INTERNATIONAL DOCTORAL SCHOOL OF THE USC

Lucía
Ramos Lage

PhD Thesis

Functional study of GEN1 in safeguarding genome stability and maintenance of embryonic stem cells

Santiago de Compostela, 2025

Doctoral Programme in Molecular Medicine

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

DOCTORAL THESIS

FUNCTIONAL STUDY OF GEN1 IN SAFEGUARDING GENOME STABILITY AND MAINTENANCE OF EMBRYONIC STEM CELLS

Lucía Ramos Lage

Supervisor/s: Diana Guallar Artal and Miguel Ángel Fidalgo Pérez

Tutora: Diana Guallar Artal

PHD PROGRAMME IN MOLECULAR MEDICINE

SANTIAGO DE COMPOSTELA - 2025

DECLARATION OF CONFLICT OF INTEREST AND AUTHORSHIP OF THE PRESENTED DATA

Lucía Ramos Lage declares that she has no conflict of interest in relation to this doctoral thesis.

The doctoral candidate hereby declares full authorship of all images and figures presented in the thesis.

Santiago de Compostela, 11th December 2025

During my doctoral research, funding was provided through grants awarded to my supervisors, Dr Diana Guallar by MICIN / AEI / FEDER (PID2022-136608OB-I00), Ramón Areces (2022-PO062) and Xunta de Galicia (ED431C2023/28, ED431C 2023/28 and ED481A-2023-138) grants and Dr Miguel Fidalgo by MICIN / AEI / FEDER (PID2019-105739GB-I00 and PID2022-143105NB-I00). Dr Diana Guallar is beneficiary of a Ramón y Cajal Award (RYC2019-027305-I).

Additionally, I, Lucía Ramos Lage, was supported by a predoctoral fellowship from the Xunta de Galicia (ED481A-2023-138).

Table of contents

1. Acknowledgements

É de ben sabido, ser agradecido…

Primero, quiero dar gracias a mis directores de tesis, Diana y Miguel. Diana, muchas gracias por abrirme las puertas del laboratorio, real y metafóricamente. Nunca pensé que esa primera entrevista para hacer un TFM a través de una videollamada nos llevaría a compartir unos 5 años de mentoría. Gracias por los IM interminables, por los mil avisos para poner controles en los experimentos, por las mini broncas para que limpiase mi bench, por las correcciones en figuras que siempre decías, bueno esto lo haces tú y siempre acababas haciendo tú misma en el momento, por apoyarme cuando tenía alguna idea que podría ser factible y avisarme cuando eran de bombero. En serio, muchas gracias por todo, me has visto crecer y has contribuido a que sea la científica que soy ahora, gracias de corazón.

Miguel, muchas gracias por también abrirme las puertas del laboratorio. Gracias a las reuniones compartidas en las que siempre salía aprendiendo algo nuevo o por lo menos a cuestionarme los por qués de todos mis experimentos. Gracias por todo el conocimiento compartido, tanto a nivel experimental como intelectual que han contribuido a mi formación.

Quiero agradecer obviamente a mis compañeros de laboratorio, que mucho han tenido que aguantar. A mis chicas de P1L3, tanto las que están como las que ya no están. CP, gracias por ser mi entrenadora Pokémon. De ti aprendí como NO se debe hacer un sándwich para un WB, pero también muchas de las técnicas básicas del labo. Gracias por creer en mi durante todos estos años cuando ni yo lo hacía y por escucharme incluso cuando ya no estabas aquí. Alba, muchas gracias por estar a mi lado, literalmente, durante toda esta etapa y ser otro punto de apoyo más en ella. Raquel, última adición de P1L3, gracias por todos los ánimos, risas e himnos creados en el laboratorio que han hecho que este último año no se hiciese tan cuesta arriba. Helena, entramos casi al mismo tiempo en el grupo y, aun así, para mí, siempre serás mi senior. Muchísimas gracias por todo en general, tanto ayudándome con experimentos como dando ánimos para no rendirme, gracias por acompañarme a cada paso de esta aventura. Raquel y Helena, escribir una tesis es mucho más divertido con vosotras a mi lado, gracias por el tiempo compartido y por compartir. Adrián, gracias por tu apoyo durante esos primeros pasos de mi tesis, rompías nuestro squad de chicas, pero sin ti P1L3 no sería lo mismo ni es lo mismo ahora. Muchísimas gracias, chicas de P1L3, sin vuestra ayuda probablemente no estaría entregando esta tesis ni sería la persona que soy ahora, os quiero.

Gracias a mis compañeros de P1L4, sin los cuales no habría un P1L3. Cris y Álex, gracias por acogerme en el labo y ayudarme cuando lo he necesitado. David, muchas gracias por la ayuda, el apoyo y las conversaciones sobre política y demás temas random que hicieron estos años más llevaderos. Celia y Lucía, gracias por estar ahí y escuchar y aguantarme. Tiago, nuestro portugués favorito, muchas gracias por los chistes malos que no hacían gracia y perturbaban, pero sobre todo por estar ahí e incluso compartir habitación de hotel. Vera, muchas gracias por todo, por ese protocolo de fraccionamiento celular, entre otras muchas cosas. Gracias por el apoyo tanto dentro como fuera del laboratorio, lágrimas compartidas dentro de los baños del CiMUS y muchísimas risas y aventuras que hicieron de la tesis también una época preciosa. Contigo aprendí muchísimo y no solo de ciencia, sino de la vida y estoy muy agradecida de poder seguir contando contigo. Adriana, muchas gracias, no solo me has apoyado durante la tesis, sino que también me has tenido que aguantar en casa. Gracias por compartir Abastos conmigo, ser la mejor pinche del mundo y soportar mis manías. Gracias por animarme cuando tengo un mal día, cantar y bailar conmigo cuando es uno bueno y apuntarte sin miramientos a tomarte una caña conmigo. Vera y Adriana, ambas sois pilares fundamentales de estos años y vaya tres patas para un banco somos, pero no puedo estar más contenta de que nuestros caminos se hayan cruzado, os quiero.

I also want to thank Dr Evi Soutoglou, my supervisor during my stay in Brighton. Thank you for welcoming me into your lab. I learnt a lot during those three months, and it was a pleasure to be surrounded by such great scientists. También gracias a Raquel y Diana, sin vosotras dos mi estancia no hubiese sido ni la mitad de productiva ni la mitad de interesante ni la mitad de divertida. Fuisteis esenciales para que aprendiese muchísimo en un periodo tan corto y que me lo pasase tan bien y sentirme tan acogida, muchas gracias de verdad.

Gracias a la Dra. Aurora Gómez Durán y, sobre todo, gracias a Diego por enseñarme y ayudarme a realizar algunos experimentos que hoy van incluidos en esta tesis. Además, gracias a toda aquella persona que no haya mencionado, pero haya colaborado de alguna manera a que este trabajo se haya podido realizar.

A pesar de no estar conmigo en el laboratorio todos los días, hay personas que se merecen un agradecimiento enorme por aguantarme. Gracias a mis chiquis de la uni: Ainoa, Deborah, Erik, Mónica, Sarah y Xènia. Me visteis crecer y hacer mis primeros pinitos en el labo, compañeros no sólo de carrera sino de la vida. Muchísimas gracias por todos los años en

Barcelona, por todas las aventuras vividas y por vivir. Vernos una vez al año no es suficiente, pero es mejor que nada y de todos modos siempre os llevo en el corazón, us estimo.

Thank you, Rachael, for your support. Vienna was the first place where I did some real work in a lab, but also the place that holds so many memories for both of us. Thank you so much for watching my vlogs and understanding me through all these years. I'm so happy that we can still call each other friends, even when we live miles apart. Love you lots.

Grazas aos de sempre. Adrián, Adriana e Kati, coñecémonos dende fai moito, fixémonos maiores, pero o que nunca cambia é que seguides estando ahí para apoiarme, nas boas e nas malas. Grazas por soportar os meus monólogos falando da miña vida no labo e por aguantarme na vida en xeral. Sodes moi importantes para min e estou moi agradecida de poder camiñar con vós ao meu lado, quérovos, aínda que non o digamos moito.

Por último e, probablemente, as persoas máis importantes sen as que esta tese non sería posible, quero agradecer a meus pais e a miña irmá. María, gracias por ser miña irmá e apoiar cada decisión, por moi malas ou boas que sexan, pero sempre estando no meu equipo. Espero a túa visita alá onde vaia. Papá e mamá, moitas gracias por absolutamente todo, sen vós isto non sería posible, literalmente. Grazas por facer realidade cada unha das miñas ideas, dende deixarme ir estudiar con 18 á outra punta de España a marchar de Erasmus aínda máis lonxe. Durante estes anos deime de conta do moito que traballastes para que eu puidera cumprir aquelo que me propuxen. Creo que tería que vivir moito máis de 100 anos para devolvervos todo o que fixestes por min e xa sabedes que igual é mellor que non dure tanto por se se me vai demasiado a pinza. Non importa a onde vaia, se a casa de baixo ou a Australia, sempre estaredes ao meu lado e sei que Cuns sempre será a miña casa. Estou moi agradecida de tervos como pais e de toda as ensinanzas que me dades, sempre serei a vosa cativa. Son quen son grazas a vós, quérovos moito.

2. Abstract

Embryonic stem cells (ESCs) possess unique properties of self-renewal and pluripotency, which depend on the maintenance of a highly stable genome. Due to their rapid proliferation and high levels of replication stress, ESCs require robust DNA repair mechanisms to safeguard genomic integrity. Structure-selective endonucleases (SSEs) play key roles in resolving complex DNA intermediates that arise during replication and recombination. Among these, GEN1 is primarily recognised as a Holliday Junction (HJ) resolvase. However, its specific contribution to genome maintenance in pluripotent cells remains unclear.

In contrast to its more secondary role in other cell types, this thesis identifies GEN1 as an essential factor for ESC survival and genomic stability. Complete loss of GEN1 compromises ESC viability, pointing out its critical function in preserving genome integrity under basal conditions in pluripotent stem cells. *Gen1*-deficient ESCs exhibit persistent DNA damage signalling without a concomitant increase in total DNA breaks, suggesting accumulation of unresolved replication or recombination intermediates rather than overt strand breaks. Cytogenetic analyses reveal increased chromosomal fusions, indicating that GEN1 is necessary to maintain chromosomal integrity through processing replication-associated DNA structures.

Colony formation assays indicate that both catalytically active and inactive forms of GEN1 can partially restore ESC colony maintenance, suggesting that GEN1 may exert additional functions beyond its nuclease activity. These findings raise the possibility that GEN1 contributes to cell survival through structural or regulatory roles, although further studies are needed to define the underlying mechanisms.

Together, these findings position GEN1 as a possible guardian of genome stability in embryonic stem cells. Beyond its canonical activity as a HJ resolvase, GEN1 performs broader roles in replication stress management and the preservation of chromosomal structure. This multifaceted function ensures faithful DNA replication and genomic integrity, which are critical for maintaining pluripotency and enabling the safe application of pluripotent cells in regenerative medicine and developmental biology. Future studies delineating the interaction network of GEN1 and its substrates will provide deeper insights into the mechanisms safeguarding ESC genomes.

3. Extended summary in Galician

3.1. RESUMO

As células nai embrionarias (ESCs, do inglés *embryonic stem cells*) posúen propiedades únicas de autorrenovación e pluripotencia, as cales dependen do mantemento dun xenoma altamente estable. Debido á súa rápida proliferación e aos elevados niveis de estrés replicativo, as ESCs requiren mecanismos robustos de reparación do ADN para garantir a integridade xenómica. As endonucleasas específicas de estrutura (SSEs, do inglés *structure specific endonucleases*) desempeñan papeis fundamentais na resolución de intermedios complexos de ADN que se xeran durante os procesos de replicación e recombinación. Entre elas, GEN1 é recoñecida principalmente como unha resolvasa de unións de Holliday. Non obstante, a súa contribución específica ao mantemento do xenoma en células pluripotentes segue sen coñecerse.

En contraste co seu papel secundario noutros tipos celulares, esta tese identifica GEN1 como un factor esencial para a supervivencia e a estabilidade xenómica das ESCs. A perda completa de GEN1 compromete a viabilidade destas células, poñendo de manifesto a súa función crítica na preservación da integridade xenómica en condicións basais. As ESCs deficientes en *Gen1* presentan unha sinalización persistente de dano no ADN sen un incremento paralelo no número total de roturas, o que suxire a acumulación de intermedios de replicación ou recombinación non resoltos, máis ca roturas evidentes das cadeas. As análises citoxenéticas revelan un incremento nas fusións cromosómicas, indicando que GEN1 é necesaria para manter a integridade cromosómica probablemente mediante o procesamento de estruturas de ADN asociadas á replicación e recombinación.

Os ensaios de formación de colonias amosan que tanto as formas cataliticamente activas como as inactivas de GEN1 poden restaurar parcialmente o mantemento de colonias de ESCs, suxerindo que GEN1 podería exercer funcións adicionais máis alá da súa actividade como nucleasa. Estes resultados abren a posibilidade de que GEN1 contribúa á supervivencia celular mediante papeis estruturais ou reguladores, aínda que son necesarios máis estudos para definir os mecanismos subxacentes.

En conxunto, estes resultados sitúan GEN1 como un posible gardián central da estabilidade xenómica nas células nai embrionarias. Alén da súa actividade canónica como resolvasa de unións de Holliday, GEN1 parece desempeñar funcións máis amplas na xestión do estrés replicativo e na preservación da estrutura cromosómica. Esta función multifacética garante unha replicación fiel do ADN e o mantemento da integridade xenómica, aspectos fundamentais

para conservar a pluripotencia e permitir a aplicación segura das ESCs na medicina rexenerativa e na bioloxía do desenvolvemento. Estudos futuros orientados a delimitar a rede de interaccións de GEN1 e os seus substratos estruturais ofrecerán unha comprensión máis profunda dos mecanismos que protexen os xenomas das ESCs.

3.2. INTRODUCIÓN

A pluripotencia defínese como a capacidade dunha célula para diferenciarse nos tres linaxes xerminais embrionarios que dan orixe a un organismo completo. Este estado aparece de maneira transitoria durante o desenvolvemento embrionario temperán, pero pode reproducirse *in vitro* mediante a derivación de células da masa celular interna do blastocisto (E3.5), coñecidas como células nai embrionarias (ESCs). Estas células caracterízanse por dúas propiedades fundamentais: a autorrenovación ilimitada e a pluripotencia, é dicir, a capacidade de xerar calquera tipo celular.

As ESCs representan un recurso clave para a medicina rexenerativa debido á súa capacidade de diferenciarse en múltiples tecidos e á súa utilidade na modelaxe de enfermidades e no cribado farmacolóxico. Non obstante, o uso de material embrionario suscita preocupacións éticas, o que motivou o desenvolvemento das células nai pluripotentes inducidas (iPSCs, do inglés *induced pluripotent stem cells*), obtidas por reprogramación de células somáticas mediante a expresión dos factores OCT4, SOX2, KLF4 e c-MYC (OSKM). Estas células, case indistinguibles das ESCs, ofrecen unha alternativa ética e personalizada para terapias rexenerativas e estudos de medicina de precisión.

En cultivo, as ESCs poden adoptar distintos estados de pluripotencia: *naïve*, *metastable* e *primed*. Estes reproducen etapas concretas do desenvolvemento embrionario no rato. O estado *naïve*, equivalente ao epiblasto preimplantacional, mantense mediante medio libre de soro suplementado con LIF e inhibidores de MEK e GSK3 (2iL). O estado *primed*, correspondente ao epiblasto posimplantacional, require FGF2 e activina A, mentres que o estado *metastable*, mantido en medio con soro e LIF, presenta unha poboación heteroxénea que oscila entre os dous anteriores. Estes estados difiren tamén en epixenética, metabolismo e potencial de desenvolvemento, reflectindo a natureza dinámica da pluripotencia.

O mantemento deste estado depende dunha rede de regulación transcricional central composta por OCT4, SOX2 e NANOG. Estes tres factores interaccionan de forma autoreguladora e cooperan para activar xenes de pluripotencia e inhibir programas de diferenciación. O equilibrio preciso na súa expresión garante a autorrenovación e a identidade pluripotente das ESCs.

As ESCs posúen ademais un ciclo celular único, caracterizado pola súa brevidade e pola redución das fases G1 e G2, o que lles permite duplicarse en menos de 12 horas. Esta arquitectura deriva dunha activación constitutiva de CDKs e dunha escaseza de inhibidores, provocando a ausencia dun control estrito no punto de verificación G1/S. Aínda así, conservan un punto de control funcional en G2/M que prevén a segregación de ADN danado. A aceleración do ciclo celular xera estrés replicativo e risco de roturas de dobre cadea (DSBs, do inglés *double-strand breaks*), mais as ESCs posúen mecanismos de reparación moi eficientes que aseguran unha baixa taxa de mutacións.

As roturas de dobre cadea poden resolverse por unión de extremos non homólogos (NHEJ, do inglés *non homologous end joining*) ou por recombinación homóloga (HR, do inglés *homologous recombination*). Esta última é esencial en ESCs, pois garante reparacións sen erros durante a replicación e mantén a integridade xenómica. Factores clave como RAD51, BRCA1/2, PALB2 e o complexo MRN son imprescindibles para a viabilidade e estabilidade destas células.

Nos estadios finais da HR, pódese producir a formación de estructuras secundaria denominadas unións de Holliday (HJ, do inglés *Holliday junctions*) que presentan unha ameaza para a estabilidade xenómica. Por iso deben ser correctamente procesadas por disolución (complexo BLM–TOP3A–RMI1/2) ou resolución mediante endonucleasas como MUS81–EME1 ou GEN1. A acción coordinada destas vías evita aberracións cromosómicas e garante unha replicación segura. A pesar dos avances existentes, os papeis de BLM, MUS81 e GEN1 en ESCs seguen a estar pouco explorados. Comprender a súa contribución é esencial para desentrañar os mecanismos que manteñen a estabilidade xenómica nas células pluripotentes e asegurar o seu uso seguro na biomedicina rexenerativa.

3.3. OBXECTIVOS E HIPÓTESE

Tendo en conta que as endonucleases específicas de estrutura (SSEs) e a helicase BLM desempeñan un papel fundamental na reparación dos intermediarios de ADN durante a reparación mediada por recombinación homóloga, e que as células pluripotentes presentan unha capacidade extraordinaria para reparar o ADN, ***formulamos a hipótese de que as SSEs desempeñan un papel relevante no mantemento da pluripotencia***.

Para poñer a proba esta hipótese, propóñese analizar o modo de acción molecular de SSEs específicas, MUS81 e GEN1, así como da disolvasa BLM, que son esenciais na resolución das unións de Holliday durante o mantemento das células pluripotentes. Para iso, abordaranse os seguintes obxectivos:

1. Determinar a contribución das SSEs á supervivencia e á estabilidade xenómica das ESCs.
2. Investigar posibles funcións non canónicas de GEN1 nas ESCs.
3. Analizar a implicación de GEN1 nas estruturas de ADN asociadas á replicación.

O estudo da relevancia das SSEs no mantemento das células pluripotentes constitúe unha oportunidade sen precedentes non só para identificar novos factores implicados na regulación da pluripotencia, senón tamén para descubrir novas vías de control da integridade xenómica. Estes coñecementos poderían, ademais, ser modulados para optimizar a tecnoloxía das iPSCs, achegándoa ao seu uso clínico efectivo.

3.4. MATERIAIS E MÉTODOS

Os métodos e materiais empregados nesta investigación inclúen unha ampla variedade de reactivos, ferramentas moleculares e modelos experimentais destinados a estudar o papel das SSEs, en particular GEN1 e MUS81, e BLM nas ESCs. As liñas celulares utilizadas mantivéronse en condicións estándar de cultivo, e realizáronse ensaios de xenotipaxe mediante illamento de ADN xenómico, PCR e secuenciación Sanger para confirmar as modificacións introducidas por edición xenética.

Para a manipulación xenética das células, empregouse o sistema CRISPR/Cas9, deseñando gRNAs específicos para *Blm*, *Gen1* e *Mus81*, así como plásmidos doadores para a introdución de etiquetas epítopo e mutantes catalíticos. Ademais, producíronse partículas lentivirais para a

expresión estable de plásmidos para o silenciamento xénico mediante shRNAs, xunto coa utilización do sistema *PiggyBac* para expresión transxénica. Para a obtención dos plásmidos purificados levouse a cabo a transformación en bacterias dos mesmos e posterior e purificación.

O estudo incluíu tamén a transfección de liñas mamíferas, ensaios de formación de colonias e tinguidura de fosfatasa alcalina para avaliar a capacidade de autorrenovación e diferenciación das ESCs. A análise do ciclo celular e da viabilidade realizouse mediante citometría de fluxo, complementada con ensaios de apoptose (Anexina V/ ioduro de propidio). Para estudar o dano e a estabilidade xenómica, leváronse a cabo ensaios de cometa, detección de micronúcleos e análises citoxenéticas.

A extracción de ARN total e a análise da expresión xénica mediante RT-qPCR permitiron cuantificar os niveis de transcrición dos xenes de interese, mentres que os extractos proteicos, tanto totais como fraccionados, foron avaliados por Western blot e inmunofluorescencia. Finalmente, os estudos de microscopía confocal, xunto coa análise bioinformática de datos derivados de expresión e interacción proteica, contribuíron a esclarecer o papel de GEN1 no mantemento da estabilidade xenómica e da pluripotencia nas ESCs.

3.5. RESULTADOS

Expresión elevada de xenes de recombinación homóloga en ESCs

As células nai embrionarias (ESCs) presentan unha expresión significativamente maior de xenes implicados na reparación de ADN mediante recombinación homóloga (HR) en comparación con células diferenciadas. Isto foi confirmado mediante análise de datos xa publicados de RNA-seq de ESCs e fibroblastos embrionarios (MEFs, do inglés *mouse embryonic fibroblasts*), estes mostran un incremento notábel de xenes clave de HR como *Mre11a*, *Brca1*, *Brca2* e *Rad51*. Pola contra, os reguladores de NHEJ mostraron expresión variable, con algúns xenes como *Xrrc6* máis altos en ESCs, mentres outros permaneceron constantes. A análise de datos proteómicos publicados en ESCs revelou que, a pesar das diferenzas de RNA, os niveis de proteína eran máis similares, suxerindo regulación post-transcricional. Ademais, a diferenciación de ESCs cara corpos embrionarios (EBs) vese como reduce progresivamente a expresión de xenes de HR, mentres que os xenes de NHEJ permaneceron inalterados, reforzando que a alta expresión de HR é unha característica específica de células pluripotentes.

Perfil de expresión das resolvasas e disolvasa de Holliday (GEN1, MUS81, BLM)

Entre as encimas encargadas de procesar intermediarios de recombinación, a disolvasa BLM mostrou unha maior expresión á das resolvasas principais, e GEN1 exprésase máis que MUS81 nas ESCs. Esta expresión mantense relativamente estable durante o desenvolvemento embrionario, con GEN1 alcanzando o máximo no trofectodermo. A diferenciación ou a adquisición de pluripotencia mediante reprogramación de MEFs reflectiu os mesmos patróns, co incremento de expresión de GEN1 asociado ao estado pluripotente.

Depleción de GEN1 compromete a supervivencia de ESCs

Mediante CRISPR/Cas9, a eliminación de GEN1 reduciu de forma significativa a capacidade de formar colonias pluripotentes, mentres que a depleción de MUS81 ou BLM non afectou notablemente o mantemento das mesmas. A análise xenotípica das colonias sobreviventes confirmou que estas correspondían a clons non editados de GEN1, indicando que a perda desta resolvase é incompatible coa supervivencia das ESCs. Ademais, a sobreexpresión da forma humana de GEN1 restaurou a formación de colonias, demostrando a especificidade do efecto da presenza de GEN1 no mantemento das ESCs

Localización subcelular de GEN1

Estudos de fraccionamento celular e inmunofluorescencia indicaron que GEN1 reside principalmente no citoplasma das ESCs, con só unha pequena fracción no núcleo. Esta localización citoplasmática mantense durante a sincronización do ciclo celular, suxerindo que GEN1 só accede ao ADN durante a mitose cando a envoltura nuclear se desorganiza.

Impacto de GEN1 sobre a pluripotencia e a diferenciación

O silenciamento de GEN1 mediante shRNAs reduciu a formación de colonias pluripotentes, aínda que as poucas sobreviventes manteron expresión de marcadores pluripotentes (NANOG e REX1). A análise transcriptómica mostrou cambios modestos con 491 xenes diferencialmente expresados, incluíndo procesos de organización da cromatina, división celular e reparación de ADN, indicando unha resposta compensatoria. Ademais, a diferenciación a corpos embrionarios revelou que a perda de GEN1 afecta principalmente á liñaxe endodérmica, sen alterar significativamente a ectodérmica ou mesodérmica.

Rol de GEN1 na reprogramación de fibroblastos

O silenciamento de GEN1 tamén reduciu a eficiencia de reprogramación de fibroblastos a células pluripotentes inducidas (iPSCs), comprometendo a activación de xenes de pluripotencia como *Nanog*, aínda que a transición mesenquima-epitelial (MET) permanecía funcional, indicando que GEN1 podería actuar tras esta transición.

