival after damage, clonogenic assays with different DSB-inducing agents were performed as described previously (Puck et al., 1956).

Double-strand breaks were produced by ionizing radiation, PARP inhibition or by treatment with topoisomerase inhibitors camptothecin (Sigma, C9911) and etoposide (VP16; Sigma, E1383). First, U2OS cells were subjected to downregulation or overexpression of indicated proteins and seeded in 6-well plates at different concentrations in triplicates. For IR, 250 and 500 cells were seeded and for PARPi, CPT and VP16 treatments, 500 and 1000 cells were seeded per well. The following day, DNA was damage by different procedures. On the one hand, cells were irradiated with doses of 0, 2 or 4 Gy using an irradiator device that emits gamma rays (BIOBEAM GM 8000, Gamma-Service Medical GmbH). On the other hand, acute treatments (1 hour) with different concentrations of PARP or topoisomerase inhibitors were used, using DMSO as control. Concentrations used for such treatments were 0.01 µM, 0.1 µM and 10 µM µM of PARP inhibitor Olaparib (AstraZeneca, AZD2281), 0.01 µM, 0.05 µM and 0.1 µM of camptothecin and 5 µM and 10 µM of etoposide. Cells were incubated with drugs or DMSO for 1 hour and then were washed twice with PBS and fresh medium was added to each well. Cells were then grown at 37ºC for 8-9 days to allow colonies formation. Once colonies were big enough, they were fixed and visualized by staining with 0.5% crystal violet (Merck, 1.15940.0025) diluted in 20% ethanol, followed by washes with water to remove the excess of staining. The plates were dried and the number of colonies per well was scored for each condition and normalized to the untreated condition for each cell type.

3.4. Microscopy

3.4.1. RPA/γH2AX foci immunofluorescence

U2OS cells downregulated or overexpressed for different proteins were seeded on coverslips (Labolan, 20012). Cells growing on coverslips were irradiated (10 Gy) or mock treated, incubated for 1 hour to allow foci formation, washed once with cold PBS and collected. Coverslips were treated with pre-extraction buffer (25 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 3 mM MgCl2, 300 mM sucrose and 0.2% Triton X-100) for 5 min on ice to remove the nucleoplasmic and cytoplasmic proteins, leaving behind the chromatin-bound and matrix-associated proteins. Cells were then washed once with cold PBS, fixed with 4% paraformaldehyde (w/v) in PBS for 15 minutes on ice, washed three times with PBS and incubated for 1 h with blocking buffer (5% FBS in PBS). Afterwards, cells were co-stained with the appropriate primary antibodies (Table M5) in blocking buffer for 2 h at room temperature, washed three times with PBS and then co-immunostained with the appropriate secondary antibodies (Table M6) in blocking buffer for 1 hour at room temperature in the dark. Coverslips were washed again with PBS, dehydrated into increasing concentrations of ethanol, dried and mounted onto glass slides using Vectashield mounting medium containing 4',6-diamidino-2-phenylindole (DAPI) (Vector Laboratories, H-1200). Samples were visualized using Leica AF6000 fluorescence microscope and a 63x objective. In each experiment, at least 200 cells were analysed.

3.4.2. BRCA1/γH2AX foci immunofluorescence

U2OS cells downregulated or overexpressed for different proteins were grown on coverslips, irradiated (10 Gy) or mock treated, incubated for 1 hour to allow foci formation, washed once with cold PBS and collected. Coverslips were treated with pre-extraction buffer (10 mM PIPES pH 6.8, 50 mM NaCl, 3 mM EDTA, 300 mM sucrose, 0.5% Triton X-100 and 1X protease inhibitor [Roche, 11873580001]) for 5 min on ice and washed once with cold PBS. Cells were subjected to fixation, blocking, staining and mounting as previously described (section 3.4.1.).

3.4.3. RIF1/γH2AX foci immunofluorescence

U2OS cells were downregulated or overexpressed for different proteins. Cells were grown on coverslips, irradiated (10 Gy) or mock-irradiated the next day and incubated for 1 hour at 37ºC for foci formation before being collected. Cells were washed once with cold PBS, fixed with 4% paraformaldehyde (w/v) in PBS for 15 min at room temperature, washed twice with PBS and permeabilized with 0.25% Triton X-100 diluted in PBS for 15 min at room temperature. Coverslips were washed twice with PBS and blocked for at least 1 hour with 5% FBS diluted in PBS. Incubation with primary and secondary antibodies, coverslips mounting and analysis were carried out as for RPA/γH2AX foci immunofluorescence (paragraph 3.4.1.), using the appropriate antibodies (Tables M5 and M6).

3.4.4. Single-molecule analysis of resection tracks (SMART)

SMART assay was performed as previously described (Huertas and Cruz-García, 2018). The different cell lines were seeded at the required density to reach 80% confluence at the time of harvest. Cells were transfected with the indicated siRNAs or plasmids and grown in the presence of 10 µg/ml bromodeoxyuridine (BrdU) for 20-24 hours. Cultures were then irradiated (10 Gy) and incubated for 1 hour at 37ºC, with the exception of miPSC that were not irradiated after treatment with BrdU. Cells were harvested using accutase, centrifuged at 400 g for 5 min, resuspended in PBS and embedded in 1 vol of 1% low melting agarose (Bio-Rad, 161-3111) diluted in PBS and poured into a casting mould to create agarose plugs. For DNA extraction, plugs were incubated in a solution of TE-50 (10 mM Tris-HCl pH 8 and 50 mM EDTA) containing 1% sarkosyl (Sigma, L5125) and 0.2 mg/ml proteinase K (Sigma, P2308) at 50ºC overnight and 6 hours more with a new preparation of the same buffer. Plugs were then washed four times with TE-50 for 10 min with gentle shaking and washed again four times with TE (10 mM Tris-HCl pH8 and 1 mM EDTA) for 5 min. Afterwards, to release DNA fibres from the agarose, each plug was melted in 2.5 ml of 50 mM 2-(N-morpholino) ethanesulfonic acid (MES) pH 5.7 by incubating 20 - 30 min at 65ºC. Then, solution was cooled down to 42ºC before adding 3 U β-agarase I (New England

Biolabs, M0392L) diluted in 100 µl MES and incubated overnight at 42ºC. To stretch DNA fibres, silanized coverslips (Genomic Vision, COV-002-RUO) were dipped into the DNA solution for 15 min at room temperature and pulled out at constant speed (300 µm/sec). Coverslips were baked for at least 2 hours at 65ºC for DNA crosslinking. For immunodetection, coverslips were fixed to slides and blocked with 1% BSA in PBS with 0.1% Triton X-100 (PBS-T) for 15 min and incubated with anti-BrdU mouse monoclonal antibody (Table M5) diluted in PBS-T for 45 min at room temperature. Slides were washed five times with PBS-T for 2 min each and then incubated with the fluorescent secondary antibody (Table M6) diluted in PBS-T for 30 min at room temperature in the dark. Finally, coverslips were washed again with PBS-T, dried, mounted with 20 µl ProLong Gold Antifade Reagent (Molecular Probes, P36930) and stored at -20ºC.

DNA fibres were observed with a Nikon Eclipse NI-E fluorescence microscope with automatized stage and a 40x objective. The images were recorded and processed with NIS ELEMENTS Nikon software. For each experiment, at least 200 DNA fibres were analysed, and their length was measured with Adobe Photoshop CS4 Extended version 11.0 (Adobe Systems Incorporated).

3.5. Array Comparative Genomic Hybridization

For array comparative genomic hybridization, CNV was detected from genomic DNA isolated from iPSC clones and hybridized to SurePrint G3 Mouse CGH Microarray 4x180K (CNV) (Agilent Technologies) following manufacturer’s instructions. CNV was identified using Agilent CytoGenomics v2.0 analysis software, following ADM-2 algorithm suggested by Agilent Technologies. The assay was carried out by the CNIO service.

4. Statistical analysis

Statistical analyses used for each experiment are specified in the pertinent figure legend. All analyses were performed with PRISM software (Graphpad Software Inc.). Statistically significant differences were labelled with one, two, or three asterisks if $p < 0.05$, $p < 0.01$ or $p < 0.001$, respectively. Biologically relevant comparisons that were not statistically significant are stated as ns.

5. Tables of materials

Table M1. Cell lines used in this thesis.

Cell line	Description	Selection	Source/Reference
MEFs-i4F	C57BL/6 primary MEFs carrying a doxycycline-inducible tetracistronic cassette encoding the four murine reprogramming factors Oct4, Sox2, Klf4, and c-Myc	No	Kind gift from M. Serrano (Abad et al., 2013)
miPSC	Mouse induced-pluripotent stem cells	No	This study
ES-D3	Embryonic stem cells	No	Kind gift of B. Soria and K. Hmadcha (CABIMER)
U2OS	Human osteosarcoma	No	ATCC-HTB-96
HEK293T	Human embryonic kidney	No	ATCC CRL-11268
U2OS DR-GFP	U2OS cell line with DR-GFP reporter integrated	Puromycin	Kind gift from Dr. S. P. Jackson
U2OS SA-GFP	U2OS cell line with SA-GFP reporter integrated	Puromycin	Generated in our lab by Ana López Saavedra using hprtSAGFP plasmid (Addgene, 41594)
U2OS EJ5-GFP	U2OS cell line with EJ5-GFP reporter integrated	Puromycin	Generated in our lab by Ana López Saavedra using pimEJ5GFP plasmid (Addgene, 44026)
U2OS SSR	U2OS cell line with SSR reporter integrated	G418	Generated in our lab by Dr. Daniel Gómez Cabello (Gomez-Cabello et al., 2013)

Table M2. siRNAs used in this thesis.

siRNA	Sense sequence (5'-3')	Source
Non-target (NT)	UGGUUUACAUGUCGACUAA	Sigma
CtIP	GCUAAAACAGGAACGAAUC	Sigma
KLF4#1	CCGAGGAGUUCAACGAUCU	Sigma
KLF4#2	GACCUGGACUUUAUUCUCU	Sigma
PRMT5	CCGCUAUUGCACCUUGGAA	Sigma

Table M3. Plasmids used in this thesis.

Plasmid	Description	Selection marker	Source/Reference
pRRL_sEF1a_ HA.NLS.SceOPT.T2A.TagBFP	Vector containing the I-SceI and BFP genes for simultaneous induction of the endonuclease and labelling of the transfected cells	Ampicillin	Addgene (31484) (Certo et al., 2011)
p8.91	Vector used for expression of lentiviral capsid proteins	Ampicillin/ Puromycin	Kind gift from Dr. Felipe Cortés Ledesma
pVSVG	Vector used for expression of lentiviral envelope proteins	Ampicillin/ Puromycin	Kind gift from Dr. Felipe Cortés Ledesma
pLKO.1 eGFP shNT	Vector used for expression of shRNA against a non-target sequence		Sigma (SHC005)
pLKO.1 eGFP shCtip	Vector used for expression of shRNA against Ctip	Ampicillin/ Puromycin	Generated in our lab by Ana López Saavedra from pLKO.1-mshCtip plasmid
pMX	Mammalian Expression, Retroviral	Ampicillin	Kind gift from Dr. S. P. Jackson
pMX-KLF4	pMX vector containing KLF4 gene	Ampicillin	Kind gift from Dr. Francisco Bedoya
pMX-KLF43K	pMX vector containing mutant R374K, R376K and R378K KLF4 gene	Ampicillin	This study

Table M4. Primers used in this thesis.

Primer name	Sequence (5’-3’)	Use
ACTB qPCR Fw	ACGAGGCCCAGAGCAAGA	RT-qPCR of Actin
ACTB qPCR Rv	GACGATGCCGTGCTCGAT	RT-qPCR of Actin
mCtIP qPCR Fw	TTCCTGCTCAAGACACCGATT	RT-qPCR of mouse Ctip
hCtIP qPCR Rv	CGTCTGAGTAGAAGGAAAACCAACT	RT-qPCR of mouse Ctip
NANOG qPCR Fw	AGCAGATGCAAGAACTCTCCTCCA	RT-qPCR of Nanog
NANOG qPCR Rv	CCGCTTGCACTTCATCCTTTGGTT	RT-qPCR of Nanog
KLF4 qPCR Fw	ACCCACACAGGTGAGAAACC	RT-qPCR of KLF4
KLF4 qPCR Rv	ATGTGTAAGGCGAGGTGGTC	RT-qPCR of KLF4
hCtIP qPCR Fw	AGAAATTGGCTTCCTGCTCAAG	RT-qPCR of CtIP
hCtIP qPCR Rv	GAAAACCAACTTCCCAAAAATTCTC	RT-qPCR of CtIP
Mut. KLF4 Fw	GGAGCCCAAGCCAAAGAAGGGAAAAAAATCGTGGCCCCGGAAAAG	Mutation R374K, R376K and R378K of KLF4 with QuickChange Multisite kit
Mut. KLF4 Rv	CTTTTCCGGGGCCACGATTTTTTTCCCTTCTTTGGCTTGGGCTCC	Mutation R374K, R376K and R378K of KLF4 with QuickChange Multisite kit

Table M5. Primary antibodies used in this thesis.

Antibody	Species	Reference/Suppliers	Application (dilution)
α-tubulin	Mouse	Sigma (T9026)	WB (1:500)
CtIP	Mouse	Kind gift from R. Baer	WB (1:500)
BrdU	Mouse	Amersham (RPN202)	SMART, Flow Cyt (1:500)
γH2Ax	Rabbit	Cell Signaling (2577L)	IF (1:500), WB (1:1000)
NANOG	Rabbit	Bethyl Laboratories (A300-397A)	WB (1:500)
CHK1	Mouse	Santa Cruz (sc-8408)	WB (1:500)
Caspase-3	Mouse	Novus (31A1067)	WB (1:50)
KLF4	Rabbit	Santa Cruz (sc-20691)	WB (1:500)
RPA32	Mouse	Abcam (ab2175)	IF (1:500)
BRCA1	Mouse	Santa Cruz (sc-6954)	IF (1:500)
RIF1	Goat	Santa Cruz (sc-55979)	IF (1:100)
TIP60	Mouse	Santa Cruz (sc-166323)	WB (1:500)
PRMT5	Rabbit	Abcam (ab31751)	WB (1:1000)

Table M6. Secondary antibodies used in this thesis.

Antibody	Species	Reference/Suppliers	Application (dilution)
Alexa Fluor 594 anti-mouse	Goat	Invitrogen (A11005)	IF, SMART (1:1000)
Alexa Fluor 488 anti-rabbit	Goat	Invitrogen (A11034)	IF (1:1000)
Alexa Fluor 647 anti-mouse	Goat	Invitrogen (A21235)	Flow Cyt (1:1000)
Alexa Fluor 594 anti-goat	Donkey	Invitrogen (A11058)	IF (1:1000)
Alexa Fluor 488 anti-rabbit	Donkey	Invitrogen (A21206)	IF (1:1000)
IRDye 680RD anti-mouse IgG (H+L)	Goat	LI-COR (926-68070)	WB (1:5000 - 1:15000)
IRDye 800CW anti-rabbit IgG (H+L)	Goat	LI-COR (926-32211)	WB (1:5000 - 1:15000)

VII. Bibliography

Bibliography

Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., Graña, O., Megías, D., Domínguez, O., Martínez, D., et al. (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature *502*, 340–345.

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

Adams, B.R., Golding, S.E., Rao, R.R., and Valerie, K. (2010a). Dynamic dependence on ATR and ATM for double-Strand break repair in human embryonic stem cells and neural descendants. PLoS One *5*, e10001.

Adams, B.R., Hawkins, A.J., Povirk, L.F., and Valerie, K. (2010b). ATM-independent, high-fidelity nonhomologous end joining predominates in human embryonic stem cells. Aging (Albany. NY). *2*, 582–596.

Afghahi, A., Telli, M.L., and Kurian, A.W. (2016). Genetics of triple-negative breast cancer: Implications for patient care. Curr. Probl. Cancer *40*, 130–140.

Aguilera, A., and Gómez-González, B. (2008). Genome instability: A mechanistic view of its causes and consequences. Nat. Rev. Genet. *9*, 204–217.

Ahmed, S.U., Carruthers, R., Gilmour, L., Yildirim, S., Watts, C., and Chalmers, A.J. (2015). Selective inhibition of parallel DNA damage response pathways optimizes radiosensitization of glioblastoma stem-like cells. Cancer Res. *75*, 4416–4428.

Ahuja, A.K., Jodkowska, K., Teloni, F., Bizard, A.H., Zellweger, R., Herrador, R., Ortega, S., Hickson, I.D., Altmeyer, M., Mendez, J., et al. (2016). A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. Nat. Commun. *7*, 10660.

Ai, W., Zheng, H., Yang, X., Liu, Y., and Wang, T.C. (2007). Tip60 functions as a potential corepressor of KLF4 in regulation of HDC promoter activity. Nucleic Acids Res. *35*, 6137–6149.

Aladjem, M.I., Spike, B.T., Rodewald, L.W., Hope, T.J., Klemm, M., Jaenisch, R., and Wahl, G.M. (1998). ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. Curr. Biol. *8*, 145–155.

Alison, M.R., and Islam, S. (2009). Attributes of adult stem cells. J. Pathol. *217*, 144–160.

Anand, R., Ranjha, L., Cannavo, E., and Cejka, P. (2016). Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. Mol. Cell *64*, 940–950.

Andres, S.N., and Williams, R.S. (2017). CtIP/Ctp1/Sae2, molecular form fit for function. DNA Repair (Amst). *56*, 109–117.

Andres, S.N., Appel, C.D., Westmoreland, J.W., Williams, J.S., Nguyen, Y., Robertson, P.D., Resnick, M.A., and Williams, R.S. (2015). Tetrameric Ctp1 coordinates DNA binding and DNA bridging in DNA double-strand-break repair. Nat. Struct. Mol. Biol. *22*, 158–166.

Aymard, F., Bugler, B., Schmidt, C.K., Guillou, E., Caron, P., Briois, S., Iacovoni, J.S., Daburon, V., Miller, K.M., Jackson, S.P., et al. (2014). Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. Nat. Struct. Mol. Biol. *21*, 366–374.

Banito, A., Rashid, S.T., Acosta, J.C., Li, S. De, Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M., et al. (2009). Senescence impairs successful reprogramming to pluripotent stem cells. Genes Dev. *23*, 2134–2139.

Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D., and Rich, J.N. (2006). Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature *444*, 756–760.

Beerman, I., Bock, C., Garrison, B.S., Smith, Z.D., Gu, H., Meissner, A., and Rossi, D.J. (2013). Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. Cell Stem Cell *12*, 413–425.

Beerman, I., Seita, J., Inlay, M.A., Weissman, I.L., and Rossi, D.J. (2014). Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. Cell Stem Cell *15*, 37–50.

Behr, R., and Kaestner, K.H. (2002). Developmental and cell type-specific expression of the zinc finger transcription factor Krüppel-like factor 4 (Klf4) in postnatal mouse testis. Mech. Dev. *115*, 167–169.

Behrens, A., Van Deursen, J.M., Rudolph, K.L., and Schumacher, B. (2014). Impact of genomic damage and ageing on stem cell function. Nat. Cell Biol. *16*, 201–207.

Bekker-Jensen, S., Lukas, C., Kitagawa, R., Melander, F., Kastan, M.B., Bartek, J., and Lukas, J. (2006). Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. J Cell Biol *173*, 195–206.

Ben-David, U., and Benvenisty, N. (2012). High prevalence of evolutionarily conserved and species-specific genomic aberrations in mouse pluripotent stem cells. Stem Cells *30*, 612–622.

Ben-David, U., Mayshar, Y., and Benvenisty, N. (2011). Large-scale analysis reveals acquisition of lineage-specific chromosomal aberrations in human adult stem cells. Cell Stem Cell *9*, 97–102.

Bennardo, N., Cheng, A., Huang, N., and Stark, J.M. (2008). Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. PLoS Genet. *4*, e1000110.

Blanpain, C., Mohrin, M., Sotiropoulou, P.A., and Passegué, E. (2011). DNA-damage response in tissue-specific and cancer stem cells. Cell Stem Cell *8*, 16–29.

Blasco, M.A., Serrano, M., and Fernandez-Capetillo, O. (2011). Genomic instability in iPS: Time for a break. EMBO J. *30*, 991–993.