Efectos sobre o metabolismo

A perda de GEN1 provocou cambios modestos en xenes implicados no metabolismo, incluíndo a fosforilación oxidativa e o ciclo TCA, pero non alterou a produción total nin mitocondrial de ATP, indicando que GEN1 non é esencial para manter a actividade metabólica basal nas ESCs.

Proliferación, ciclo celular e integridade cromosómica

As ESCs con GEN1 silenciado mostraron menor proliferación sen aumento significativo de apoptose. As exploracións de metafase revelaron aumento de fusións cromosómicas, mentres que os sitios fráxiles comúns (CFSs, do inglés *Common Fragile Sites*) non sufriron cambios significativos na súa expresión. Ademais, a presenza de danos en ADN aumentou significativamente, incluíndo durante mitose, con acumulación de marcadores de reparación como 53BP1, demostrando que GEN1 é crítico para manter a integridade xenómica.

Función independente da actividade catalítica e resposta ao estrés de replicación

A sobreexpresión da forma humana de GEN1, cataliticamente inactiva, restaurou a formación de colonias, indicando que a función de GEN1 na supervivencia das ESCs é independente da súa actividade de resolución de Holliday. Ademais, o silenciamiento de GEN1 proporcionou resistencia específica a inhibidores de topoisomerase I como camptotecina, suxerindo un papel na resposta a intermediarios de replicación asociados a TOP1, pero non a TOP2.

En conxunto, os resultados destacan que GEN1 é esencial para a supervivencia, mantemento da pluripotencia, e resposta ao dano de ADN nas ESCs. A súa función vai máis alá da resolución catalítica de intermediarios de recombinación, influíndo na expresión xenómica e na diferenciación endodérmica, así como na eficiencia de reprogramación somática. Estes achados subliñan a importancia de GEN1 como factor clave en células pluripotentes e abren oportunidades para explorar mecanismos non catalíticos na regulación da integridade xenómica.

3.6. DISCUSIÓN

As células nai embrionarias (ESCs) destacan pola súa capacidade de autorrenovación e diferenciación, o que require unha integridade xenómica moi precisa para evitar a transmisión de mutacións ás células filla. Para protexer o seu xenoma, as ESCs dependen de vías de reparación do ADN altamente eficaces que contrarresten o estrés de replicación e manteñan a estabilidade xenómica. Dentro deste contexto, as endonucleases específicas de estrutura (SSEs) desempeñan un papel central ao resolver intermediarios de ADN xerados durante a replicación e a recombinación homóloga. Aínda que a súa importancia está recoñecida, a contribución específica de cada SSE ao mantemento do xenoma en ESCs non está completamente definida. Entender estes mecanismos é esencial non só para explicar a pluripotencia, senón tamén para garantir a integridade xenómica das ESCs empregadas en investigación e medicina rexenerativa.

O obxectivo desta tese foi determinar o papel das SSEs no mantemento da pluripotencia usando ESCs como modelo. Os resultados apuntan a GEN1 como a principal resolvase nas ESCs, sendo a súa eliminación completa incompatible coa supervivencia destas células. Isto suxire que GEN1 podería ter un papel crucial na protección da estabilidade xenómica, aínda que a natureza exacta deste papel require máis estudo. A análise de bases de datos públicas confirmou que GEN1 é a resolvase máis expresada en ESCs e esencial para a supervivencia celular. Ademais, observouse a regulación á baixa de marcadores endodérmicos durante a diferenciación a corpos embrionarios (EBs) sen GEN1, suxerindo unha posible relación entre inestabilidade xenómica e especificación de liñaxes, aínda que esta conexión require confirmación experimental.

As ESCs deficientes en GEN1 mostraron un defecto de proliferación non explicado pola apoptose, acompañado dun atraso na transición de M a G1, o que suxire dificultades na progresión do ciclo celular. Ademais, observaronse aberracións cromosómicas e aumento de marcadores de dano xenómico (γH2AX e 53BP1), a pesar da presenza doutros resolvases. Esta inestabilidade non se traduciu nunha desregulación das transposicións, habitualmente usadas como marcadores de inestabilidade xenómica.

Os experimentos de rescate mostraron que tanto as formas cataliticamente activas como inactivas de GEN1 poden soste-la supervivencia das ESCs, indicando un posible papel independente da súa actividade nucleasa. Isto apunta a funcións non canónicas que poderían

involucrar o procesamento de intermediarios de replicación, como roturas de dobre cadea asociadas á replicación, aínda que estas funcións requiren confirmación experimental.

A perda de GEN1 non parece alterar significativamente a distribución global do ciclo celular en ESCs asíncronas, pero a análise de células sincronizadas revelou un atraso leve na transición M a G1. Dado que as ESCs en medio SL carecen dun control funcional do punto de verificación G1/S, poden progresar á fase S con dano xenómico, e a entrada en mitose podería activar o punto de verificación G2/M, resultando en detención ou formas alternativas de morte celular, como catástrofe mitótica ou necrose. Estas rutas de morte constitúen hipóteses baseadas en observacións de cambios na expresión de xenes mitocondriais, e requiren experimentación adicional para ser confirmadas.

O estudo citoxenético mostrou un aumento nas fusións cromosómicas en ESCs sen GEN1, sen un incremento claro de roturas, suxerindo un tipo particular de inestabilidade estrutural. Isto podería reflectir unha adaptación das células pluripotentes para minimizar a fragmentación do ADN a cambio de rearranxos estruturais, aínda que esta interpretación segue sendo especulativa. A análise de expresión xenómica tamén suxire que a perda de GEN1 afecta xenes en sitios fráxiles comúns, indicando un posible impacto funcional que require confirmación experimental.

Experimentos de rescate adicionais indican que a función de GEN1 en ESCs non se limita á resolución de xuncións de Holliday. A capacidade da forma inactiva de soste-la supervivencia suxire roles non enzimáticos posiblemente relacionados co procesamento de intermediarios de replicación. Ensaios con axentes xenotóxicos apuntan a un papel específico de GEN1 fronte a intermediarios xerados por TOP1, mentres que non se observou un efecto similar con TOP2, o que constitúe unha hipótese interesante para futuras investigacións. Ensaios de reparación mediada por recombinación homóloga fronte a unión de extremos non homólogos poderían clarificar se a ausencia de GEN1 altera a elección de vía de reparación baixo estrés de replicación.

En resumo, os resultados destacan a GEN1 como un regulador clave da estabilidade xenómica en ESCs, con funcións tanto canónicas como non canónicas. A súa presenza nuclear persistente, o seu papel non redundante na resolución de intermediarios de replicación e a súa capacidade de mantelas ESCs independentemente da actividade nucleasa distinguen a súa función nas células pluripotentes da observada en células somáticas. A posible relación entre

estrés de replicación non resolto e alteración na especificación de liñaxes subliña a importancia de vías robustas de reparación do ADN, especialmente cando se consideran aplicacións en medicina rexenerativa. Estudos futuros poderían definir os substratos de ADN de GEN1, mapear a súa rede de interaccións proteicas e analizar como regula a elección da vía de reparación, permitindo comprender mellor os mecanismos que protexen o xenoma das ESCs.

3.7. CONCLUSIÓNS

1. GEN1 é esencial para a supervivencia das ESCs, xa que a perda completa de GEN1 é incompatible coa viabilidade das células.
2. A perda de GEN1 provoca estrés xenómico crónico, con sinalización persistente de dano no ADN probablemente debido a intermediarios de ADN non resoltos.
3. A función de GEN1 nas ESCs non se explica completamente pola súa actividade resolvase de HJ, xa que a versión cataliticamente inactiva de GEN1 rescata o fenotipo de perda de función.
4. GEN1 exprésase a niveis altos e dinámicos nos estados pluripotentes, diminuíndo durante a diferenciación e aumentando de novo durante a reprogramación.
5. GEN1 é principalmente citoplasmática, pero unha pequena fracción mantense no compartimento nuclear das ESCs.
6. A perda de GEN1 durante a saída da pluripotencia prexudica a diferenciación ao endodermo.
7. GEN1 é necesario para unha reprogramación eficiente das células somáticas.
8. A perda de GEN1 compromete a estabilidade cromosómica nas ESCs, dando lugar a un aumento das fusións cromosómicas.
9. A perda de función de GEN1 confire vantaxes de supervivencia fronte ao estrés replicativo causado pola inhibición da topoisomerasa I.

4. Introduction

4.1. EMBRYONIC STEM CELLS: BIOLOGICAL BASIS AND THERAPEUTIC POTENTIAL

Pluripotency is defined as the capacity for a cell to differentiate into all three embryonic germ layers that give rise to an organism. While pluripotency *in vivo* is only a transient state that characterises the cells from the inner cell mass (ICM) of the preimplantation blastocyst, it can be captured *in vitro* by the derivation of cells from the ICM of the mid-blastocyst at embryonic day 3.5 (E3.5) [1]. These cells are called embryonic stem cells (ESCs) (**Figure I1**). ESCs are a unique cell type that display two defining characteristics: pluripotency, as explained above, and self-renewal, the capacity to proliferate indefinitely *in vitro* without losing its properties. Nevertheless, not all derived ESCs exhibit the same pluripotent characteristics due to differences of the culture conditions in which they are cultured [2-4]. Hence, different states of pluripotency might arise *in vitro*, associated with different differentiation potentials, and they will be discussed in the following section.

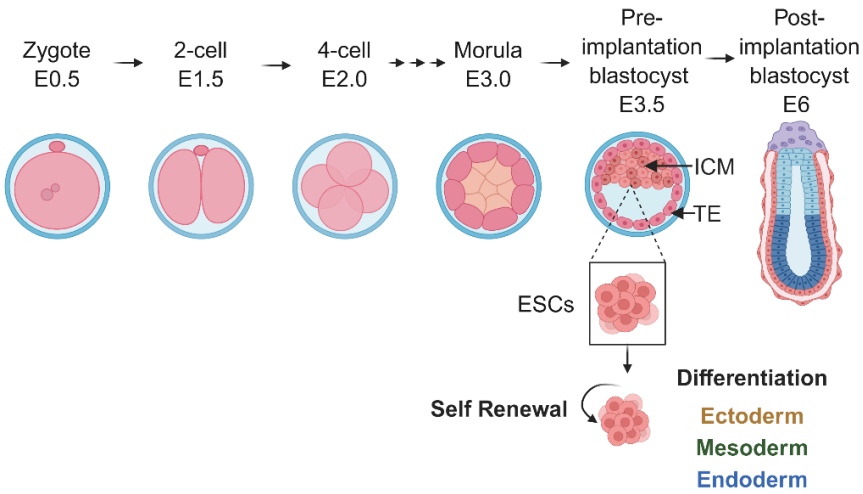

Figure I1. Different stages during the mouse early development. Mouse embryogenesis begins with the zygote (E0.5) and progresses through a series of cleavage divisions to the 2-cell (E1.5) and 4-cell (E2.0) stages, followed by the morula (approximately 8-cell, E3.0) and the pre-implantation blastocyst (E3.5). The blastocyst comprises the ICM, which gives rise to the embryo proper, and the trophectoderm (TE), which contributes to the formation of extraembryonic tissues. Upon implantation (E6), the blastocyst undergoes further differentiation. ESCs derived from the ICM, possess the dual capacity for long-term self-renewal and differentiation into the three germ layers: ectoderm, mesoderm, and endoderm.

ESCs hold great potential in regenerative medicine due to their special combination of properties. Their differentiation capacity provides a versatile source for tissue repair and replacement [5], as they can give rise to any cell type of interest. Their ability to maintain stemness through continuous division enables the production of large quantities of cells for therapy, which is crucial for producing sufficient material for transplantation or tissue

engineering. ESCs also serve as invaluable tools for modelling embryonic development and disease, allowing the study of differentiation processes and the generation of patient-specific cell types for disease modelling and drug screening [6]. In addition, they can be used to generate specialised cells for tissue engineering and organoid formation, opening new avenues for organ repair and regenerative treatments [7].

Although ESCs offer a wide range of applications as stated above, the use of embryonic material always raises ethical concerns that limit their clinical application. Nowadays, pluripotent stem cells (PSCs) can also be obtained *in vitro* by reprogramming somatic cells to induced pluripotent cells (iPSCs) following a so-called somatic cell reprogramming (SCR) process. In a pioneering study in 2006, Yamanaka and colleagues were the first scientists able to obtain iPSCs from mouse embryonic fibroblasts (MEFs) or adult fibroblasts by the overexpression of four transcription factors: OCT4, SOX2, KLF4 and c-MYC, referred to as OSKM [8, 9]. These four proteins effectively rearrange the transcriptional landscape and chromatin state, among other cellular properties, to convert a somatic into an embryonic-like cell. This groundbreaking discovery, which granted Yamanaka, together with Sir. Gurdon the Nobel Prize in Physiology or Medicine in 2012, has allowed the generation of patient-specific pluripotent cells, establishing a foundation for personalised regenerative therapies with minimal risk of immune rejection [10].

However, the regulatory pathways that sustain pluripotency and enable their differentiation into specific tissues, processes that are fundamental for tissue replacement therapies (regenerative medicine) and personalised interventions (precision medicine), are not yet fully understood. Additionally, these two cell types are considered paradigms of "young cells", and thus their study is of great importance for understanding the aging process and for designing novel strategies to promote healthy and active ageing.

4.2. DISTINCT STATES IN ESCS: FROM NAÏVE TO PRIMED PLURIPOTENCY

The culture conditions used for ESCs maintenance can lead to different pluripotency states, including naïve, metastable and primed [2, 3]. These states represent a continuum in the pluripotency state and can be assimilated to *in vivo* developmental stages corresponding to the preimplantation epiblast (E3.5) and post-implantation epiblast (E6) for the naïve and primed pluripotent states, respectively (**Figure I2**). The cells from the primed state are named epiblast

stem cells (EpiSCs) [11]. In between those states we can find a metastable state which shows a more heterogenous population [2].

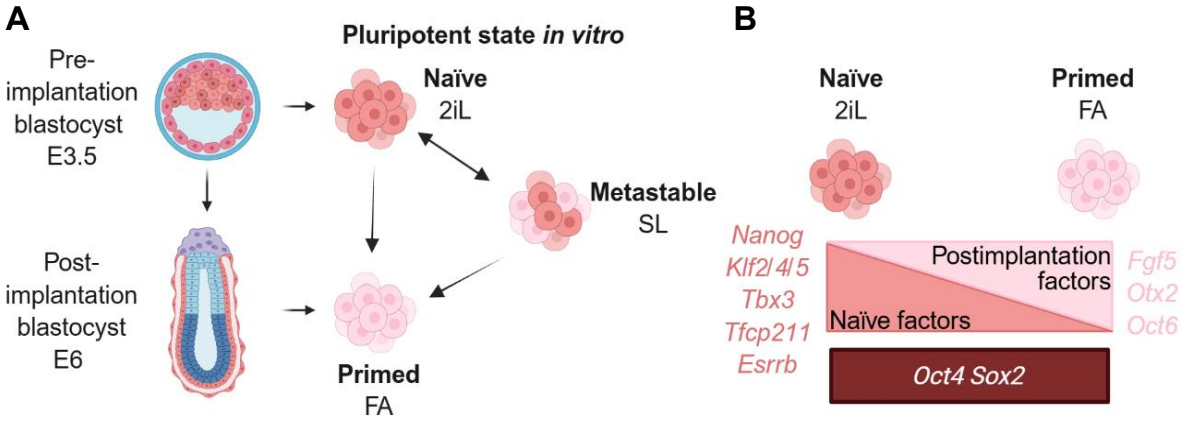

Figure I2. ESCs derived from the ICM can resemble either pre-implantation or post-implantation epiblast cells depending on culture conditions. (A) Naïve culture conditions, which preserve characteristics of the pre-implantation blastocyst, are maintained by cultivating ESCs in medium containing 2i and LIF (2iL). In contrast, primed conditions, established using FGF2 and activin A (FA) medium, promote the acquisition of features characteristic of post-implantation epiblast cells. A heterogeneous population can be obtained by culturing ESCs in serum/LIF (SL) medium. **(B)** Distinct molecular markers are used to identify these states, including *Nanog* as an example of a naïve marker and *Fgf5* as an example of a primed marker.

In the case of the naïve pluripotent state, the serum-free medium is supplemented with the leukaemia inhibitory factor (LIF) which supports the self-renewal capacity of the ESCs [12]. This protein activates a signalling cascade which leads to the activation of the signal transducer and activator of transcription 3 (STAT3) [13]. This transcription factor (TF) activation results in the induction of expression of target genes essential for maintaining pluripotency and self-renewal in ES cells, including *Klf4* and *Gbx3* [14, 15]. However, this is not sufficient to maintain the naïve state, or so-called "ground state", of ESCs in culture without undergoing differentiation. Therefore, the media also contains two inhibitors (2i): a MEK inhibitor (i.e. PD0325901) and a GSK3 inhibitor (i.e. CHIR99021) [16]. The MEK/ERK signalling pathway is antagonistic to self-renewal and normally leads to activation of lineage specification programs. Hence, its inhibition promotes the self-renewal capacity and blocks the differentiation of cells to stabilise the naïve ESCs [17]. GSK3 is another antagonist of ESCs self-renewal and therefore, its inhibition also favours maintenance of undifferentiated naïve ESCs by repressing TCF3 through stabilisation of β-catenin [2, 18]. For that reason, the culture of naïve cells is supported by medium containing LIF and 2i (2iL medium). In primed pluripotent cells, their maintenance *in vitro* is supported by FGF2 and activin A (FA medium)

[3, 4]. In this case, FGF2 activates the MEK/ERK pathway which, in contrast to the naïve state, supports the pluripotency network by keeping the expression of core pluripotency factors as OCT4, SOX2 and NANOG [19]. Activin A activates the SMAD2/3 signalling, which cooperates with NANOG and OCT4 to sustain the pluripotency gene expression. Moreover, it inhibits the differentiation toward the neuroectoderm of the primed pluripotent cells [20]. Therefore, together these two supplements help to maintain pluripotency in primed cells.

Somewhere in between those states, we can find metastable pluripotent stem cells which are culture in serum with LIF (SL) medium. This is the classical condition for keeping ESCs in culture before the development of the 2iL medium technology [21]. As for naïve ESCs, LIF acts as a self-renewal supporter and in this case the serum will provide multiple growth factors. Actually, this serum supplementation contributes with additional signals that supress differentiation, such as BMP4, which helps blocking neural differentiation, but also contains FGF2, which can promote differentiation [19, 22]. This creates a dynamic and heterogenous population in which cells self-renew but fluctuate between naïve- and primed- like states.

The culture conditions are not the only difference between these pluripotency states. Even though *in vitro* they share the two unique characteristics of canonical ESCs (self-renewal and pluripotency), this does not hold true for their *in vivo* potential. Naïve cells can produce germline chimeras, whereas primed cells require injection into more developmentally advanced (post-implantation/gastrula-stage) embryos to do so, suggesting that primed cells have less potency than naïve [3, 23]. These states differ not only in their developmental capacity, but also in their epigenetic profile and basal metabolism (summarised in the table below (**Table 1**)) as well as in their cell cycle length (further discussed in **section 4.4)** [24-26].

Table 1. Comparison of the three pluripotent states in culture.

State	Medium	Key Pathways	Epigenetic profiles	Metabolism
Metastable	SL	LIF/STAT3 → pluripotency FGF/ERK → differentiation BMP/SMAD1/5 → blocks neural fate	Partial DNA hypomethylation, heterogeneous chromatin (mixed open and compact states)	Mixed glycolysis + Oxidative phosphorylation
Naïve	2iL	LIF/STAT3 → pluripotency MEK/ERK inhibited → blocks differentiation GSK3 inhibited → pluripotency	Global hypomethylation, open chromatin	Glycolysis + Oxidative phosphorylation
Primed	FA	FGF/ERK + Activin/SMAD2/3 → pluripotency LIF ineffective	Higher DNA methylation, more compact chromatin	Glycolysis

The coexistence of distinct pluripotent states highlights the dynamic nature of pluripotency, sustained by the coordinated interplay of transcriptional and signalling regulators that together establish the pluripotency network.

4.3. THE PLURIPOTENCY NETWORK IN ESCS

The pluripotent state is maintained through an intrinsic network governed by TFs. There are three master regulators of pluripotency: OCT4, SOX2 and NANOG. OCT4 is one of the first TFs discovered to be essential for ESCs [27, 28]. OCT4 depletion is embryonic lethal and incompatible with self-renewal of ESCs [29]. On the other hand, overexpression of this factor can induce differentiation [30], and therefore, a correct balance in the expression levels of this protein is key to maintain the pluripotent population. Another factor crucial for ESC maintenance is SOX2. Similar to OCT4, SOX2 suppression leads to loss of pluripotency [31]. Moreover, SOX2 interacts with OCT4 at the protein level and can enhance *Oct4* transcription [32]. SOX2 is expressed throughout the mouse preimplantation period [33]. Overexpression of SOX2 also predisposes ESCs to differentiate as observed with OCT4 [34, 35]. The third core pluripotency regulator is NANOG, whose expression is more restricted to the naïve state compared to OCT4 and SOX2 [36]. It is also important for self-renewal and colony formation capacity of ESCs [37, 38]. In contrast to OCT4 and SOX2, ESCs can survive without NANOG in 2iL culture conditions [39]. Nonetheless, NANOG maintains a stable undifferentiated state in ESCs and it is necessary for normal embryonic development [36]. OCT4, SOX2, and NANOG form a core pluripotency transcriptional network in ESCs, where each factor not only regulates downstream target genes but also positively regulates the expression of the others (**Figure I3**). OCT4 and SOX2 physically bind to regulatory regions of the *Nanog* gene to maintain its transcription, while NANOG, in turn, can enhance the expression of *Oct4* and *Sox2*, forming a positive feedback loop [40]. This interconnected autoregulatory circuit stabilises the pluripotent state, ensuring self-renewal and preventing differentiation, and allows stem cells to respond dynamically to extrinsic signals while maintaining their undifferentiated identity.

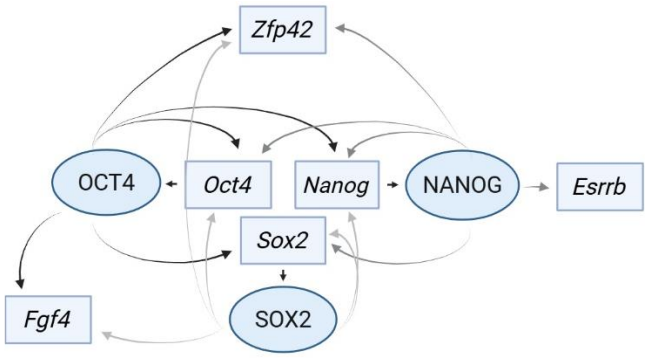

Figure I3. The pluripotency network is mainly governed by the three master regulators: OCT4, NANOG and SOX2. They regulate their own expression as well as the others expression to keep a balance (i.e. *Fgf4, Zfp42* and *Esrrb*, among others). Proteins are circled and genes are squared.

4.4. CELL CYCLE REGULATION IN ESCS

Another unique characteristic of ESCs is their cell cycle profile and regulation. In somatic cells, the cell cycle is driven by the sequential activation of cyclin-dependent kinases (CDKs) bound to their regulatory cyclins, ensuring the accurate duplication and segregation of genetic material. In G1 phase, mitogenic signals induce D-type cyclins (D1–D3), which activate CDK4 and CDK6 to initiate phosphorylation of the retinoblastoma protein (RB1) and partially relieve repression of E2F transcription factors [41]. Later in G1, cyclin E–CDK2 completes RB1 inactivation, committing cells to S phase, where DNA replication is initiated by cyclin E–CDK2 in conjunction with DBF4-dependent kinase and sustained by cyclin A2–CDK2, later paired with cyclin A2–CDK1 [41, 42]. As cells transition through G2 phase, cyclin A–CDK1 activity prepares the genome for mitosis, culminating in M phase, which is driven by nuclear accumulation of cyclin B–CDK1 [43]. Cyclin B–CDK1 orchestrates chromosome condensation, nuclear envelope breakdown, and spindle assembly to ensure faithful segregation of genetic material. Throughout the cycle, CDK inhibitors of the INK4 family (p15, p16, p18, p19) and the CIP/KIP family (p21, p27, p57) modulate cyclin–CDK activity, fine-tuning cell cycle progression and maintaining genomic integrity [44]. Collectively, the coordinated activity of cyclin–CDK complexes establishes a robust framework that drives proliferation while safeguarding the fidelity of genome duplication and transmission.

Mouse ESCs (mESCs) exhibit a distinctive cell cycle organisation that distinguishes them from somatic cells. One of the most prominent features of ESCs is their abbreviated cell cycle, which is approximately twice as fast as that of differentiated proliferative somatic cells, enabling completion of DNA replication and cytokinesis in less than 12 h (**Figure I4**) [45, 46].

This accelerated proliferation is primarily attributable to a shorter G1 phase, resulting in a high fraction of cells residing in S phase. The molecular underpinnings of this unique cell cycle profile involve a combination of constitutive activation of CDKs and cyclins, together with a pronounced deficiency in cyclin-dependent kinase inhibitors such as p21 and p27 [47]. This configuration permits continuous CDK activity, effectively bypassing the conventional regulatory checkpoints that operate in somatic cells. A critical target of CDKs is the retinoblastoma tumour suppressor protein (pRB) which functions as a repressor of E2F transcription factors [48]. These TFs are required for the transition from G1-to-S phase.

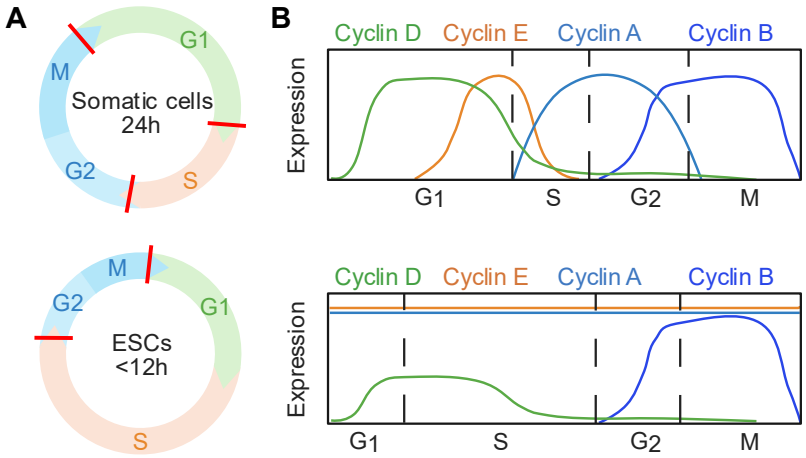

Figure I4. Differences in cell cycle length and cyclins expression in somatic cells and ESCs. (A) The overall length of the cell cycle differs between these two cell types, primarily due to the markedly shortened gap phases (G1 and G2) in ESCs. Moreover, ESCs in SL conditions lack the G1/S checkpoint, checkpoints are marked with a red line. **(B)** In somatic cells, cyclin expression is tightly regulated and oscillates in a phase-dependent manner. By contrast, ESCs maintain elevated levels of cyclins throughout the cycle, facilitating rapid progression and efficient transition between cell cycle phases.

In ESCs, high CDK activity maintains RB in a hyperphosphorylated and therefore inactive state, thereby releasing E2F and promoting transcription of genes required for DNA synthesis. Consequently, ESCs surpass the functional G1/S checkpoint facilitating rapid progression through the G1 phase and contributing to their overall abbreviated cell cycle. Importantly, ESCs are able to move through this checkpoint even when DNA damage is present, mainly because they express high levels of CDK activators (i.e. CDC25A, which activates CDK2) [49]. This ability is supported by the coordinated pluripotency network that maintains ESCs identity. However, they do possess an active G2 checkpoint through ATM activation which can temporarily arrest the ESCs at G2/M stage [50]. The above mechanisms operate in SL cultured ESCs, but when cells are culture in naïve-like 2iL medium, they exhibit a partial restoration of

cell cycle regulation. In this context, there is a reduction in CDKs activity involved in pRB hyperphosphorylation. It allows pRB to adopt its hypophosphorylated and therefore active state, reinstating the G1/S checkpoint [51]. Notably, reactivation of checkpoint control under these conditions is associated with an increased propensity for differentiation, highlighting the intrinsic link between cell cycle dynamics and the maintenance of pluripotency.

Importantly, the abbreviated cell cycle of ESCs does not reflect an accelerated rate of DNA replication. The duration of S phase is largely comparable between ESCs and somatic cells [47, 52]. Rather, the observed reduction in total cell cycle length arises from the shortening of the gap phases, G1 and G2 [53]. Thus, ESCs achieve rapid proliferation primarily through the modulation of regulatory mechanisms controlling the initiation and progression of cell cycle phases, rather than through alterations in DNA synthesis kinetics. Importantly, the abbreviated gap phases that characterise the cell cycle of ESCs, while essential for sustaining their pluripotent state, have been proposed to increase replication stress [53, 54]. This replication stress arises due to the rapid proliferation capacity of ESCs and can result in the accumulation of DNA lesions, including double-strand breaks (DSBs) (**Figure I5**).

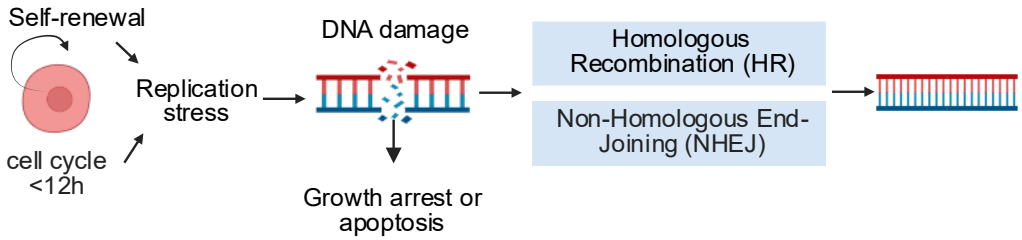

Figure I5. Features of ESCs contribute to elevated replication stress, which can generate DNA damage in the form of DSBs. These DSBs can be repaired through multiple pathways, including homologous recombination (HR) and non-homologous end-joining (NHEJ).

DSBs are among the most cytotoxic forms of DNA damage, with the potential to trigger growth arrest, apoptosis, or genomic instability if not properly repaired. Importantly, replication stress in ESCs is not only a consequence of their fast cycling but is also intimately linked to their developmental state. Self-renewing ESCs employ distinct mechanisms to tolerate and resolve replication-associated DNA damage, whereas the onset of differentiation alters cell cycle dynamics and DNA repair pathway choice, thereby influencing the level and consequences of replication stress [55]. Although the rapid proliferation and shortened gap phases of ESCs generate replication stress and can lead to DSBs, this does not translate into elevated mutation rates [52, 56, 57]. This feature is observed not only in ESCs but also in iPSCs,

where reprogramming to a pluripotent state has been shown to reduce mutation rates [58]. Pluripotent cells possess highly efficient DNA repair mechanisms that safeguard genomic integrity. DSBs are resolved primarily through two canonical pathways: homologous recombination (HR), which uses a sister chromatid as a template to ensure error-free repair, and non-homologous end joining (NHEJ) (**Figure I5**). Together, these mechanisms enable ESCs to minimise the accumulation of mutations despite the intrinsic replication stress associated with their rapid cell cycle [59].

4.5. MECHANISMS FOR DOUBLE-STRAND BREAK REPAIR IN ESCS

As previously mentioned, ESCs efficiently repair DSBs using two main pathways: HR and NHEJ. There are many differences between HR and NHEJ repair methods involving different proteins and complexes (**Figure I6**). NHEJ repairs DSB without requiring a homologous template, instead ligating the broken DNA ends after minimal processing. The pathway is initiated when the XRCC6/XRCC5 heterodimer rapidly binds DNA ends, protecting them from degradation. This will recruit the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to the broken site to regulate DNA end processing and repair. When DNA end structures are incompatible with direct ligation, DCLRE1C and polymerases μ and λ will process the DNA end and fill gaps, respectively. Finally, DNA ligation will be performed by LIG4 with the help of XRCC4 and NHEJ1. This concerted process ensures efficient but potentially error-prone repair, particularly in the absence of compatible DNA termini [60].

In HR, the process starts by recognition of the break and nucleolytic resection of the 5′ strand in both sites of the break to generate 3′ single-stranded DNA (ssDNA) overhangs (**Figure I6**). The MRE11–RAD50–NBS1 (MRN) complex acts as an early sensor and promotes initial end processing, cooperating with CtIP to initiate limited end resection. After end resection generates 3′ ssDNA overhangs, RPA rapidly coats the ssDNA to stabilise it and prevent secondary structure formation. BRCA1, often in cooperation with CtIP and PALB2, promotes end resection and creates a recombination-permissive environment. BRCA2, together with PALB2 and DSS1, then mediates the replacement of RPA with RAD51, stabilising the RAD51 nucleoprotein filament. This filament catalyses homology search and strand invasion, generating a displacement loop (D-loop) that initiates homologous pairing and DNA synthesis.

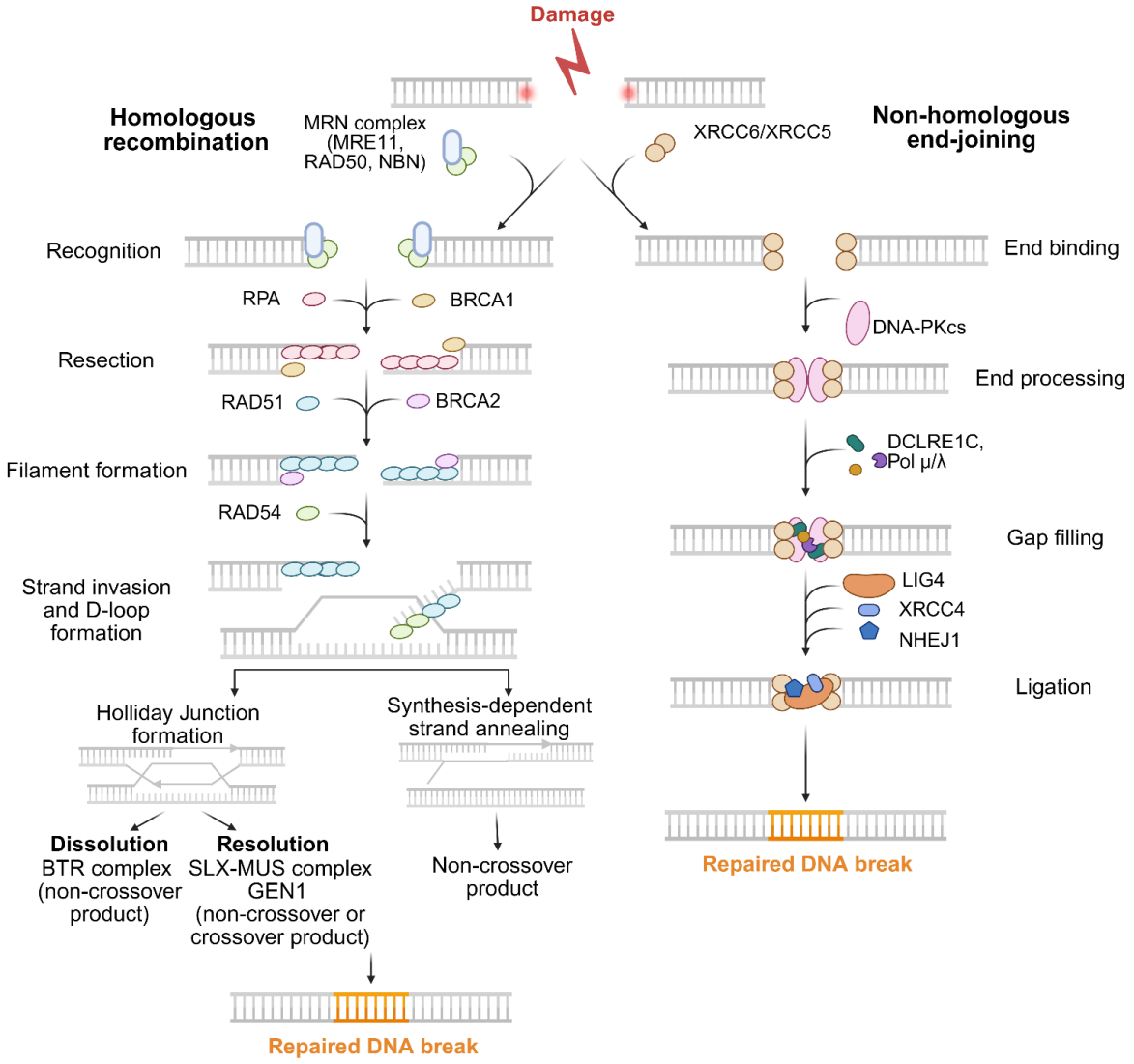

Figure I6. The two primary repair pathways, HR and NHEJ, employ distinct sets of proteins to repair the DNA, resulting in either error-free or modified repair. HR is a more prolonged process, during which numerous secondary DNA structures form and must be resolved to complete repair.

Then, RAD54 interacts with RAD51 filament and the D-loop to promote its stabilisation and homology search. Subsequently, RAD54 can remodel chromatin facilitating DNA synthesis around the D-loop. At this stage, repair may proceed via the SDSA pathway: the D-loop is disrupted, releasing the newly synthesised strand, which anneals to the complementary ssDNA at the second DNA end, leading to gap filling and repair completion, and producing exclusively non-crossover products. Alternatively, the second DNA end may be captured and extended, occasionally yielding joint-molecule intermediates that form double Holliday junctions (dHJs). The dHJ are processed by two general routes: (i) dissolution by the BLM–TOP3A–RMI1–RMI2 (BTR) complex [61], or (ii) nuclease-mediated resolution by structure-specific

endonucleases (resolvases) such as GEN1 or the MUS81–EME1 complex acting with SLX1–SLX4 [62, 63]. While BTR involves a helicase which mediates convergent branch migration to decadenate the junction and produce exclusively non-crossover products, resolution by endonucleases MUS81 or GEN1 includes HJs cleavage and can generate either crossover or non-crossover products depending on the orientation of the cut [64].

In summary, HR and NHEJ differ in how they process DNA breaks, the speed of repair, and the accuracy of the repair, a difference of special importance in ESCs. Because ESCs serve as the foundation for all cells in the future organism, any mutation they acquire would be propagated to every descendant cell. To safeguard the developing embryo, ESCs rely heavily on HR during S and G2 phases, minimising the risk of passing deleterious mutations or genomic instability [65]. HR provides error-free repair by using the homologous sequence of the sister chromatid as a template, whereas NHEJ directly ligates the broken DNA ends, which can introduce deletions or insertions at the break site. Although HR is a high-fidelity process it is also more time-consuming and it requires that the cell is in S or G2, ESCs preferentially use it to avoid transmitting genetic errors [59].

4.6. HOMOLOGOUS RECOMBINATION IN ESCS: KNOWNS AND UNKNOWNS

As previously discussed, ESCs experience elevated levels of replication stress due to their unusually rapid cell cycle, abbreviated G1 phase, and reliance on continuous proliferation [53]. Under these conditions, HR plays a central role in safeguarding genome stability by providing a high-fidelity mechanism to repair replication-associated DNA double-strand breaks. Consequently, ESCs are particularly vulnerable to the loss of HR factors with absence of these regulators resulting in unresolved DNA lesions, accumulation of chromosomal aberrations, and ultimately compromised self-renewal and developmental potential [59, 66]. The importance of HR for ESC maintenance is therefore twofold: it minimises mutation rates, thereby preserving pluripotency, and it ensures cellular survival under the constant pressure of replication stress. In fact, most core HR proteins have been shown to be essential for normal embryonic development with the knockout (KO) being indispensable for survival (**Table 2**).

Disruption of any component of the MRN complex impairs the initial recognition and processing of DNA DSBs, leading to chromosome instability, failure to maintain genomic integrity, and widespread cell death in the early embryo. Similarly, BRCA1 and BRCA2 are critical for the recruitment and assembly of DNA repair complexes. Therefore, mice lacking

these genes die early in development and ESCs become nonviable or highly genome unstable. Downstream effectors like PALB2, RAD51 and the RAD51C/XRCC3 paralogs are equally indispensable: their loss blocks homologous pairing and strand invasion steps, prevents proper chromosome repair and leads to failed cell proliferation or embryonic lethality. These findings underscore that the core HR pathway is not merely advantageous but absolutely required for both viability of pluripotent stem cells and successful mammalian development.

Table 2. Function of HR proteins and evidence for their essentiality in mouse development and ESCs

Protein	Function	Lethality Evidence	Key References
MRE11, RAD50, NBS1	DSB sensing and initiation of DNA end resection.	Mice: KO for each die early in utero. ESCs: Show severe genome instability or are compromised.	[67-70]
BRCA1	Promotes DNA end resection and recruits repair machinery.	Mice: Null embryos die early in development. ESCs: Require it for viability.	[71, 72]
BRCA2	Loads RAD51 onto ssDNA and stabilises the filament.	Mice: Deficiency is embryonic lethal. ESCs: Cannot maintain genome integrity without it.	[73, 74]
PALB2	Bridges BRCA1 and BRCA2 and stabilises RAD51.	Mice: Loss is embryonically lethal. ESCs: Fail to survive.	[75, 76]
RAD51	Catalyses strand invasion and homology search.	Mice: Die at the preimplantation stage. ESCs: Fail to survive.	[77, 78]
RAD51C / XRCC3	Stabilise RAD51 filaments and aid in later HR stages.	Mice: KO mice die embryonically. ESCs: Essential for proliferation and survival.	[79-81]

4.7. Mechanisms of Holliday Junctions processing and genome maintenance in ESCs

Despite significant advances in understanding how ESCs detect DNA damage and initiate HR repair, much less is known about the later stages of the pathway. In particular, the processing of Holliday Junction (HJ) intermediates that can arise during HR, which represents a decisive step in determining crossover versus non-crossover outcomes, remains incompletely characterised in the pluripotent state. The proteins responsible for HJ dissolution and resolution are well studied in somatic contexts [82, 83], yet their precise contributions in ESCs are only beginning to be elucidated. The activities of HJ-resolving proteins are tightly regulated in both space and time. In somatic cells, the BTR complex (BLM–TOPIIIα–RMI1/2) predominantly mediates HJ dissolution during S and G2 phases, representing the preferred pathway to prevent crossover formation and chromosomal abnormalities [61]. As a complementary mechanism, MUS81–EME1 is activated by cell cycle-dependent phosphorylation in late S and G2/M,

enabling cleavage of HJs and other branched intermediates that arise during replication [84, 85]. Finally, the nuclease GEN1 serves as a crucial fail-safe pathway. Its activity is tightly restricted during interphase by cytoplasmic sequestration, ensuring it only gains access to the nucleus upon nuclear envelope breakdown in mitosis [86]. The ortholog version in yeast, Yen1 is also regulated by its phosphorylation status, in which the phosphorylation affect the substrate binding and inhibits its nuclear import whereas its dephosphorylation can promote nuclear entry and enzymatic activation [87-89]. Nevertheless, human GEN1 shows no effect on its capacity to resolve HJ through phosphorylation [86]. Therefore, its activity is only regulated by the cell cycle due to the nuclear export signal that it contains. This allows GEN1 to resolve any remaining HJs just before chromosome segregation. The coordinated and hierarchical action of these distinct dissolution and cleavage pathways ensures that potentially toxic DNA intermediates are safely and accurately eliminated, underscoring the vital importance of spatiotemporal control in DNA repair and the preservation of genomic integrity.

In addition to HJ, structure-specific resolvases can process a variety of branched DNA intermediates (**Figure I7**). The endonuclease complex MUS81–EME1, for example, cleaves stalled and collapsed replication forks, generating double-strand breaks (DSBs) that promote replication restart [83, 90]. MUS81–EME1 also processes 3′ flaps formed during DNA repair pathways such as synthesis-dependent strand annealing (SDSA) and break-induced replication (BIR), as well as nicked HJs and D-loops [91-93]. D-loops structures arise when a 3’ single-stranded end invades a homologous template during HR, for instance at collapsed forks or programmed DSBs. Similarly, 5’ flap intermediates are generated during lagging-strand maturation and long-patch repair processes, where strand displacement synthesis produces displaced 5’ termini that must be removed. As for MUS81, human GEN1 activity is not limited to HJ resolution, it can also cleave 5′ flaps and replication forks, as shown by *in vitro* studies [94, 95]. In budding yeast, its ortholog Yen1 has been demonstrated to possess additional, non-canonical nuclease activities, including a 5’-to-3’ exonuclease activity on branched DNA and the ability to cleave D-loops, particularly when acting with *Mus81* in *in vitro* assays [96, 97].

Whether mammalian GEN1 also possesses such non-canonical activities remains to be determined; if so, this would broaden our understanding of its roles in genome maintenance beyond its known substrates.

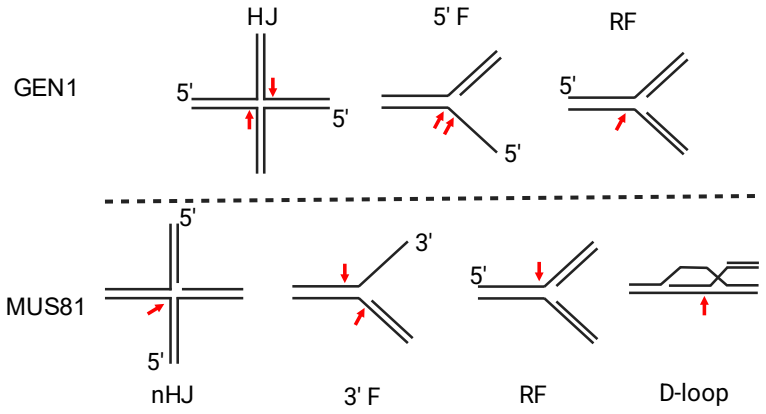

Figure I7. The structure selective endonucleases GEN1 and MUS81 can cleave other DNA structures apart from HJ. The substrates for GEN1 include Holliday junctions (HJ), 5' flaps (5'F) and replication forks (RF) *in vitro*. In the case of MUS81, the substrates are nicked HJ, 3' flaps (3'F), RF and D-loops.

During early development, BLM is essential for mice viability, as its depletion in mice results in embryonic lethality at day E13.5 [98]. On the other hand, *Mus81* KO mice are viable, but exhibit reduced fertility and heightened sensitivity to DNA-damaging agents [99]. Finally, *Gen1*-deficient mice are viable, but display defects in kidney development [100, 101]. Importantly, the simultaneous loss of MUS81 and GEN1 results in embryonic lethality, as no double KO mice or late-stage embryos are recovered *in vivo* [100, 102]. While the precise developmental stage of lethality has not been exhaustively determined, these findings underscore the essential and redundant functions of these resolvases during embryogenesis. Together, these observations indicate that BLM, MUS81, and GEN1 contribute both independently and cooperatively to genome maintenance during early development.

While most studies of these three proteins have been conducted in somatic cells, in ESCs, work using *Blm* KO lines have demonstrated that BLM deficiency leads to chromosomal instability and increased loss of heterozygosity, highlighting its central role in safeguarding the pluripotent genome [103-105]. In the case of MUS81, one study showed that its depletion in ESCs resulted in the accumulation of replication-associated DNA lesions and hypersensitivity to DNA crosslinking agents, indicating a key role in replication-associated DNA repair [106]. However, the impact of GEN1 loss in ESCs remains unexplored. Although the roles of BLM and MUS81 in pluripotent cells has been somehow addressed through these limited KO studies, the specific contribution of GEN1 to DNA repair in ESCs represent a critical gap in current knowledge. Given the importance of maintaining strict genome stability to avoid mutations detrimental to stem cells, it is of utmost importance to have repair pathways that ensure error-

free DNA repair. A deeper mechanistic understanding of late HR stages is essential, as it will shed light on how ESCs maintain long-term genomic stability while sustaining their unique proliferative and developmental properties.

5. Objectives and hypothesis

Given that Structure-Specific Endonucleases (SSEs) and the helicase BLM play a critical role in the repair of DNA intermediates during homologous recombination-mediated DNA repair, and pluripotent cells exhibit an extraordinary capacity for DNA repair, ***we hypothesise that SSEs play a significant role in the maintenance of pluripotency and/or its acquisition through somatic cell reprogramming***. To test this hypothesis, we propose to dissect the molecular mode of action of specific SSEs, MUS81 and GEN1, as well as the dissolvase BLM, which are instrumental in the resolution of HJs in the maintenance of pluripotent cells. To achieve this, the following objectives will be addressed:

1. Determine the contribution of SSEs and BLM to ESCs survival and genomic stability.
2. Investigate possible non-canonical functions of GEN1 in ESCs.
3. Explore the involvement of GEN1 in replication-associated DNA structures.

Exploring the relevance of SSEs and BLM in the maintenance of pluripotent cells represents an unprecedented opportunity not only to discover novel factors involved in pluripotency regulation but also new pathways for genomic integrity control. These insights could potentially be modulated to improve iPSC technology, bringing it closer to clinical application.

6. Materials and Methods

6.1. REAGENTS, KITS AND EQUIPMENT

All reagents, kits and equipment used in this thesis are listed on **Supplementary Tables 1**, **2** and **3**.

6.2. CELL LINES

The cell lines used in this thesis are listed in the table below (**Table 3**).

Table 3. Cell lines used in this thesis.