Bogomazova, A.N., Lagarkova, M.A., Tskhovrebova, L. V., Shutova, M. V., and Kiselev, S.L. (2011). Error-prone nonhomologous end joining repair operates in human pluripotent stem cells during late G2. Aging (Albany. NY). *3*, 584–596.

Branzei, D., and Foiani, M. (2008). Regulation of DNA repair throughout the cell cycle. Nat. Rev. Mol. Cell Biol. *9*, 297–308.

Caldecott, K.W. (2008). Single-strand break repair and genetic disease. Nat. Rev. Genet. *9*, 619–631.

Cannan, W.J., and Pederson, D.S. (2016). Mechanisms and Consequences of Double-strand DNA Break Formation in Chromatin. J. Cell. Physiol. *231*, 3.

Ceccaldi, R., Rondinelli, B., and D'Andrea, A.D. (2016). Repair Pathway Choices and Consequences at the Double-Strand Break. Trends Cell Biol. *26*, 52–64.

Certo, M.T., Ryu, B.Y., Annis, J.E., Garibov, M., Jarjour, J., Rawlings, D.J., and Scharenberg, A.M. (2011). Tracking genome engineering outcome at individual DNA breakpoints. Nat. Methods *8*, 671–676.

Cervantes, R.B., Stringer, J.R., Tischfield, J.A., Stambrook, P.J., and Shao, C. (2002). Embryonic stem cells and somatic cells differ in mutation frequency and type. Proc. Natl. Acad. Sci. *99*, 3586–3590.

Chambers, S.M., Shaw, C.A., Gatza, C., Fisk, C.J., Donehower, L.A., and Goodell, M.A. (2007). Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. PLoS Biol. *5*, 1750–1762.

Chang, C.H., Zhang, M., Rajapakshe, K., Coarfa, C., Edwards, D., Huang, S., and Rosen, J.M. (2015). Mammary Stem Cells and Tumor-Initiating Cells Are More Resistant to Apoptosis and Exhibit Increased DNA Repair Activity in Response to DNA Damage. Stem Cell Reports *5*, 378–391.

Chapman, J.R., Taylor, M.R.G., and Boulton, S.J. (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. Mol. Cell *47*, 497–510.

Chapman, J.R., Barral, P., Vannier, J.B., Borel, V., Steger, M., Tomas-Loba, A., Sartori, A.A., Adams, I.R., Batista, F.D., and Boulton, S.J. (2013). RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection. Mol. Cell *49*, 858–871.

Chen, L., Nievera, C.J., Lee, A.Y.-L., and Wu, X. (2008). Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. J. Biol. Chem. *283*, 7713–7720.

Chen, P.-L., Liu, F., Cai, S., Lin, X., Li, A., Chen, Y., Gu, B., Lee, E.Y.-H.P., and Lee, W.-H. (2005a). Inactivation of CtIP Leads to Early Embryonic Lethality Mediated by G1 Restraint and to Tumorigenesis by Haploid Insufficiency. Mol. Cell. Biol. *25*, 3535–3542.

Chen, X., Johns, D.C., Geiman, D.E., Marban, E., Dang, D.T., Hamlin, G., Sun, R., and Yang, V.W. (2001). Krüppel-like Factor 4 (Gut-enriched Krüppel-like Factor) Inhibits Cell Proliferation by Blocking G1/S Progression of the Cell Cycle. J. Biol. Chem. *276*, 30423–30428.

Chen, X., Whitney, E.M., Gao, S.Y., and Yang, V.W. (2003). Transcriptional profiling of Krüppel-like factor 4 reveals a function in cell cycle regulation and epithelial differentiation. J. Mol. Biol. *326*, 665–677.

Chen, Z.Y., Wang, X., Zhou, Y., Offner, G., and Tseng, C.C. (2005b). Destabilization of Krüppel-like factor 4 protein in response to serum stimulation involves the ubiquitin-proteasome pathway. Cancer Res. *65*, 10394–10400.

Cheung-Ong, K., Giaever, G., and Nislow, C. (2013). DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. Chem. Biol. *20*, 648–659.

Chiambaretta, F., De Graeve, F., Turet, G., Marceau, G., Gain, P., Dastugue, B., Rigal, D., and Sapin, V. (2004). Cell and tissue specific expression of human Krüppel-like transcription factors in human ocular surface. Mol. Vis. *10*, 901–909.

Chiang, K., Zielinska, A.E., Shaaban, A.M., Sanchez-Bailon, M.P., Jarrold, J., Clarke, T.L., Zhang, J., Francis, A., Jones, L.J., Smith, S., et al. (2017). PRMT5 Is a Critical Regulator of Breast Cancer Stem Cell Function via Histone Methylation and FOXP1 Expression. Cell Rep. *21*, 3498–3513.

Chiruvella, K.K., Liang, Z., and Wilson, T.E. (2013). Repair of double-strand breaks by end joining. Cold Spring Harb. Perspect. Biol. *5*, a012757.

Cho, Y.G., Song, J.H., Kim, C.J., Nam, S.W., Yoo, N.J., Lee, J.Y., and Park, W.S. (2007). Genetic and epigenetic analysis of the KLF4 gene in gastric cancer. APMIS *115*, 802–808.

Choi, H., Roh, J., Choi, H., and Roh, J. (2018). Role of Klf4 in the regulation of apoptosis and cell cycle in rat granulosa cells during the periovulatory period. Int. J. Mol. Sci. *20*, 87.

Ciccia, A., and Elledge, S.J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. Mol. Cell *40*, 179–204.

Clarke, T.L., Sanchez-Bailon, M.P., Chiang, K., Reynolds, J.J., Herrero-Ruiz, J., Bandeiras, T.M., Matias, P.M., Maslen, S.L., Skehel, J.M., Stewart, G.S., et al. (2017). PRMT5-Dependent Methylation of the TIP60 Coactivator RUVBL1 Is a Key Regulator of Homologous Recombination. Mol. Cell *65*, 900-916.e7.

Cottarel, J., Frit, P., Bombarde, O., Salles, B., Négrel, A., Bernard, S., Jeggo, P.A., Lieber, M.R., Modesti, M., and Calsou, P. (2013). A noncatalytic function of the ligation complex during nonhomologous end joining. J. Cell Biol. *200*, 173–186.

Cruz-García, A., López-Saavedra, A., and Huertas, P. (2014). BRCA1 accelerates CtIP-mediated DNA-end resection. Cell Rep. *9*, 451–459.

Cullingford, T.E., Butler, M.J., Marshall, A.K., Tham, E.L., Sugden, P.H., and Clerk, A. (2008). Differential regulation of Krüppel-like factor family transcription factor expression in neonatal rat cardiac myocytes: Effects of endothelin-1, oxidative stress and cytokines. Biochim. Biophys. Acta - Mol. Cell Res. *1783*, 1229–1236.

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.

Dang, D.T., Chen, X., Feng, J., Torbenson, M., Dang, L.H., and Yang, V.W. (2003). Overexpression of Krüppel-like factor 4 in the human colon cancer cell line RKO leads to reduced tumorigenecity. Oncogene *22*, 3424–3434.

De Bont, R., and van Larebeke, N. (2004). Endogenous DNA damage in humans: A review of quantitative data. Mutagenesis *19*, 169–185.

de Klein, A., Muijtjens, M., van Os, R., Verhoeven, Y., Smit, B., Carr, A.M., Lehmann, A.R., and Hoeijmakers, J.H. (2000). Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. Curr. Biol. *10*, 479–482.

Desai, A., Webb, B., and Gerson, S.L. (2014). CD133+ cells contribute to radioresistance via altered regulation of DNA repair genes in human lung cancer cells. Radiother. Oncol. *110*, 538–545.

Dev, H., Chiang, T.-W.W., Lescale, C., de Krijger, I., Martin, A.G., Pilger, D., Coates, J., Sczaniecka-Clift, M., Wei, W., Ostermaier, M., et al. (2018). Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. Nat. Cell Biol. *20*, 954–965.

Diehn, M., Cho, R.W., Lobo, N.A., Kalisky, T., Dorie, M.J., Kulp, A.N., Qian, D., Lam, J.S., Ailles, L.E., Wong, M., et al. (2009). Association of reactive oxygen species levels and radioresistance in cancer stem cells. Nature *458*, 780–783.

Du, J.X., McConnell, B.B., and Yang, V.W. (2010). A small ubiquitin-related modifier-interacting motif functions as the transcriptional activation domain of Krüppel-like factor 4. J. Biol. Chem. *285*, 28298–28308.

Dubin, M.J., Stokes, P.H., Sum, E.Y.M., Williams, R.S., Valova, V.A., Robinson, P.J., Lindeman, G.J., Glover, J.N.M., Visvader, J.E., and Matthews, J.M. (2004). Dimerization of CtIP, a BRCA1- and CtBP-interacting protein, is mediated by an N-terminal coiled-coil motif. J. Biol. Chem. *279*, 26932–26938.

Duquette, M.L., Zhu, Q., Taylor, E.R., Tsay, A.J., Shi, L.Z., Berns, M.W., and McGowan, C.H. (2012). CtIP Is Required to Initiate Replication-Dependent Interstrand Crosslink Repair. PLoS Genet. *8*, e1003050.

Dynan, W.S., and Yoo, S. (1998). Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. Nucleic Acids Res. *26*, 1551–1559.

Edel, M.J., Menchon, C., Menendez, S., Consiglio, A., Raya, A., and Izpisua Belmonte, J.C. (2010). Rem2 GTPase maintains survival of human embryonic stem cells as well as enhancing reprogramming by regulating p53 and cyclin D1. Genes Dev. *24*, 561–573.

El-Karim, E.A., Hagos, E.G., Ghaleb, A.M., Yu, B., and Yang, V.W. (2013). Krüppel-like factor 4 regulates genetic stability in mouse embryonic fibroblasts. Mol. Cancer *12*, 89.

Escribano-Díaz, C., Orthwein, A., Fradet-Turcotte, A., Xing, M., Young, J.T.F., Tkáč, J., Cook, M.A., Rosebrock, A.P., Munro, M., Canny, M.D., et al. (2013). A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice. Mol. Cell *49*, 872–883.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells. Cell Stem Cell *6*, 71–79.

Eun, K., Ham, S.W., and Kim, H. (2017). Cancer stem cell heterogeneity: origin and new perspectives on CSC targeting. BMB Rep. *50*, 117–125.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Evans, P.M., Zhang, W., Chen, X., Yang, J., Bhakat, K.K., and Liu, C. (2007). Krüppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation. J. Biol. Chem. *282*, 33994–34002.

Ezashi, T., Telugu, B.P.V.L., Alexenko, A.P., Sachdev, S., Sinha, S., and Roberts, R.M. (2009). Derivation of induced pluripotent stem cells from pig somatic cells. Proc. Natl. Acad. Sci. *106*, 10993–10998.

Feinberg, M.W., Cao, Z., Wara, A.K., Lebedeva, M.A., SenBanerjee, S., and Jain, M.K. (2005). Kruppel-like factor 4 is a mediator of proinflammatory signaling in macrophages. J. Biol. Chem. *280*, 38247–38258.

Felgentreff, K., Du, L., Weinacht, K.G., Dobbs, K., Bartish, M., Giliani, S., Schlaeger, T., DeVine, A., Schambach, A., Woodbine, L.J., et al. (2014). Differential role of nonhomologous end joining factors in the generation, DNA damage response, and myeloid differentiation of human induced pluripotent stem cells. Proc. Natl. Acad. Sci. *111*, 8889–8894.

Ferralli, J., Chiquet-Ehrismann, R., Degen, M., Ferralli, J., Chiquet-Ehrismann, R., and Degen, M. (2014). KLF4α stimulates breast cancer cell proliferation by acting as a KLF4 antagonist. Oncotarget *7*, 45608–45621.

Filion, T.M., Qiao, M., Ghule, P.N., Mandeville, M., Van Wijnen, A.J., Stein, J.L., Lian, J.B., Altieri, D.C., and Stein, G.S. (2009). Survival responses of human embryonic stem cells to DNA damage. J. Cell. Physiol. *220*, 586–592.

Fletcher, O., Johnson, N., Orr, N., Hosking, F.J., Gibson, L.J., Walker, K., Zelenika, D., Gut, I., Heath, S., Palles, C., et al. (2011). Novel breast cancer susceptibility locus at 9q31.2: Results: of a genome-wide association study. J. Natl. Cancer Inst. *103*, 425–435.

Folmes, C.D.L., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., Perez-Terzic, C., and Terzic, A. (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metab. *14*, 264–271.

Foray, N., Marot, D., Gabriel, A., Randrianarison, V., Carr, A.M., Perricaudet, M., Ashworth, A., and Jeggo, P. (2003). A subset of ATM- and ATR-dependent phosphorylation events requires the BRCA1 protein. EMBO J. *22*, 2860–2871.

Forment, J. V., Jackson, S.P., and Pellegrini, L. (2015). When two is not enough: A CtIP tetramer is required for DNA repair by homologous recombination. Nucleus *6*, 344–348.

Foster, K.W., Frost, A.R., McKie-Bell, P., Lin, C.Y., Engler, J.A., Grizzle, W.E., and Ruppert, J.M. (2000). Increase of GKLF messenger RNA and protein expression during progression of breast cancer. Cancer Res. *60*, 6488.

Foster, K.W., Liu, Z., Nail, C.D., Li, X., Fitzgerald, T.J., Bailey, S.K., Frost, A.R., Louro, I.D., Townes, T.M., Paterson, A.J., et al. (2005). Induction of KLF4 in basal keratinocytes blocks the proliferation–differentiation switch and initiates squamous epithelial dysplasia. Oncogene *24*, 1491–1500.

Fruman, D.A., Ferl, G.Z., An, S.S., Donahue, A.C., Satterthwaite, A.B., and Witte, O.N. (2002). Phosphoinositide 3-kinase and Bruton’s tyrosine kinase regulate overlapping sets of genes in B lymphocytes. Proc. Natl. Acad. Sci. *99*, 359–364.

Fusco, C., Reymond, A., and Zervos, A.S. (1998). Molecular Cloning and Characterization of a Novel Retinoblastoma-Binding Protein. Genomics *51*, 351–358.

Garcia, V., Phelps, S.E.L., Gray, S., and Neale, M.J. (2011). Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. Nature *479*, 241–244.

Garrett-Sinha, L.A., Eberspaecher, H., Seldin, M.F., and De Crombrugghef, B. (1996). A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. J. Biol. Chem. *271*, 31384–31390.

Geiman, D.E. (2002). Transactivation and growth suppression by the gut-enriched Kruppel-like factor (Kruppel-like factor 4) are dependent on acidic amino acid residues and protein-protein interaction. Nucleic Acids Res. *28*, 1106–1113.

Genois, M.-M., Paquet, E.R., Laffitte, M.-C.N., Maity, R., Rodrigue, A., Ouellette, M., and Masson, J.-Y. (2014). DNA Repair Pathways in Trypanosomatids: from DNA Repair to Drug Resistance. Microbiol. Mol. Biol. Rev. *78*, 40–73.

Ghaleb, A.M., and Yang, V.W. (2017). Krüppel-like factor 4 (KLF4): What we currently know. Gene *611*, 27–37.

Ghaleb, A.M., Katz, J.P., Kaestner, K.H., Du, J.X., and Yang, V.W. (2007). Krüppel-like factor 4 exhibits antiapoptotic activity following γ-radiation-induced DNA damage. Oncogene *26*, 2365–2373.

Ghule, P.N., Medina, R., Lengner, C.J., Mandeville, M., Qiao, M., Dominski, Z., Lian, J.B., Stein, J.L., Van Wijnen, A.J., and Stein, G.S. (2011). Reprogramming the pluripotent cell cycle: Restoration of an abbreviated G1 phase in human induced pluripotent stem (iPS) cells. J. Cell. Physiol. *226*, 1149–1156.

Godmann, M., Kosan, C., and Behr, R. (2010). Krüppel-like factor 4 is widely expressed in the mouse male and female reproductive tract and responds as an immediate early gene to activation of the protein kinase A in TM4 Sertoli cells. Reproduction *139*, 771–782.

Gomez-Cabello, D., Jimeno, S., Fernández-Ávila, M.J., and Huertas, P. (2013). New Tools to Study DNA Double-Strand Break Repair Pathway Choice. PLoS One *8*, 1–9.

González, F., Georgieva, D., Vanoli, F., Shi, Z., Ludwig, T., Jasin, M., and Huangfu, D. (2013). Homologous Recombination DNA Repair Genes Play a Critical Role in Reprogramming to a Pluripotent State. Cell Rep. *3*, 651–660.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63–67.

Grandela, C., Pera, M.F., and Wolvetang, E.J. (2007). p53 is required for etoposide-induced apoptosis of human embryonic stem cells. Stem Cell Res. *1*, 116–128.

Gravel, S., Chapman, J.R., Magill, C., and Jackson, S.P. (2008). DNA helicases Sgs1 and BLM promote DNA double-strand break resection. Genes Dev. *22*, 2767–2772.

Grawunder, U., Wilm, M., Wu, X., Kulesza, P., Wilson, T.E., Mann, M., and Lieber, M.R. (1997). Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature *338*, 492–495.

Gu, B., and Chen, P.L. (2009). Expression of PCNA-binding domain of CtIP, a motif required for CtIP localization at DNA replication foci, causes DNA damage and activation of DNA damage checkpoint. Cell Cycle *8*, 1409–1420.

Gunasekharan, V.K., Li, Y., Andrade, J., and Laimins, L.A. (2016). Post-Transcriptional Regulation of KLF4 by High-Risk Human Papillomaviruses Is Necessary for the Differentiation-Dependent Viral Life Cycle. PLoS Pathog. *12*, e1005747.

Gupta, R., Somyajit, K., Narita, T., Maskey, E., Stanlie, A., Kremer, M., Typas, D., Lammers, M., Mailand, N., Nussenzweig, A., et al. (2018). DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. Cell *173*, 972-988.e23.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. *10*, 622–640.

Hagos, E.G., Ghaleb, A.M., Dalton, W.B., Bialkowska, A.B., and Yang, V.W. (2009). Mouse embryonic fibroblasts null for the Krüppel-like factor 4 gene are genetically unstable. Oncogene *28*, 1197–1205.

Hamard, P.-J., Santiago, G.E., Liu, F., Karl, D.L., Martinez, C., Man, N., Mookhtiar, A.K., Duffort, S., Greenblatt, S., Verdun, R.E., et al. (2018). PRMT5 regulates DNA repair by controlling the alternative splicing of key histone-modifying enzymes. Cell Rep. *24*, 2643–2657.

Hartlerode, A.J., and Scully, R. (2009). Mechanisms of double-strand break repair in somatic mammalian cells. Biochem. J. *423*, 157–168.

Hindley, C., and Philpott, A. (2013). The cell cycle and pluripotency. Biochem. J. *451*, 135–143.

Hoeijmakers, J.H.J. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366–374.

Hoeijmakers, J.H.J. (2009). DNA damage, aging, and cancer. N. Engl. J. Med. *361*, 1475–1485.

Hong, Y., and Stambrook, P.J. (2004). Restoration of an absent G1 arrest and protection from apoptosis in embryonic stem cells after ionizing radiation. Proc. Natl. Acad. Sci. *101*, 14443–14448.

Hu, D., Gur, M., Zhou, Z., Gamper, A., Hung, M.C., Fujita, N., Lan, L., Bahar, I., and Wan, Y. (2015). Interplay between arginine methylation and ubiquitylation regulates KLF4-mediated genome stability and carcinogenesis. Nat. Commun. *6*, 8419.

Hu, W., Hofstetter, W.L., Li, H., Zhou, Y., He, Y., Pataer, A., Wang, L., Xie, K., Swisher, S.G., and Fang, B. (2009). Putative tumor-suppressive function of Krüppel-like factor 4 in primary lung carcinoma. Clin. Cancer Res. *15*, 5688–5695.

Huang, C.C., Liu, Z., Li, X., Bailey, S., Nail, C.D., Foster, K.W., Frost, A.R., Ruppert, J.M., and Lobo-Ruppert, S.M. (2005). KLF4 and PCNA identify stages of tumor initiation in a conditional model of cutaneous squamous epithelial neoplasia. Cancer Biol. Ther. *4*, 1401–1408.

Huertas, P. (2010). DNA resection in eukaryotes: Deciding how to fix the break. Nat. Struct. Mol. Biol. *17*, 11–16.