Name	Source	Description
CJ7 (mouse ESCs)	Dr Jianlong Wang (Columbia University, USA)	Derived from ICM of 129/SvlmJ male embryos
GEN1^{3xFL} ESCs (mouse ESCs)	This study. Background: CJ7	CRISPR-Cas9 knock-in of 3 flag tags at the C-terminal end of the *Gen1* gene
W3T3 iOSKM/rtTA	Dr Miguel Fidalgo (Universidade de Santiago de Compostela, Spain)	3T3 immortalised fibroblasts transduced with a doxycycline-inducible OSKM cassette and constitutively express rtTA
HEK 293T	ATCC (CRL-3216)	Immortalised human embryonic kidney cell line containing the SV40 T antigen

6.3. CELL CULTURE AND TREATMENTS

First, gelatine-coated plates were prepared using 0.1% gelatine (Sigma-Aldrich, G9391) diluted in milli-Q water. Mouse ESCs were cultured on gelatine-coated plates. 0.1% (w/v) gelatine (Sigma-Aldrich, G9391) was prepared in milli-Q water and autoclaved for sterilisation. Plates were coated with it and incubated for at least 30 min at 37ºC in a cell incubator. After the incubation period, the gelatine was removed by aspiration prior to seeding the cells. ESCs were cultured using embryonic stem cells medium (ESM^{LIF}), containing High-Glucose Dulbecco's modified Eagle Medium (Corning, 10-013-CVR), 15% heat inactivated fetal bovine serum (Corning Media Tech, 35-079-CV), 100 U/mL penicillin/streptomycin (Gibco, 15140-122), 1% nucleoside mix [adenosine (Sigma-Aldrich, A4036), cytidine (Sigma-Aldrich, C4654), guanosine (Sigma-Aldrich, G6264), thymidine (Sigma-Aldrich, T1895) and uridine (Sigma-Aldrich, U3003)], 2 mM L-glutamine (Gibco, 25030024), 1x Non-Essential amino acids solution (Gibco, 11140050), 0.1 mM β-mercaptoethanol (Sigma-Aldrich, M6250) and 10^3 U/mL homemade Leukaemia Inhibitory Factor (LIF). For passaging of the cells, cells were trypsinised with 0.05% trypsin-EDTA solution (Hyclone, SH30236.01) for 7 min at 37ºC and

neutralised with medium containing 10% of FBS. ESCs were passaged every other day at a split ratio of 1:8. Cells were kept in culture in a humidified incubator at 37ºC and 5% CO_2.

Murine W3T3 and human HEK 293T cells were cultured on untreated plates in medium containing High-Glucose Dulbecco's modified Eagle Medium (Corning, 10-013-CVR), 10% heat inactivated fetal bovine serum (Corning Media Tech, 35-079-CV), 100 U/mL penicillin/streptomycin (Gibco, 15140-122), and 2 mM L-glutamine (Gibco, 25030024). Cells were passaged every two days at a split ratio of 1:5 and kept in culture in a humidified incubator at 37ºC and 5% CO_2.

For long term preservation, cells were collected by trypsinisation followed by centrifugation at 300 x *g* for 5 min at room temperature (RT). Then, they were resuspended in freezing medium consisting of 90% cell culture medium (containing fetal bovine serum) and 10% dimethyl sulfoxide (DMSO; Corning, 25-950-CQC) and kept in cryovials at -80ºC for a maximum of 48 h. Then, vials were transferred to liquid nitrogen tanks for long term storage.

For inducing replication stress in ESCs, 0.3 µM of aphidicolin (APH; Sigma-Aldrich, A4487) were used for 24 h followed by the protocol described in section 5.20 for metaphase spreads preparations. APH is a DNA polymerase inhibitor that at low doses can slow replication and produce mild replication stress [107].

6.4. TRANSFECTION OF MAMMALIAN CELLS

For cell transfection, the jetOPTIMUS transfection reagent (Polyplus, 409-01) was used following the manufacturer's instructions. Briefly, 15,000 cells per cm^2 were seeded the day prior to transfection in order to reach 60-80% confluency for transfection. On the day of the transfection, medium was changed before the procedure. In the case of a 6-well plate, 2 µg of DNA were diluted in 200 µL of jetOPTIMUS buffer and briefly mixed by vortexing. Then, 2 µL of jetOPTIMUS reagent was added to the mix and vortexed again. The transfection mix was incubated for 10 min at RT and then added dropwise to the cells. The following day, cells were trypsinised and seeded in medium containing the appropriate antibiotic to select the transfected cells. Media were changed every day until the control cells (i.e. untransfected cells cultured in presence of antibiotic) were all dead.

6.5. LENTIVIRAL PRODUCTION AND CELL TRANSDUCTION

6.5.1. Generation of lentiviral particles

Lentiviral pLKO shRNAs plasmids (see **section 6.26.4**) were encapsulated in lentiviral particles as previously described [108]. Briefly, HEK 293T cells were transfected with three separate plasmids required for viral production: two packaging vectors (i.e., psPAX2 and pMD2.G) and the pLKO shRNA-containing vector. Cells were seeded at a density of $8x10^6$ cells per dish of 150 mm. Next day, the media were replaced and cells were transfected using polyethylenimine (PEI; Sigma-Aldrich, 408727). DNA (10 µg of each packaging vector and 20 µg of the pLKO-shRNA plasmid) was diluted in 150 mM NaCl to a final volume of 1.6 mL in a 15 ml falcon tube. In 15 mL tubes, 8 µL of 5x PEI (22.5 mg/mL) was diluted in 150 mM NaCl and mixed by gentle vortexing. The PEI-containing solution was then added dropwise to the DNA-containing solution, vortexed and incubated for 30 min at RT. The transfection mix was then added dropwise to the HEK 293T cells. The following day, media were refreshed and supernatants containing the viruses were collected at 48- and 72-h post-transfection. For concentration of the lentiviral particles, collected media were centrifuged at 300 x *g* for 5 min to remove cell debris. The supernatant was then centrifuged at 3,500 x *g* for 30 min at 4ºC using Amicon® Ultra-15 centrifuge filters with a 0.22 µm pore size (Millipore, UFC9030). The concentrated lentiviral particles were aliquoted and stored at -80ºC until use.

6.5.2. Cell transduction

ESCs were trypsinised and collected by centrifugation at 300 x *g* for 5 min. The cell pellet was resuspended in ESM^{LIF} supplemented with 8 ug/mL of polybrene (Sigma-Aldrich, H9268) and divided into the corresponding number of 1.5 ml tubes. Concentrated lentiviral particles were added to each tube in the appropriate amount, ensuring the added volume did not exceed 10% of the total final volume. Then, tubes were incubated in the cell incubator for 1 h with gentle agitation every 10 min to enhance infection efficiency. The infection suspension was then plated and cultured for 24 h under standard conditions supplemented with 8 ug/mL polybrene. The next day, infected cells were selected with 2 ug/mL puromycin (Fisher Scientific, 10054207). Fresh ESM^{LIF} supplemented with the corresponding antibiotic were changed daily until all control non-infected cells died.

6.6. Cell proliferation and viability assessment

For cell proliferation assay, 15,000 cells per cm^2 were seeded. After two days of culture, cells were trypsinised and reseeded in a 1:6 split ratio. The remaining cells were counted using a Neubauer chamber (Hirschmann, 8100103). This process was repeated every other day. To determine the total cell accumulation during the assay, dilutions and split ratios were taken into account.

For assessing the proportion of alive and dead cells, trypsinised cells were diluted 1:1 with 0.4% trypan blue solution (VWR Chemicals, K940). Cells which were stained blue were considered dead cells and unstained cells, alive cells.

6.7. Generation of a knock-in cell line using CRISPR/Cas9

For the generation of the GEN1^{3xFL} ESC line, wild type mouse ESCs were transfected (see **section 5.4**) with an sgRNA against the C-terminal region of *Gen1*, together with a donor DNA of the 3xFL-P2A-NeoR cassette which contained two homology arms identical to the 3´end of the *Gen1 locus* (see **section 6.26** for plasmid generation of sgRNA and donor DNAs). Cells were selected with G418 sulphate (neomycin analogue, Gibco, 11811023) until selection control had died completely. Then, ESC colonies were picked and gDNA and protein were extracted for genotyping and verifying correct DNA editing and knocked-in protein expression.

6.8. Immunofluorescence

For immunofluorescence, cells were seeded in 12-wells plates containing coverslips previously coated with 50 µg/mL poly-D-Lysine (Gibco™, A3890401). After 12 h post seeding, the plate containing the coverslips was placed on ice and the wells were washed with PBS. The PBS was aspirated and pre-extraction was performed with cold 0.2% Triton-X (Thermo Scientific, 327371000) in PBS for 2 min. The pre-extraction was aspirated, and coverslips were washed with cold PBS. Cells were fixed using 4% PFA (Sigma-Aldrich, P6148) for 10 min at RT. Cells were washed twice with PBS after fixation. Cells were permeabilised with 0.2% Triton-X in PBS for 5 min and then washed twice with PBS. Blocking was done using 5% BSA (Fisher Bioreagents, BP9702) in PBS for 1 h at room temperature. Then, cells were incubated with the primary antibody of choice diluted in 1% BSA in PBS for 1 h at RT in a humidity chamber. Cells were washed three times with 0.1% Tween in PBS for 5 min each time. Protected from light, cells were incubated with the appropriate secondary antibody in 1% BSA in PBS for 1 h

at room temperature in a humidity chamber. Next, cells were washed three times with 0.1% Tween in PBS for 5 min each time. Then, they were washed once with PBS for 5 min followed by another wash with milli-Q water for another 5 min. The coverslips were air-dried completely at RT. Coverslips were mounted on the slides using Vectashield® containing 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) (VectorLabs, H-1200-10), to stain nuclei. Coverslips were sealed with nail varnish and air-dried. Images were acquired with ZEISS LSM 880 confocal microscope using a 40x objective. Intensity of each marker was measured using CellProfiler software and normalised to DAPI intensity.

6.9. MICRONUCLEI DETECTION

Cells were seeded on coverslips previously coated with 50 µg/mL poly-D-Lysine (Gibco, A3890401). The plate containing the coverslips was placed on ice and washed twice with PBS. Cells were fixed with 4% PFA for 10 min at RT. Cells were washed twice with PBS after fixation at RT. Cells were permeabilised with 0.2% Triton-X in PBS for 5 min and then washed twice with PBS. Cells were stained with 1 µg/mL DAPI (Sigma-Aldrich, D9542). Then, they were washed once with PBS for 5 min followed by another wash with milli-Q water for another 5 min. The coverslips were air-dried completely at RT. Coverslips were mounted on slides using Vectashield® (Vector Labs, H-1000-10) mounting medium. They were sealed with nail varnish and air-dried. Images were acquired using a ZEISS LSM 880 confocal microscope. The number of micronuclei was scored after checking at least 500 cells per condition.

6.10. METAPHASE SPREADS PREPARATIONS

First, cells were accumulated in metaphase by treatment with 0.2 µg/mL demecolcine (Sigma-Aldrich, D1925) for two h in regular culture conditions (see **section 6.3**). Afterwards, the media were collected, and cells were washed with PBS. Both the medium and PBS were collected. Cells were trypsinised and the enzymatic reaction was stopped by adding the collected media and PBS. Cells were spun at 400 x *g* for 3 min at RT. Then, cell pellet was washed with PBS and spun again. The cell pellet was resuspended in PBS. For the hypotonic shock, prewarmed 0.03 M trisodium citrate [dilution of 2 M sodium chloride (Fisher Scientific, 10316943) and 0.3 M Sodium Citrate Dihydrate (Fisher Scientific, 10797024)] was used for 30 min in a 37ºC water bath. Then, cells were centrifuged at 700 x *g* for 3 min at 4ºC. The supernatant was discarded and the cells were fixed with 10 mL of methanol-acetic acid solution (3:1 ratio) drop by drop

with gentle shaking. After, they were centrifuged at 700 x *g* for 3 min at 4ºC. This fixation step was performed three times. The slides for the spread of the fixed cells were pretreated with acetic acid for 10 s for metaphase fixation to the slides. The fixed cell suspension was dropped onto pretreated slides and air-dried. Slides were stained with Giemsa solution (Sigma-Aldrich, GS500) diluted 1:20 in milli-Q water for 20 min. Then, slides were washed with milli-Q water for 2 min and then air-dried overnight at RT. DePeX mounting medium (VWR/BDH, 361254D) was used to mount the slides. Images were taken using the Leica DM4B and Zeiss AxioObserver Z1 epifluorescence microscopy system. Fifty metaphases were scored per condition.

6.11. COLONY FORMATION ASSAY

Cells were seeded in low density for assessing their clonogenic capacity. 800 cells per well of a 6-well plate or 400 cells per well of a 12-well plate coated with 0.1% gelatine were seeded. Media were changed every day for 4-5 days, when colonies were stained for alkaline phosphatase (AP) activity, an enzyme highly expressed in undifferentiated cells and commonly used as a marker of pluripotency. Staining was performed using an Alkaline Phosphatase kit (Sigma-Aldrich, 86R-1KT) following the manufacturer's instructions. Prior to use, the fixation and staining solutions were prepared as indicated in the table below (**Table 4**).

Table 4. Recipes for the solutions used for AP staining of a 6-well plate.

Fixation solution		
Reagent	**Source/Reference**	**Volume (mL)**
Citrate	Sigma-Aldrich, 915	2.5
Acetone	Fisher Scientific SL, 10314930	6.5
37% Formaldehyde	Sigma-Aldrich, F8775	0.8
Staining solution		
Reagent	**Source/Reference**	**Volume (mL)**
Sodium Nitrite Solution	Sigma-Aldrich, 86R	0.2
FRV-Alkaline Solution	Sigma-Aldrich, 86R	0.2
Mix by gentle inversion. Let it set for 2 min and added it to deionised water		
Deionised water		9
Mix		
Naphtol AS-Bl Alkaline Solution	Sigma-Aldrich, 86R	0.2

Plates containing the cultured cells were washed with Dulbecco's Phosphate Buffered Saline with $MgCl_2$ and $CaCl_2$ (Sigma-Aldrich, D8662). Cells were then treated with fixation solution for 30 s at RT. The fixation solution was removed, and plates were washed with deionised water for 1 min at RT. Cells were stained with the staining solution for 15 min in the

dark at RT, followed by washing again with deionised water for 2 min. Plates were air-dried at RT. AP-positive colonies were counted under a phase-contrast microscope. Undifferentiated (AP-positive), partially differentiated (partially AP-positive stained) or undifferentiated (AP negative) were scored.

6.12. ALKALINE COMET ASSAY

Cells from the corresponding conditions were harvested and counted. 20,000 cells per condition were taken and centrifuged at 300 x *g* for 5 min. Cells were resuspended in 120 µL of 0.5% agarose in PBS and applied to a slide previously coated with 1.5% agarose in PBS. After solidification, slides were incubated in cold Final Lysis solution freshly prepared from the stock lysis solution (recipe in **Table 5**) for one h at 4ºC, protected from light. Afterwards, the slides were transferred to a gel tank filled with final electrophoresis buffer (recipe in **Table 5**). They were incubated for 30 min, and then a current was applied at 25 V for 25 min. The slides were rinsed with neutralisation buffer (recipe in **Table 5**) for 15 min at RT. Then, they were drained and dipped into 100% ethanol for 10 min and air-dried. Slides were stained with 200 µL of 1 µM DAPI solution, covered and sealed with nail varnish. Several images were acquired using a Leica DM4B microscope and Leica LAS X software. Comet scoring was performed using the OpenComet software (v1.3.1) in ImageJ. At least 100 comets were scored per condition, and the olive tail moment, indicative of the comet tail, was used as the representative metric.

Table 5. Alkaline comet assay solutions recipe.

Stock Lysis solution (pH=10)	
Reagent	**Final concentration**
NaCl	2.5 M
EDTA	100 mM
Tris	10 mM
Final Lysis solution (for 100 mL)	
Stock Lysis solution	89 mL
Triton X-100	1 mL
DMSO	10 mL
Final Electrophoresis solution	
NaOH	300 mM
EDTA 200	1 mM
Neutralisation buffer (pH 7.5)	
Tris	0.4 M

6.13. EMBRYOID BODY ASSAY

To examine the role of *Gen1* in early differentiation to the three germ layers (i.e., ectoderm, mesoderm and endoderm) *in vitro*, we employed an embryoid body (EB) formation assay as previously reported [109]. EBs were generated through the spontaneous aggregation of ESCs in suspension culture. For this, ESCs were seeded in low-attachment Petri dishes (Corning, 3472) at a density of 155,000 cells per ml using standard pluripotency medium lacking LIF. Every two days, the culture medium was carefully replaced to avoid losing the EBs while aspirating. Briefly, the medium was collected into centrifuge tubes, and the Petri dishes were rinsed with DPBS (Corning, 21-031-CV). The pooled DPBS–medium mixture was left undisturbed at room temperature for 15 min to allow EB structures to settle, minimising disruption. The supernatant was then gently aspirated, and the EBs were resuspended in fresh LIF-free medium before being transferred to new Petri dishes. RNA samples were harvested on days 0, 3, and 6 of EB formation to monitor induction of differentiation by RT-qPCR (see **sections 6.20** and **6.21**).

6.14. REPROGRAMMING ASSAY

To assess the impact of *Gen1* loss of function on SCR toward iPSCs, we employed the mouse fibroblast line W3T3 harbouring a doxycycline-inducible *Oct4-Sox2-Klf4-cMyc* (i.e., OSKM) polycistronic cassette. *Gen1* expression was first silenced by knockdown (KD), and the selected cells were seeded at a density of 800 cells per well in 6-well plates. The following day, cultures were switched to ESM^{LIF} and treated with 2 µg/mL doxycycline (DOX) to induce the expression of the OSKM reprogramming factors. Media were refreshed daily, maintaining DOX treatment for 14 days. Emerging colonies were subsequently stained using a commercial AP detection kit (see **section 6.11**), and RNA was harvested for RT-qPCR analysis (see **sections 6.20** and **6.21**).

6.15. SYNCHRONISATION OF CELLS

For synchronisation in G1/S, a double thymidine (Thy) block method was employed [110]. Briefly, cells were seeded at a density of 30,000 cells per cm^2 in regular culture conditions. Four h after seeding, 2 mM Thy (Sigma-Aldrich, T1895) was added to the culture. After 16 h of treatment, the media were replaced with fresh, Thy-free ESM^{LIF} for 8 h. Following this incubation, 2 mM Thy was added again for another 16 h. Cells were then released from the

block by replacing the Thy-containing medium with fresh ESM^{LIF}, and samples were collected at different time points post-release for cell cycle analysis (see **section 6.16**).

6.16. CELL CYCLE ANALYSIS

For the analysis of the cell cycle of ESCs, cells were collected by trypsinisation after the desired treatment (e.g. synchronisation of cells) at the chosen time point. Cells in suspension were centrifuged for 5 min at 300 x *g* and washed with PBS. After removing the PBS, cells were fixed with ice-cold 70% ethanol, drop by drop with gentle shaking and then incubated for 30 min at 4ºC. After fixation, cells were tempered at room temperature and centrifuged again at 500 x *g* for 5 min. The cell pellet was washed by resuspension in PBS and centrifugation at 500 x *g* for 5 min, twice. Afterwards, the cell pellet was resuspended in the staining solution containing 50 µg/mL Propidium Iodide (PI; Sigma-Aldrich, P4864) and 10 µg/mL RNase A (Thermo Scientific, EN0531) in PBS to remove all RNA, which can also be stained by PI and thus give background noise. Samples were transferred to FACS tubes and incubated at 37ºC for 30 min in the dark. Samples were analysed by flow cytometry using a FACSCalibur (BD Biosciences) equipped with and 488 nm laser and FL2 (585/42 nm) detector. The data generated were analysed using FlowJo software using Watson pragmatic model.

6.17. CELL VIABILITY ASSAY

Cell viability was assessed using the MTT assay (3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide; Sigma-Aldrich, M5655). ESCs were seeded at a density of 5,000 cells per well in 96-well plates. The following day, culture medium was replaced with fresh medium containing the genotoxic treatment of choice (**Table 6**). After 24 h, cells were transferred to treatment-free medium for another 24 h. Viability was then evaluated by incubating cultures in ESM^{LIF} supplemented with 0.5 mg/mL MTT for 4 h under standard conditions in the dark. Following incubation, the medium was removed by aspirating, and the resulting formazan crystals were solubilised in a 1:1 solution of isopropanol and DMSO by shaking the plates at 225 rpm for 15 min at RT. Absorbance was measured at 570 nm and 690 nm to correct for background noise using a Mithras LB 940 reader (Berthold Technologies). Cell viability, inferred from mitochondrial-dependent formazan production, was calculated by normalising absorbance values to vehicle-treated controls.

Table 6. Genotoxic reagents for viability assay in ESCs.

Treatment	Brand	Reference
4-Nitroquinoline 1-oxide (4-NQO)	Sigma-Aldrich	N8141
Bleomycin	Selleck Chemicals	S1214
Camptothecin (CPT)	Selleck Chemicals	S1288
Cisplatin	Sigma-Aldrich	P4394
Etoposide (Eto)	Sigma-Aldrich	E1383
Hydroxyurea (HU)	Sigma-Aldrich	H8627
Olaparib	Selleck Chemicals	S1060

6.18. ANNEXIN V APOPTOSIS ASSAY

Cell apoptosis was assessed using the Dead Cell Apoptosis Kit with Annexin V Alexa Fluor 488 (AF488) and PI (Fisher Scientific, Molecular Probes™ V13245) according to the manufacturer's protocol. ESCs were transfected as described in **section 6.4**. At the indicated time point, cells were counted and resuspended at 10^5 cells per 100 µL of assay buffer. Stained cells were analysed by flow cytometry in a CytoFLEX S (Beckman Coulter) using the 488 nm laser with B525/40 BP for Annexin V (AF488), and the 561 nm laser with Y690/50 BP for PI. Data were gated to exclude cell debris and to distinguish Annexin $V^{+/-}$ and $PI^{+/-}$ populations, including double-positive cells, using FlowJo software.

6.19. ATP MEASUREMENTS ASSAY

The total and mitochondrial ATP levels of ESCs in sh*Luci* or sh*Gen1* conditions were determined following published protocols [111]. ATP levels were measured using the kit CellTiter-Glo Luminescent Cell Viability Assay (Promega, G7570) according to the manufacturer's instructions. Briefly, cells in cultured were incubated for 2 h in measurement medium (156 mM NaCl, 3 mM KCl, 2 mM $MgSO_4$, 1.25 mM KH_2PO_4, 2 mM $CaCl_2$, 20 mM HEPES, pH 7.35) with either 5 mM glucose (to measure total ATP production) or 5 mM of D-deoxyglucose and 110 mg/mL pyruvate (to measure mitochondrial ATP production). Afterwards, plates were equilibrated for 15 min at RT. Cells were lysed with the mix of the substrate and buffer for 2 min. The lysates were incubated at RT for 10 min to stabilise the luminescence signal. Samples were measured at 560 nm using a CLARIOstar Plus (BMG Labtech) microplate reader.

6.20. RNA ISOLATION

Total RNA was extracted using two different methods. Either EZNA Total RNA Kit I (Omega Bio-Tek, R6834-02) or TRIzol reagent (Invitrogen, 12044977) was used following the respective manufacturer's instructions. In brief, 600 µL of TRIzol reagent per well of a 6 well plate, cells were scraped and transferred into centrifuge tubes. Chloroform (FisherScientific, 10488400) was then added at 200 µL per mL of TRIzol and the mixture was vigorously shaken for 15 s. The suspension was moved to Phasemaker tubes (Invitrogen, A33248), incubated at RT for 3 min, and centrifuged at 12,000 x *g* and 4ºC for 15 min. The upper aqueous phase containing the RNA was transferred to a fresh tube containing isopropanol (Fisher Scientific, 11398461) at 500 µL per mL of TRIzol. Samples were gently mixed, incubated for 10 min at RT and centrifuged again at 12,000 x *g* and 4ºC for 10 min. The precipitated RNA pellet was then washed with 1 mL of ice-cold ethanol (Sigma-Aldrich, 108543), centrifuged at 7500 x *g* and 4ºC for 5 min and supernatant was removed. The RNA pellet was air-dried for upt to 5 min and dissolved in 20-50 uL of UltraPure DNase/RNase-Free distilled water (Invitrogen, 10977-035). RNA concentration and quality were determined by NanoDrop 2000 Spectrophotometer (Thermo Scientific, ND2000). RNA was stored at -80ºC until use.

6.21. ANALYSIS BY REVERSE TRANSCRIPTION QUANTITATIVE POLYMERASE CHAIN REACTION (RT-QPCR)

First, complementary DNA (cDNA) was synthesised from total RNA using the qSCRIPT cDNA SuperMix kit (Quantabio, 95048) following the manufacturer's instructions. Briefly, 1 ug of RNA was used as template mixed with 4 µL of the qScript cDNA SuperMix to a final volume of 20 µL. Samples were vortex gently to mix contents. Reverse transcription was performed using a SimpliAmp Thermocycler System (Applied Biosystems, A24811) following the next protocol: 5 min at 25ºC, 30 min at 42ºC and 5 min at 85ºC. Quantitative PCR (qPCR) was performed using the PowerUp SYBR Green Master Mix (Applied Biosystems, A25742) with gene-specific primers (listed in **Supplementary Table 4**), together with the appropriate cDNA sample. qPCR reactions were performed in 384-well skirted PCR microplates (Axygen, PCR-384M2-C) sealed with microplate sealing film and tapes (Axygen, UC-500) in a QuantStudio 5 qPCR system (Applied Biosystems, A33621) using the program indicated in **Table 7**.

Table 7. qPCR thermocycler conditions.

Temperature	Time	Number of cycles
95 ºC	10 min	1
95 ºC	15 s	40
55 ºC	11 s	
72 ºC	40 s	
95 ºC	5 s	1
65 ºC	1 min	Increase 5ºC/min until reaching 97ºC
97 ºC	15 s	1
4 ºC	hold	-

For the analysis, the average threshold cycle (Ct) was determined from triplicates, and the data were analysed using the delta-delta Ct (ΔΔCT) method [112]. Expression levels were normalised to the values of a housekeeping gene (e.g., *Actb*) and expressed as relative expression levels to the indicated control samples (e.g., WT or sh*Luci*).

6.22. GENOMIC DNA ISOLATION

Genomic DNA (gDNA) was extracted using the Centra Puregene Cell Kit (Qiagen, 158388). Approximately 10^6 cells were resuspended in 200 µL of Cell Lysis Solution (Qiagen, 158908) and incubated for 1 h at room temperature (RT). Then, 1 µL of RNase A (1 mg/mL, Thermo Scientific, EN0531) was added to the samples and they were incubated for 1 h at 37ºC. Samples were incubated for 5 min on ice to cool down and 100 µL of Protein Precipitation Solution (Qiagen, 158912) was added to each sample. Samples were vortexed for 20 s and then centrifuged for 10 min at RT and 16,000 x *g* to remove the protein fraction. The supernatant was added to tubes containing 200 µL of isopropanol (Fisher Scientific, 11398461) for DNA precipitation and mixed by inverting the tube fifty times. Samples were centrifuged at 1,6000 x *g* for 5 min at RT. The supernatant was discarded, and the pellet was washed with 200 µL 70% ethanol (Sigma-Aldrich, 108543). Samples were centrifuged again at 16,000 x *g* for 1 min. The supernatant was discarded and allowed to air-dry for 5 min. Finally, the DNA pellet was rehydrated with 50 µL of DNA Hydration solution (Qiagen, 158916) and samples were incubated for 1 h at 65ºC to properly dissolve the DNA followed by overnight incubation at RT. Genomic DNA concentration and purity were determined by NanoDrop 2000 Spectrophotometer (Thermo Scientific, ND2000). DNA was stored at 4ºC until use.

6.23. CELL GENOTYPING

The genotype of GEN1^{3xFL} ESCs was confirmed by PCR using gDNA extracted from individual clones (see **section 6.22**). Primers were designed so they would lead to the amplification of two products of different sizes, a smaller band of 1,619 bp for the WT genotype and a bigger one of 2,598 bp for the KI allele. DreamTaq Green PCR Master Mix (Thermo Scientific, K1081) was used for the PCR. Results were visualised by electrophoresis on a 1% agarose gel prepared with 1xTAE buffer and stained with EtBr. Sequences of the genotyping primers are listed in **Supplementary Table 4**.

For genotyping ESCs transfected to generate GEN1 KO via CRISPR/Cas9, PCR was performed using gDNA extracted from individual clones (see **section 6.22**). Three primers, two forward and one reverse, were designed to generate two PCR products with different sizes: 181 and 226 bp. DreamTaq Green PCR Master Mix (Thermo Scientific™, K1081) was used for the PCR. Results were visualised in a 2.5% agarose/synergel in 1xTBE buffer with EtBr to detect the DNA.

6.24. PROTEIN EXTRACTION

6.24.1. Whole-cell protein extracts

Whole-cell protein extracts were obtained by cell lysis with RIPA Lysis Buffer (Santa Cruz Biotechnology, sc-24948) with freshly added protease and phosphatase inhibitors (Santa Cruz Biotechnology, sc-24948). Trypsinised cells were resuspended with RIPA at 100 µL per two million cells and incubated on ice for 30 min with vortexing every 10 min. Samples were centrifuged at 10,000 x *g* for 10 min at 4ºC to remove cell debris. The supernatant containing the soluble protein fraction was transferred to a new tube. Protein concentration was determined with Pierce BCA protein assay kit (Thermo Scientific, 23227) following the manufacturer's instructions. Samples were stored at -80ºC until use.

6.24.2. Subcellular fractionation protein extracts

Subcellular localisation of the desired proteins was analysed using the Subcellular Protein Fractionation kit for Cultured Cells (Thermo Scientific, 78840) following the manufacturer's instructions. Cultured cells were harvested and counted to obtain between 1-10 million of cells per condition before proceeding with the protocol. The extracts obtained included cytoplasmic, membrane, nuclear soluble, chromatin-bound and cytoskeletal protein fractions. Protein

concentration was determined with Pierce BCA protein assay kit (Thermo Scientific, 23227) following the manufacturer's protocol, and independent calibration curves were obtained for each fraction, by adding the appropriate buffer to the protein standards. Samples were stored at -80ºC until use.

6.25. WESTERN BLOT

For Western blot analysis, homemade Laemmli buffer [40% glycerol (Sigma-Aldrich, G5516), 240 mM Tris-HCl pH 6.8, 8% SDS (Fisher Scientific, 10607443), 0.04% bromophenol blue (Sigma-Aldrich, B8026) and 5% β-mercaptoethanol (Sigma-Aldrich, M-6250)] was added to protein lysates and heated for 5 min at 95ºC for protein denaturation. Denatured protein samples were loaded onto wells of homemade 8-12% acrylamide/bis-acrylamide (Sigma-Aldrich, A3574) gels or commercial Novex WedgeWell 4-20% tris-Glycine Gel (Invitrogen, XP04205BOX). Proteins were resolved by electrophoresis for 1.5 h with a voltage of 95V. Then, proteins were transferred to polyvinylidene difluoride (PVDF) membranes which were activated with methanol (Fisher Scientific, 10675112). For semi-dry transfer, Immun-Blot PVDF membrane (BioRad, 1620177) were used with an intensity of 0.2 A for 1h 45 min. For wet transfer, Amersham Hybond PVDF membrane (Cytiva, 10600023) were employed, and the transference was carried out at 4ºC overnight at 35 V. Afterwards, membranes were washed with milli-Q water and stained for 5 min at RT with 0.1% Ponceau S solution containing Ponceau S reagent (Sigma-Aldrich, P3504),) and 5% acetic acid glacial (Fisher Scientific, 10744361) diluted in water for 5 min. Then, membranes were washed again with milli-Q water to remove unspecific staining.

For the blotting, the membranes were rinsed with 1xTBS buffer (Fisher Bioreagents, BP2471-1) supplemented with 0.1% Tween® 20 (Sigma-Aldrich, P1379) (TBS-T) three times for 5 min each one, followed by blocking with 5% skimmed milk (Central Lechera Asturiana) in TBS-T for 1 h at RT. Membranes were washed again 3 times with gentle shaking for 15 min each time in TBS-T. Afterwards, they were incubated with specific primary antibodies diluted in 0.5% skimmed milk in TBS-T for 1 h or overnight at 4ºC, depending on the antibody instructions (see **Supplementary Table 6**). Then, membranes were washed again 15 min with TBS-T for three times. After, they were incubated with HRP-conjugated secondary antibodies in 0.5% skimmed milk in TBS-T for 45 min at 4ºC. Finally, the membranes were rinsed three times for 15 min each time in TBS-T and protein signals were detected by incubation with SuperSignal

West Pico PLUS Chemiluminescent substrate (Thermo Scientific, 34580) or Immobilon Western Chemiluminescent HRP substrate (Sigma-Aldrich, WBKLS0100) followed by image processing using Chemidoc imaging system (BioRad, 12003153). Dry membranes were stored in transparent plastic foil for potential future use.

6.26. PLASMID CLONING

All used oligonucleotides and plasmids are listed on **Supplementary Tables 4** and **5**, respectively.

6.26.1. Cloning of protein-coding sequences into expression plasmids

The coding DNA sequence of the human *GEN1* gene (NM_001130009) was obtained from a pENTRY plasmid kindly provided by Dr Miguel G Blanco's laboratory (Universidade de Santiago de Compostela, Spain). The coding sequence was cloned into a *piggybac* expression plasmid using the Gateway cloning method. Briefly, the pENTRY plasmid containing the gene was incubated with the pDEST vector (pBASE-Empty Vector-BSDr, provided by Dr Jose Silva group (University of Cambridge, UK) and the Gateway LR Clonase II Enzyme mix (Invitrogen, 11791020) for 1 h at RT. Then, 10 µL of the recombination reaction was transformed into competent *Stbl3* bacterial cells for plasmid isolation (see **section 6.28**).

6.26.2. Cloning of sgRNAs into a Cas9-expressing plasmid

To KO the function of HJ resolvases/dissolvase in ESCs using the CRISPR/Cas9 system, two sgRNAs targeting each gene of interest were designed with the CRISPOR online tool (https://crispor.gi.ucsc.edu/) taking into account the highest specificity score minimising the number of off-targets. The sgRNA oligos were purchased to Merck as 100 µM solutions in milli-Q water. First, 2.5 uM of the forward and reverse sgRNAs oligos were used for annealing and phosphorylation using T4 PNK (Thermo Scientific, 10120670). Samples were incubated in a thermocycler for 30 min at 37ºC, 5 min at 95ºC, the temperature was reduced by 5ºC every min until reaching 25ºC, followed by stabilisation at 4ºC. The sgRNA duplexes were diluted 1:100 for ligation and cloned into the pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene #62988) backbone. Prior to ligation, the backbone was digested with the restriction enzyme *BbsI* (Thermo Scientific, 10569110) to generate compatible cohesive ends and purified by electrophoresis on a 0.7% agarose gel using the GenElute gel extraction kit (Sigma-Aldrich, NA1111). Oligo duplex and backbone were ligated using T4 DNA ligase (Thermo Scientific,

10786591) following the manufacturer's protocol and then transformed into *Stbl3* competent bacterial cells for plasmid isolation (see **section 6.28**). The integration of the guides was verified by Sanger sequencing at StabVida.

6.26.3. Cloning of shRNAs into transfection plasmids

For gene silencing, we used short hairpin RNAs (shRNAs). Oligonucleotides containing the desired target sequence were designed based on public repository, Broad Institute GPP (https://portals.broadinstitute.org/gpp/public/). Preferably, we chose those shRNAs with the highest proven efficiency and those that target the UTR of the gene of interest. The two complementary oligonucleotides per shRNA of interest were annealed in a reaction containing 2.5 µM of each oligonucleotide and 1x NEB buffer 2 (New England BioLabs, B7002S). For that purpose, the reaction was incubated in a thermocycler for 10 min at 95ºC, 10 min at 70ºC and then reducing the temperature 1ºC every 2.5 min until reaching 25ºC and stabilising them at 4ºC. Since the resulting dsDNA duplexes possess cohesive restriction enzyme sites of *AgeI* (Thermo Scientific, 10860241) and *EcoRI* (Thermo Scientific, 10374340), they were cloned into a pHSVg plasmid for shRNA expression under U6 promoter. The vector was digested with *AgeI* (Thermo Scientific, 10860241) and *EcoRI* (Thermo Scientific, 10374340) at the same time for 15 min at 37ºC, resolved by electrophoresis on a 0.7% agarose gel and purified using the GenElute gel extraction kit (Sigma-Aldrich, NA1111). Then, the dsDNA duplexes were subcloned into the vector using the T4 DNA ligase (Thermo Scientific, EL0016). The ligation product was transformed into *Stbl3* competent bacterial cells for plasmid isolation (see **section 6.28**). The following day, ampicillin-resistant bacterial clones containing the shRNA-expressing plasmids were verified by colony PCR using the DreamTaq Green PCR Master Mix (Thermo Scientific, K1081) using the following thermocycler conditions (**Table 8**).

Table 8. Thermocycler conditions for the colony PCR using the DreamTaq Green PCR Master Mix.

Temperature (ºC)	Time	Number of cycles
95	2 min	1
95	30 s	30
60	30 s	
68	30 s	
68	5 min	1
4	hold	-

PCR amplicons were analysed by electrophoresis using a 2.5% (w/v) agarose (Fisher Bioreagents, 10776644) with synergel (Diversified Biotech, DIVESYN-100) and ethidium

bromide (EtBr; Fisher Bioreagents, BP1302-10) and visualised using the BioRad GelDoc XR+ system (BioRad, 1708195EDU). Positive colonies were selected for isolation of plasmid DNA (see **section 6.28**). Finally, the integration of the shRNAs was verified by Sanger sequencing at StabVida.

6.26.4. Cloning of shRNAs into pLKO lentiviral vectors

The pLKO.1 empty vector (pLKO.pim kindly provided by Dr Ihor R. Lemischka group (former professor at Icahn School of Medicine at Mount Sinai, USA) was digested with *AgeI* (Thermo Scientific, 10860241) and *EcoRI* (Thermo Scientific, 10374340) restriction enzymes at 37ºC for 15 min. The digested plasmid was resolved by electrophoresis on a 0.7% agarose gel and purified using the GenElute gel extraction kit (Sigma-Aldrich, NA1111) according to the manufacturer's instructions. Annealed oligonucleotide duplexes (obtained as in **section 6.26.3**) were ligated into the digested pLKO.1 vector using the T4 DNA ligase (Thermo Scientific™, EL0016). The ligation product was transformed into *Stbl3* competent bacterial cells for plasmid isolation (see **section 6.28**). Proper cloning of the shRNAs was validated by PCR and Sanger sequencing at StabVida.

6.26.5. Generation of a donor plasmid for knock-in 3xFL-P2A-NeoR cassette

For the generation of the donor plasmid for the knock-in (KI) at the C-terminal of *Gen1 locus*, we used the GeneArt Gibson Assembly HiFi Master Mix (Invitrogen, A46627) following the manufacturer's instructions. Four DNA fragments were generated for assembly. The insert fragment, containing 3xFL-P2A-NeoR, was amplified from the pFETCH-donor (Addgene #63934) with oligos designed to incorporate overlaps to the homology arms, using the Platinum SuperFi II PCR Master Mix (Invitrogen, 12369010). The two 800 bp homology arms were amplified from mouse CJ7 ESCs gDNA, with primers containing overlaps matching both the insert and the plasmid backbone. For the backbone, the pFETCH-donor plasmid was digested with *BbsI* (Thermo Scientific, 10569110) and *BsaI* (Thermo Scientific, 10232740) restriction enzymes. All amplified or digested products were separated on a 0.7% agarose gel in 1xTAE and extracted using the GenElute gel extraction kit (Sigma-Aldrich, NA1111). Equimolar amounts of the four fragments were incubated together in the presence of Gibson Master Mix for assembly, after which 10 µL of the reaction was transformed into *Stbl3* competent cells for plasmid isolation (see **section 6.28**). Recombinants were screened by diagnostic cut with *XbaI* (Thermo Scientific, 10796601) following manufacturer's instructions and resolved on a 0.8%

agarose in 1xTAE at 85 V. Correct recombinants were identified by the presence of two bands at 2123 and 5009 bp.

6.27. SITE-DIRECTED MUTAGENESIS

To generate the catalytic mutant D157A in human GEN1, we performed site-directed mutagenesis to switch a A>C resulting in an amino acid substitution (D157A). Two complementary oligonucleotides were designed with the desired point mutation in the middle of the sequence (see **Supplementary Table 4**). We used the Platinum SuperFi II PCR Master Mix (Invitrogen, 12369010) for the mutagenesis PCR using the pENTRY plasmid containing the coding sequence for human GEN1 (NM_001130009) which was kindly provided by Dr Miguel G Blanco (Universidade de Santiago de Compostela, Spain). The PCR product was treated with *DpnI* (Thermo Scientific, 10809410) to remove the original plasmid. Afterwards, the PCR plasmid was transformed into *Stbl3* competent bacteria for plasmid isolation (see **section 6.28**). The presence of the desired mutation was confirmed by Sanger sequencing at StabVida.

6.28. ISOLATION OF PLASMID DNA

6.28.1. Bacterial transformation

Plasmid DNA was incubated with competent bacterial cells (*DH5α* or *Stbl3*) at a volume not exceeding 10% of the total cell suspension volume and homogenised by gentle rotation of the tube. The cell suspension was incubated on ice for 30 min followed by a heat-shock at 37ºC for 30 s for the *DH5α* bacterial strain or 42ºC for 1 min for *Stbl3* bacterial strains. Immediately afterwards, the transformed cells were incubated for 5 min on ice. Then, 450 µL of homemade Super Optimal broth with Catabolite repression (SOC; 20 g Tryptone, 5 g yeast extract and 0.5 g NaCl (Fisher Scientific, 10316943) dissolved in a L of distilled water) medium was added to the transformation mix. Bacteria were allowed to recover for one h at 37ºC with shaking at 225 rpm. After incubation, 100 µL of the mix was spread into each plate of Luria broth (LB)-agar with the appropriate antibiotic for selection. The plates were incubated overnight at 37ºC to allow for growth of transformed bacteria.

6.28.2. Plasmid DNA extraction

One single resistant colony was picked from the LB-agar plate with antibiotic and inoculated into a tube containing 8 mL of LB medium with the desired antibiotic for plasmid selection. The culture was kept at 37ºC and shaking 225 rpm overnight to allow for bacteria growth. The next day, the plasmid DNA was isolated from the bacterial culture using the GenElute plasmid miniprep kit (Sigma-Aldrich, NA0150) following the manufacturer's instructions. For larger plasmid DNA yields, bacterial cultures of 100 mL were grown and plasmid DNA was extracted using the PureLink HiPure Plasmid Filter Midiprep kit (Invitrogen, K210015) according to the manufacturer's instructions. The plasmid DNA concentration and quality were determined using a NanoDrop 2000 Spectrophotometer (Thermo Scientific, ND-2000).

6.29. RNA SEQUENCING

For total RNA sequencing of ESCs transfected with sh*Gen1* or sh*Luci* as control, total RNA was extracted from TRIzol as described in **section 6.20**. RNA quality was verified using an Agilent Bioanalyzer 2100 by NOVOGENE, followed by library construction with Ribo-Zero® kit (Illumina, 20040525). RNA-seq libraries were sequenced on a HiSeq 4000 platform using a 150 bp paired-end, unstranded protocol.

Raw data was analysed in our laboratory by Alba Cortés Coego (differential gene expression analysis, DEG), and Fernando Gálvez (transposable element analysis), following optimised pipelines. For DEG, sequencing adaptors were removed using the software Trimmomatic (v0.39) and the quality of the reads was assessed using FastQC (v0.11.5) software. For read alignment to the mouse genome (Ensembl GRCm39) and removal of sequencing duplicates, the software packages STAR (v2.7.6) and Picard-tools (v2.27.2) were used. After the filtering, reads were converted into gene counts using HTSeq (v0.13.5). For the analysis of the gene counts, we used R (v4.2.2) and, specifically, the DESeq2 (v1.32.0) package. The resulting data were annotated with gene ID, chromosome, start and end position, strand and transcript type with the BiomaRT (v2.52.0) package.

For TE analysis, the TEtranscripts software package was used [113]. Initially, reads were trimmed and quality filtered using the fastp program. Subsequently, they were aligned to the mouse genome GRCm39 with the STAR aligner. TE transcripts were quantified using

TEtranscripts. TE transcript counts were adjusted for library size to facilitate statistical modelling and were assessed for differential expression using DESeq2, as previously mentioned. Sample-specific size factors were applied to account for variations in library size during TE analyses. Additionally, SQuIRE (v0.9.9.92) was used for *locus*-specific TE expression [114]. Reference repeat files were downloaded and purged of redundant entries using SQuIRE Fetch and SQuIRE Clean commands, respectively. Quality control of raw RNA-seq reads was assessed using FastQC (v0.11.5) and were trimmed for adapters and low-quality sequences using Trim Galore (v0.6.7) with a Phred quality score threshold of >20. Filtered reads were aligned to the mouse reference genome (GRCm39/mm39) using SQuIRE Map, which internally calls STAR (v2.5.3) to generate coordinate-sorted BAM files. Quantification was performed at both gene (RefSeq) and *locus*-specific TE levels using SQuIRE Count. Differential expression analysis was conducted with SQuIRE Call, which leverages DESeq2 (v1.32.0). TEs with Benjamini-Hochberg adjusted p-values < 0.05 were considered significantly deregulated.

6.30. External datasets

All published external datasets used for different expression analysis are listed on **Supplementary Table 7**.

6.31. Gene ontology and KEGG pathway analysis

The functional annotation tool of The Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) was used for gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis using the misregulated genes obtained after RNA-seq analysis. The GO terms and KEGG pathways used are listed on **Supplementary Table 8**.

6.32. Statistical analysis

GraphPad Prism or R were used for statistical analysis. The corresponding test which was used for determining the statistical significance is specified in each figure legend. An FDR or p-value below 0.05 was considered statistically significant.

6.33. BIBLIOGRAPHIC RESOURCES AND USE OF AI-ASSISTED TOOLS

PubMed (https://pubmed.ncbi.nlm.nih.gov/) was consulted as the main bibliographic source for this manuscript. Bibliographic references were managed using EndNote Web (https://web.endnote.com/). In the preparation of this thesis, AI-assisted technology, specifically ChatGPT (https://chatgpt.com/), was used to support the refinement of English language expression. All text generated with the assistance of this tool was subsequently reviewed, edited, and verified by me. I assume full responsibility for the accuracy, integrity, and final content of this doctoral thesis.

6.34. SOFTWARE AND ALGORITHMS

Table 9. List of softwares and algorithms

Software	Source	Hyperlink
BiomaRT v2.52.0	[115]	https://bioconductor.org/packages/
Cluster v3.0	Stanford University	http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm
CRISPOR	[116]	http://crispor.tefor.net/
DAVID	LHRI laboratory	https://david.ncifcrf.gov/tools.jsp
DESeq2 v1.32.0	[117]	https://bioconductor.org/packages/release/
FastQC v0.11.5	Babraham Institute	https://www.bioinformatics.babraham.ac.uk/%20projects/fastqc
FlowJo v7.6.1	FlowJo, LLC	https://www.flowjo.com/
Genetic Perturbation Platform (GPP)	Broad Institute	https://portals.broadinstitute.org/gpp/public/
GraphPad Prism 9	GraphPad	https://www.graphpad.com/
HTseq-count v0.13.5	[118]	https://htseq.readthedocs.io/en/master/
Image J v1.33	National Institutes of Health	https://imagej.nih.gov/ij
Leica LAS X v3.4.2.18368	Leica Microsystems	https://www.leica-microsystems.com/
OpenComet	[119]	https://cometbio.org/index.html
Picard-tools v2.27.2	Broad Institute	https://broadinstitute.github.io/picard/
R (v4.2.2)	R Foundation for Statistical Computing, Vienna, Austria.	https://www.r-project.org/
Sequencher 5.4	Gene Codes Corporation	http://www.genecodes.com/
SerialCloner v2.6.1	Serial Cloner	http://serialbasics.free.fr/Serial_Cloner.html
SQuIRE (v0.9.9.92)	[114]	https://github.com/wyang17/SQuIRE
STAR v2.7.6	[120]	https://github.com/alexdobin/STAR
TEtranscripts	[113]	https://hammelllab.labsites.cshl.edu/software/
The Universal Protein Resource (UniProt)	The UniProt consortium and host institutions EMBL-EBI, SIB and PIR	https://www.uniprot.org/
Trim Galore (v0.6.7)	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
Trimmomatic v0.39	[121]	http://www.usadellab.org/

7. Results

7.1. HJ PROCESSORS, BLM AND GEN1, SHOW ELEVATED RNA EXPRESSION IN ESCS

Several studies have previously shown that pluripotent cells display higher expression of proteins involved in the HR DNA repair pathway, compared to more differentiated cells [65]. In order to further confirm this observation, we used a published RNA-seq dataset of ESCs and MEFs as pluripotent and differentiated/somatic cells, respectively [122]. Using gene ontology (GO) terms for genes involved in HR and NHEJ regulators (GO: 0000724 and GO: 0006303, respectively), we observed that ESCs express higher levels of genes related to the HR pathway compared to somatic cells (**Figure R1A**), as previously described by others [65, 123]. As an example, increased levels of the main HR players *Mre11a*, *Brca1*, *Brca2* or *Rad51* (**Figure I6**) could be observed in ESCs compared to MEFs (**Figure R1B**). In the case of NHEJ, there was no general increase in expression of all its regulators, with specific genes expressed at higher levels in pluripotent cells (e.g., *Xrrc6* and *Dclre1c*) while others showed no changes in RNA level (e.g., *Lig4*, *Xrcc4* and *Nhej1*). In contrast, we observed an increase in RNA levels of genes related to NHEJ in differentiated cells, suggesting that a prominent role of HR over NHEJ is a pluripotent-specific feature.

We then used a liquid chromatography-tandem mass spectrometry from a published data of embryonic stem cells [124] to address the protein abundance levels of these regulators at the protein level (**Figure R1C**). We observed that, although RNA expression levels vary considerably, for instance, *Brca1* shows approximately threefold higher expression than *Brca2* and *Rad51*, no changes at protein levels were observed. This indicates that differences in RNA abundance do not directly translate into differences in protein expression, suggesting the involvement of post-transcriptional regulatory mechanisms that differentially modulate genes involved in DNA repair. These findings highlight that analysis at the RNA level alone does not provide a complete picture of their regulation.

The different expression patterns of HR genes and NHEJ in pluripotent versus differentiated cells are further confirmed by pluripotency exit (i.e., differentiation) and acquisition (i.e., somatic cell reprogramming).Taking advantage of a published dataset of ESCs differentiation to embryoid bodies (EBs; [125]), which enables ESCs to spontaneously differentiate into derivatives of all three germ layers, we observed that exit from pluripotency leads to a progressive decrease in expression of HR proteins, whereas the NHEJ-related genes remain unchanged (**Figure R1D**). In agreement with this, the analysis of a published dataset of

MEFs reprogramming [126] showed that the acquisition of pluripotency by SCR of MEFs leads to a higher expression of all the repair-related genes, reinforcing the idea that higher expression levels of repair proteins are a specific characteristic of pluripotent cells (**Figure R1E**).