Huertas, P., and Cruz-García, A. (2018). Single Molecule Analysis of Resection Tracks. Genome Instab. Methods Protoc. *1672*, 147–154.

Huertas, P., and Jackson, S.P. (2009). Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. J. Biol. Chem. *284*, 9558–9565.

Huertas, P., Cortés-Ledesma, F., Sartori, A.A., Aguilera, A., and Jackson, S.P. (2008). CDK targets Sae2 to control DNA-end resection and homologous recombination. Nature *455*, 689–692.

Hui, Hongxiang; Tang, Yongming; Hu, Min; Zhao, X. (2011). Stem Cells: General Features and Characteristics. In Stem Cells in Clinic and Research, pp. 3–20.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., and Benvenisty, N. (2000). Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. Mol. Med. *6*, 88–95.

Jackson, S.P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis *23*, 687–696.

Jackson, A.L., and Loeb, L.A. (2001). The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *477*, 7–21.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Jacquet, K., Fradet-Turcotte, A., Avvakumov, N., Lambert, J.P., Roques, C., Pandita, R.K., Paquet, E., Herst, P., Gingras, A.C., Pandita, T.K., et al. (2016). The TIP60 Complex Regulates Bivalent Chromatin Recognition by 53BP1 through Direct H4K20me Binding and H2AK15 Acetylation. Mol. Cell *62*, 409–421.

Jean, J.C., George, E., Kaestner, K.H., Brown, L.A.S., Spira, A., and Joyce-Brady, M. (2013). Transcription Factor Klf4, Induced in the Lung by Oxygen at Birth, Regulates Perinatal Fibroblast and Myofibroblast Differentiation. PLoS One *8*, e54806.

Jeon, H.M., Sohn, Y.W., Oh, S.Y., Kim, S.H., Beck, S., Kim, S., and Kim, H. (2011). ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9*-mediated suppression of SOX2. Cancer Res. *71*, 3410–3421.

Jerónimo, C., Bastian, P.J., Bjartell, A., Carbone, G.M., Catto, J.W.F., Clark, S.J., Henrique, R., Nelson, W.G., and Shariat, S.F. (2011). Epigenetics in prostate cancer: Biologic and clinical relevance. Eur. Urol. *60*, 753–766.

Jossen, R., and Bermejo, R. (2013). The DNA damage checkpoint response to replication stress: A Game of Forks. Front. Genet. *4*, 26.

Kashiwagi, H., Shiraishi, K., Sakaguchi, K., Nakahama, T., and Kodama, S. (2018). Repair kinetics of DNA double-strand breaks and incidence of apoptosis in mouse neural stem/progenitor cells and their differentiated neurons exposed to ionizing radiation. J. Radiat. Res. *59*, 261–271.

Kastan, M.B., and Bartek, J. (2004). Cell-cycle checkpoints and cancer. Nature *432*, 316–323.

Kawai-Kowase, K., Ohshima, T., Matsui, H., Tanaka, T., Shimizu, T., Iso, T., Arai, M., Owens, G.K., and Kurabayashi, M. (2009). PIAS1 mediates TGFβ-induced SM α-actin gene expression through inhibition of KLF4 function-expression by protein sumoylation. Arterioscler. Thromb. Vasc. Biol. *29*, 99–106.

Kee, Y., and D'Andrea, A.D. (2010). Expanded roles of the Fanconi anemia pathway in preserving genomic stability. Genes Dev. *24*, 1680–1694.

Kim, M.O., Kim, S.H., Cho, Y.Y., Nadas, J., Jeong, C.H., Yao, K., Kim, D.J., Yu, D.H., Keum, Y.S., Lee, K.Y., et al. (2012). ERK1 and ERK2 regulate embryonic stem cell self-renewal through phosphorylation of Klf4. Nat. Struct. Mol. Biol. *19*, 283–290.

Kinoshita, T., Nagamatsu, G., Kosaka, T., Takubo, K., Hotta, A., Ellis, J., and Suda, T. (2011). Ataxia-telangiectasia mutated (ATM) deficiency decreases reprogramming efficiency and leads to genomic instability in iPS cells. Biochem. Biophys. Res. Commun. *407*, 321–326.

Kreso, A., and Dick, J.E. (2014). Evolution of the cancer stem cell model. Cell Stem Cell *14*, 275–291.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature *367*, 645–648.

Lee, Y., and McKinnon, P.J. (2007). Responding to DNA double strand breaks in the nervous system. Neuroscience *145*, 1365–1374.

Lemaître, C., Grabarz, A., Tsouroula, K., Andronov, L., Furst, A., Pankotai, T., Heyer, V., Rogier, M., Attwood, K.M., Kessler, P., et al. (2014). Nuclear position dictates DNA repair pathway choice. Genes Dev. *28*, 2450–2463.

Li, F., He, Z., Shen, J., Huang, Q., Li, W., Liu, X., He, Y., Wolf, F., and Li, C.-Y. (2010a). Apoptotic Caspases Regulate Induction of iPSCs from Human Fibroblasts. Cell Stem Cell *7*, 508–520.

Li, Q., Seo, J.H., Stranger, B., McKenna, A., Pe'Er, I., Laframboise, T., Brown, M., Tyekucheva, S., and Freedman, M.L. (2013). Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. Cell *152*, 633–641.

Li, S., Chen, P.L., Subramanian, T., Chinnadurai, G., Tomlinson, G., Osborne, C.K., Sharp, Z.D., and Lee, W.H. (1999). Binding of CtIP to the BRCT repeats of BRCA1 involved in the transcription regulation of p21 is disrupted upon DNA damage. J. Biol. Chem. *274*, 11334–11338.

Li, S., Ting, N.S.Y., Zheng, L., Chen, P.L., Ziv, Y., Shiloh, Y., Lee, E.Y.H.P., and Lee, W.H. (2000). Functional link of BRCA1 and ataxia telangiectasia gone product in DNA damage response. Nature *406*, 210–215.

Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H., and Ding, S. (2009).

Generation of Rat and Human Induced Pluripotent Stem Cells by Combining Genetic Reprogramming and Chemical Inhibitors. Cell Stem Cell *4*, 16–19.

Li, Y., McClintick, J., Zhong, L., Edenberg, H.J., Yoder, M.C., and Chan, R.J. (2005). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. Blood *105*, 635–637.

Li, Z., Zhao, J., Li, Q., Yang, W., Song, Q., Li, W., and Liu, J. (2010b). KLF4 promotes hydrogen-peroxide-induced apoptosis of chronic myeloid leukemia cells involving the bcl-2/bax pathway. Cell Stress Chaperones *15*, 905–912.

Liang, G., and Zhang, Y. (2013). Genetic and epigenetic variations in iPSCs: Potential causes and implications for application. Cell Stem Cell *13*, 149–159.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., et al. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat. Cell Biol. *10*, 731–739.

Liao, S., Tammaro, M., and Yan, H. (2016). The structure of ends determines the pathway choice and Mre11 nuclease dependency of DNA double-strand break repair. Nucleic Acids Res. *44*, 5689–5701.

Lim, K.H., Kim, S.R., Ramakrishna, S., and Baek, K.H. (2014). Critical lysine residues of Klf4 required for protein stabilization and degradation. Biochem. Biophys. Res. Commun. *443*, 1206–1210.

Lim, Y.C., Roberts, T.L., Day, B.W., Harding, A., Kozlov, S., Kijas, A.W., Ensbey, K.S., Walker, D.G., and Lavin, M.F. (2012). A Role for Homologous Recombination and Abnormal Cell-Cycle Progression in Radioresistance of Glioma-Initiating Cells. Mol. Cancer Ther. *11*, 1863–1872.

Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. Nat. Cell Biol. *7*, 165–171.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709–715.

Liu, F., and Lee, W.-H. (2006). CtIP Activates Its Own and Cyclin D1 Promoters via the E2F/RB Pathway during G1/S Progression. Mol. Cell. Biol. *26*, 3124–3134.

Liu, H., Zhu, F., Yong, J., Zhang, P., Hou, P., Li, H., Jiang, W., Cai, J., Liu, M., Cui, K., et al. (2008). Generation of Induced Pluripotent Stem Cells from Adult Rhesus Monkey Fibroblasts. Cell

Stem Cell *3*, 587–590.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT Method. Methods *25*, 402–408.

López-Saavedra, A., Gómez-Cabello, D., Domínguez-Sánchez, M.S., Mejías-Navarro, F., Fernández-Ávila, M.J., Dinant, C., Martínez-Macías, M.I., Bartek, J., and Huertas, P. (2016). A genome-wide screening uncovers the role of CCAR2 as an antagonist of DNA end resection. Nat. Commun. *7*, 12364.

Lukas, J., and Bartek, J. (2004). Watching the DNA repair ensemble dance. Cell *118*, 666–668.

Lukas, J., Lukas, C., and Bartek, J. (2004). Mammalian cell cycle checkpoints: Signalling pathways and their organization in space and time. DNA Repair (Amst). *3*, 997–1007.

Mahmoudi, S., and Brunet, A. (2012). Aging and reprogramming: a two-way street. Curr. Opin. Cell Biol. *24*, 744–756.

Makharashvili, N., and Paull, T.T. (2015). CtIP: A DNA damage response protein at the intersection of DNA metabolism. DNA Repair (Amst). *32*, 75–81.

Makharashvili, N., Tubbs, A.T., Yang, S.H., Wang, H., Barton, O., Zhou, Y., Deshpande, R.A., Lee, J.H., Lobrich, M., Sleckman, B.P., et al. (2014). Catalytic and Noncatalytic Roles of the CtIP Endonuclease in Double-Strand Break End Resection. Mol. Cell *54*, 1022–1033.

Malik, N., and Rao, M.S. (2013). A review of the methods for human iPSC derivation. In Methods in Molecular Biology, pp. 23–33.

Maréchal, A., and Zou, L. (2013). DNA Damage Sensing by the ATMand ATR Kinases. Cold Spring Harb. Perspect Biol *5*, a012716.

Marión, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149–1153.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. *18*, 4261–4269.

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.

Maynard, S., SWISTOWSKA, A.M., LEE, J.W., LIU, Y., LIU, S.-T., A, ALEXANDRE BETTENCOURT DA CRUZ, MAHENDRA RAO, NADJA C. DE SOUZA-PINTO, XIANMIN ZENG, B., and BOHR, V.A. (2008). Human Embryonic Stem Cells Have Enhanced Repair of Multiple Forms of DNA Damage. Stem Cells *26*, 2266–2274.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell Stem Cell *7*, 521–531.

McKinnon, P.J. (2009). DNA repair deficiency and neurological disease. Nat. Rev. Neurosci. *10*, 100–112.

McVey, M., and Lee, S.E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet. *24*, 529–538.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature *455*, 770–774.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Moehrle, B.M., Nattamai, K., Brown, A., Florian, M.C., Ryan, M., Vogel, M., Bliederhaeuser, C., Soller, K., Prows, D.R., Abdollahi, A., et al. (2015). Stem Cell-Specific Mechanisms Ensure Genomic Fidelity within HSCs and upon Aging of HSCs. Cell Rep. *13*, 2412–2424.

Mohrin, M., Bourke, E., Alexander, D., Warr, M.R., Barry-Holson, K., Le Beau, M.M., Morrison, C.G., and Passegué, E. (2010). Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. Cell Stem Cell *7*, 174–185.

Momcilovic, O., Knobloch, L., Fornsaglio, J., Varum, S., Easley, C., and Schatten, G. (2010). DNA damage responses in human induced pluripotent stem cells and embryonic stem cells. PLoS One *5*, e13410.

Momčilović, O., Choi, S., Varum, S., Bakkenist, C., Schatten, G., and Navara, C. (2009). Ionizing Radiation Induces Ataxia Telangiectasia Mutated-Dependent Checkpoint Signaling and G $_2$

But Not G $_1$ Cell Cycle Arrest in Pluripotent Human Embryonic Stem Cells. Stem Cells *27*, 1822–1835.

Morales-Martinez, M., Valencia-Hipolito, A., Vega, G.G., Neri, N., Nambo, M.J., Alvarado, I., Cuadra, I., Duran-Padilla, M.A., Martinez-Maza, O., Huerta-Yepez, S., et al. (2019). Regulation of Krüppel-Like Factor 4 (KLF4) expression through the transcription factor Yin-Yang 1 (YY1) in non-Hodgkin B-cell lymphoma. Oncotarget *10*, 2173–2188.

Morrical, S.W. (2016). DNA-Pairing and Annealing Processes HR HDR. Cold Spring Harb Perspect Biol *7*, a016444.

Müller, L.U.W., Milsom, M.D., Harris, C.E., Vyas, R., Brumme, K.M., Parmar, K., Moreau, L.A., Schambach, A., Park, I.-H., London, W.B., et al. (2012). Overcoming reprogramming resistance of Fanconi anemia cells. Blood *119*, 5449–5457.

Murina, O., von Aesch, C., Karakus, U., Ferretti, L.P., Bolck, H.A., Hänggi, K., and Sartori, A.A. (2014). FANCD2 and CtIP cooperate to repair DNA interstrand crosslinks. Cell Rep. *7*, 1030–1038.

Nagano, O., Okazaki, S., and Saya, H. (2013). Redox regulation in stem-like cancer cells by CD44 variant isoforms. Oncogene *32*, 5191–5198.

Nakahara, Y., Northcott, P.A., Li, M., Kongkham, P.N., Smith, C., Yan, H., Croul, S., Ra, Y.-S., Eberhart, C., Huang, A., et al. (2015). Genetic and Epigenetic Inactivation of Kruppel-like Factor 4 in Medulloblastoma. Neoplasia *12*, 20–27.

Nevins, J.R. (2001). The Rb/E2F pathway and cancer. Hum. Mol. Genet. *10*, 699–703.

Nicoletti, I., Migliorati, G., Pagliacci, M.C., Grignani, F., and Riccardi, C. (1991). A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry. J. Immunol. Methods *139*, 271–279.

Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P., and Kowalczykowski, S.C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes Dev. *25*, 350–362.

Noordermeer, S.M., Adam, S., Setiaputra, D., Barazas, M., Pettitt, S.J., Ling, A.K., Olivieri, M., Álvarez-Quilón, A., Moatti, N., Zimmermann, M., et al. (2018). The shieldin complex mediates 53BP1-dependent DNA repair. Nature *560*, 117–121.

Ogiwara, H., and Kohno, T. (2012). CBP and p300 Histone Acetyltransferases Contribute to

Homologous Recombination by Transcriptionally Activating the BRCA1 and RAD51 Genes. PLoS One *7*.

Oliver, L., Hue, E., Séry, Q., Lafargue, A., Pecqueur, C., Paris, F., and Vallette, F.M. (2013). Differentiation-related response to DNA breaks in human mesenchymal stem cells. Stem Cells *31*, 800–807.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. Cell *132*, 631–644.

Pandya, A.Y., Talley, L.I., Frost, A.R., Fitzgerald, T.J., Trivedi, V., Chakravarthy, M., Chhieng, D.C., Grizzle, W.E., Engler, J.A., Krontiras, H., et al. (2004). Nuclear Localization of KLF4 Is Associated with An Aggressive Phenotype in Early-Stage Breast Cancer. Clin. Cancer Res. *10*, 2709–2719.

Panigada, M., Porcellini, S., Sutti, F., Doneda, L., Pozzoli, O., Consalez, G.G., Guttinger, M., and Grassi, F. (1999). GKLF in thymus epithelium as a developmentally regulated element of thymocyte-stroma cross-talk. Mech. Dev. *811*, 103–113.

Paudyal, S.C., Li, S., Yan, H., Hunter, T., and You, Z. (2017). Dna2 initiates resection at clean DNA double-strand breaks. Nucleic Acids Res. *45*, 11766–11781.

Paull, T.T., and Gellert, M. (1998). The 3′ to 5′ exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. Mol. Cell *1*, 969–979.

Petermann, E., and Helleday, T. (2010). Pathways of mammalian replication fork restart. Nat. Rev. Mol. Cell Biol. *11*, 683–687.

Peterson, S.E., and Loring, J.F. (2014). Genomic instability in pluripotent stem cells: Implications for clinical applications. J. Biol. Chem. *289*, 4578–4584.

Peterson, S.E., Li, Y., Wu-Baer, F., Chait, B.T., Baer, R., Yan, H., Gottesman, M.E., and Gautier, J. (2013). Activation of DSB Processing Requires Phosphorylation of CtIP by ATR. Mol. Cell *49*, 657–667.

Pfeiffer, P., Goedecke, W., and Obe, G. (2000). Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. Mutagenesis *15*, 289–302.

Philipsen, S., and Suske, G. (1999). A tale of three fingers: The family of mammalian Sp/XKLF transcription factors. Nucleic Acids Res. *27*, 2991–3000.

Phillips, D.H., Hewer, A., Martin, C.N., Garner, R.C., and King, M.M. (1988). Correlation of DNA adduct levels in human lung with cigarette smoking. Nature *336*, 790–792.

Pierce, A.J., Johnson, R.D., Thompson, L.H., and Jasin, M. (1999). XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev. *13*, 2633–2638.

Pietras, E.M., Warr, M.R., and Passegué, E. (2011). Cell cycle regulation in hematopoietic stem cells. J. Cell Biol. *195*, 709–720.

Polato, F., Callen, E., Wong, N., Faryabi, R., Bunting, S., Chen, H.-T., Kozak, M., Kruhlak, M.J., Reczek, C.R., Lee, W.-H., et al. (2014). CtIP-mediated resection is essential for viability and can operate independently of BRCA1. J. Exp. Med. *211*, 1027–1036.

Polo, S., and Jackson, S. (2011). Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. Genes Dev. *25*, 409–433.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617–1632.

Porter, A.G., and Jänicke, R.U. (1999). Emerging roles of caspase-3 in apoptosis. Cell Death Differ. *6*, 99–104.

Preiss, A., Rosenberg, U.B., Kienlin, A., Seifert, E., and Jäckle, H. (1985). Molecular genetics of Krüppel, a gene required for segmentation of the Drosophila embryo. Nature *313*, 27–32.

Prendergast, Á.M., Cruet-Hennequart, S., Shaw, G., Barry, F.P., and Carty, M.P. (2011). Activation of DNA damage response pathways in human mesenchymal stem cells exposed to cisplatin or γ-irradiation. Cell Cycle *10*, 3768–3777.

Price, B.D., and D’Andrea, A.D. (2013). Chromatin remodeling at DNA double-strand breaks. Cell *152*, 1344–1354.

Przetocka, S., Porro, A., Bolck, H.A., Walker, C., Lezaja, A., Trenner, A., von Aesch, C., Himmels, S.-F., D’Andrea, A.D., Ceccaldi, R., et al. (2018). CtIP-Mediated Fork Protection Synergizes with BRCA1 to Suppress Genomic Instability upon DNA Replication Stress. Mol. Cell *72*, 568-582.e6.

Puck, T.T., Marcu, P.I., and Cieciura, S.J. (1956). Clonal growth of mammalian cells in vitro; growth characteristics of colonies from single HeLa cells with and without a feeder layer. J. Exp. Med.

103, 273–283.

Qi, W., Chen, H., Xiao, T., Wang, R., Li, T., Han, L., and Zeng, X. (2016). Acetyltransferase p300 collaborates with chromodomain helicase DNA-binding protein 4 (CHD4) to facilitate DNA double-strand break repair. Mutagenesis *31*, 193–203.

Qi, X., Li, Y., Zhang, Y., Xu, T., Lu, B., Fang, L., Gao, J., Yu, L., Zhu, D., Yang, B., et al. (2019). KLF4 functions as an oncogene in promoting cancer stem cell-like characteristics in osteosarcoma cells. Acta Pharmacol. Sin. *40*, 546–555.

Qin, H., Yu, T., Qing, T., Liu, Y., Zhao, Y., Cai, J., Li, J., Song, Z., Qu, X., Zhou, P., et al. (2007). Regulation of apoptosis and differentiation by p53 in human embryonic stem cells. J. Biol. Chem. *282*, 5842–5852.

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.

Rothkamm, K., Krüger, I., Thompson, L.H., Kru, I., and Lo, M. (2003). Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle. Mol. Cell. Biol. *23*, 5706–5715.

Rousseau, L., Etienne, O., Roque, T., Desmaze, C., Haton, C., Mouthon, M.A., Bernardino-Sgherri, J., Essers, J., Kanaar, R., and Boussin, F.D. (2012). In vivo importance of homologous recombination DNA repair for mouse neural stem and progenitor cells. PLoS One *7*, e37194.

Rowland, B.D., Bernards, R., and Peeper, D.S. (2005). The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. Nat. Cell Biol. *7*, 1074–1082.

Rubio, D., Garcia-Castro, J., Martín, M.C., De La Fuente, R., Cigudosa, J.C., Lloyd, A.C., and Bernad, A. (2005). Spontaneous human adult stem cell transformation. Cancer Res. *65*, 3035–3039.

Ruiz, S., Panopoulos, A.D., Herrerías, A., Bissig, K.-D., Lutz, M., Berggren, W.T., Verma, I.M., and Belmonte, J.C.I. (2011). A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. Curr. Biol. *21*, 45.

Ruiz, S., Lopez-Contreras, A.J., Gabut, M., Marion, R.M., Gutierrez-Martinez, P., Bua, S., Ramirez, O., Olalde, I., Rodrigo-Perez, S., Li, H., et al. (2015). Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. Nat. Commun.

6, 1–8.

San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of Eukaryotic Homologous Recombination. Annu. Rev. Biochem. *77*, 229–257.

Sarig, R., Rivlin, N., Brosh, R., Bornstein, C., Kamer, I., Ezra, O., Molchadsky, A., Goldfinger, N., Brenner, O., and Rotter, V. (2010). Mutant p53 facilitates somatic cell reprogramming and augments the malignant potential of reprogrammed cells. J. Exp. Med. *207*, 2127–2140.

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.

Savic, V., Yin, B., Maas, N.L., Bredemeyer, A.L., Carpenter, A.C., Helmink, B.A., Yang-Iott, K.S., Sleckman, B.P., and Bassing, C.H. (2009). Formation of Dynamic γ-H2AX Domains along Broken DNA Strands Is Distinctly Regulated by ATM and MDC1 and Dependent upon H2AX Densities in Chromatin. Mol. Cell *34*, 298–310.

Schaeper, U., Subramanian, T., Lim, L., Boyd, J.M., and Chinnadurai, G. (1998). Interaction between a Cellular Protein That Binds to the C-terminal Region of Adenovirus E1A (CtBP) and a Novel Cellular Protein Is Disrupted by E1A through a Conserved PLDLS Motif. J. Biol. Chem. *273*, 8549–8552.

Schmidt, C.K., Galanty, Y., Sczaniecka-Clift, M., Coates, J., Jhujh, S., Demir, M., Cornwell, M., Beli, P., and Jackson, S.P. (2015). Systematic E2 screening reveals a UBE2D-RNF138-CtIP axis promoting DNA repair. Nat. Cell Biol. *17*, 1458–1470.

Serrano, L., Liang, L., Chang, Y., Deng, L., Maulion, C., Nguyen, S., and Tischfield, J.A. (2010). Homologous Recombination Conserves DNA Sequence Integrity Throughout the Cell Cycle in Embryonic Stem Cells. Stem Cells Dev. *20*, 363–374.

Sfeir, A., and Symington, L.S. (2015). Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? Trends Biochem. Sci. *40*, 701–714.

Shi, G., and Jin, Y. (2010). Role of Oct4 in maintaining and regaining stem cell pluripotency. Stem Cell Res. Ther. *1*, 39.

Shibata, A. (2017). Regulation of repair pathway choice at two-ended DNA double-strand breaks. Mutat. Res. Mol. Mech. Mutagen. *803–805*, 51–55.

Shibata, A., Moiani, D., Arvai, A.S., Perry, J., Harding, S.M., Genois, M.M., Maity, R., van Rossum-

Fikkert, S., Kertokalio, A., Romoli, F., et al. (2014). DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct MRE11 Nuclease Activities. Mol. Cell *53*, 7–18.

Shie, J.-L. (2000). Gut-enriched Kruppel-like factor represses cyclin D1 promoter activity through Sp1 motif. Nucleic Acids Res. *28*, 2969–2976.

Shields, J.M., and Yang, V.W. (1997). Two potent nuclear localization signals in the gut-enriched Kruppel- like factor define a subfamily of closely related Kruppel proteins. J. Biol. Chem. *272*, 18504–18507.

Shields, J.M., and Yang, V.W. (1998). Identification of the DNA sequence that interacts with the gut-enriched Kruppel-like factor. Nucleic Acids Res. *26*, 796–802.

Shields, J.M., Christy, R.J., and Yang, V.W. (1996). Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. J. Biol. Chem. *271*, 20009–20017.

Shimada, H., Nakada, A., Hashimoto, Y., Shigeno, K., Shionoya, Y., and Nakamura, T. (2010). Generation of canine-induced pluripotent stem cells by retroviral transduction and chemical inhibitors. Mol. Reprod. Dev. *7*, 2.

Shiotani, B., and Zou, L. (2009). Single-Stranded DNA Orchestrates an ATM-to-ATR Switch at DNA Breaks. Mol. Cell *33*, 547–558.

Shiras, A., Chettiar, S.T., Shepal, V., Rajendran, G., Prasad, G.R., and Shastry, P. (2007). Spontaneous Transformation of Human Adult Nontumorigenic Stem Cells to Cancer Stem Cells Is Driven by Genomic Instability in a Human Model of Glioblastoma. Stem Cells *25*, 1478–1489.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688–690.

Song, E., Ma, X., Li, H., Zhang, P., Ni, D., Chen, W., Gao, Y., Fan, Y., Pang, H., Shi, T., et al. (2013). Attenuation of Krüppel-Like Factor 4 Facilitates Carcinogenesis by Inducing G1/S Phase Arrest in Clear Cell Renal Cell Carcinoma. PLoS One *8*, e67758.

Soria-Bretones, I., Cepeda-García, C., Checa-Rodriguez, C., Heyer, V., Reina-San-Martin, B., Soutoglou, E., and Huertas, P. (2017). DNA end resection requires constitutive sumoylation of CtIP by CBX4. Nat. Commun. *8*, 113.

Sotiropoulou, P.A., Candi, A., Mascré, G., De Clercq, S., Youssef, K.K., Lapouge, G., Dahl, E., Semeraro, C., Denecker, G., Marine, J.C., et al. (2010). Bcl-2 and accelerated DNA repair mediates resistance of hair follicle bulge stem cells to DNA-damage-induced cell death. Nat. Cell Biol. *12*, 572–582.

Sotiropoulou, P.A., Karambelas, A.E., Debaugnies, M., Candi, A., Bouwman, P., Moers, V., Revenco, T., Rocha, A.S., Sekiguchi, K., Jonkers, J., et al. (2013). Brca1 deficiency in skin epidermis leads to selective loss of hair follicle stem cells and their progeny. Genes Dev. *27*, 39–51.

Stark, J.M., Pierce, A.J., Oh, J., Pastink, A., and Jasin, M. (2004). Genetic Steps of Mammalian Homologous Repair with Distinct Mutagenic Consequences. Mol. Cell. Biol. *24*, 9305–9316.

Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. Oncogene *21*, 8320–8333.

Steger, M., Murina, O., Hühn, D., Ferretti, L.P., Walser, R., Hänggi, K., Lafranchi, L., Neugebauer, C., Paliwal, S., Janscak, P., et al. (2013). Prolyl isomerase PIN1 regulates DNA double-strand break repair by counteracting DNA end resection. Mol. Cell *50*, 333–343.

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.

Stucki, M., Clapperton, J.A., Mohammad, D., Yaffe, M.B., Smerdon, S.J., and Jackson, S.P. (2005). MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. Cell *5*, 378–391.

Suvorova, I.I., Grigorash, B.B., Chuykin, I.A., Pospelova, T. V., and Pospelov, V.A. (2016). G1 checkpoint is compromised in mouse ESCs due to functional uncoupling of p53-p21waf1 signaling. Cell Cycle *15*, 52–63.

Symington, L.S. (2014). End resection at double-strand breaks: Mechanism and regulation. Cold Spring Harb. Perspect. Biol. *6*, a016436.

Tahmasebi, S., Ghorbani, M., Savage, P., Yan, K., Gocevski, G., Xiao, L., You, L., and Yang, X.J. (2013). Sumoylation of Krüppel-like factor 4 inhibits pluripotency induction but promotes adipocyte differentiation. J. Biol. Chem. *288*, 12791–12804.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872.

Talmasov, D., Zhang, X., Yu, B., Nandan, M.O., Bialkowska, A.B., Elkarim, E., Kuruvilla, J., Yang, V.W., and Ghaleb, A.M. (2015). Krüppel-like factor 4 is a radioprotective factor for the intestine following γ-radiation-induced gut injury in mice. Am. J. Physiol. Liver Physiol. *308*, G121–G138.

Thomson, J.A. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tian, Y., Luo, A., Cai, Y., Su, Q., Ding, F., Chen, H., and Liu, Z. (2010). MicroRNA-10b promotes migration and invasion through KLF4 in human esophageal cancer cell lines. J. Biol. Chem. *285*, 7986–7994.

Tichy, E.D., Pillai, R., Stambrook, P.J., Schwemberger, S.J., Liang, L., Tischfield, J., Babcock, G.F., and Deng, L. (2010). Mouse Embryonic Stem Cells, but Not Somatic Cells, Predominantly Use Homologous Recombination to Repair Double-Strand DNA Breaks. Stem Cells Dev. *19*, 1699–1711.

Tilgner, K., Neganova, I., Moreno-Gimeno, I., Al-Aama, J.Y., Burks, D., Yung, S., Singhapol, C., Saretzki, G., Evans, J., Gorbunova, V., et al. (2013). A human iPSC model of Ligase IV deficiency reveals an important role for NHEJ-mediated-DSB repair in the survival and genomic stability of induced pluripotent stem cells and emerging haematopoietic progenitors. Cell Death Differ. *20*, 1089–1100.

Ton-That, H., Kaestner, K.H., Shields, J.M., Mahatanankoon, C.S., and Yang, V.W. (1997). Expression of the gut-enriched Kruppel-like factor gene during development and intestinal tumorigenesis. FEBS Lett. *419*, 239–243.

Trujillo, K.M., Yuan, S.S.F., Lee, E.Y.H.P., and Sung, P. (1998). Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. J. Biol. Chem. *273*, 21447–21450.

Unno, J., Itaya, A., Taoka, M., Sato, K., Tomida, J., Sakai, W., Sugasawa, K., Ishiai, M., Ikura, T., Isobe, T., et al. (2014). FANCD2 binds CtIP and regulates DNA-end resection during DNA interstrand crosslink repair. Cell Rep. *7*, 1039–1047.

Vahidi Ferdousi, L., Rocheteau, P., Chayot, R., Montagne, B., Chaker, Z., Flamant, P., Tajbakhsh, S., and Ricchetti, M. (2014). More efficient repair of DNA double-strand breaks in skeletal muscle stem cells compared to their committed progeny. Stem Cell Res. *13*, 492–507.

Varum, S., Rodrigues, A.S., Moura, M.B., Momcilovic, O., Easley IV, C.A., Ramalho-Santos, J., van Houten, B., and Schatten, G. (2011). Energy metabolism in human pluripotent stem cells and their differentiated counterparts. PLoS One *6*, e20914.

Vitale, I., Manic, G., De Maria, R., Kroemer, G., and Galluzzi, L. (2017). DNA Damage in Stem Cells. Mol. Cell *66*, 306–319.

Walter, D., Lier, A., Geiselhart, A., Thalheimer, F.B., Huntscha, S., Sobotta, M.C., Moehrle, B., Brocks, D., Bayindir, I., Kaschutnig, P., et al. (2015). Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. Nature *520*, 549–552.

Wang, B., Zhao, M. zhi, Cui, N. peng, Lin, D. dan, Zhang, A. yi, Qin, Y., Liu, C. yun, Yan, W. tao, Shi, J. hong, and Chen, B. ping (2015). Krüppel-like factor 4 induces apoptosis and inhibits tumorigenic progression in SK-BR-3 breast cancer cells. FEBS Open Bio *5*, 147–154.

Wang, H., Shi, L.Z., Wong, C.C.L., Han, X., Hwang, P.Y.H., Truong, L.N., Zhu, Q., Shao, Z., Chen, D.J., Berns, M.W., et al. (2013a). The Interaction of CtIP and Nbs1 Connects CDK and ATM to Regulate HR-Mediated Double-Strand Break Repair. PLoS Genet. *9*, e1003277.

Wang, H., Li, Y., Truong, L.N., Shi, L.Z., Hwang, P.Y.H., He, J., Do, J., Cho, M.J., Li, H., Negrete, A., et al. (2014a). CtIP maintains stability at common fragile sites and inverted repeats by end resection-independent endonuclease activity. Mol. Cell *54*, 1012–1021.

Wang, K., Zhang, T., Dong, Q., Nice, E.C., Huang, C., and Wei, Y. (2013b). Redox homeostasis: The linchpin in stem cell self-renewal and differentiation. Cell Death Dis. *4*, e537.

Wang, L., Du, Y., Ward, J.M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y., Zhou, B., Han, L., Fargo, D.C., et al. (2014b). INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. Cell Stem Cell *14*, 575–591.

Wang, Z., Kong, J., Wu, Y., Zhang, J., Wang, T., Li, N., Fan, J., Wang, H., Zhang, J., and Ling, R. (2017). PRMT5 determines the sensitivity to chemotherapeutics by governing stemness in breast cancer. Breast Cancer Res. Treat. *168*, 1–12.

Ward, I.M., and Chen, J. (2001). Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. J. Biol. Chem. *276*, 47759–47762.

Wei, D., Gong, W., Kanai, M., Schlunk, C., Wang, L., Yao, J.C., Wu, T.T., Huang, S., and Xie, K. (2005). Drastic down-regulation of krüppel-like factor 4 expression is critical in human gastric cancer development and progression. Cancer Res. *65*, 2746–2754.

White, J., and Dalton, S. (2005). Cell Cycle Control of Embryonic Stem Cells Cell. Stem Cell Rev. *1*, 131–138.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R.C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C.F., Eshkind, L., Bockamp, E., et al. (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. Cell *135*, 1118–1129.

Wohlbold, L., and Fisher, R.P. (2009). Behind the wheel and under the hood: Functions of cyclin-dependent kinases in response to DNA damage. DNA Repair (Amst). *8*, 1018–1024.

Wolf, B.B., Schuler, M., Echeverri, F., and Green, D.R. (1999). Caspase-3 Is the Primary Activator of Apoptotic DNA Fragmentation via DNA Fragmentation Factor-45/Inhibitor of Caspase-activated DNase Inactivation. J. Biol. Chem. *274*, 30651–30656.

Wong, A.K.C., Ormonde, P.A., Pero, R., Chen, Y., Lian, L., Salada, G., Berry, S., Lawrence, Q., Dayananth, P., Ha, P., et al. (1998). Characterization of a carboxy-terminal BRCA1 interacting protein. Oncogene *17*, 2279–2285.

Wu, S., Shi, Y., Mulligan, P., Gay, F., Landry, J., Liu, H., Lu, J., Qi, H.H., Wang, W., Nickoloff, J.A., et al. (2007). A YY1-INO80 complex regulates genomic stability through homologous recombination-based repair. Nat. Struct. Mol. Biol. *14*, 1165–1172.

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., and Kosik, K.S. (2009). MicroRNA-145 Regulates OCT4, SOX2, and KLF4 and Represses Pluripotency in Human Embryonic Stem Cells. Cell *137*, 647–658.

Yang, W.T., and Zheng, P.S. (2014). Promoter hypermethylation of KLF4 inactivates its tumor suppressor function in cervical carcinogenesis. PLoS One *9*, e88827.

Yang, Y., and Bedford, M.T. (2013). Protein arginine methyltransferases and cancer. Nat. Rev. Cancer *13*, 37–50.

Yang, D., Scavuzzo, M.A., Chmielowiec, J., Sharp, R., Bajic, A., and Borowiak, M. (2016). Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. Sci. Rep. *6*, 21264.

Yao, S., Chen, S., Clark, J., Hao, E., Beattie, G.M., Hayek, A., and Ding, S. (2006). Long-term self-renewal and directed differentiation of human embryonic stem cells in chemically defined conditions. Proc. Natl. Acad. Sci. *103*, 6907–3912.

Yasuda, T., Kagawa, W., Ogi, T., Kato, T.A., Suzuki, T., Dohmae, N., Takizawa, K., Nakazawa, Y., Genet, M.D., Saotome, M., et al. (2018). Novel function of HATs and HDACs in homologous recombination through acetylation of human RAD52 at double-strand break sites.

Ye, B., Liu, B., Hao, L., Zhu, X., Yang, L., Wang, S., Xia, P., Du, Y., Meng, S., Huang, G., et al. (2018). Klf4 glutamylation is required for cell reprogramming and early embryonic development in mice. Nat. Commun. *9*, 1261.

Yeo, J.E., Lee, E.H., Hendrickson, E.A., and Sobeck, A. (2014). CtIP mediates replication fork recovery in a FANCD2-regulated manner. Hum. Mol. Genet. *23*, 3695–3705.

Yet, S.F., McA'Nulty, M.M., Folta, S.C., Yen, H.W., Yoshizumi, M., Hsieh, C.M., Layne, M.D., Chin, M.T., Wang, H., Perrella, M.A., et al. (1998). Human EZF, a Kruppel-like zinc finger protein, is expressed in vascular endothelial cells and contains transcriptional activation and repression domains. J. Biol. Chem. *273*, 1026–1031.

Yoon, H.S., and Yang, V.W. (2004). Requirement of Krüppel-like Factor 4 in Preventing Entry into Mitosis following DNA Damage. J. Biol. Chem. *279*, 5035–5041.

Yoon, H.S., Chen, X., and Yang, V.W. (2003). Krüppel-like Factor 4 Mediates p53-dependent G1/S Cell Cycle Arrest in Response to DNA Damage. J. Biol. Chem. *278*, 2101–2105.

Yoon, H.S., Ghaleb, A.M., Nandan, M.O., Hisamuddin, I.M., Dalton, W.B., and Yang, V.W. (2005). Krüppel-like factor 4 prevents centrosome amplification following γ-irradiation-induced DNA damage. Oncogene *24*, 4017–4025.

Yori, J.L., Johnson, E., Zhou, G., Jain, M.K., and Keri, R.A. (2010). Kruppel-like factor 4 inhibits epithelial-to-mesenchymal transition through regulation of E-cadherin gene expression. J. Biol. Chem. *285*, 16854–16863.

Yori, J.L., Seachrist, D.D., Johnson, E., Lozada, K.L., Abdul-Karim, F.W., Chodosh, L.A., Schiemann, W.P., and Keri, R.A. (2011). Krüppel-like factor 4 inhibits tumorigenic progression and metastasis in a mouse model of breast cancer. Neoplasia *13*, 601–610.

You, Z., and Bailis, J.M. (2010). DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. Trends Cell Biol. *20*, 402–409.

You, Z., Shi, L.Z., Zhu, Q., Wu, P., Zhang, Y.W., Basilio, A., Tonnu, N., Verma, I.M., Berns, M.W., and Hunter, T. (2009). CtIP Links DNA Double-Strand Break Sensing to Resection. Mol. Cell *36*, 954–969.

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. Mol. Cell. Biol. *24*, 9478–9486.

Yu, F., Li, J., Chen, H., Fu, J., Ray, S., Huang, S., Zheng, H., and Ai, W. (2011). Kruppel-like factor 4 (KLF4) is required for maintenance of breast cancer stem cells and for cell migration and invasion. Oncogene *30*, 2161–2172.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. Science *318*, 1917–1920.

Yu, X., Wu L.C., Bowcock, A.M., Aronheim, A., and Baer, R. (1998). The C-terminal (BRCT) domains of BRCA1 interact in vive with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. J. Biol. Chem. *273*, 25388–25392.

Yu, X., Fu, S., Lai, M., Baer, R., and Chen, J. (2006). BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. Genes Dev. *20*, 1721–1726.

Yuan, J., and Chen, J. (2009). N terminus of CtIP is critical for homologous recombination-mediated double-strand break repair. J. Biol. Chem. *284*, 31746–31752.

Yuan, M., Eberhart, C.G., and Kai, M. (2014). RNA binding protein RBM14 promotes radio-resistance in glioblastoma by regulating DNA repair and cell differentiation. Oncotarget *5*, 2820–2826.