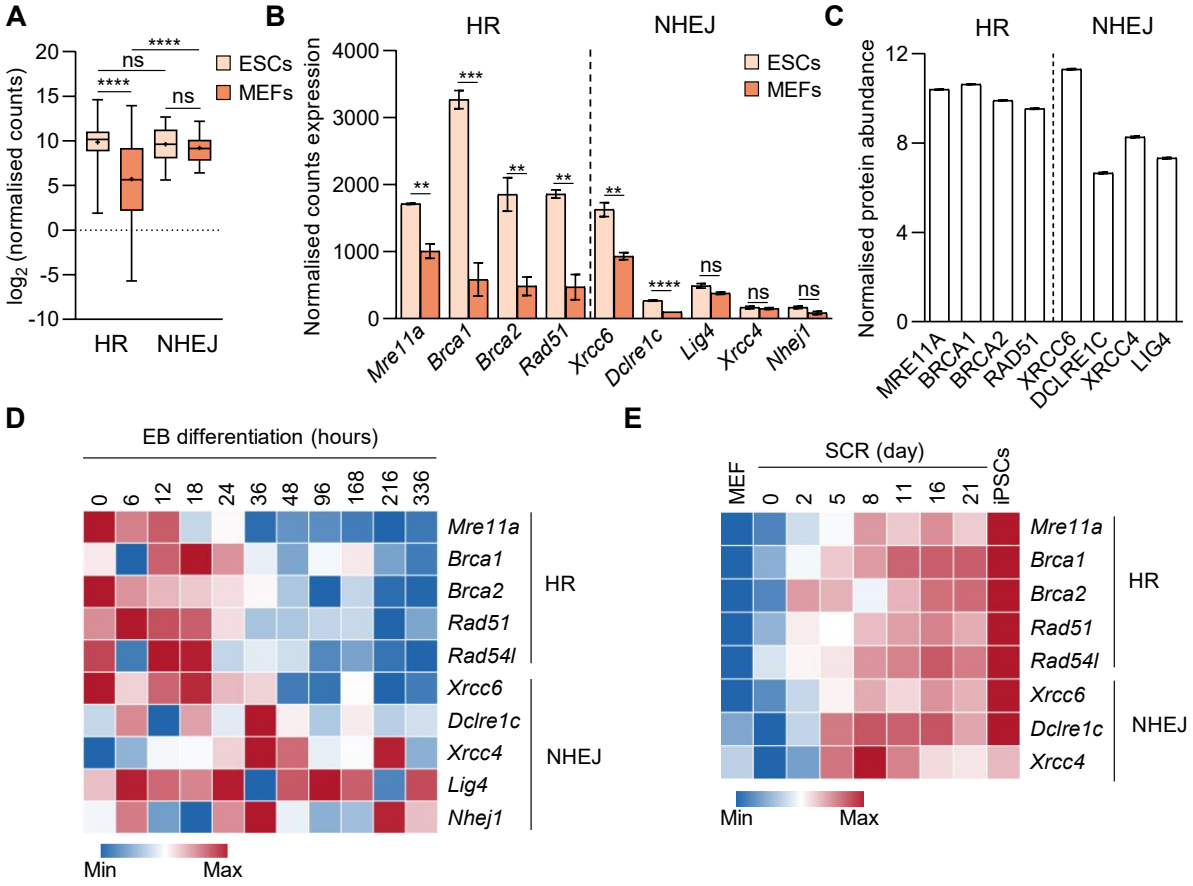

Figure R1. ESCs express higher levels of homologous recombination proteins compared to non-homologous end joining proteins. (A) Box plot depicting the of RNA expression levels of genes included in the HR (GO:0000724, n=160) and NHEJ (GO:0006303, n=63) gene ontology pathways, from a published RNA-seq dataset (GSE69823). **(B)** Normalised RNA expression levels of the indicated DNA repair proteins from a published RNA-seq dataset (GSE69823). **(C)** Normalised protein abundance in mouse ESCs of the indicated HR and NHEJ factors identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS) from a published dataset (ProteomeXchange: PXD033001). **(D)** and **(E)** Heatmap of the expression pattern of *Blm*, *Mus81* and *Gen1* during exit from pluripotency by EB differentiation (GSE3749) and acquisition of pluripotency by somatic cell reprogramming (SCR; GSE21757) using published datasets, respectively. Statistical significance was determined by unpaired t-test: *: p-value<0.05; **: p-value<0.01; ***: p-value<0.001; ****: p-value<0.0001 and ns: non-significant.

During HR, several DNA repair intermediates arise, presenting a potential damaging impact if they are not properly processed. In particular, HJs represent the most toxic DNA structure that can arise during HR repair of DSBs and several pathways are activated for their removal as it was described in **section 4.7**. Given that HJ processing enzymes have not been

studied in detail in the context of pluripotency, we wondered how the HJ dissolvase *Blm* and the resolvases *Mus81* and *Gen1* were expressed in pluripotent versus differentiated cells. Using the same dataset as **Figure R1A**, we observed that the dissolvase *Blm* is expressed at higher levels than the main resolvases, with *Gen1* showing slightly higher expression than *Mus81* (**Figure R2A**). This result is in line with previous observations in differentiated human cells, which favour the dissolution pathway over the resolution pathway to avoid crossover products [127, 128]. *Mus81* shows similar levels in pluripotent and differentiated cells, unlike to *Gen1* which is higher in pluripotent cells, suggesting the possibility that the main resolvase in pluripotent cells is GEN1 instead of MUS81. This is also true for the protein abundance levels in ESCs even though there is only a slight but significant change (**Figure R2B**).

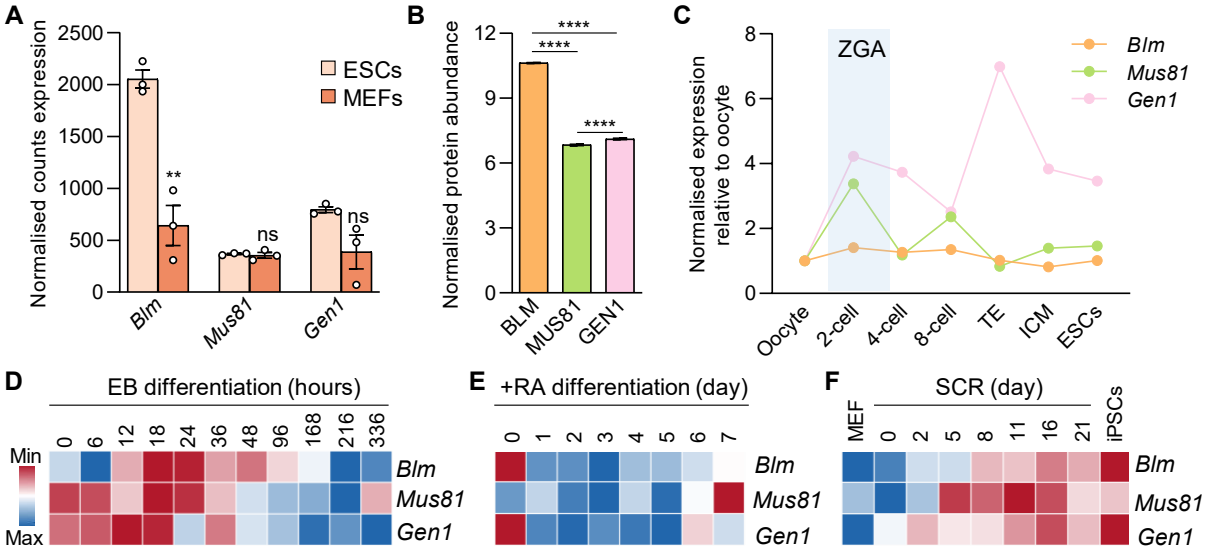

Figure R2. The expression levels of the dissolvase and resolvases of HJ oscillate depending on the differentiation state of the cells. (A) RNA Expression and of HJ resolvases and dissolvase in iPSCs versus MEFs using normalised expression from from a published RNA-seq dataset (GSE69823). **(B)** Protein abundance of those enzymes in ESCs (ProteomeXchange: PXD033001). **(C)** Relative RNA expression pattern during early development from oocyte to ESCs (GSE22182). TE stands for trophectoderm. **(D)** and **(E)** Heatmap of the expression pattern of *Blm*, *Mus81* and *Gen1* during exit from pluripotency by EB differentiation (GSE3749) and RA differentiation [129], respectively. **(F)** Heatmap of the RNA expression pattern of *Blm*, *Mus81* and *Gen1* during entry of pluripotency by somatic cell reprogramming (GSE21757). Statistical significance was determined by unpaired t-test for A and one-way ANOVA using Tukey's multiple comparison test for B: *: p-value<0.05; **: p-value<0.01; ***: p-value<0.001; ****: p-value<0.0001 and ns: non-significant.

Having established the expression profiles of *Blm*, *Mus81*, and *Gen1* in pluripotent versus differentiated cells, we next examined their expression during early development. Using a published RNA-seq dataset of early embryonic development, we found that *Gen1* expression

exhibited the most pronounced variation, whereas *Blm* and *Mus81* levels remained relatively stable (**Figure R2C**). During zygote genome activation (ZGA), the expression of both resolvases is induced, and then decreases in subsequent cleavage divisions (**Figure R2C**). *Gen1* expression was highest in the trophectoderm (TE) relative to other embryonic states. *Mus81* expression was elevated at the 2-cell and 8-cell stages, although overall levels remained comparable to those observed in oocytes. *Blm* expression, in contrast, showed minimal fluctuation throughout development.

We next examined the expression of these proteins during differentiation to all three germ layers (EB differentiation) or to neuroectoderm (retinoic acid, RA) using published datasets [125, 129]. Upon exit from pluripotency, both EB and RA differentiation protocols led to a decrease in the expression of all three enzymes (**Figures R2D** and **R2E**). Conversely, their expression increased during the acquisition of pluripotency (**Figure R2F**). Hence, HJ processing enzymes follow a similar pattern to other HR related genes (**Figure R1E**), with decreased expression during differentiation and increased levels upon pluripotency acquisition, suggesting a more prevalent function in pluripotent versus differentiated cells.

7.2. GEN1 DEPLETION IS INCOMPATIBLE WITH THE SURVIVAL OF ESCS

Although the KO of BLM and MUS81 in ESCs has been previously described, their roles in maintaining self-renewal and pluripotency, assessed by alkaline phosphatase (AP) staining as a marker of pluripotent colonies, remain unexplored. Considering the high expression of GEN1, BLM and MUS81 in pluripotent stem cells (PSCs) and the essential role of the HR repair pathway, we sought to determine which of these factors is most critical for ESC pluripotency and self-renewal. Therefore, we performed CRISPR/Cas9 loss-of-function assays using two distinct sgRNAs targeting the transcription start site of the three genes encoding these HJ regulators (**Figure R3A**). To assess whether ESCs can be maintained in the absence of these enzymes, we conducted a colony formation assay (CFA) with AP staining (**Figure R3B**). The representative images of the wells after staining show the colony formation capacity of ESCs is only affected upon *Gen1* depletion, while the depletion of *Mus81* and *Blm* does not greatly affect ESC colony numbers (**Figure R3C**). This was further confirmed by manually counting AP-positive (AP+) colonies, as ESCs with sgRNAs targeting *Gen1* gave rise to a reduced number of AP+ colonies compared to mock control, which was not observed when using either of the sgRNAs targeting *Blm* or *Mus81* genes (**Figure R3D**). Our initial results strongly

suggested GEN1 plays a predominant role in ESC maintenance, an observation not previously documented. Based on these findings, we focused the subsequent experiments specifically on dissecting the functional requirement of GEN1 in ESCs.

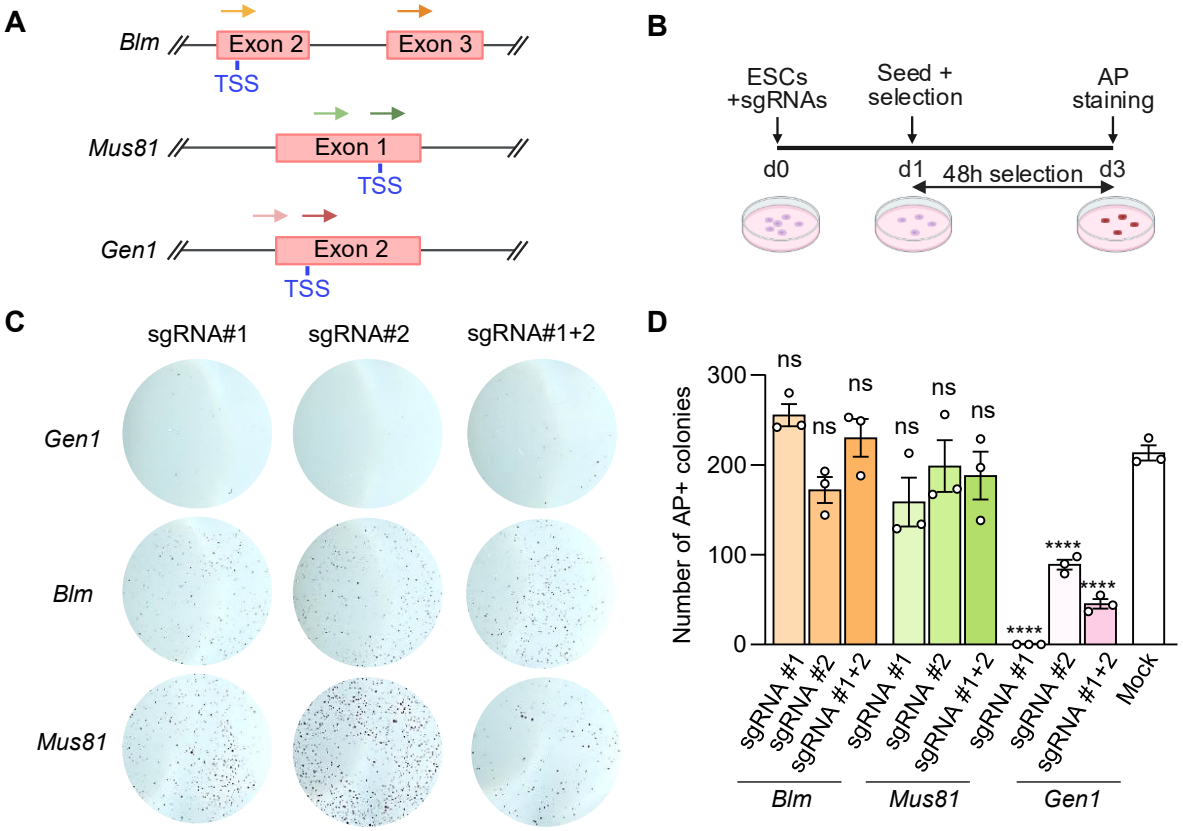

Figure R3. *Gen1* depletion is incompatible with ESCs maintenance *in vitro*. (A) Representative scheme of the genomic regions of *Blm*, *Mus81* and *Gen1* which were tagged by the sgRNAs. TSS: transcription start site indicated in blue. **(B)** Scheme of the KO strategy using Cas9-sgRNAs against *Blm*, *Gen1* and *Mus81*. **(C)** Representative images of the wells after AP staining at day 3. **(D)** Number of AP+ colonies obtained for each condition. Statistical significance was determined by unpaired t-test comparing to mock conditions****: p-value<0.0001 and ns: non-significant.

Given that *Gen1* depletion greatly reduced AP+ colony formation, but a small amount of AP+ colonies could still be formed, we wanted to know whether the thriving colonies belonged to unsuccessful CRISPR/Cas9 depletion of the *Gen1* gene. Therefore, we proceeded with transfecting the sgRNAs targeting the *Gen1* TSS and, after selection, we seeded the surviving cells at low density to select clones for genotyping (**Figure R4A**). For the genotyping PCR, we designed a set of three primers, two forward (one annealing to the TSS in exon 2) and one reverse. This strategy allowed the detection of alterations (mainly genomic deletions or insertions) by PCR (**Figure R4B**). PCR of wild-type (WT) ESCs revealed two bands of the expected sizes (i.e., 181 bp and 226 bp), corresponding to the WT *locus*. We observed that all

clones which were able to self-renew even after being treated with *Gen1* sgRNA displayed band patterns compatible with WT genotype, which would explain their survival despite CRISPR/Cas9 treatment (**Figure R4C**).

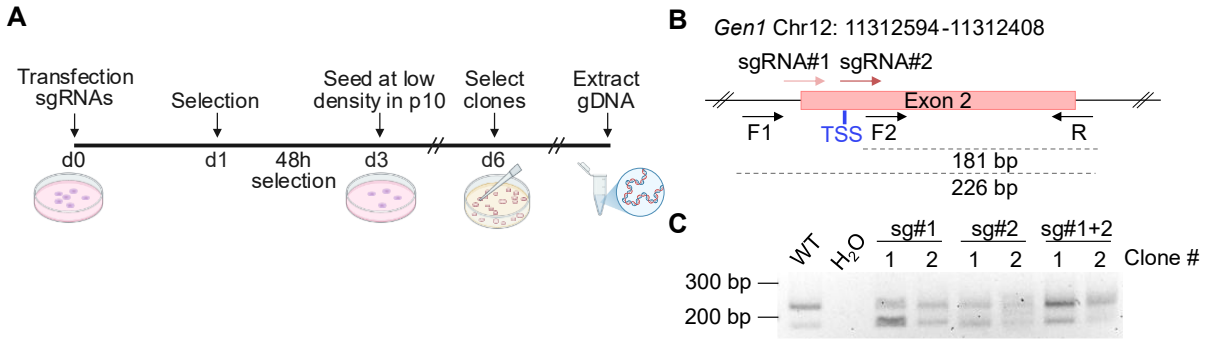

Figure R4. *Gen1* KO clones cannot be obtained in ESCs. (A) Scheme of the KO strategy using Cas9-sgRNAs against *Gen1* to generate a KO clone of this gene in ESCs. **(B)** Genotyping strategy for the verification of the KO of *Gen1* in ESCs using three primers (two forward, F1 and F2 and one reverse, R) which would generate two PCR products consisting of 181 and 226 bp for the WT form. **(C)** Image of 2.5% agarose/synergel showing PCR products obtained for the genotyping of different clones obtained from the transfections using sgRNAs#1, #2 and #1 and #2 against *Gen1*.

To further confirm that GEN1 depletion was responsible for the loss of ESCs in culture, we performed a rescue experiment transfecting the cells with the human ortholog of GEN1. The human and murine GEN1 proteins display a high degree of similarity of their catalytic domains (we defined the catalytic domain as amino acids between 1-458 and did not consider the disordered domain), with a sequence homology close to 85% (**Figure R5A**). First, we generated an ESC line ectopically overexpressing a full-length human GEN1 protein tagged with three FLAG tags (hGEN1^{3xFL}) or an empty vector (EV) as a control (**Figure R5B**). Given the lack of a good commercial antibody against hGEN1, this strategy allowed us to analyse the expression of the ectopically overexpressed protein by Western blot. The hGEN1 overexpression did not induce noticeable changes in morphology or colony-forming capacity in those cell lines overexpressing GEN1, as seen by AP staining in bulk cell cultures (**Figure R5C**). We then transfected the sgRNAs against the *Gen1* gene (sg*Gen1*#1 or sg*Gen1*#2) or a non-targeting sgRNA control (sg*NT*) in the overexpressing cell line (**Figure R5D**). Of note, the sgRNAs targeting the endogenous *Gen1* gene were not predicted to target the overexpressing hGEN1^{3xFL} construct as it lacks their target sequence. While in cells with only endogenous *Gen1* levels (+empty vector, EV), CRISPR/Cas9 sgRNAs against *Gen1* reduced AP+ colonies compared to control (sg*NT*) (**Figures R5E** and **F**), the expression of hGEN1 (hGEN1^{3xFL}) was sufficient to rescue the loss of AP+ colonies seen upon the endogenous *Gen1* CRISPR depletion (**Figures**

R5E and **F**). Thus, we concluded that the loss of PSC self-renewal and/or maintenance observed following CRISPR/Cas9-mediated targeting of *Gen1* was indeed a direct consequence of *Gen1* loss-of-function rather than an off-target effect.

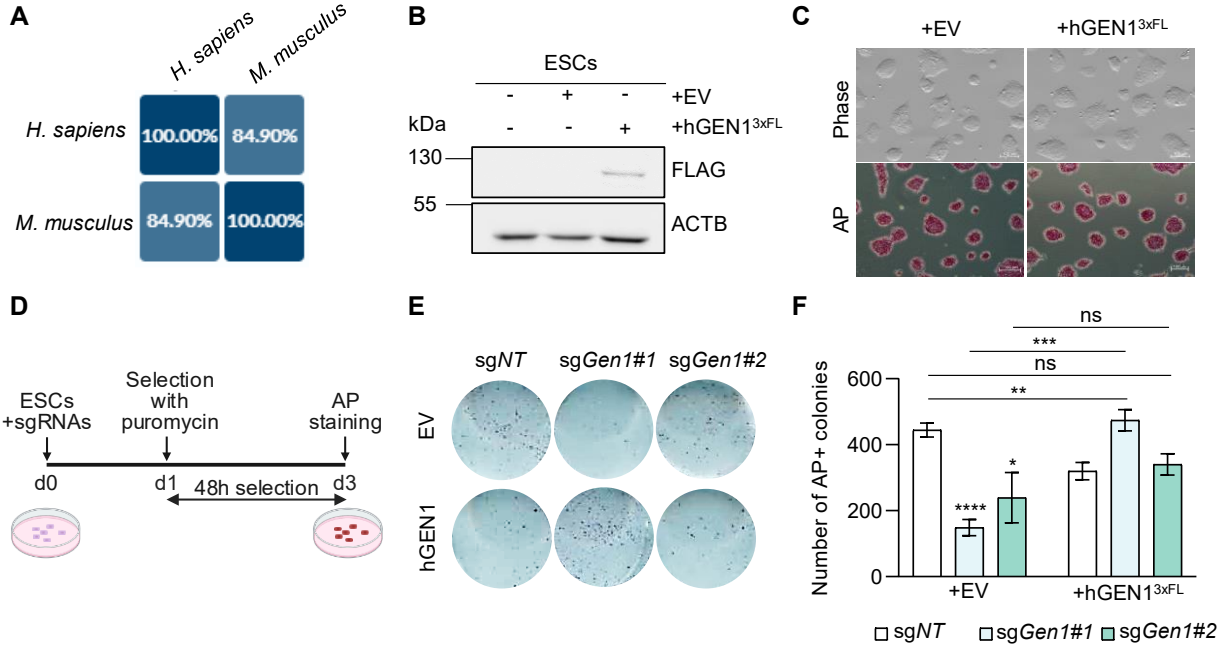

Figure R5. The presence of human GEN1 rescues the loss of the number of AP+ colonies when ESCs are depleted from mouse GEN1. (A) Percentage of sequence similarity between the catalytic domain of GEN1 from *Homo sapiens* and *Mus Musculus*. **(B)** Detection of expression of hGEN1 in the overexpressing cell line of ESCs. **(C)** Representative microscopy image in phase contrast of WT ESCs and ESCs overexpressing hGEN1 with or without AP staining. **(D)** Scheme of the KO strategy using the sgRNAs against mouse *Gen1*. E) Representative image of the AP staining of the different conditions used and **(F)** bar diagram with the count of those colonies. The empty vector (EV) with the sgRNA against non-targeting (NT) condition was used as control. Statistical significance was determined by unpaired t-test comparing against sg*NT* of EV condition: *: p-value<0.05; **: p-value<0.01; ***: p-value<0.001; ****: p-value<0.0001 and ns: non-significant.

7.3. GEN1 IS LOCALISED AT THE CYTOPLASM OF ESCS

Since *Gen1* KO ESCs could not be generated, we opted to study the phenotype of *Gen1* loss-of-function by performing KD using small hairpin RNAs (shRNAs). Due to the lack of a good commercial antibody against our protein of interest, first, we generated a KI cell line targeting *Gen1 locus* with a construct at its 3' end, which resulted in the expression of a tagged GEN1 at the C-terminal domain with three FLAG tags (GEN1^{3xFL}) (**Figure R6A**). For that purpose, we employed a CRISPR/Cas9-mediated KI strategy by co-transfecting a donor plasmid encoding the desired tag with an antibiotic resistance cassette for selection, together with a single-guide RNA (sgRNA) targeting the 3′ end of the *Gen1 locus*. To identify clones harbouring the KI

construct, we developed a genotyping approach capable of distinguishing WT, heterozygous, and homozygous KI alleles, thereby enabling the selection of clones with the construct in homozygosity. The genotyping strategy was based on PCR amplification using primers flanking the KI region. This approach generated a PCR product of 2,598 bp for the KI allele, whereas the WT allele yielded a 1,619 bp product (**Figure R6B**). After transfection of the plasmids and selection with G418 antibiotic, we picked several clones for validation. We selected a positive clone which presented a unique band using our genotyping PCR strategy (**Figure R6C**), and a band of the expected size by Western blot using an anti-FLAG antibody (**Figure R6D**).

GEN1 has been described as a cytoplasmic protein in both mice and humans, accessing DNA only during mitosis when the nuclear envelope breaks down [86]. Its localisation in pluripotent stem cells, however, has not been established. To examine GEN1 distribution in ESCs, we performed subcellular fractionation and analysed the cytoplasmic, nuclear, and chromatin-bound fractions by Western blot. Although we detected GEN1 in the cytoplasmic fraction, using GAPDH as a cytoplasmic marker, we could also detect this resolvase in the nucleus (marked by HDAC1 presence) in asynchronous cells. Additionally, a faint band could also be observed in chromatin-bound fraction, where histone H3 was used as the reference marker. These results show that GEN1 is found in both the cytoplasmic and nuclear compartments in ESCs, with only a minor fraction bound to the chromatin (**Figure R6E**). Importantly, this was not due to cross-contamination of subcellular fractions during the extraction protocol, as specific markers of nucleus (HDAC1) and cytoplasm (GAPDH) were present only in their respective compartments.

To corroborate these observations, we conducted immunofluorescence analysis. The FLAG-tagged GEN1 protein was labelled with Alexa Fluor 488 (green), and cell nuclei were counterstained with DAPI (blue), which stains DNA. Consistent with the fractionation results, GEN1 was localised in the cytoplasm of ESCs, although a small fraction of the signal co-localised with the DAPI-stained nuclei (**Figure R6F**). Together, these results indicate that GEN1 resides in the cytoplasm of ESCs but can also be present in the nucleus or associated with chromatin.

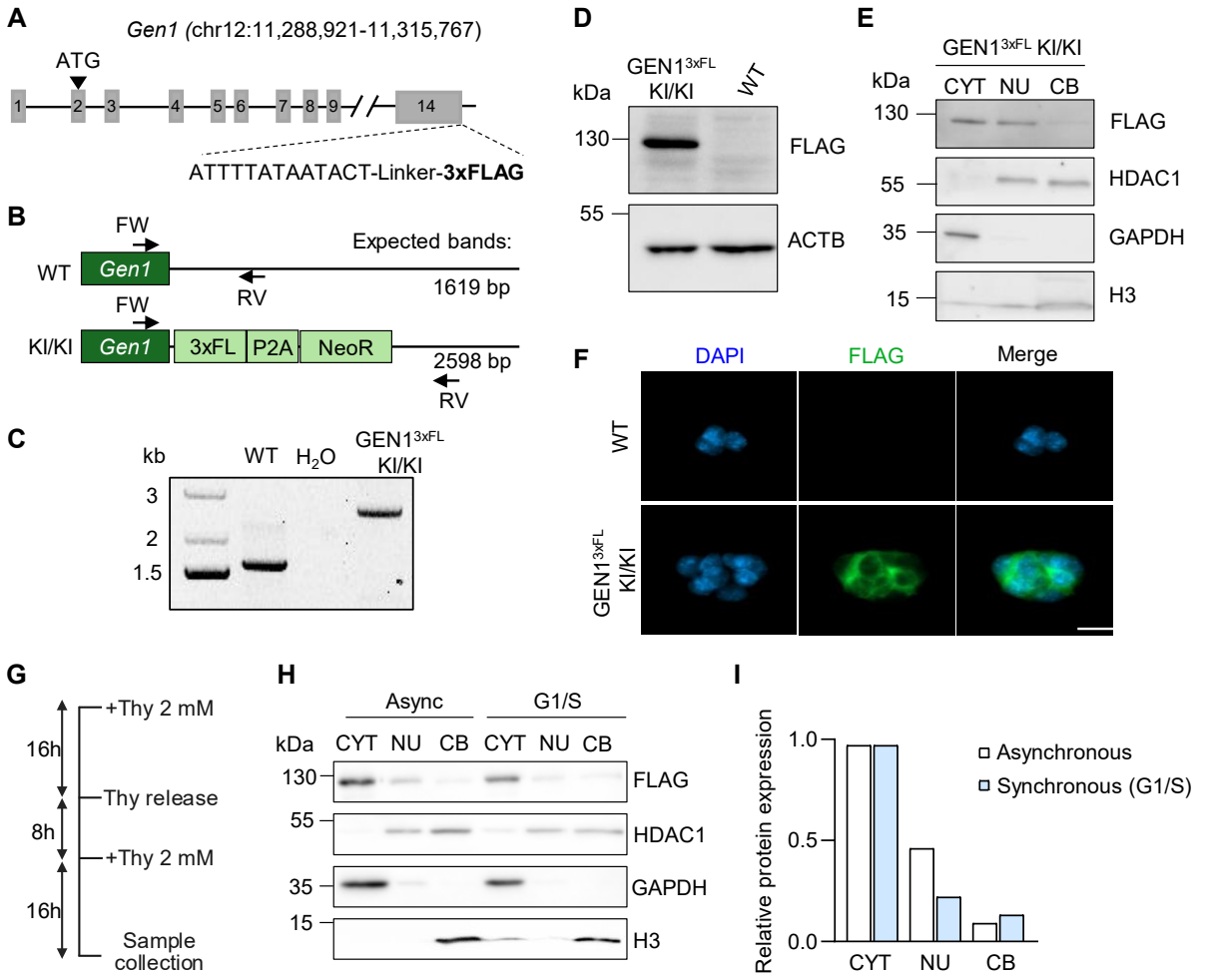

Figure R6. Generation of a cell line with GEN1 tagged with 3xFL at C-terminal. (A) Scheme of genomic region of *Gen1* tagged with the 3 Flag tag. **(B)** Scheme of the genotyping PCR to distinguish between KI/KI cell line and wild type cell line. **(C)** PCR products for the wild type and GEN1^{3xFL} ESCs. **(D)** Western blot of whole cell extracts and **(E)** subcellular fractioning of the GEN1^{3xFL} ESCs, respectively. Antibodies against FLAG (tag for GEN1), HDAC1 (marker for nuclear (NU) compartment), GAPDH (marker for cytoplasm (CYT)), H3 (marker for chromatin-bound (CB) fraction) and ß-actin (ACTB as loading marker) were used. **(F)** Immunofluorescence against FLAG in wild type and GEN1^{3xFL} ESCs. DAPI was used for staining cellular nucleus. **(G)** Scheme for the experimental approach followed for the synchronisation of cells in G1/S with the double thymidine (Thy) block. **(H)** Western blot of the subcellular fractioning in asynchronous and G1/S synched GEN1^{3xFL} ESCs. The same markers were used as in **(E)**. **(I)** Quantification of the protein levels of GEN1 at the different compartments relative to asynchronous samples.

Access of GEN1 to DNA is restricted to mitosis due to its nuclear export signal (NES), which actively excludes GEN1 from the nucleus during interphase. Intriguingly, when analysing asynchronous ESCs, a detectable fraction of GEN1 was present in the nucleus, suggesting that its export may not be fully efficient. To clarify this, we synchronised ESCs at the G1/S transition, where the nuclear envelope remains intact and NES-mediated export is expected to be active, to examine whether GEN1 persists in the nucleus during this phase.

Considering this cell cycle-dependent regulation, we sought to determine whether its subcellular distribution is influenced by cell cycle progression in ESCs. Cells were arrested and synchronised via a double thymidine block. Both asynchronous and G1/S-synchronised populations were collected for subcellular fractionation (**Figure R6G**). Consistent with previous observations, GEN1 was predominantly cytoplasmic in both populations, however, the nuclear fraction was reduced in G1/S-synchronised cells compared to asynchronous cells (**Figures R6H** and **I**) [86]. These results suggest that while NES-dependent exclusion of GEN1 from the nucleus is robust during interphase, a small nuclear fraction can still be detected and its presence may be linked to cell cycle position. This would raise the possibility that short G1 phases in ESCs or transitional nuclear envelope states allow limited nuclear retention.

7.4. *Gen1* silencing compromises the self-renewal capacity of ESCs

Once we had achieved the generation of the GEN1^{3xFL} KI cell line, we used shRNAs to deplete its expression and characterise its role in ESCs. We performed a CFA to verify whether the phenotype was recapitulated relative to that observed upon CRISPR/Cas9-mediated loss-of-function (**Figures R3** and **R7A**). Indeed, we observed that cells where *Gen1* was silenced (sh*Gen1*) generated fewer pluripotent AP+ colonies (**Figures R7B** and **C**), similar to what was observed upon CRISPR/Cas9 targeting (**Figure R3D**). Although some cells with sh*Gen1* could proliferate and form colonies, those colonies looked smaller compared to the control transfected with sh*Luci*, as it can be observed for tdTomato+ ESCs (**Figures R7D** and **E**).

While the colonies formed by *Gen1*-silenced ESCs were positive for the pluripotency marker alkaline phosphatase, we wanted to further examine the expression of key pluripotency markers by Western blot. For that purpose, we checked the levels of the pluripotency factors NANOG and REX1, together with GEN1 to validate its levels upon the treatment with shRNAs. We observed almost no expression of GEN1^{3xFL} upon *Gen1* KD at 72 h post-transfection, showing that the shRNAs had worked efficiently (**Figure R7F**). On the other hand, we observed that *Gen1*-silenced cells showed no major differences in their expression of NANOG or REX1 pluripotency markers, confirming that the colonies with reduced *Gen1* expression were still pluripotent (**Figure R7F**).

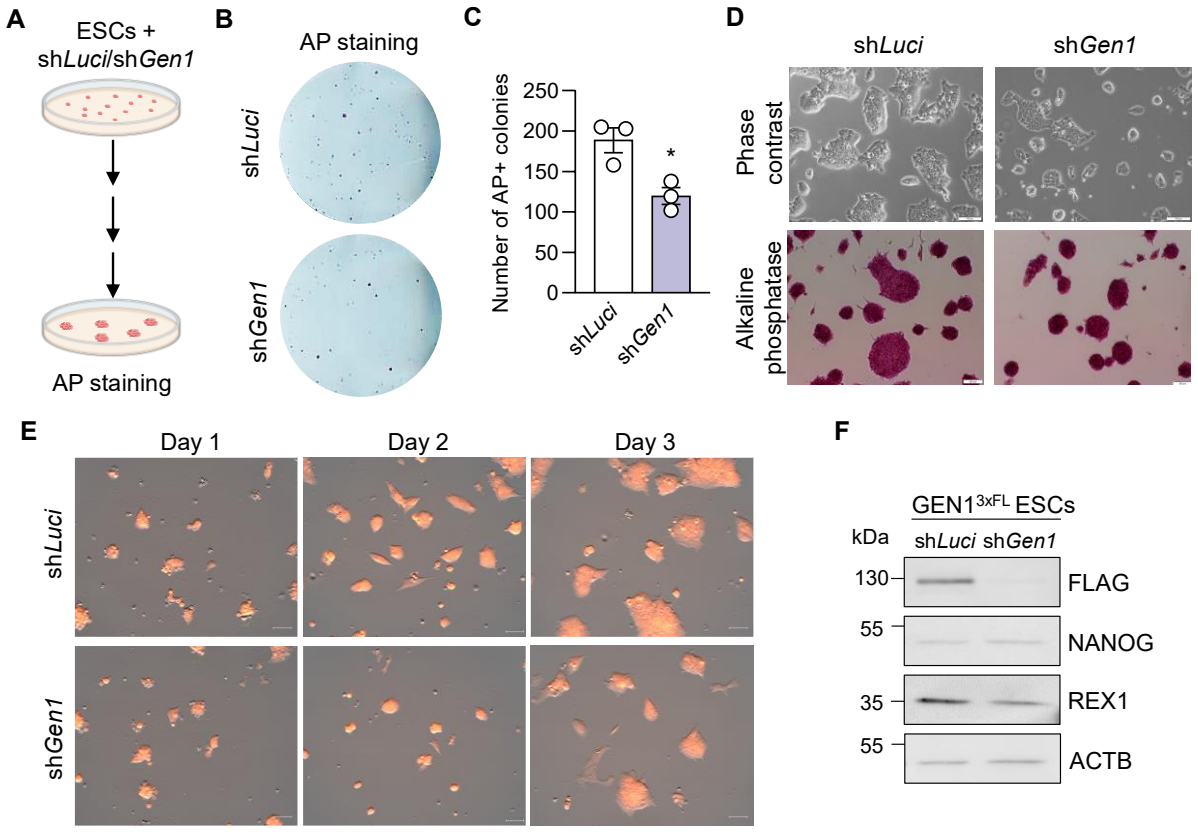

Figure R7. *Gen1* silencing compromises the capacity of self-renewal of ESCs. (A) Scheme for the experimental approach for the silencing of *Gen1* in ESCs, using sh*Luci* as control. **(B)** Representative images of the well after AP staining and **(C)** colony count of ESCs. **(D)** Representative image of ESCs transfected with sh*Luci* or sh*Gen1* pre- and post- alkaline phosphatase staining. **(E)** Representative phase contrast images of tdTomato+ cells silenced for *Gen1*, sh*Luci* as control, at different days post transfection. **(F)** Western blot image using antibodies against FLAG as marker for GEN1 expression and pluripotency markers NANOG and REX1. ß-actin (ACTB) was used as loading control. Statistical significance was determined by unpaired t-test: *: p-value<0.05.

7.5. KNOCKDOWN OF *GEN1* LEADS TO MODEST ALTERATIONS IN THE ESCS TRANSCRIPTOME

We then asked if the loss of *Gen1* led to transcriptional changes that could cause to this observed defect on self-renewal. For this purpose, we extracted total RNA from ESCs silenced for *Gen1*, sh*Luci* as a control, and samples were analysed by RNA-seq (**Figure R8A**). A principal component analysis showed that the sample duplicates clustered together (**Figure R8B**), indicating high reproducibility and consistency within each experimental group. This suggests that technical variability was minimal and that observed differences between groups are likely driven by true biological effects.

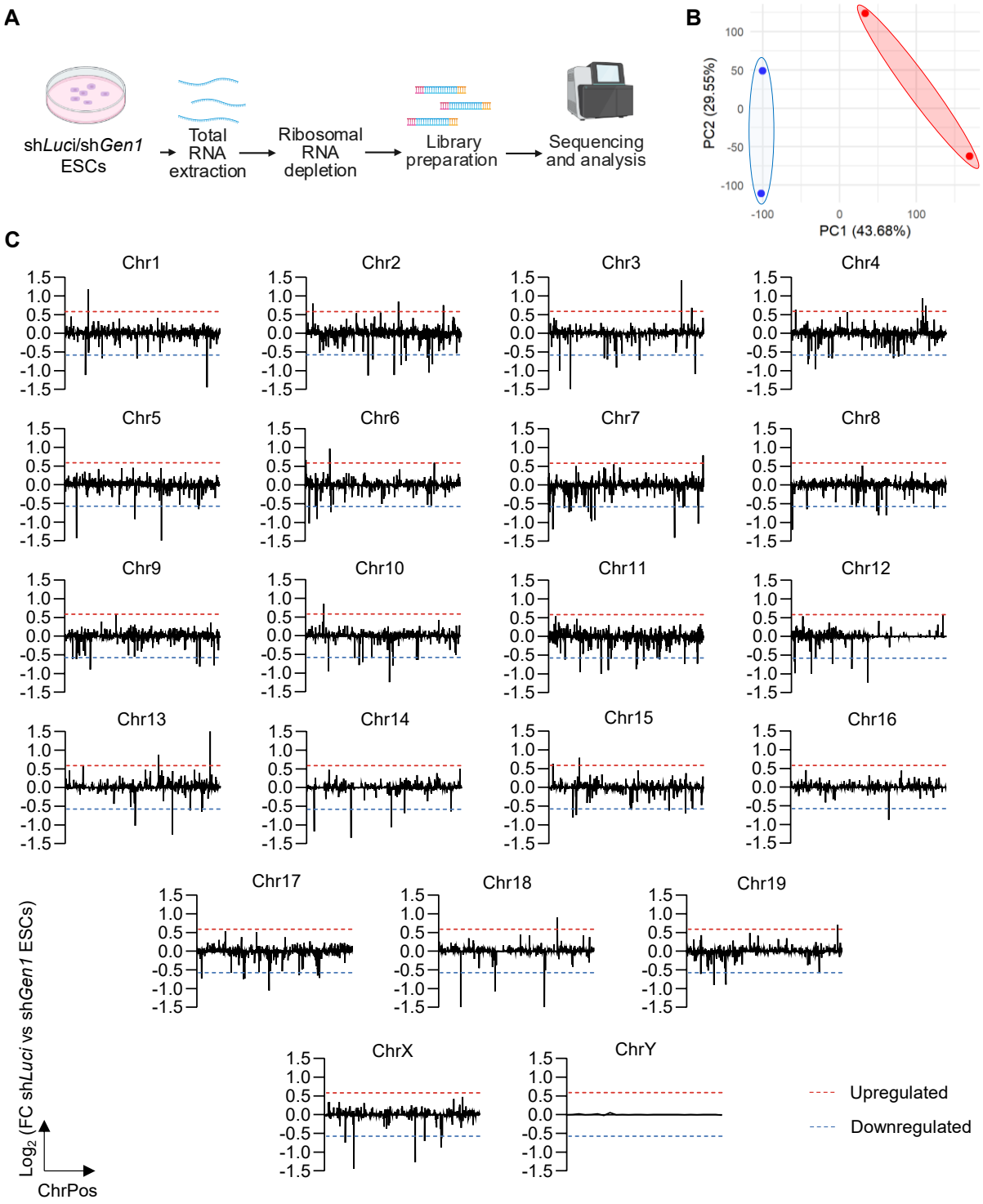

Figure R8. ***Gen1*** **silencing leads to discrete changes in the transcriptional landscape of ESCs. (A)** Representative scheme of the procedure for the RNA sequencing. **(B)** Principal component analysis of the transcriptomic profiles of the indicated sh*Luci* and sh*Gen1* ESCs. **(C)** Graphic representation of the FC RNA expression in sh*Gen1* vs sh*Luci* across all mouse chromosomes. The dotted line considers there is a FC above (red line) or below (blue line) 1.5 comparing silencing *Gen1* condition to *Luci*.

We next analysed the differentially expressed genes (DEGs), using a significance criterion of false discovery rate (FDR) smaller than 0.05. After mapping the DEGs across all chromosomes, we found no clear enrichment in any particular autosome, indicating that expression changes were scattered and not location-dependent (**Figure R8C**). We detected a total of 491 transcripts which were deregulated upon KD of *Gen1* in ESCs, from which 169 were upregulated and 322 were downregulated which met the criteria of FDR<0.05 (**Figure R9A**).

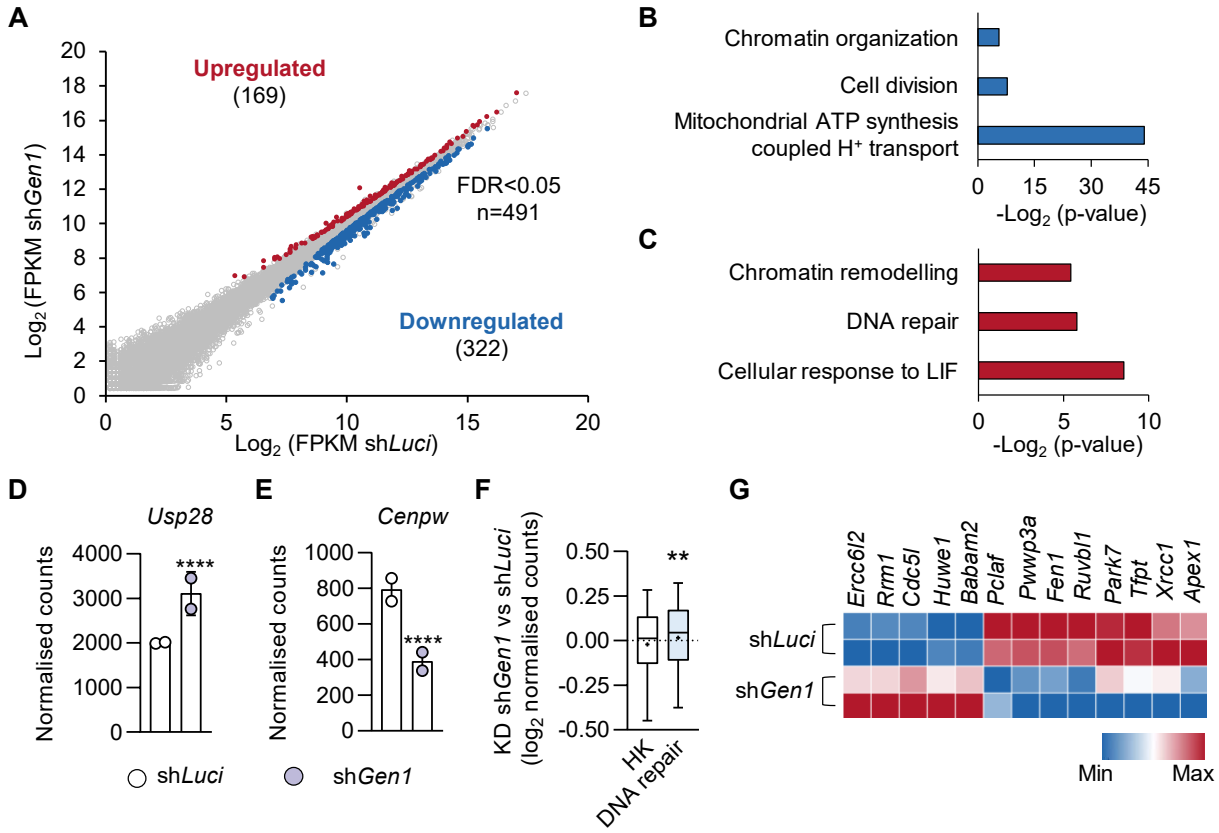

Figure R9. Several biological processes are affected by the misregulation of genes upon silencing of *Gen1* in ESCs. (A) Scatter plot of gene expression values (expressed as Log_2(FPKM)) obtained by RNA-seq analysis comparing the expression of sh*Luci* vs sh*Gen1*. The genes which met the criteria of FDR<0.05 are shown in red for upregulated genes and in blue for downregulated genes (n=2). **(B)** Gene ontology analysis of biological processes affected by the downregulation or **(C)** upregulation of genes in ESCs silenced for *Gen1*. Example of misregulated genes which belong to GO related to biological process as **(D)** DNA repair, *Usp28* or **(E)** cell division, *Cenpw*. **(F)** Box-and-whiskers plot showing expression levels (Log_2 normalised counts) in sh*Luci* vs sh*Gen1* ESCs for genes related to the gene ontology term DNA repair (GO:0006281, n=380) and using a dataset for housekeeping (HK) genes as control from public sources [130]. **(G)** Heatmap showing some genes comprehended inside the DNA repair GO term with FDR<0.05. Statistical significance was determined by the RNA-seq analysis (****: FDR<0.0001) for D) and E) and unpaired t-test for F): **: p-value<0.01.

To further interpret the functional relevance of these transcriptional changes, we performed GO enrichment analysis. Downregulated genes were significantly enriched in biological processes associated with chromatin organisation, cell division (i.e. *Cenpw* (**Figure R9E**)), and mitochondrial ATP production (**Figure R9B**). Upregulated transcripts encoded proteins involved in biological processes concerning chromatin remodelling, DNA repair (i.e. *Usp28* (**Figure R9D**)), and cellular response to LIF (**Figure R9C**). Consistent with the established role of GEN1 in resolving Holliday junctions during homologous recombination, several genes associated with DNA repair were significantly upregulated upon *Gen1* KD in ESCs, suggesting a potential compensatory transcriptional response to maintain genome integrity (**Figures R9F** and **G**).

7.6. METASTABLE ESCS EXHIBIT INCREASED GEN1 EXPRESSION

Because *Gen1* KD affected genes involved in DNA repair, chromatin organisation and, specifically, pluripotency pathways, we next aimed to verify whether *Gen1* expression varies across distinct pluripotent states. As mentioned in **section 2.2**, ESCs can exist in metastable, primed, or naïve states in culture, each differing in self-renewal, differentiation potential, and genome maintenance requirements [131] as well as cell cycle control [51]. Naïve cells cultured in 2iL conditions display a longer G1 phase compared to metastable cells [51]. Measuring GEN1 levels across these states would allow us to determine whether its expression is constant or dynamically regulated, providing insight into how GEN1 may contribute to the transcriptional and functional effects observed following its depletion. Therefore, we decided to look at the RNA and protein expression levels at the different pluripotent states (**Figure R10A**).

First, using published datasets, we observed a higher expression of *Gen1* in SL conditions (metastable state) compared to the 2iL condition (naïve state), whereas the opposite pattern was observed for the other resolvase *Mus81* and no significant differences were detected for *Blm* (**Figure R10B**). In order to verify these observations, we cultured ESCs in SL, 2iL or FA media. We observed a higher expression of *Gen1* in metastable SL conditions, compared to 2iL, both at the RNA (**Figure R10C**) and protein levels for the GEN1^{3xFL} ESC line extracts (**Figure R10D**), confirming the results from the published dataset. These findings demonstrate that *Gen1* expression is higher under metastable SL conditions compared to naïve 2iL, at both the

transcript and protein levels, confirming that *Gen1* abundance in ESCs is sensitive to the pluripotent state and culture environment.