Zeman, M.K., and Cimprich, K.A. (2014). Causes and Consequences of Replication Stress. Nat. Cell Biol. *16*, 2.

Zhang, P., Andrianakos, R., Yang, Y., Liu, C., and Lu, W. (2010). Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. J. Biol. Chem.

285, 9180–9189.

Zhang, W., Geiman, D.E., Shields, J.M., Dang, D.T., Mahatan, C.S., Kaestner, K.H., Biggs, J.R., Kraft, A.S., and Yang, V.W. (2000). The gut-enriched Kruppel-like factor (Kruppel-like factor 4) mediates the transactivating effect of p53 on the p21(WAF1)/(Cip)1 promoter. J. Biol. Chem. *275*, 18391–18398.

Zhao, W., Hisamuddin, I.M., Nandan, M.O., Babbin, B.A., Lamb, N.E., and Yang, V.W. (2004). Identification of Krüppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer. Oncogene *23*, 395–402.

Zhou, B.B.S., and Elledge, S.J. (2000). The DNA damage response: Putting checkpoints in perspective. Nature *408*, 433–439.

Zhou, G., Meng, S., Li, Y., Ghebre, Y.T., and Cooke, J.P. (2016). Optimal ROS Signaling Is Critical for Nuclear Reprogramming. Cell Rep. *15*, 919–925.

Zhou, Q., Hong, Y., Zhan, Q., Shen, Y., and Liu, Z. (2009). Role for Kruppel-Like Factor 4 in Determining the Outcome of p53 Response to DNA Damage. Cancer Res. *69*, 8284–8292.

Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 Helicase and Two Nucleases Dna2 and Exo1 Resect DNA Double-Strand Break Ends. Cell *134*, 981–994.

Zimmermann, M., Lottersberger, F., Buonomo, S.B., Sfeir, A., and De Lange, T. (2013). 53BP1 regulates DSB repair using Rif1 to control 5′ end resection. Science *339*, 700–704.

Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542–1548.

VIII. Publications

Stem Cell Reports
Article
ISSCR
OPEN ACCESS

CtIP-Specific Roles during Cell Reprogramming Have Long-Term Consequences in the Survival and Fitness of Induced Pluripotent Stem Cells

Daniel Gómez-Cabello,[1,4,5,*] Cintia Checa-Rodríguez,[1,2,4] María Abad,[3,6] Manuel Serrano,[3] and Pablo Huertas[1,2,5,*]
[1]Andalusian Center for Molecular Biology and Regenerative Medicine (CABIMER), Seville 41092, Spain
[2]Department of Genetics, University of Seville, Seville 41012, Spain
[3]Tumour Suppression Group, Spanish National Cancer Research Centre (CNIO), Madrid 28029, Spain
[4]Co-first author
[5]Co-senior author
[6]Present address: Vall d'Hebron Institute of Oncology (VHIO) Hospital Vall d'Hebron, Barcelona 08035, Spain
*Correspondence: daniel.gomez@cabimer.es (D.G.-C.), pablo.huertas@cabimer.es (P.H.)
http://dx.doi.org/10.1016/j.stemcr.2016.12.009

SUMMARY

Acquired genomic instability is one of the major concerns for the clinical use of induced pluripotent stem cells (iPSCs). All reprogramming methods are accompanied by the induction of DNA damage, of which double-strand breaks are the most cytotoxic and mutagenic. Consequently, DNA repair genes seem to be relevant for accurate reprogramming to minimize the impact of such DNA damage. Here, we reveal that reprogramming is associated with high levels of DNA end resection, a critical step in homologous recombination. Moreover, the resection factor CtIP is essential for cell reprogramming and establishment of iPSCs, probably to repair reprogramming-induced DNA damage. Our data reveal a new role for DNA end resection in maintaining genomic stability during cell reprogramming, allowing DNA repair fidelity to be retained in both human and mouse iPSCs. Moreover, we demonstrate that reprogramming in a resection-defective environment has long-term consequences on stem cell self-renewal and differentiation.

INTRODUCTION

The ability to generate induced pluripotent stem cells (iPSCs) has been heralded to have great potential in regenerative medicine and research (Yu et al., 2007; Takahashi and Yamanaka, 2006). However, this potential is currently under debate, due to evidence that iPSCs can acquire DNA damage and genomic instability during the reprogramming process (Ruiz et al., 2015; Liang and Zhang, 2013; Gore et al., 2011; Mayshar et al., 2010). In fact, even just expressing the reprogramming factors, regardless of the methodology used to generate them, causes DNA damage, mainly by replication stress (Ruiz et al., 2015; Gonzalez et al., 2013; Tilgner et al., 2013). It is critical, however, to obtain "safe" iPSCs that are genetically identical to their parent cells for clinical use. An essential prerequisite for this is to obtain a thorough understanding about how the DNA repair machinery acts in these cells.

Several pieces of evidence suggest that pluripotent stem cells need more active DNA repair pathways than somatic differentiated cells (Rocha et al., 2013). Supporting this view, members of the DNA damage response (DDR) have been shown to prevent genomic instability in iPSCs (Hong et al., 2009; Kawamura et al., 2009; Li et al., 2009). Indeed, proteins involved in the repair of DNA double-strand breaks (DSBs), in both homologous recombination (HR) and non-homologous end joining (NHEJ), have a relevant role in reprogramming efficiency (Gonzalez et al., 2013; Ruiz et al., 2013; Tilgner et al., 2013). HR is required for an error-free repair of DSBs, using homologous sequences (normally from the sister chromatid) (Heyer et al., 2010), and for restarting replication forks stalled during replication stress (Petermann and Helleday, 2010). In contrast, NHEJ competes with HR for DSB repair in a more error-prone pathway (Gomez-Cabello et al., 2013; Huertas, 2010; Lieber, 2008). DNA end resection is a key event that regulates the DSB repair pathway choice between NHEJ and HR. This mechanism generates single-strand DNA (ssDNA) by 5′ to 3′ degradation at both sides of a break (Huertas and Jackson, 2009; Jackson and Bartek, 2009). Although resected DNA is an obligate substrate for HR, it blocks NHEJ (Heyer et al., 2010). CtIP is a major player in the decision between HR and NHEJ as it allows for ssDNA formation, precluding binding of the NHEJ machinery to DNA breaks (Huertas, 2010). DNA end resection is highly regulated by multiple signals, including cell-cycle-dependent CtIP phosphorylation (Sartori et al., 2007). Cells depleted of CtIP fail to repair DNA DSBs by HR, are sensitive to DNA damaging agents, and accumulate chromosomal aberrations in response to DNA damage (Sartori et al., 2007; Huertas and Jackson, 2009).

Here, we identified DNA end resection as one of the predominant mechanisms required for cell reprogramming in human and mouse iPSCs (miPSCs), most likely based on its function in repairing replication-induced DNA insults. In addition, we have determined that CtIP is essential for reprogramming efficiency in both organisms. Cells deficient for CtIP underwent apoptosis rather than reprogramming,

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

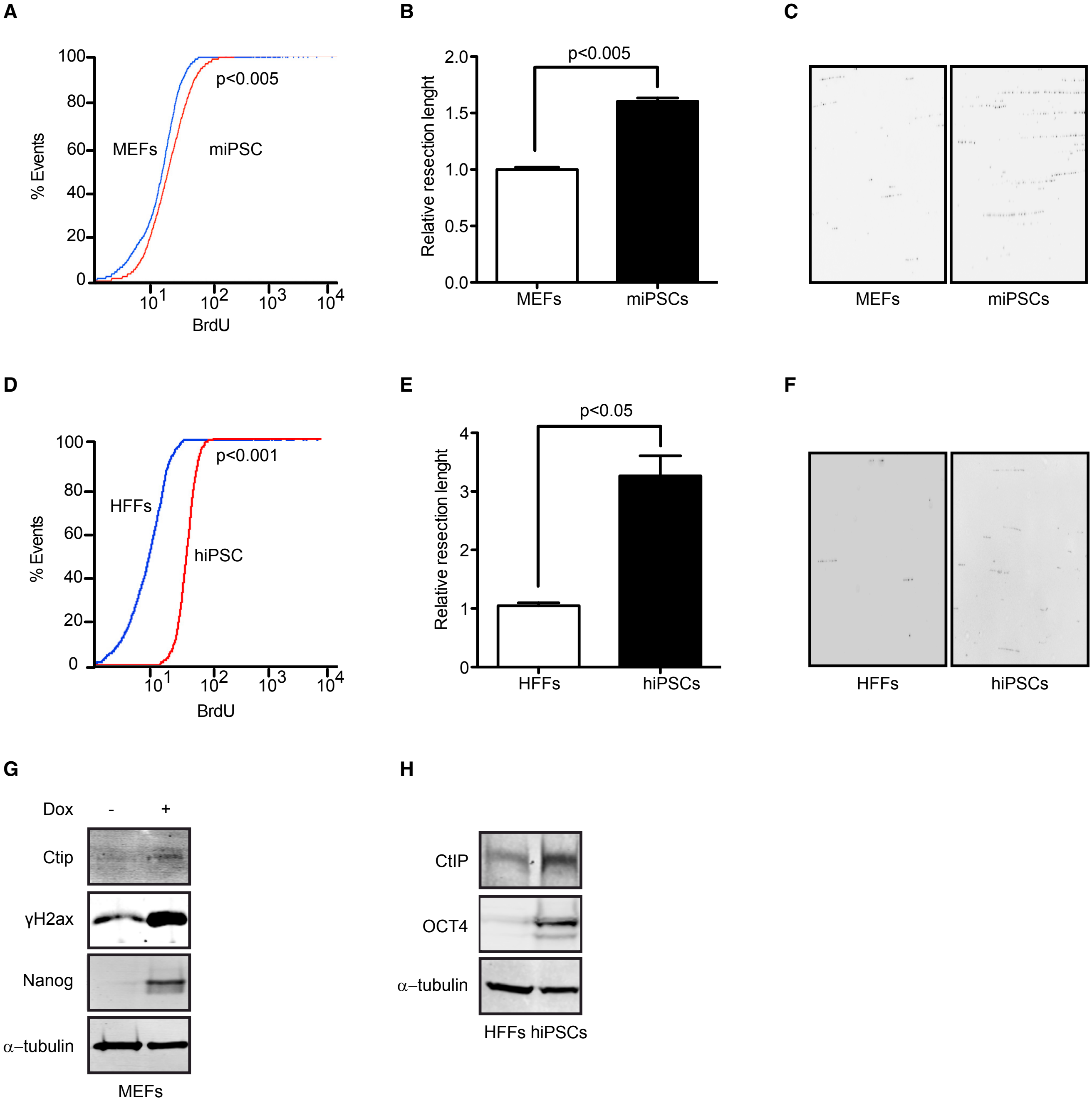


Figure 1. DNA End Resection Is an Essential Mechanism for Cell Reprogramming

(A) FACS analysis of BrdU exposed by DNA end resection in MEFs and their respective reprogrammed cells (miPSCs). p Values were calculated using the Kolmogorov-Smirnov test. At least three independent experiments were performed. Representative histogram is shown.

(B) Resected DNA length obtained by SMART technique in MEFs and miPSCs. Error bars indicate ±SEM of three independent experiments.

(C) Representative images of DNA fibers visualized with the anti-BrdU antibody.

(D) Same as (A) except using human foreskin fibroblasts (HFFs) and the human iPSCs (hiPSCs) derived from them. At least three independent experiments were performed. Representative histogram is shown.

(E) Same as (B) except using human cells. Error bars indicate ±SEM of three independent experiments.

(F) Same as (C) except using human cells.

(legend continued on next page)

probably due to an unbearable load of DNA damage during the process. Those cells that could reprogram with low levels of CtIP acquired a burden in terms of genomic mutations that compromised their long-term survival and ability to differentiate again. We suggest that CtIP has a genomic stability protector role during reprogramming. Exploiting such a role could contribute to creating genetically stable iPSCs that meet clinical safety standards for use in regenerative medicine.

RESULTS

DNA End Resection Increases in Mouse and Human iPSCs

Cell reprogramming per se, by the expression of reprogramming factors (OCT4, SOX2, KLF4, and c-MYC), increases DNA damage and genetic instability, mainly by replication stress (Ruiz et al., 2015). Here, we used a previously reported mouse embryonic fibroblast (MEF) cell line bearing the doxycycline-inducible system of mouse genes of *Klf4*, *Oct4*, *Sox2*, and *c-Myc* to carry out the reprogramming process (Abad et al., 2013). First, we analyzed cellular levels of DNA end resection in MEFs and their corresponding iPSCs generated by doxycycline treatment. We developed a new strategy for a readout of DNA end resection based on bromodeoxyuridine (BrdU) detection by fluorescence-activated cell sorting (FACS) analysis using native conditions. In contrast to standard proliferation assays using BrdU incorporation, this assay is based on a BrdU epitope that is hidden in double-stranded DNA, and thereby unavailable to anti-BrdU antibodies under native conditions. Critically, the assay is non-responsive to DNA replication, and the epitope is only exposed after formation of ssDNA by resection. This novel method demonstrated that miPSCs had more exposed BrdU than primary MEFs not treated with doxycycline, showing that a higher amount of endogenously occurring breaks were resected in reprogrammed cells (Figure 1A). We further confirmed that this increased BrdU signal intensity was indeed due to canonical DNA end resection, as it disappeared when the key resection factor CtIP was depleted (Figure S1A).

These results demonstrated that the DNA end resection process was activated in miPSCs in the absence of exogenous damage, most likely due to replication stress and DNA damage generated during cell reprogramming. We hypothesized that this resection activation reflects not only an increased number of breaks being processed, but also a higher processivity of the resection machinery itself. Thus, we analyzed whether the length of resected DNA was longer in miPSCs than in MEFs, using a high-resolution technique to measure the length of resected DNA in individual DNA fibers (Cruz-Garcia et al., 2014). We demonstrated that miPSCs generated significantly longer tracks of ssDNA compared with the primary differentiated parent cells (Figures 1B and 1C). A 50% increase in the median length of resected DNA was observed in pluripotent cells with respect their MEF control (Figures 1B and 1C). Again, we could demonstrate that this was caused by activation of the canonical resection machinery, as this increased length of ssDNA depended on CtIP activity (Figure S1C). Strikingly, the number of lesions and the amount of resected DNA following reprogramming to iPSCs was equivalent to that seen after treating primary cells with high doses of exogenous damage (Figures S1B and S1C), in agreement with the idea that this process represents a severe challenge for genomic integrity.

To address whether the activation of resection during cell reprogramming was evolutionarily conserved, we investigated whether DNA end processing also increases during reprogramming of primary human cells. We used four retroviral vectors bearing one of the OSKM factors (*OCT4*, *SOX2*, *KLF4*, or c-*MYC*) to generate human iPSCs (hiPSCs) from human foreskin fibroblasts (HFFs). Similar to miPSCs, hiPSCs showed both increased ssDNA-BrdU exposure, indicating a higher number of resected DSBs (Figure 1D) and longer resected tracks (Figures 1E and 1F) than the parental HFF somatic cells, despite the lack of any exogenous source of DNA damage. In fact, this effect in hiPSCs was more exacerbated than in miPSCs, as hiPSCs had up to a 3-fold gain on the length of resected DNA. All of our results point toward a hyper-activation of DNA resection during reprogramming in mouse and human cells, which likely minimizes the impact of the DNA damage caused during the process.

CtIP Levels Increase in miPSCs and hiPSCs

CtIP activates DNA end processing in HR and is known to be a major regulator of DNA end resection (Cruz-Garcia et al., 2014; Wang et al., 2013; Nakamura et al., 2010; Huertas and Jackson, 2009; Huertas et al., 2008; Sartori et al., 2007). Indeed, we observed that CtIP was required for the resection hyper-activation observed in miPSCs (Figures

---

(G) MEFs and miPSCs were immunoblotted to analyze the indicated proteins. At least three independent experiments were performed. A representative western blot is shown.

(H) Same as (G) except showing protein levels in HFFs and hiPSCs. At least three independent experiments were performed. A representative western blot is shown.

See also Figures S1–S3.

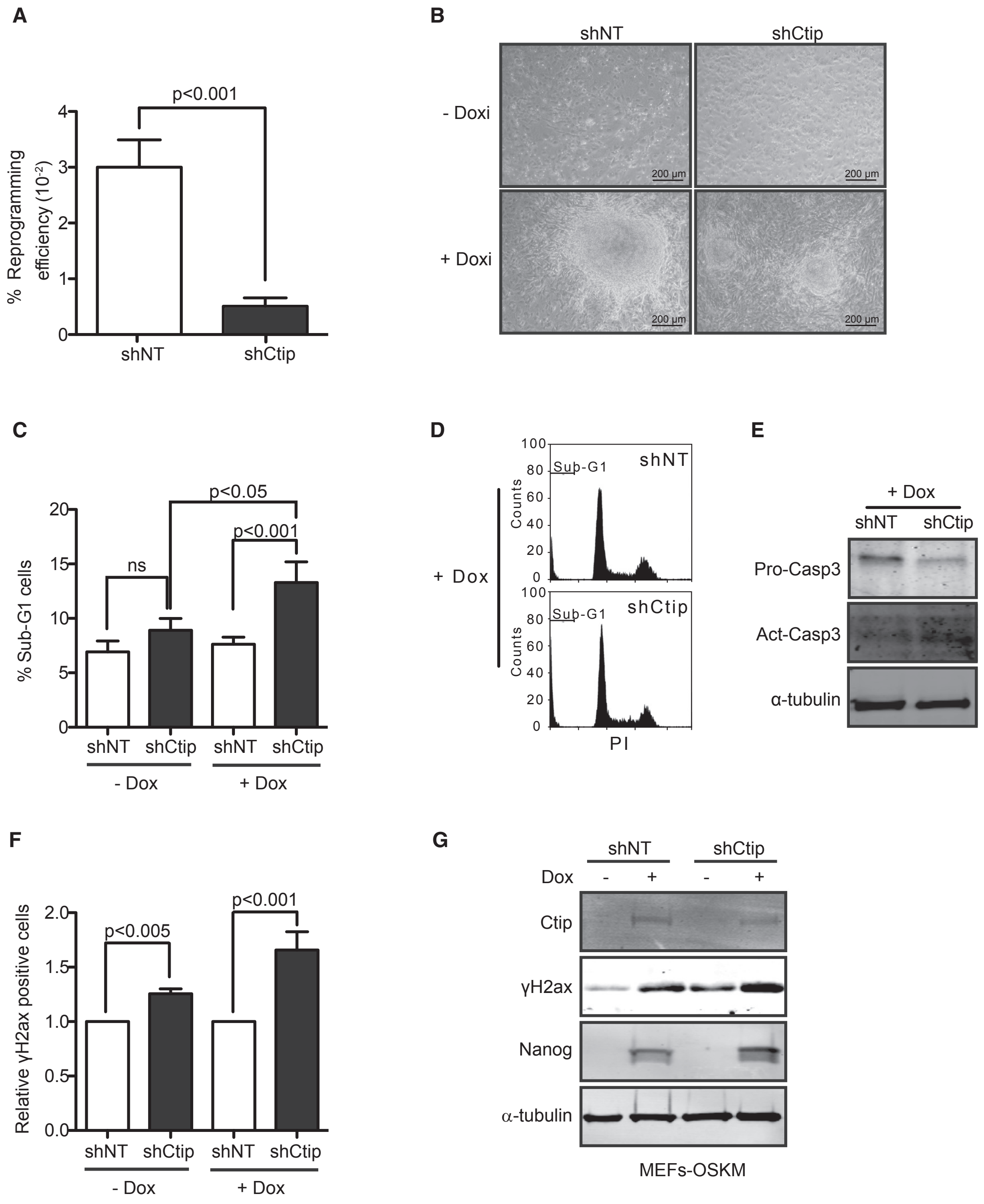


Figure 2. CtIP Deficiency Impairs Mouse Cell Reprogramming

(A) Doxycycline-inducible OSKM MEFs were transduced with shCtIP and shNT lentivirus prior to induction with doxycycline. Reprogramming efficiency was analyzed by counting the number of colonies in triplicate of each biological sample. Error bars indicate ±SEM of a biological triplicate.

(B) Representative pictures of miPSC morphology.

(C) The sub-G1 peak was quantified in MEFs (−Dox) or reprogrammed MEFs at day 10 (+Dox) by FACS analysis in cells downregulated for CtIP (black bars) or expressing a control shRNA (white bars). Error bars indicate ±SEM of three independent experiments.

(D) Representative cell-cycle plots of reprogrammed MEFs in each condition. At least three independent experiments were performed.

(legend continued on next page)

S1A and S1C). Congruently, we observed that CtIP expression and protein levels increased in miPSCs with respect to primary MEFs, concurrent with the expression of the pluripotency marker Nanog (Figures 1G and S2A–S2C). We confirmed by immunoblot that hiPSCs also incremented the expression and protein levels of CtIP in a concomitant manner to NANOG and OCT4 (Figures 1H, S2D, and S2E). These results clearly suggested that the hyper-active DNA resection in iPSCs requires CtIP upregulation during the reprogramming process, most likely to manage an increased load of DSBs that have to be repaired by HR. Thus, both mouse and human iPSCs showed intrinsic increased levels of CtIP and DNA end resection compared with their parent cells.

CtIP Is Required for Efficient Cell Reprogramming for Both Mouse and Human Cells

To determine the relevance of CtIP and its upregulation during cell reprogramming, we forced the induction process in cells with reduced levels of the CtIP protein. For this, we transduced MEFs with short hairpin RNA (shRNA) lentivirus against CtIP or control shRNA (shNon-Target [shNT]), which also contained the OSKM factors under doxycycline induction, and then analyzed them for reprogramming (Figure S2A). Three weeks after induction, we observed a dramatic and significant reduction of the number of colonies with stem cell-like morphology in CtIP-depleted cells compared with shNT control cells (Figure 2A). Moreover, CtIP-depleted colonies were smaller than control ones (Figure 2B). In agreement with results published previously (Chen et al., 2005), and as a control of the CtIP depletion effects in primary MEFs, we observed less proliferation in CtIP-downregulated MEFs (Figure S3A). However, this reduced proliferation was not associated with an increase in sub-G1 phase cells (Figure 2C, –Dox), even though, as expected, CtIP deficiency provoked a significant increase in DNA damage, as measured by γH2ax (Figures 2F and 2G). Interestingly, in agreement with the idea that unrepaired DNA damage during cell reprogramming triggers apoptosis as a consequence of genomic instability, we observed that reprogrammed mouse cells expressing an shRNA against CtIP showed a strong increase in sub-G1 cells 10 days after doxycycline induction (Figures 2C and 2D), compared with either reprogrammed cells expressing a control shRNA or MEFs depleted for CtIP. Further, we observed an increase of proteolytic cleavage and activation of caspase-3 in reprogrammed MEFs with short hairpin CtIP (shCtIP) at day 10 (Figure 2E), confirming that the increased levels of sub-G1 were indeed due to induced apoptosis.

We also detected that the amount of the DNA damage marker γH2ax increased slightly but significantly in cells containing shCtIP versus those with shNT, as shown by both FACS analysis and immunoblot (Figures 2F and 2G). Even though this difference was also observed in MEFs, it was more intense in reprogrammed cells, in agreement with CtIP playing a more critical role in repairing DNA damage during cell reprogramming. This suggests that excessive damage prompted the observed induction of apoptosis. Pluripotent status was confirmed by the observation of Nanog protein levels (Figures 2G, S2A, and S2C).

Efficient reprogramming requires functional repair pathways to favor an error-free dedifferentiation. In agreement with this, DDR was activated during reprogramming process, as shown by the increase in Chk1, γH2ax, and CtIP protein levels (Figures 2G and S2A). Reprogrammed cells that had downregulated CtIP also had increased DNA damage levels compared with reprogrammed control cells at 10 days after doxycycline induction (Figure 2G), supporting the idea that replication stress causes DNA damage and Chk1 activation during reprogramming (Ruiz et al., 2015).

Congruently, hiPSCs obtained from HFF cells by expression of OSKM factors in the presence of an shRNA against CtIP showed an increased sub-G1 peak (Figures 3A and 3B), an enrichment of the activated form of caspase-3 (Figure 3C) and elevated numbers of γH2AX-positive cells (Figure 3D), when compared with the same cells reprogramed bearing a control shRNA. In addition, expressing shCtIP constitutively during reprogramming completely blocked iPSC colony generation, while expressing control shRNA had no effects (not shown). We wondered whether this strong effect was due to cell death caused by a lack of CtIP in HFF cells (e.g., an excessive depletion of CtIP) or reflected a problem during reprogramming. Strikingly, CtIP depletion in HFF cells without OSKM expression neither changed their proliferation rate (Figure S3B) nor significantly induced DNA damage or apoptosis, measured as sub-G1 cells (Figures 3A and 3D). Thus, we conclude that the lack of iPSCs was not caused by problems in HFF cells,

(E) Immunoblot of proactive and active/cleavage caspase-3 in reprogrammed MEFs at day 10 (+Dox) bearing the indicated shRNAs. A representative western blot is shown of three independent experiments.
(F) FACS quantification of cells positive for γH2ax. Other details as in (C).
(G) The indicated proteins were immunodetected in samples from OSKM-induced MEFs after 10 days of continuous doxycycline treatment. A representative western blot is shown of three independent experiments.
t Test statistical analyses in (A), (C), and (F) were performed using at least three independent experiments.
See also Figures S2 and S3.

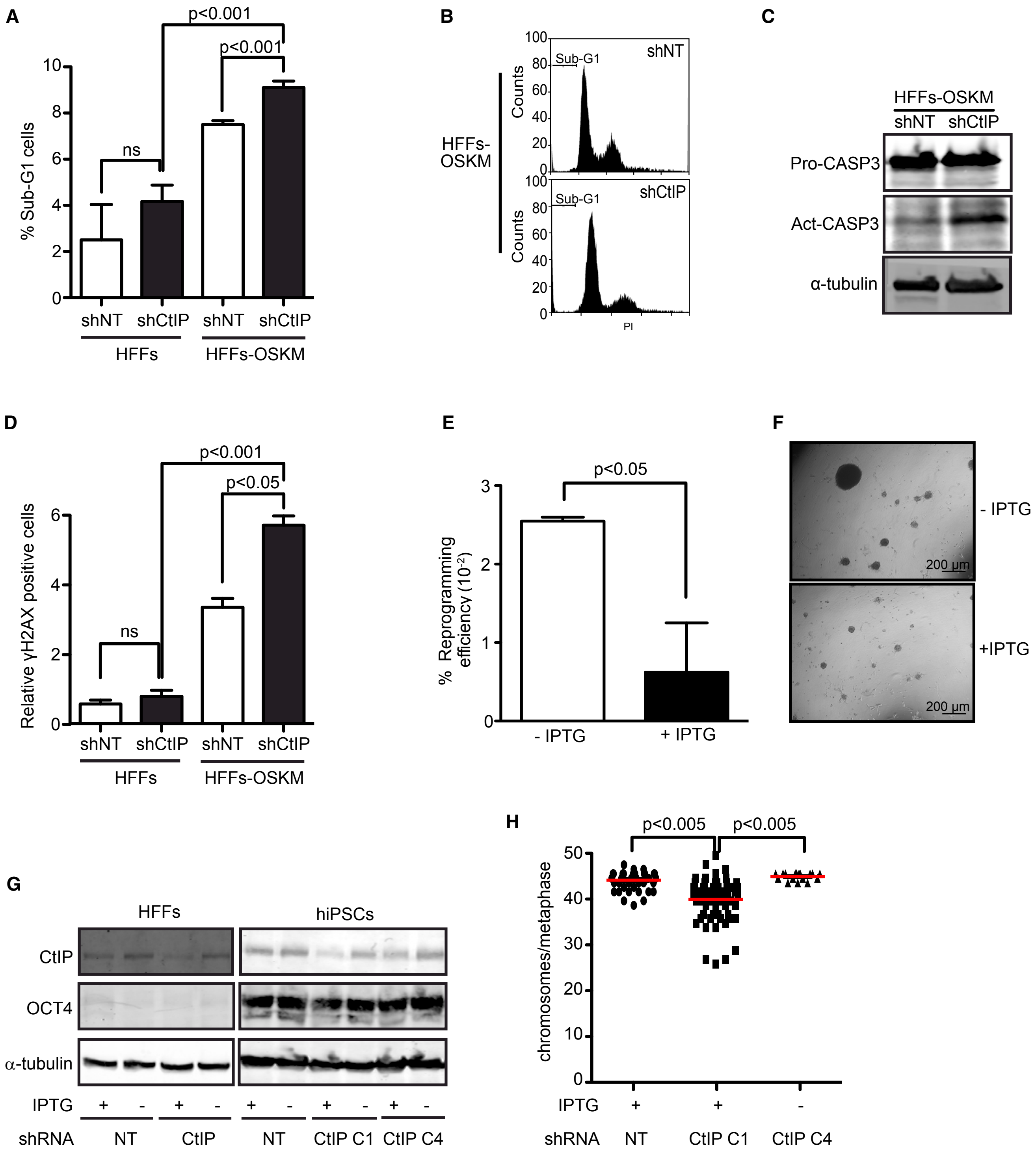


Figure 3. CtIP Depletion Impairs Human Cell Reprogramming

(A) Percentage of cells in sub-G1 phase in HFF cells or those reprogrammed to iPSCs bearing either the control shRNA shNT (white bars) or shCtIP (black bars). Error bars indicate ±SEM of three independent experiments. p Values were calculated using the t test. ns, not significant.

(B) Representative FACS plots of cell-cycle analysis in reprogrammed HFF cells of each condition. At least three independent experiments were performed.

(legend continued on next page)

but by a critical role of CtIP in human cell reprogramming. To confirm this relevant role, we next generated iPSCs with limited CtIP downregulation by using an isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible shRNA (shCtIP-IPTG) and titrating IPTG concentration. Indeed, and similar to miPSCs, HFF cells partially depleted of CtIP by IPTG induction also formed fewer colonies, confirming that they had an impaired reprogramming efficiency (Figure 3E). In addition, colonies partially deficient for CtIP were smaller than control colonies (Figure 3F). Indeed, although we were able to expand several clones from the iPSCs with an IPTG-induced control shRNA (shNon-target [shNT] clones), we only successfully expanded one clone of iPSCs bearing the shCtIP-IPTG under IPTG induction (CtIP C1 clone; Figure 3G). Interestingly, we could recover CtIP expression on this clone by removing IPTG. As an additional control, we reprogrammed HFF cells containing shCtIP-IPTG to iPSCs in the absence of IPTG (C4 clone) (Figure 3G). Hence, by adding IPTG at different time points, we could compare distinct states of HFF cells after reprogramming: (1) reprogrammed in the absence of CtIP (C1 clone); (2) reprogrammed in the presence of CtIP and depleted of CtIP once pluripotency had been established (C4 clone); and (3) always with CtIP (Figure 3G). Analyzing these reprogrammed clones for chromosome number variation, we found that the C1 clone (reprogrammed without CtIP) had a significant increase in chromosomal aberrations, measured as aneuploidy, compared with those from the other two conditions (Figure 3H). This is in agreement with our results showing increased DNA damage markers when CtIP is absent during reprogramming (Figure 3D). We provide strong evidence that CtIP plays a role in avoiding the genomic instability generated specifically during cell reprogramming.

Normal CtIP Levels during Reprogramming Are Required for Maintenance and Differentiation of iPSCs

Due to the difficulty of expanding iPSC colonies from MEFs or HFFs reprogrammed in the absence of CtIP, we wondered whether the genomic instability generated during reprogramming in CtIP-downregulated cells affected self-renewal and differentiation attributes. In fact, it is known that deficiency in BRCA1, another protein involved in HR and DNA end resection that interacts with CtIP, impairs maintenance of iPSCs in culture (Gonzalez et al., 2013). To deepen our understanding about the role of CtIP, we selected several miPSC clones obtained in the presence of shCtIP or a control shRNA. Taking advantage of the presence of the *GFP* gene on the plasmid, we analyzed the permanence of GFP cells in the colonies as a proxy for the presence of the shRNA targeting CtIP. We determined that all control shNT-harboring iPSC colonies maintained higher numbers of GFP cells than shCtIP iPSCs, suggesting that the cells that continued to grow and to maintain cell pluripotency had a tendency to lose the cassette containing GFP and the shRNA against CtIP (Figures 4A and 4B). Strikingly, this effect was specific for iPSC cells that had been reprogrammed in the presence of shCtIP and was not observed in primary MEFs with shCtIP (Figure S4). Thus, collectively, miPSC colonies reprogrammed with CtIP depletion had a natural selection favoring cells with normal CtIP expression (Figure 4C). This could explain the high CtIP levels in cells bearing shCtIP (Figure S2A). To study the kinetics of loss of shCtIP expression, we sorted GFP-positive miPSCs generated from cells containing either shNT-GFP or shCtIP-GFP and analyzed them for GFP disappearance during growth (Figure 4D). We found that GFP cells were rapidly purified from the cell population containing shCtIP-GFP compared with control cells, with a reduction in the number of GFP-positive cells of up to 50% after only three passages. We next performed a self-renewal assay using shCtIP- or shNT-miPSCs after enriching the populations of shRNA-bearing cells by cell sorting, using GFP as a marker. Again, reprogrammed cells depleted for CtIP formed less-viable colonies than control cells (Figures 4E and 4G). To differentiate whether this effect of CtIP deficiency was due to defects gained during cell reprogramming or those acquired during the pluripotent state of the reprogrammed cells, we transducted

(C) Reprogrammed HFF cells bearing the indicated shRNAs at day 10 were immunoblotted to analyze for the proactive and active forms of caspase-3. A representative western blot is shown of three independent experiments.

(D) Percentage of positive cells for γH2AX. Other details as in (A).

(E) Reprogramming efficiency of human cells containing an IPTG-inducible shRNA against CtIP (shCtIP-IPTG) or the control shNT-IPTG was analyzed in three independent experiments. Error bars indicate ±SEM. Statistical significance was performed using a t test.

(F) Representative images of hiPSC morphology in cells harboring the inducible shCtIP-IPTG obtained in the presence or absence of IPTG, as indicated.

(G) Western blot analysis against the CtIP and OCT4 proteins in HFF cells and hiPSC clones bearing the indicated shRNAs and in the presence (+) or absence (−) of IPTG. A representative western blot is shown of three independent experiments.

(H) Analysis of chromosomes per metaphase in iPSCs harboring IPTG-inducible shCtIP or shNT and reprogrammed in the presence (+) or absence (−) of IPTG. Tukey's multiple comparison test was used with n > 75.

See also Figure S2.

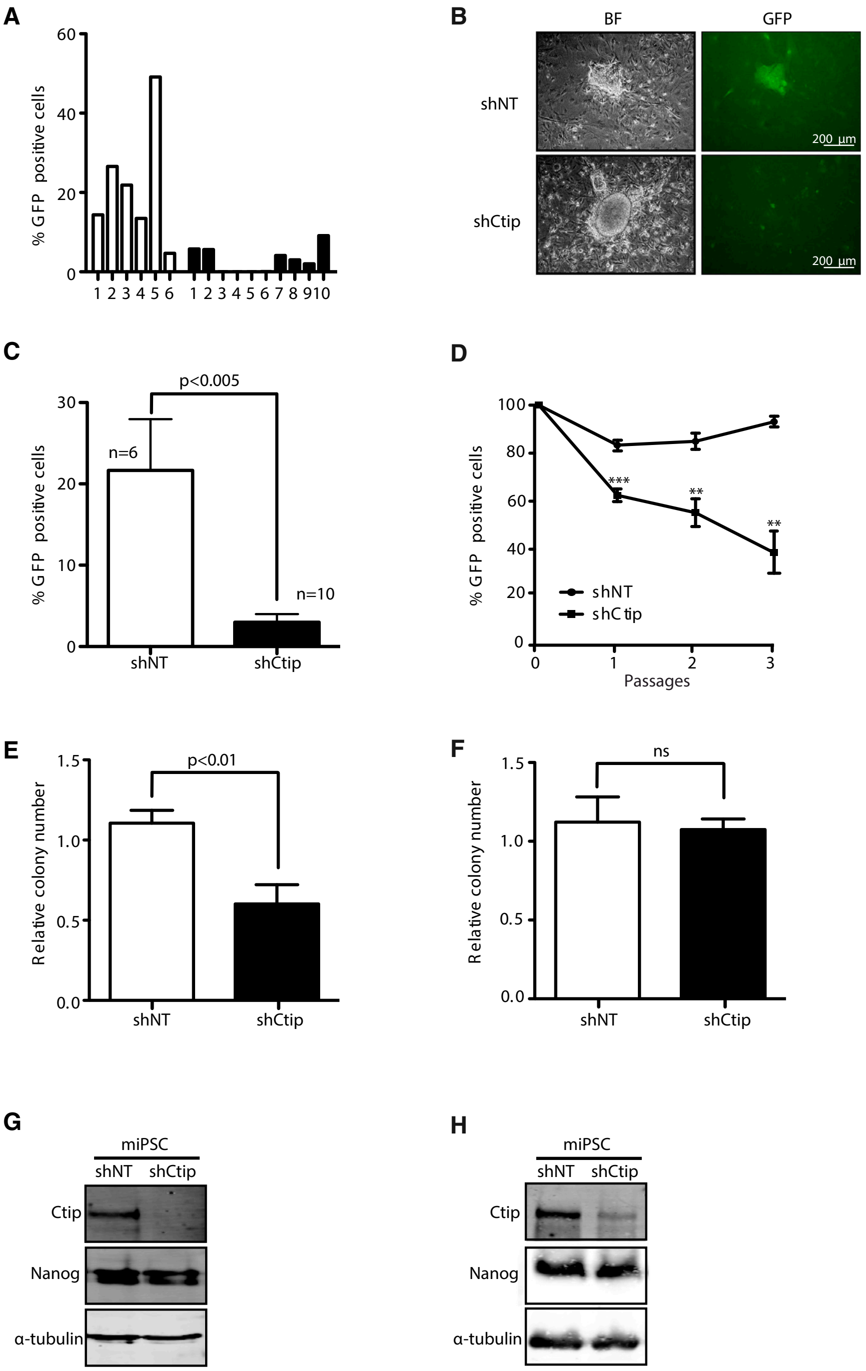


Figure 4. The Levels of CtIP During Reprogramming Affect iPSC Maintenance

(A) Percentage of GFP-positive cells in iPSC colonies isolated after MEFs reprogramming in the presence (white bars) or absence (black bars) of CtIP.

(B) Representative bright field (BF) and GFP images of miPSCs obtained by microscopy.

(C) Average percentage of GFP cells in several clones of shNT and shCtIP iPSCs. Error bars indicate ±SEM of different clones ($n = 6$ and $n = 10$). Statistical analysis was performed using a t test.

(D) Percentage of iPSCs harboring shCtIP-GFP or shNT-GFP during subsequence passages of the cells. The average and SD of five independent experiments is plotted. t test analysis for each passage is shown. $^{**}p < 0.01$, $^{***}p < 0.005$.

(E) The same amount of miPSCs reprogrammed either in the presence (white bars) or absence (black bars) of CtIP were seeded at low density and the number of each colony was measured. The relative number of colonies formed from three independent experiments is plotted. The t test was performed to compare both conditions.

(F) Same as (E) but with cells reprogrammed in the presence of CtIP, and then transduced with an shRNA against CtIP (black bars) or control shNT (white bars). Other details as in (E). ns, not significant.

(G) miPSCs reprogrammed in the presence of shCtIP or control shNT were immunoblotted to analyze the indicated proteins. A representative western blot is shown of three independent experiments.

(H) Same as in (G) but with cells transduced with an shRNA against CtIP after reprogramming.

See also Figures S4 and S5.

already-reprogrammed iPSCs with either an shRNA against CtIP or against a control sequence and analyzed them for self-renewal by colony formation. Interestingly, in this scenario, CtIP downregulation did not modify self-renewal capacity compared with control cells (Figures 4F and 4H). Along the same lines, the murine D3 embryonic stem cell line (ES-D3) was not affected by CtIP depletion in self-renewal experiments (Figures S5A–S5C).

As self-renewal was compromised by inherited defects aroused upon cell reprogramming in the absence of CtIP, we wondered whether the differentiation process could also be jeopardized in a similar way. Despite the difficulty in expanding CtIP-deficient miPSCs, we were able to force mouse embryonic body differentiation from several clones, by removing the growth factor, leukemia inhibitory factor (LIF) for 6 days. We observed that miPSCs reprogrammed

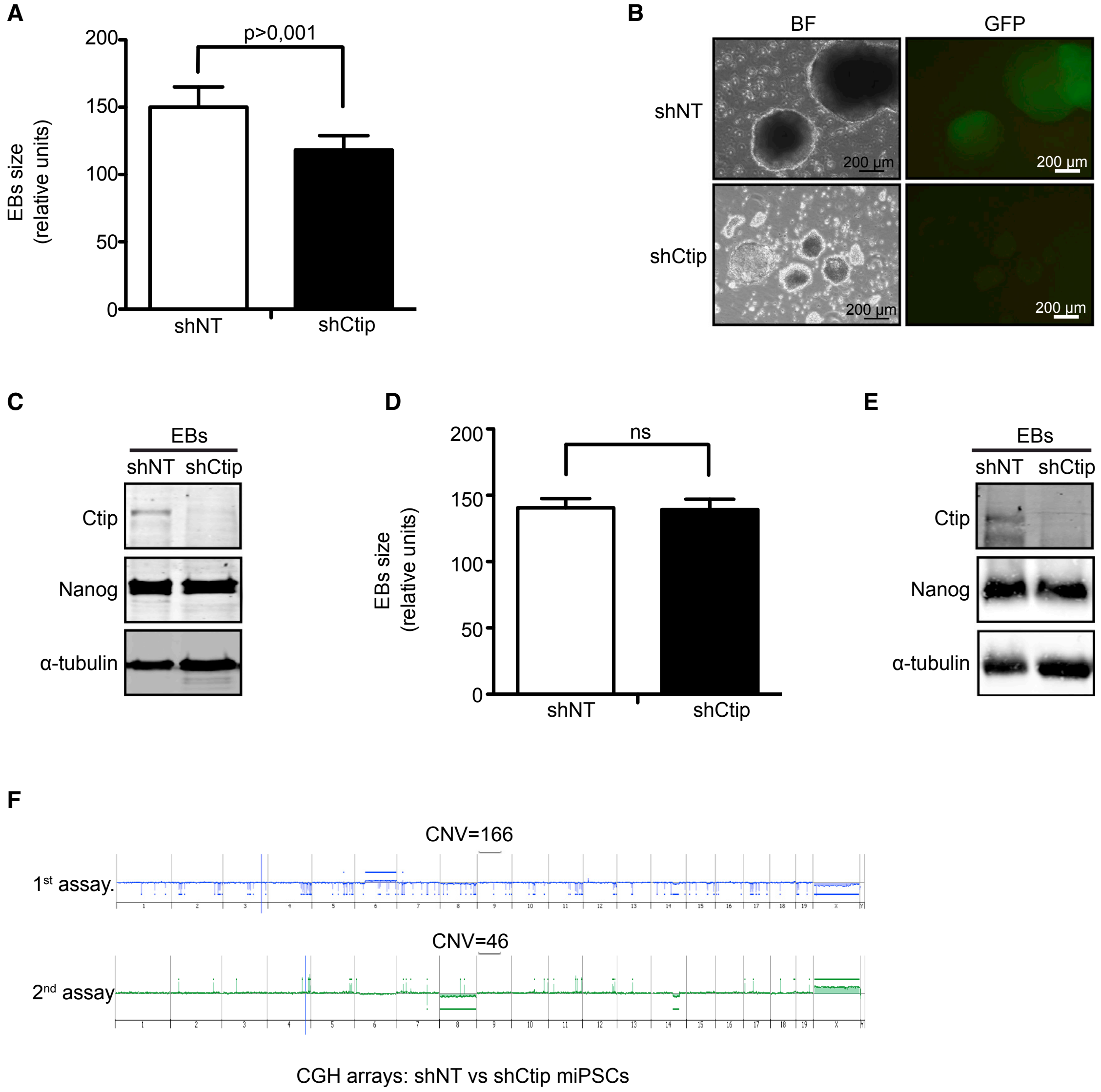


Figure 5. CtIP Is Essential for Differentiation and Genomic Stability Maintenance in iPSCs

(A) Median size of embryonic bodies (EBs) generated at 6 days after spontaneous differentiation of miPSCs reprogrammed in the presence of shCtIP or control shNT. At least 300 EBs of each condition were measured. The average and SD of five independent experiments is plotted. Statistical significance was obtained by a t test.

(B) Representative images of EBs using BF and GFP filters.

(C) Western blot analysis of indicated proteins in EBs generated from miPSCs reprogrammed in the presence of shCtIP or control shNT. A representative western blot is shown of three independent experiments.

(D) Same as in (A), but with cells reprogrammed in the presence of CtIP and then transduced with shCtIP or shNT as indicated. ns, not significant.

(E) Same as in (C) but with cells transduced with shCtIP or shNT after reprogramming.

(F) Copy number variation (CNV) of genomic DNA between iPSCs obtained in the presence or absence of CtIP during cell reprogramming. Two independent experiments (the first and second assays) are shown. CNVs were detected using SurePrint G3 Unrestricted CGH (4 × 180 K) arrays. See also Figure S5.

under CtIP depletion also generated embryonic bodies (EBs) smaller than their respective control cell reprogrammed in the presence of the control shNT, hence showing a clear deficiency in differentiation (Figures 5A–5C). We then tested if, as for self-renewal, this defect was a consequence of problems inherent to reprogramming in a CtIP-defective environment rather than a consequence of the loss of an active role of CtIP in stem cell differentiation. Indeed, CtIP downregulation in already-reprogrammed miPSCs (Figures 5D and 5E) or ES-D3 cells (Figures S5D–S5F) did not affect their differentiation capacity. Collectively, these data confirmed that genomic instability created during cell reprogramming under CtIP deficiency is the main cause of impairment of self-renewal and the differentiation process of miPSCs.

To strengthen our hypothesis, we analyzed copy number variation (CNV) between early passes of shNT- or shCtIP-iPSC clones (two of each). We found a large difference in the number of CNV (>50 CNV) between miPSCs reprogrammed with CtIP depletion and their respective control cells with shNT (Figure 5F).

DISCUSSION

Genomic instability is one of the biggest concerns in the potential clinical use of iPSCs (Rocha et al., 2013). The challenging field of regenerative medicine requires an in-depth understanding of the causes and consequences of genetic abnormalities that arise during the reprogramming process (Studer et al., 2015; Tabar and Studer, 2014; Buganim et al., 2013). Here, we study mouse and human reprogramming to iPSCs, to elucidate the role of DNA end resection as a relevant mechanism for avoiding genomic instability in this process. We demonstrate that expression of the CtIP protein, a key protein in DNA end resection (Sartori et al., 2007), is upregulated during the formation of iPSCs and is required for efficient reprogramming. A CtIP deficiency during reprogramming not only drastically impairs the reprogramming process but also endangers the future of the reprogrammed cells by critically limiting the maintenance of their pluripotency state and their further differentiation to EBs. All of these effects are likely caused by the genomic aberrations acquired by cells during reprogramming. These severe genomic consequences correlate with the well-established roles of CtIP in DNA resection, HR, and DSB repair pathway choice (Sartori et al., 2007; Huertas, 2010; Gomez-Cabello et al., 2013; López-Saavedra et al., 2016), rather than reflecting a novel role of CtIP in the reprogramming process. Intriguingly, and unexpectedly, the drastic effects of these CtIP roles are highly specific to the cell-reprogramming process: CtIP depletion in already-established iPSCs or ES cells does not reduce the ability of the cells to self-renew or differentiate into EBs. We hypothesize that this phenomenon is related to the load of endogenously induced DNA damage in these different situations. Cell reprogramming severely increases replication stress and therefore DNA damage (Ruiz et al., 2015) (see also Figures 1G, 2F, 2D, 3D, and S2A), while endogenous DNA damage in fibroblasts or iPSCs (under normal cell culture conditions) is low. Thus, these differences in the amount of DNA damage could explain why CtIP is essential during iPSC formation but not for maintenance of iPSCs or fibroblasts. Indeed, CtIP is essential for cell viability in cells that have been exposed to mutagens that result in high levels of DNA damage (Sartori et al., 2007; Huertas and Jackson, 2009) or chromosomal aberrations (Huertas and Jackson, 2009). Strikingly, and in agreement with this idea, our data suggest that cell reprogramming in wild-type human and mouse fibroblasts causes an increase in DNA resection that is comparable with high doses of ionizing radiation (Figure S1).

DDR and DNA repair genes have been shown to control genetic stability during cell reprogramming (Lu et al., 2016; Rocha et al., 2013; Hong et al., 2009; Kawamura et al., 2009; Li et al., 2009). Along those lines, we have now demonstrated that DNA end resection, a key process in DSB repair by HR, is hyper-activated in cells undergoing reprogramming compared with the parental somatic cells or already-differentiated cells. During reprogramming, cells not only have an increase in the amount of breaks that are resected, but also a gain in processivity, measured as the length of resected DNA. The most likely explanation for upregulation of DNA end resection is the occurrence of DNA damage during reprogramming. This idea is supported by an increase in the γH2AX marker levels after expression of reprogramming factors, as observed by us and others (Ruiz et al., 2015; Gonzalez et al., 2013; Tilgner et al., 2013). Replication stress, due to the presence of a stalled replication fork, is a major generator of DNA damage, and this is commonly resolved mainly by HR or microhomology-mediated end joining (Petermann and Helleday, 2010; Aguilera and Gomez-Gonzalez, 2008). Hence, the occurrence of DNA damage by replication stress during iPSC development could explain both the upregulation and essential role of DNA end resection in reprogramming.

Here we show that CtIP, a bona fide regulator of DNA end resection (Huertas and Jackson, 2009; Huertas et al., 2008; Sartori et al., 2007), is upregulated during cell reprogramming, and that it is essential for this process. Similarly, CHK1 has been suggested to be pertinent to generating iPSCs (Ruiz et al., 2015). CtIP is known to be a key player in maintaining genomic stability, and we reasoned that CtIP could be appropriately activated to repair DSBs generated by the presence of reprogramming factors. Selective CtIP depletion during mouse and human cell reprogramming interferes with iPSC generation and triggers apoptosis. In agreement with these data, BRCA1, which accelerates DNA resection through its interaction with CtIP (Cruz-Garcia et al., 2014), is also required for successful reprogramming (Gonzalez et al., 2013). In addition, other proteins linked to HR repair, such as BRCA2 and RAD51, are critical for cell reprogramming, at least in mice (Gonzalez et al., 2013).

We also observed that CtIP depletion during cell reprogramming in mouse and human cells hampers the growth and maintenance of their derived iPSCs. As a matter of fact, the human C1 clone obtained from cells bearing an inducible shRNA in the presence of IPTG showed the same long-term problems in viability regardless of the restoration of CtIP expression (data not shown, see Figure 3G for CtIP expression recovery). Likewise, Brca1-deficient MEFs have problems for cell reprogramming, and the derived iPSCs are unable to establish colonies (Gonzalez et al., 2013). This clearly differentiates proteins involved in early steps of HR, such as CtIP and BRCA1 in resection, from those affecting later steps, such as BRCA2 and RAD51, which are in fact dispensable for iPSC colony expansion (Gonzalez et al., 2013). We reasoned that iPSCs are defective for DNA end resection, which is critical for choosing between HR and NHEJ (Gomez-Cabello et al., 2013). As mentioned before, we suggest that both pluripotent and differentiated cells can likely repair their DNA by both NHEJ and HR, such that eliminating one of them only has a mild effect on cell viability unless an exogenous source of damage is present. However, during cell reprogramming, recombination is the only mechanism able to deal with the endogenous damage. We postulate that this is due to the nature of the DNA lesion, as replication stress will readily cause the appearance of one-ended DNA DSBs and will require recombination to restore replication. In the absence of CtIP, those breaks would be erroneously repaired, inducing the observed chromosomal abnormalities and CNV differences (Figures 3H and 5F). Indeed, this seems to be a recurrent mechanism that causes cells undergoing reprogramming to acquire a high degree of chromosomal instability. Even by analyzing cells from a single clone that differentiated in the absence of CtIP (the C1 clone), we can observe that each metaphase has a different number of chromosomes (Figure 3H). Thus, in the absence of CtIP, DSBs created by replication stress would be repaired through more mutagenic repair pathways, thereby increasing mutagenesis and chromosomal rearrangements (Bunting et al., 2010). This is consistent with an increase in DNA damage and genetic instability and would ultimately lead to apoptosis during reprogramming.

We could not expand the only clone we were able to obtain from CtIP-depleted hiPSCs (C1), but did establish a few clones from CtIP-depleted miPSCs, albeit with low efficiency. These clones showed high levels of alterations in chromosome number and CNV with respect to iPSCs generated under normal CtIP levels. These data are consistent with early embryonic lethality (E4.0) observed for *CtIP* knockout mice, which shows a slightly elevated apoptosis (Chen et al., 2005). Although this lethality has been previously associated with the retinoblastoma protein, our data support the idea that the resection function of CtIP is required for embryonic viability (Polato et al., 2014). Curiously, we found that established pluripotent cells, namely miPSCs and ES-D3 cells, did not require CtIP protein (at least not in the absence of an exogenous source of DNA damage). Cell reprogramming seems to be one such source of internal stress, so it is possible that other situations arise during normal embryogenesis in which CtIP, and specifically its resection activity, are essential.

Our data and those from previous reports suggest that cells from patients with deficient DDR and DNA repair pathways could not be efficiently reprogrammed and used in regenerative medicine. However, more in-depth investigations are needed to clarify which genes are specifically required to avoid genomic instability during the reprogramming process, to grant this powerful tool a future as a clinical standard procedure. Indeed, we suggest that, when studying self-renewal and differentiation of iPSCs, it is of capital importance to discriminate between the actual roles of repair proteins on those processes and the inherent, long-term consequences of genomic instability caused by reprogramming in a repair-defective environment.

EXPERIMENTAL PROCEDURES

Cell Cultures

HEK293T, SNL, and C57BL/6 primary MEFs carrying a doxycycline-inducible tetracistronic cassette encoding the four murine reprogramming factors Oct4, Sox2, Klf4, and c-Myc (provided by M. Serrano) (Abad et al., 2013) were grown in DMEM (Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich), 100 units/mL penicillin and 100 g/mL streptomycin (Sigma-Aldrich). HFF cells were growth in DMEM (Gibco) supplemented with 20% FBS (ATCC, LGC Promochem) and 100 units/mL penicillin (Sigma-Aldrich). miPSCs were cultured on gelatin-coated plates with DMEM/F12 + GlutaMAX (Gibco) supplemented with 20% knockout serum replacement (Gibco), 1,000 U/mL LIF (Millipore), 1% non-essential amino acids, 0.1 mM β-mercaptoethanol, and 100 units/mL penicillin (Sigma-Aldrich). hiPSCs were cultured on SNL feeders with DMEM/F12 + GlutaMAX (Gibco) supplemented with 20% knockout serum replacement (Gibco), 10 ng/mL basic fibroblast growth factor (Miltenyi Biotec), 1% non-essential amino acids, 0.1 mM β-mercaptoethanol, and 100 units/mL penicillin (Sigma-Aldrich).

Retroviral and Lentiviral Production

Retroviral and lentiviral particles were produced in HEK293T cells as described previously (Gomez-Cabello et al., 2013) using the plasmids listed in Table S1. Lentiviruses harboring shRNA vectors (Sigma) targeting human CtIP (CAG AAG GAT GAA GGA CAG TTT), mouse CtIP (GCA AGG TTT ACA AGT CAA AGT), and a non-target sequence (GCG CGA TAG CGC TAA TAA TTT) were used.

Mouse and Human iPSC Reprogramming

For miPSCs, MEFs were seeded at 1.5 × 10^5 cells per well of a 12-well plate. MEFs were cultured in miPSC medium supplemented with 1 μg/mL doxycycline to induce the expression of OSKM and promote reprogramming. Medium was changed every 24 hr for 21 days or until iPSCs colonies appeared. iPSCs were then expanded in 6-well gelatin-coated plates and in iPSC medium without doxycycline. For hiPSCs, HFFs with low number of passes were incubated with retroviral supernatants containing OCT4, c-MYC, SOX2, and KLF4 three times for 48 hr each time. Cells were then plated on irradiated (45 Gy) SNL feeders, the medium was changed 1 day later to iPSC medium, and cells were incubated for a further 21–30 days. For IPTG induction reprogramming, HFFs were transduced with an IPTG-inducible shCtIP lentivirus or the respective shNT lentivirus as a control. After 48 hr, iPSC regular medium with 1 mM IPTG (Sigma) was added to start the reprogramming process and maintained during the whole process. Reprogramming efficiency was calculated as the number of colonies normalized to the number of cells seeded.

Single-Molecule Analysis of Resection Tracks

iPSCs and differentiated cells (MEFs and HFFs) were seeded in 6-well plates at the required density to reach 80% confluence at the time of harvest. Cells were grown in the presence of 10 μM BrdU (GE Healthcare) for 24 hr and then harvested. Single-molecule analysis of resection tracks (SMART) was performed as described previously (Cruz-Garcia et al., 2014).

Immunoblotting

Protein extracts were prepared in Laemmli buffer (4% SDS/20% glycerol, 120 mM Tris-HCl [pH 6.8]). Proteins were resolved by SDS-PAGE, transferred to polyvinylidene fluoride (Millipore) membrane and visualized by immunoblotting. Western blot analysis used the antibodies listed in Table S2. Results were visualized and quantified using an Odyssey Infrared Imaging System (LI-COR Biosciences).

Flow-Cytometric Analysis of DNA End Resection

MEF cells, HFF cells, hiPSCs, and miPSCs were prepared for FACS analysis as follows: cells were grown in the presence of 10 μM BrdU (GE Healthcare) for 16–18 hr and then detached using Accutase (eBioscence). Cells grown in absence of BrdU were used as FACS-negative control. Cells were fixed with 4% paraformaldehyde for 10 min at 4°C, permeabilized with 0.1% Triton X-100 in PBS, washed in PBS, and then blocked with 5% FBS in PBS. After blocking, cells were incubated with an anti-BrdU mouse monoclonal antibodies (Table S2) for 1–2 hr at room temperature, and then with the appropriate secondary antibody (Table S2) for 30 min at room temperature. Additional control cells without primary antibody were used to set up FACS conditions. Cells were then washed and resuspended in PBS. Samples were analyzed with a BD FACSCalibur flow cytometer (BD Biosciences, Ref: 342975). At least 10,000 events were recorded for each sample.

Flow-Cytometric Analysis Cell Cycle

Mouse and hiPSCs were grown in 6-well plates. After 2 days, Accutase was added to remove the cells, which were then fixed with 70% ethanol at 4°C for at least 1 day. Cells were then washed and resuspended in PBS. Samples were incubated with 1 mg/mL propidium iodide and 10 mg/mL RNase (Sigma) for 20 min prior to FACS analysis. Samples were analyzed with a BD FACSCalibur flow cytometer (BD Biosciences). At least 10,000 events were recorded for each sample.

Karyotyping

hiPSCs were grown for 48 hr on gelatinized 6-well plates. At 3 hr prior to cell collection, the medium was changed and supplemented with 0.1 μg/mL demecolcine (Sigma, D7385) for 2 hr. Cells were washed, detached with Accutase, and centrifuged at 200 × *g* for 5 min at 4°C. iPSCs were then resuspended in hypotonic KCl solution (0.56%) and incubated at 37°C for 10 min. Cells underwent two rounds of fixation in methanol:glacial acetic acid (3:1) and centrifugation, after which they were resuspended in fixation solution and stained with DAPI. DAPI-stained chromosomes from at least 75 cells were counted for each condition.

EBs

miPSCs growing in gelatinized 6-well plates were detached with Accutase, counted, and replated onto ultra-low attachment 6-well plate with iPSC regular medium without LIF for 3–4 days. EBs were analyzed for size and number through microscopic images using Adobe Photoshop CS6 (Adobe Systems Incorporated).

Array Comparative Genomic Hybridization

For mouse array comparative genomic hybridization, CNV was detected from genomic DNA isolated from iPSC clones and hybridized to SurePrint G3 Human High-Resolution 1 × 1 M Microarrays (CNV) (Agilent Technologies) following manufacturer's instructions. CNV was identified using Agilent CytoGenomics v2.0 analysis software, following ADM-2 algorithm suggested by Agilent Technologies.

All iPSC clones were obtained at early passes after reprogramming and colony selection.

ACCESSION NUMBERS

The accession number for the microarray data reported in this paper is GEO: GSE90888.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2016.12.009.

AUTHOR CONTRIBUTIONS

D.G-C. and C.C-R. generated and characterized mouse and hiPSCs. OSKM-inducible MEFs were generated by M.A. and M.S. D.G.-C. and P.H. supervised the work and wrote the paper.

ACKNOWLEDGMENTS

We wish to thank the cytometer unit of CABIMER for technical support, Diana Aguilar for critical reading of the manuscript, and

Veronica Raker for style corrections. This work was funded by an R + D + I grant from the Spanish Ministry of Economy and Competitivity (SAF2013-43255-P) and an ERC Starting Grant (DSBRECA).

Received: July 14, 2016
Revised: December 7, 2016
Accepted: December 8, 2016
Published: January 5, 2017

REFERENCES

Abad, M., Mosteiro, L., Pantoja, C., Canamero, M., Rayon, T., Ors, I., Grana, O., Megias, D., Dominguez, O., Martinez, D., et al. (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature *502*, 340–345.

Aguilera, A., and Gomez-Gonzalez, B. (2008). Genome instability: a mechanistic view of its causes and consequences. Nat. Rev. Genet. *9*, 204–217.

Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. Nat. Rev. Genet. *14*, 427–439.

Bunting, S.F., Callen, E., Wong, N., Chen, H.T., 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.

Chen, P.L., Liu, F., Cai, S., Lin, X., Li, A., Chen, Y., Gu, B., Lee, E.Y., and Lee, W.H. (2005). Inactivation of CtIP leads to early embryonic lethality mediated by G1 restraint and to tumorigenesis by haploid insufficiency. Mol. Cell. Biol. *25*, 3535–3542.

Cruz-Garcia, A., Lopez-Saavedra, A., and Huertas, P. (2014). BRCA1 accelerates CtIP-mediated DNA-end resection. Cell Rep. *9*, 451–459.

Gomez-Cabello, D., Jimeno, S., Fernandez-Avila, M.J., and Huertas, P. (2013). New tools to study DNA double-strand break repair pathway choice. PLoS One *8*, e77206.

Gonzalez, F., Georgieva, D., Vanoli, F., Shi, Z.D., Stadtfeld, M., Ludwig, T., Jasin, M., and Huangfu, D. (2013). Homologous recombination DNA repair genes play a critical role in reprogramming to a pluripotent state. Cell Rep. *3*, 651–660.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63–67.

Heyer, W.D., Ehmsen, K.T., and Liu, J. (2010). Regulation of homologous recombination in eukaryotes. Annu. Rev. Genet. *44*, 113–139.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132–1135.

Huertas, P. (2010). DNA resection in eukaryotes: deciding how to fix the break. Nat. Struct. Mol. Biol. *17*, 11–16.

Huertas, P., and Jackson, S.P. (2009). Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. J. Biol. Chem. *284*, 9558–9565.

Huertas, P., Cortes-Ledesma, F., Sartori, A.A., Aguilera, A., and Jackson, S.P. (2008). CDK targets Sae2 to control DNA-end resection and homologous recombination. Nature *455*, 689–692.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Izpisua Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature *460*, 1140–1144.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136–1139.

Liang, G., and Zhang, Y. (2013). Genetic and epigenetic variations in iPSCs: potential causes and implications for application. Cell Stem Cell *13*, 149–159.

Lieber, M.R. (2008). The mechanism of human nonhomologous DNA end joining. J. Biol. Chem. *283*, 1–5.

López-Saavedra, A., Gómez-Cabello, D., Domínguez-Sánchez, M.S., Mejías-Navarro, F., Fernández-Ávila, M.J., Dinant, C., Martínez-Macías, M.I., Bartek, J., and Huertas, P. (2016). A genome-wide screening uncovers the role of CCAR2 as an antagonist of DNA end resection. Nat. Commun. *7*, 12364.

Lu, J., Li, H., Baccei, A., Sasaki, T., Gilbert, D.M., and Lerou, P.H. (2016). Influence of ATM-mediated DNA damage response on genomic variation in human induced pluripotent stem cells. Stem Cells Dev. *25*, 740–747.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell Stem Cell *7*, 521–531.

Nakamura, K., Kogame, T., Oshiumi, H., Shinohara, A., Sumitomo, Y., Agama, K., Pommier, Y., Tsutsui, K.M., Tsutsui, K., Hartsuiker, E., et al. (2010). Collaborative action of Brca1 and CtIP in elimination of covalent modifications from double-strand breaks to facilitate subsequent break repair. PLoS Genet. *6*, e1000828.

Petermann, E., and Helleday, T. (2010). Pathways of mammalian replication fork restart. Nat. Rev. Mol. Cell Biol. *11*, 683–687.

Polato, F., Callen, E., Wong, N., Faryabi, R., Bunting, S., Chen, H.T., Kozak, M., Kruhlak, M.J., Reczek, C.R., Lee, W.H., et al. (2014). CtIP-mediated resection is essential for viability and can operate independently of BRCA1. J. Exp. Med. *211*, 1027–1036.

Rocha, C.R., Lerner, L.K., Okamoto, O.K., Marchetto, M.C., and Menck, C.F. (2013). The role of DNA repair in the pluripotency and differentiation of human stem cells. Mutat. Res. *752*, 25–35.

Ruiz, S., Gore, A., Li, Z., Panopoulos, A.D., Montserrat, N., Fung, H.L., Giorgetti, A., Bilic, J., Batchelder, E.M., Zaehres, H., et al. (2013). Analysis of protein-coding mutations in hiPSCs and their possible role during somatic cell reprogramming. Nat. Commun. *4*, 1382.

Ruiz, S., Lopez-Contreras, A.J., Gabut, M., Marion, R.M., Gutierrez-Martinez, P., Bua, S., Ramirez, O., Olalde, I., Rodrigo-Perez, S., Li, H., et al. (2015). Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. Nat. Commun. *6*, 8036.

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.

Studer, L., Vera, E., and Cornacchia, D. (2015). Programming and reprogramming cellular age in the era of induced pluripotency. Cell Stem Cell *16*, 591–600.

Tabar, V., and Studer, L. (2014). Pluripotent stem cells in regenerative medicine: challenges and recent progress. Nat. Rev. Genet. *15*, 82–92.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Tilgner, K., Neganova, I., Moreno-Gimeno, I., Al-Aama, J.Y., Burks, D., Yung, S., Singhapol, C., Saretzki, G., Evans, J., Gorbunova, V., et al. (2013). A human iPSC model of Ligase IV deficiency reveals an important role for NHEJ-mediated-DSB repair in the survival and genomic stability of induced pluripotent stem cells and emerging haematopoietic progenitors. Cell Death Differ. *20*, 1089–1100.

Wang, H., Shi, L.Z., Wong, C.C., Han, X., Hwang, P.Y., Truong, L.N., Zhu, Q., Shao, Z., Chen, D.J., Berns, M.W., et al. (2013). The interaction of CtIP and Nbs1 connects CDK and ATM to regulate HR-mediated double-strand break repair. PLoS Genet. *9*, e1003277.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

Please cite this article in press as: Jachimowicz et al., UBQLN4 Represses Homologous Recombination and Is Overexpressed in Aggressive Tumors, Cell (2019), https://doi.org/10.1016/j.cell.2018.11.024

UBQLN4 Represses Homologous Recombination and Is Overexpressed in Aggressive Tumors

Ron D. Jachimowicz,[1,2,3,*] Filippo Beleggia,[3,5,18] Jörg Isensee,[6,18] Bhagya Bhavana Velpula,[1,2,18] Jonas Goergens,[3] Matias A. Bustos,[7] Markus A. Doll,[4,8] Anjana Shenoy,[2] Cintia Checa-Rodriguez,[9] Janica Lea Wiederstein,[4] Keren Baranes-Bachar,[1,2] Christoph Bartenhagen,[10,11] Falk Hertwig,[12,13,14] Nizan Teper,[1,2] Tomohiko Nishi,[7] Anna Schmitt,[3] Felix Distelmaier,[15] Hermann-Josef Lüdecke,[5,16] Beate Albrecht,[16] Marcus Krüger,[4,11] Björn Schumacher,[4,8] Tamar Geiger,[2] Dave S.B. Hoon,[7] Pablo Huertas,[9] Matthias Fischer,[10,11] Tim Hucho,[6] Martin Peifer,[11,17] Yael Ziv,[1,2,19,*] H. Christian Reinhardt,[3,4,11,19,20,*] Dagmar Wieczorek,[5,16,19,*] and Yosef Shiloh[1,2,19,*]

[1]The David and Inez Myers Laboratory for Cancer Genetics, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel
[2]Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel
[3]Clinic I of Internal Medicine, University Hospital Cologne, Cologne 50931, Germany
[4]Cologne Excellence Cluster on Cellular Stress Response in Aging-Associated Diseases, University of Cologne, Cologne, Germany
[5]Institute of Human Genetics, Heinrich-Heine-University, Düsseldorf, Germany
[6]Department of Anesthesiology and Intensive Care Medicine, Experimental Anesthesiology and Pain Research, University Hospital Cologne, Cologne 50931, Germany
[7]Department of Translational Molecular Medicine, Division of Molecular Oncology, John Wayne Cancer Institute at Providence Saint John's Health Center, Santa Monica, CA, USA
[8]Institute for Genome Stability in Aging, Cologne, Germany
[9]Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER, Universidad de Sevilla-CSIC-Universidad Pablo de Olavide and Department of Genetics, University of Sevilla, Sevilla 41092, Spain
[10]Department of Experimental Pediatric Oncology, University Hospital Cologne, Cologne, Germany
[11]Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany
[12]Department of Pediatric Oncology and Hematology, Charité, Berlin, Germany
[13]German Cancer Consortium, Germany
[14]Berlin Institute of Health, Germany
[15]Department of General Pediatrics, Neonatology and Pediatric Cardiology, University Hospital, Heinrich-Heine-University, Düsseldorf 40225, Germany
[16]Institute of Human Genetics, University Clinic Duisburg-Essen, Essen, Germany
[17]Department of Translational Genomics, University of Cologne, Cologne, Germany
[18]These authors contributed equally
[19]These authors contributed equally
[20]Lead contact
*Correspondence: ron.jachimowicz@uk-koeln.de (R.D.J.), yaelz@post.tau.ac.il (Y.Z.), christian.reinhardt@uk-koeln.de (H.C.R.), dagmar.wieczorek@hhu.de (D.W.), yossih@post.tau.ac.il (Y.S.)
https://doi.org/10.1016/j.cell.2018.11.024

SUMMARY

Genomic instability can be a hallmark of both human genetic disease and cancer. We identify a deleterious *UBQLN4* mutation in families with an autosomal recessive syndrome reminiscent of genome instability disorders. UBQLN4 deficiency leads to increased sensitivity to genotoxic stress and delayed DNA double-strand break (DSB) repair. The proteasomal shuttle factor UBQLN4 is phosphorylated by ATM and interacts with ubiquitylated MRE11 to mediate early steps of homologous recombination-mediated DSB repair (HRR). Loss of UBQLN4 leads to chromatin retention of MRE11, promoting non-physiological HRR activity *in vitro* and *in vivo*. Conversely, UBQLN4 overexpression represses HRR and favors non-homologous end joining. Moreover, we find UBQLN4 overexpressed in aggressive tumors. In line with an HRR defect in these tumors, UBQLN4 overexpression is associated with PARP1 inhibitor sensitivity. UBQLN4 therefore curtails HRR activity through removal of MRE11 from damaged chromatin and thus offers a therapeutic window for PARP1 inhibitor treatment in UBQLN4-overexpressing tumors.

INTRODUCTION

In response to genotoxic stress, cells activate a signaling network, collectively referred to as the DNA damage response (DDR). The DDR activates cell-cycle checkpoints, DNA repair pathways, and, if damage is beyond repair capacity, triggers cell-death pathways (Reinhardt and Yaffe, 2013). Following DNA double-strand breaks (DSBs), the DDR is primarily activated by the proximal kinase ataxia telangiectasia mutated (ATM), which phosphorylates a plethora of substrates, such as KAP-1, CHK2, p53, MRE11, RAD50, NBS1, and others (Shiloh and Ziv, 2013).

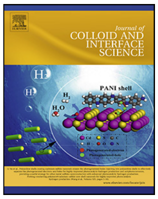
The main DSB repair pathways are the canonical non-homologous end joining (c-NHEJ) pathway and homologous

Journal of Colloid and Interface Science 521 (2018) 197–205

Contents lists available at ScienceDirect

Journal of Colloid and Interface Science

journal homepage: www.elsevier.com/locate/jcis



Regular Article

Importance of hydrophobic interactions in the single-chained cationic surfactant-DNA complexation

Manuel López-López [b], Pilar López-Cornejo [a], Victoria Isabel Martín [a], Francisco José Ostos [a], Cintia Checa-Rodríguez [c], Rosario Prados-Carvajal [c], José Antonio Lebrón [a], Pablo Huertas [c], María Luisa Moyá [a],*

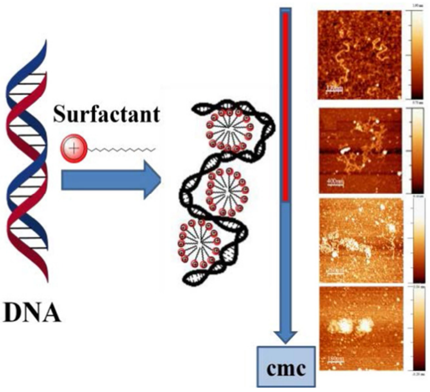
[a] *Department of Physical Chemistry, University of Seville, C/ Profesor García González 1, 41012 Seville, Spain*

[b] *Department of Chemical Engineering, Physical Chemistry and Materials Science, Faculty of Experimental Sciences, Campus de El Carmen, Avda. de las Fuerzas Armadas s/n, 21071 Huelva, Spain*

[c] *Department of Genetics, University of Seville and Andalusian Center for Molecular Biology and Regenerative Medicine-CABIMER, University of Seville-CSIC-University Pablo de Olavide, 41092, Spain*

GRAPHICAL ABSTRACT

A decrease in the critical micelle concentration favors the surfactant-DNA complexation.



A decrease in the critical micelle concentration favors the surfactant-DNA complexation

ARTICLE INFO

Article history:
Received 26 January 2018
Revised 10 March 2018
Accepted 14 March 2018
Available online 15 March 2018

Keywords:
DNA
Single-chained cationic surfactants
Complexation

ABSTRACT

The goal of this work was to understand the key factors determining the DNA compacting capacity of single-chained cationic surfactants. Fluorescence, zeta potential, circular dichroism, gel electrophoresis and AFM measurements were carried out in order to study the condensation of the nucleic acid resulting from the formation of the surfactant-DNA complexes. The apparent equilibrium binding constant of the surfactants to the nucleic acid, K_{app}, estimated from the experimental results obtained in the ethidium bromide competitive binding experiments, can be considered directly related to the ability of a given surfactant as a DNA compacting agent. The plot of $\ln(K_{app})$ vs. $\ln(\mathrm{cmc})$, cmc being the critical micelle concentration, for all the bromide and chloride surfactants studied, was found to be a reasonably good linear

---

* Corresponding author.
E-mail address: moya@us.es (M.L. Moyá).

https://doi.org/10.1016/j.jcis.2018.03.048
0021-9797/© 2018 Elsevier Inc. All rights reserved.

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-017-00183-6 OPEN

DNA end resection requires constitutive sumoylation of CtIP by CBX4

Isabel Soria-Bretones[1,2], Cristina Cepeda-García[2], Cintia Checa-Rodriguez[1,2], Vincent Heyer[3,4,5,6], Bernardo Reina-San-Martin[3,4,5,6], Evi Soutoglou[3,4,5,6] & Pablo Huertas[1,2]

DNA breaks are complex DNA lesions that can be repaired by two alternative mechanisms: non-homologous end-joining and homologous recombination. The decision between them depends on the activation of the DNA resection machinery, which blocks non-homologous end-joining and stimulates recombination. On the other hand, post-translational modifications play a critical role in DNA repair. We have found that the SUMO E3 ligase CBX4 controls resection through the key factor CtIP. Indeed, CBX4 depletion impairs CtIP constitutive sumoylation and DNA end processing. Importantly, mutating lysine 896 in CtIP recapitulates the CBX4-depletion phenotype, blocks homologous recombination and increases genomic instability. Artificial fusion of CtIP and SUMO suppresses the effects of both the non-sumoylatable CtIP mutant and CBX4 depletion. Mechanistically, CtIP sumoylation is essential for its recruitment to damaged DNA. In summary, sumoylation of CtIP at lysine 896 defines a subpopulation of the protein that is involved in DNA resection and recombination.

---

[1] Departamento de Genética, Universidad de Sevilla, Sevilla 41080, Spain. [2] Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER, Universidad de Sevilla-CSIC-Universidad Pablo de Olavide, Sevilla 41092, Spain. [3] Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch 67404, France. [4] Institut National de la Santé et de la Recherche Médicale U964, Illkirch 67404, France. [5] Centre National de Recherche Scientifique UMR7104, Illkirch 67404, France. [6] Université de Strasbourg, Illkirch 67081, France. Correspondence and requests for materials should be addressed to P.H. (email: pablo.huertas@cabimer.es)

www.impactjournals.com/oncotarget/ Oncotarget, 2017, Vol. 8, (No. 16), pp: 27380-27392

Research Paper

Targeting the centriolar replication factor STIL synergizes with DNA damaging agents for treatment of ovarian cancer

Noa Rabinowicz[1,2], Lingegowda S. Mangala[3,4], Kevin R. Brown[5], Cintia Checa-Rodriguez[6], Asher Castiel[1,2], Oren Moskovich[1,2], Giulia Zarfati[1], Luba Trakhtenbrot[1], Adva Levy-Barda[1], Dahai Jiang[3,4], Cristian Rodriguez-Aguayo[4,7], Sunila Pradeep[3], Yael van Praag[1], Gabriel Lopez-Berestein[4,7], Ahuvit David[1,2], Ilya Novikov[9], Pablo Huertas[6], Robert Rottapel[10], Anil K. Sood[3,4,8], Shai Izraeli[1,2,11]

[1] Cancer Research Center, Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel

[2] Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[3] Department of Gynecologic Oncology, MD Anderson Cancer Center, Houston, Texas, USA

[4] Center for RNA Interference and Non-Coding RNA, MD Anderson Cancer Center, Houston, Texas, USA

[5] Donnelly Centre and The Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, Canada

[6] Department of Genetics, University of Sevilla and Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER), Sevilla, Spain

[7] Department of Experimental Therapeutics, MD Anderson Cancer Center, Houston, Texas, USA

[8] Department of Cancer Biology, MD Anderson Cancer Center, Houston, Texas, USA

[9] Biostatistical Unit, Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel

[10] Princess Margaret Cancer Center, University Health Network, Toronto, Ontario, Canada

[11] The Gene Development and Environment Pediatric Research Institute, Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel

Correspondence to: *Shai Izraeli,* ***email:*** *sizraeli@sheba.health.gov.il*

Keywords: *STIL, centrosomes, ovarian cancer, DNA damage, genomic instability*

Received: *August 11, 2016* ***Accepted:*** *February 20, 2017* ***Published:*** *March 10, 2017*

Copyright: Rabinowicz et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

ABSTRACT

Advanced ovarian cancer is an incurable disease. Thus, novel therapies are required. We wished to identify new therapeutic targets for ovarian cancer. ShRNA screen performed in 42 ovarian cancer cell lines identified the centriolar replication factor STIL as an essential gene for ovarian cancer cells. This was verified *in-vivo* in orthotopic human ovarian cancer mouse models. STIL depletion by administration of siRNA in neutral liposomes resulted in robust anti-tumor effect that was further enhanced in combination with cisplatin. Consistent with this finding, STIL depletion enhanced the extent of DNA double strand breaks caused by DNA damaging agents. This was associated with centrosomal depletion, ongoing genomic instability and enhanced formation of micronuclei. Interestingly, the ongoing DNA damage was not associated with reduced DNA repair. Indeed, we observed that depletion of STIL enhanced canonical homologous recombination repair and increased BRCA1 and RAD51 foci in response to DNA double strand breaks. Thus, inhibition of STIL significantly enhances the efficacy of DNA damaging chemotherapeutic drugs in treatment of ovarian cancer.