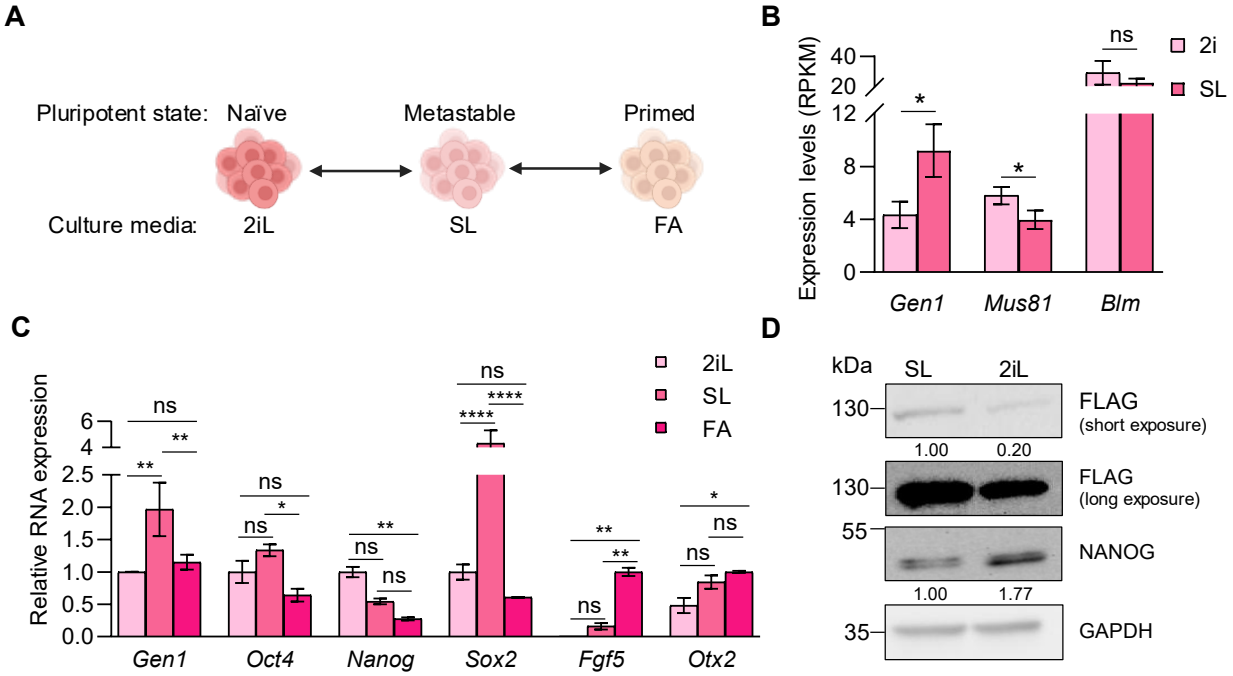

Figure R10. GEN1 has higher expression levels in SL conditions than in 2iLIF conditions. (A) Scheme with the different pluripotent state that exist *in vitro* with their corresponding media conditions. **(B)** RNA expression levels of the resolvases and dissolvase of HJs in 2i and SL conditions from published dataset (GSE23943). **(C)** Relative RNA expression levels of *Gen1* in the different culture media. **(D)** Protein expression of GEN1 in SL and 2iL culture media. Relative protein quantification SL condition is shown underneath the blot for FLAG as marker of GEN1 expression and NANOG as marker of naïve pluripotent state. Statistical significance was determined by unpaired t-test for **(B)** and two-way ANOVA using Tukey's comparison test for **(C)**: *: p-value<0.05; **: p-value<0.01; ***: p-value<0.001; ****: p-value<0.0001; ns: non-significant.

7.7. *GEN1* SILENCING ALTERS ENDODERMAL DIFFERENTIATION POTENTIAL OF ESCS

Previous results from this doctoral thesis indicate that GEN1 could be the main resolvase in ESCs, showing a high expression in the metastable state (**Figure R10**), and its depletion compromising cell survival (**Figure R3**). Therefore, given that differentiation is a defining characteristic of ESCs, we next assessed whether *Gen1* KD could affect their differentiation potential. As an initial step, we performed an *in silico* analysis of misregulated genes in *Gen1* KD ESCs to assess potential enrichment of germ layer–specific markers. While endodermal markers were significantly affected compared to controls, a smaller number of ectodermal and mesodermal markers were also perturbed, suggesting a selective impact of *Gen1* depletion on differentiation-associated genes expression (**Figures R11A** and **B**). To evaluate whether the

transcriptional changes observed upon *Gen1* KD result in functional defects in differentiation, we performed an embryoid body (EB) formation assay upon *Gen1* depletion (**Figure R11C**). By comparing EB formation and lineage-specific marker expression between control and *Gen1*-silenced ESCs, we aimed to determine whether GEN1 is required for proper lineage specification and exit from pluripotency. We observed that the markers for pluripotency (i.e., *Oct4* and *Nanog*), mesoderm (i.e., *Brachyury/T* and *Mixl*) and ectoderm (i.e., *Sox1* and *Otx2*) showed similar expression patterns to the sh*Luci* control (**Figure R11D**). This is in line with the RNA-seq results which did not show major differences between the reference and the perturbed sample for those germ layers. In contrast, the endoderm markers *Gata6* and *Sox17* showed impaired induction upon *Gen1* KD (**Figure R11E**) in the last day of EB differentiation (**Figures R11F** and **G**). These observations suggest that GEN1 is specifically required for proper endodermal differentiation in ESCs, indicating a selective role in guiding lineage commitment during early differentiation.

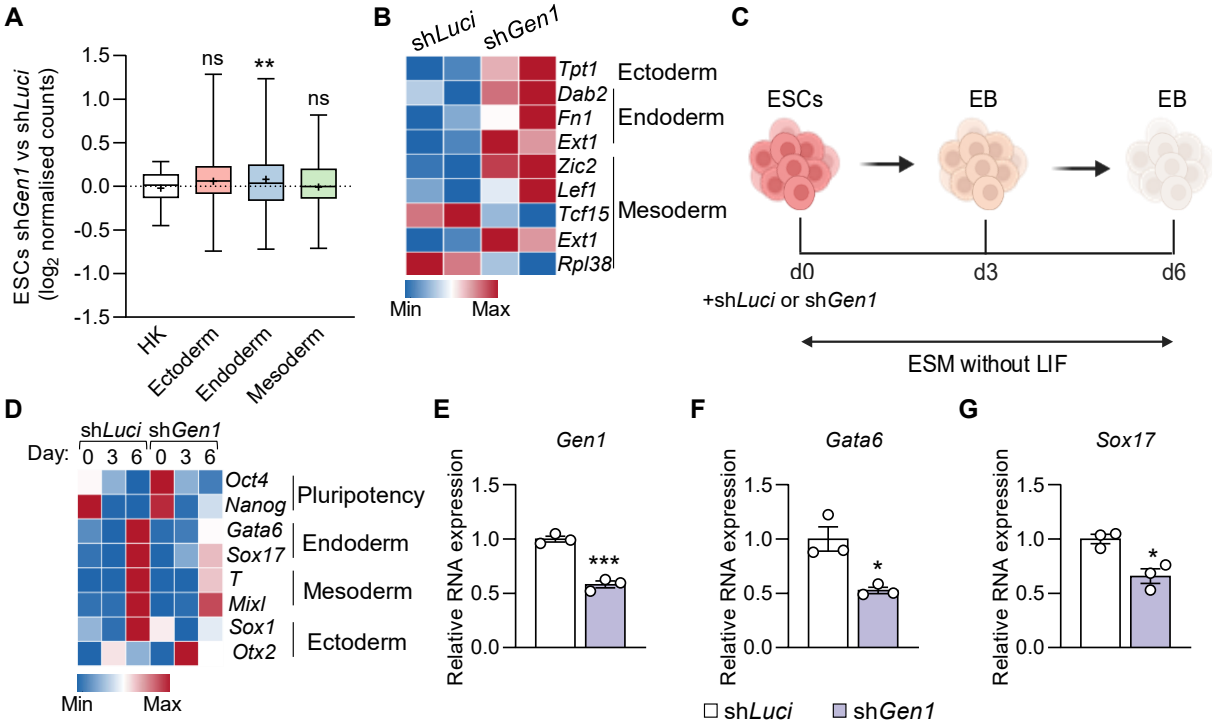

Figure R11. *Gen1* knockdown impairs the induction of endodermal markers during embryoid body differentiation of ESCs. (A) Box-and-whiskers plot showing expression levels (Log_2 normalised counts) in sh*Luci* vs sh*Gen1* ESCs for genes related to the following GO terms: ectoderm development (GO: 0007398, n=16), endoderm development (GO: 0007492, n=63) and mesoderm differentiation (GO:0048333, n=107). A list of housekeeping genes (HK, n=4491) was used as control. **(B)** Heatmap showing genes from the GO terms used before which have an FDR<0.05 and to which germ layer they belong. **(C)** Scheme of the embryoid body formation assay in KD conditions of *Gen1*, sh*Luci* as control. **(D)** Heatmap showing relative expression levels of different markers for pluripotency and the germ layers at differentiation day 0, 3 and 6. Relative RNA expression levels of **(E)** *Gen1*

and **(F)** and **(G)** endoderm markers *Gata6* and *Sox17* at day 6 of differentiation. Statistical significance was determined by unpaired t-test: *: p-value<0.05; **: p-value<0.01; ***: p-value<0.001; ns: non-significant.

7.8. *GEN1* KNOCKDOWN LIMITS FIBROBLAST REPROGRAMMING POTENTIAL

Having observed that GEN1 is required for ESCs (**Figure R3**) and that its loss-of-function also influences the differentiation, we next asked whether its role extended to the acquisition of pluripotency. To address this, we examined the impact of *Gen1* silencing on the efficiency of somatic cell reprogramming (SCR) into induced pluripotent stem cells (iPSCs). We used immortalised fibroblasts stably transduced with a doxycycline-inducible *Oct4-Sox2-Klf4-Myc* (i.e., OSKM) transgene and with sh*Gen1* and sh*Luci* as control. Cells were maintained in pluripotency medium (ESM^{LIF}) supplemented with doxycycline for 14 days (**Figure R12A**). On day 14 of SCR, AP staining revealed a marked reduction in the pluripotent colonies generated from fibroblasts in which *Gen1* was silenced compared to sh*Luci* control, indicating that GEN1 is required for iPSCs generation (**Figure R12B**).

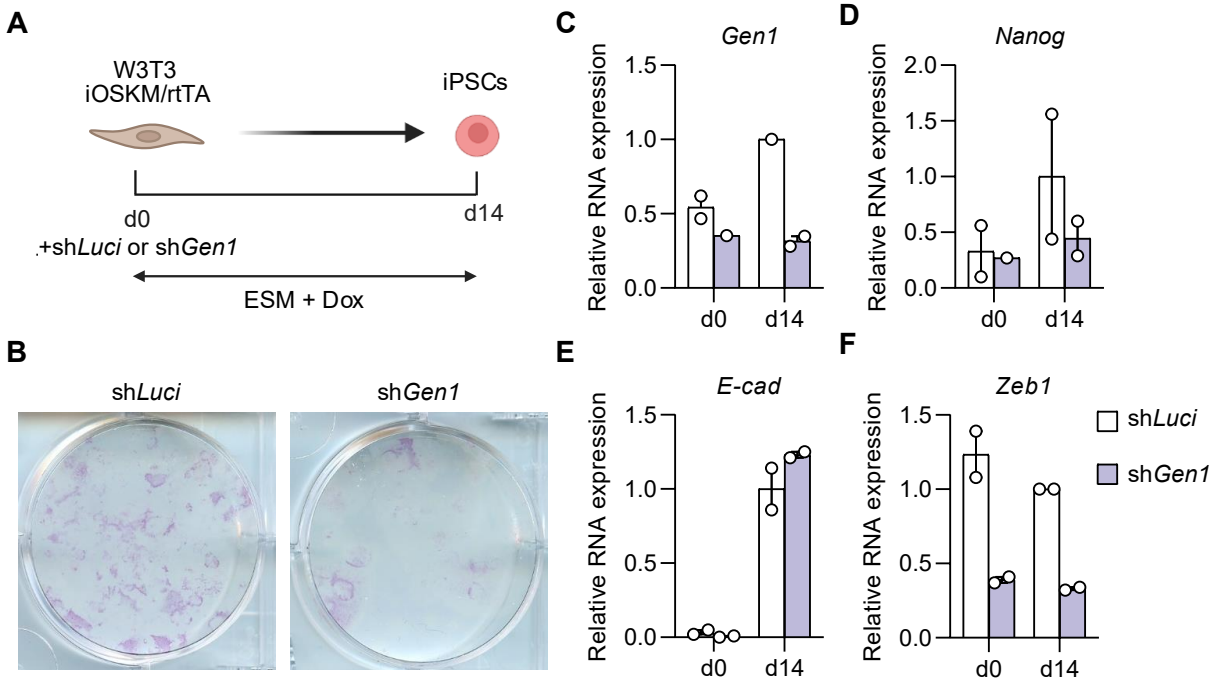

Figure R12. *Gen1* silencing affects the reprogramming capacity of fibroblasts. (A) Scheme of the reprogramming experiment with fibroblast with a doxycycline inducible system of the OSKM factors infected with shRNAs against *Gen1* and *Luciferase* as control. **(B)** Representative image after AP staining of reprogrammed wells for the two conditions. RNA expression levels of **(C)** *Gen1*, **(D)** pluripotent marker *Nanog*, **(E)** epithelial marker *E-cadherin* (*E-cad*) and mesenchymal marker *Zeb1*.

To validate our observations at the molecular level, we evaluated the expression of specific markers (**Figure R12C**). We observed that the pluripotency marker *Nanog* was strongly

upregulated in control (sh*Luci*) cells by day 14 as expected, whereas *Gen1* KD cells did not allow for *Nanog* induction (**Figure R12D**), suggesting that GEN1 is required for the activation of the pluripotency network during SCR. *E-cadherin*, a marker of epithelial identity and mesenchymal-to-epithelial transition (MET), increased similarly in both conditions by day 14 (**Figure R12E**), indicating that the MET proceeds normally despite *Gen1* depletion. In contrast, the mesenchymal regulator *Zeb1*, which represses epithelial genes and antagonises pluripotency, showed reduced expression in *Gen1* KD cells at both day 0 and day 14 compared to controls, with no significant changes over time. These suggests that MET is not impaired by *Gen1* absence, since E-cadherin induction proceeded normally. This supports the interpretation that the role of GEN1 in reprogramming lies downstream of MET, particularly in the robust activation of pluripotency genes. Thus, together with my previous results, GEN1 would be required not only for the maintenance of pluripotent cells but also for their generation through SCR.

7.9. GEN1 LOSS DOES NOT AFFECT ATP PRODUCTION IN ESCS

Taking a deeper look at the RNA-seq data, KEGG pathway analysis revealed significant deregulation of genes involved in oxidative phosphorylation among other pathways such as metabolism and neurodegeneration (**Figure R13A**). Therefore, we next asked whether *Gen1* silencing impacts specific metabolic pathways in ESCs. Examination of GO terms related to the tricarboxylic acid (TCA) cycle (GO: 0006099), oxidative phosphorylation (GO: 0006119) and serine metabolism (GO: 0009069) confirmed significant misregulation of these pathways at the RNA level upon *Gen1* depletion (**Figure R13B**). Specifically, expression of several genes encoding mitochondrial respiratory chain complexes was also affected (**Figure R13C**).

Furthermore, ESCs in the metastable state are known to use a mixed metabolism of glycolysis and oxidative phosphorylation [132], and persistent DNA damage has been shown to deregulate oxidative phosphorylation [133]. To determine whether these transcriptional changes translated into functional alterations in energy production, we measured total and mitochondrial ATP levels using a luciferase-based bioluminescence assay (**Figures R13D** and **E**). Our results revealed no significant differences between control and *Gen1*-depleted ESCs, neither in total nor in mitochondrial ATP levels (**Figure R13F**). Taken together, these results suggest that despite the transcriptional deregulation of genes coding for metabolic regulators upon *Gen1* silencing, overall cellular energy production remains unaffected, (**Figures R13A**

and **B**) indicating that GEN1 is not essential for maintaining basal metabolic activity in metastable ESCs.

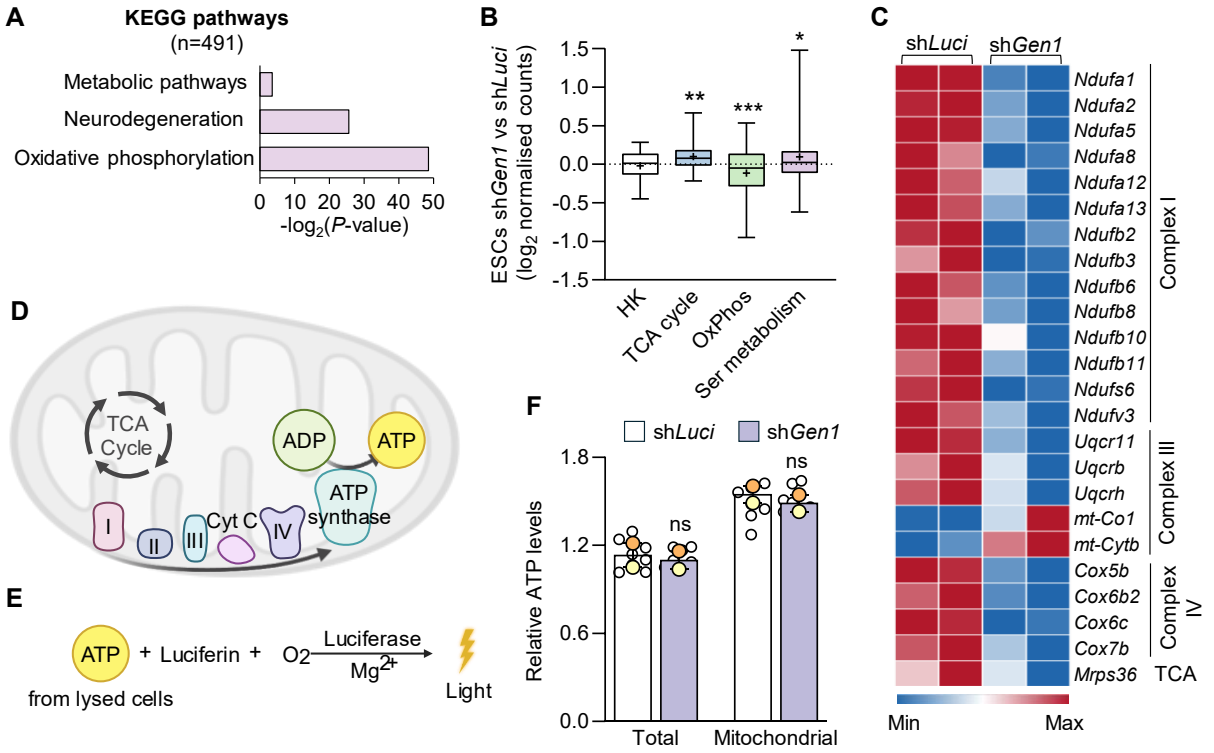

Figure R13. The silencing of *Gen1* does not affect the levels of ATP in ESCs. (A) KEGG pathways analysis of deregulated genes in sh*Gen1* ESCs (n=491 with FDR<0.05). **(B)** Box-and-whiskers plot showing expression levels (Log_2 normalised counts) in sh*Luci* vs sh*Gen1* ESCs for genes related to the following GO terms: TCA cycle (GO: 0006099, n=29), oxidative phosphorylation (OxPhos; GO: 0006119, n=93) and serine (Ser) metabolism (GO:0009069, n=31). A list of housekeeping genes (HK, n=4491) was used as control. **(C)** Heatmap showing genes from the GO terms used before which have an FDR<0.05 and to which mitochondrial complex they belong. **(D)** Representative schematic illustrating the measurement of intracellular ATP levels via a luciferase-based bioluminescence assay following cell lysis and luciferin addition. **(E)** Quantification of ATP levels total and mitochondrial normalised to the total protein levels and to sh*Luci* condition. Statistical significance was determined by unpaired t-test: *: p-value<0.05; **: p-value<0.01; ***: p-value<0.001; ns: non-significant.

7.10. *GEN1* KNOCKDOWN REDUCES THE PROLIFERATION OF ESCS

ESCs are distinguished by their high proliferative capacity in culture [1]. To investigate whether impaired proliferation contributes to the loss of ESCs upon *Gen1* silencing, we first compared the growth kinetics of sh*Gen1* and sh*Luci* control cells (**Figure R14A**). The proliferation assay revealed that *Gen1*-silenced ESCs exhibited a reduced proliferation rate compared to controls (**Figure R14B**). Importantly, trypan blue exclusion showed no significant differences in cell viability, indicating that the reduction in cell numbers is primarily due to impaired expansion

and not to increased late-stage cell death (**Figure R14C**). However, this method may underestimate early apoptotic events that preserve membrane integrity. To further assess early cell death, annexin V staining was performed and revealed no increase in apoptosis at 24, 48, or 72 h after KD (**Figures R14D** and **E**). This would support the idea that impaired proliferation is the primary cause of ESC loss.

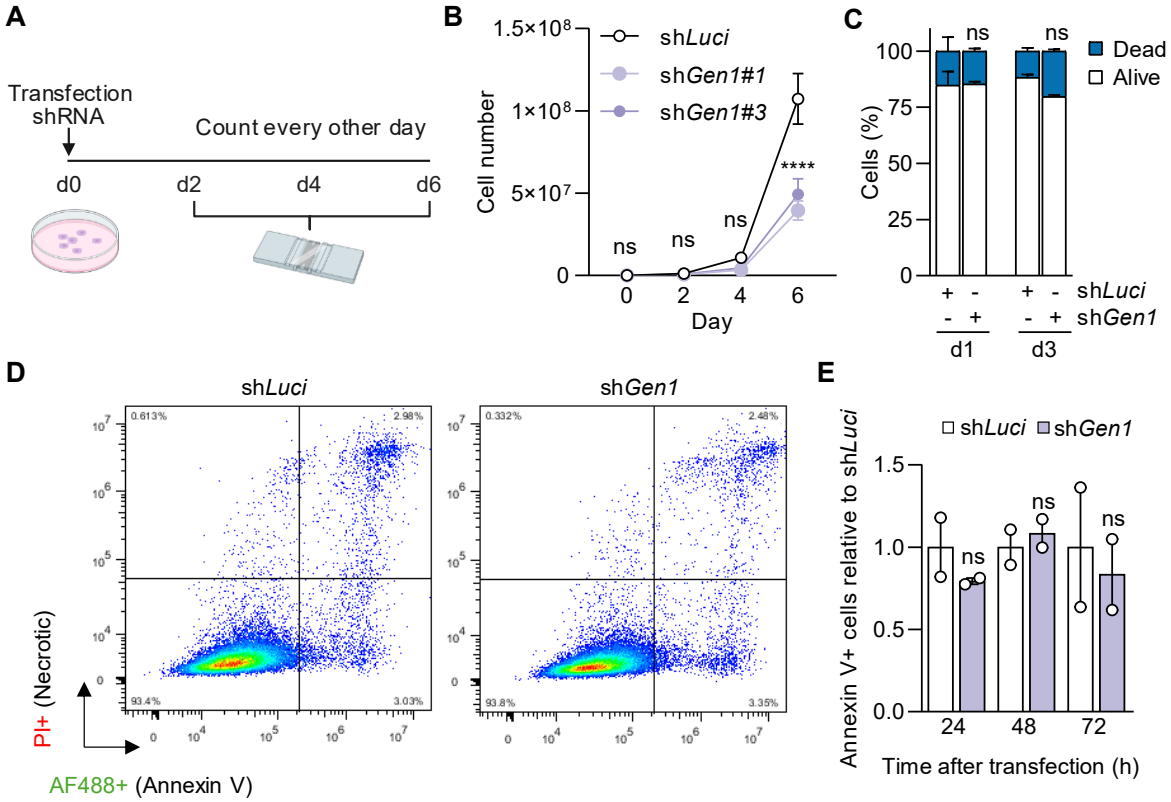

Figure R14. *Gen1* silencing produces a lower proliferation rate in ESCs. (A) Scheme of the experimental approach followed for the determination of cell proliferation of ESCs. **(B)** Growth curve of ESCs silenced for *Gen1* using two different shRNAs against *Gen1* and *Luci* as control. **(C)** Bar plot showing the corresponding percentages of alive (white) and dead (blue) cells in sh*Luci* vs sh*Gen1* condition of ESCs. **(D)** Representative scatter plot images after annexin V assay for detecting apoptotic cells (annexin V+ cells). **(E)** Quantification of annexin V positive cells in sh*Luci* vs sh*Gen1* in ESCs at different hours post transfection (n=2). Statistical significance was determined by unpaired t-test: ****: p-value<0.0001; ns: non-significant.

Given these results, we next examined whether cell cycle alterations could contribute to the reduced proliferation of ESCs upon *Gen1* KD (**Figure R7**). Analysis of bulk populations by flow cytometry showed no significant changes in the distribution of cells across G1, S, or G2/M cell cycle phases (**Figures R15A** and **B**).

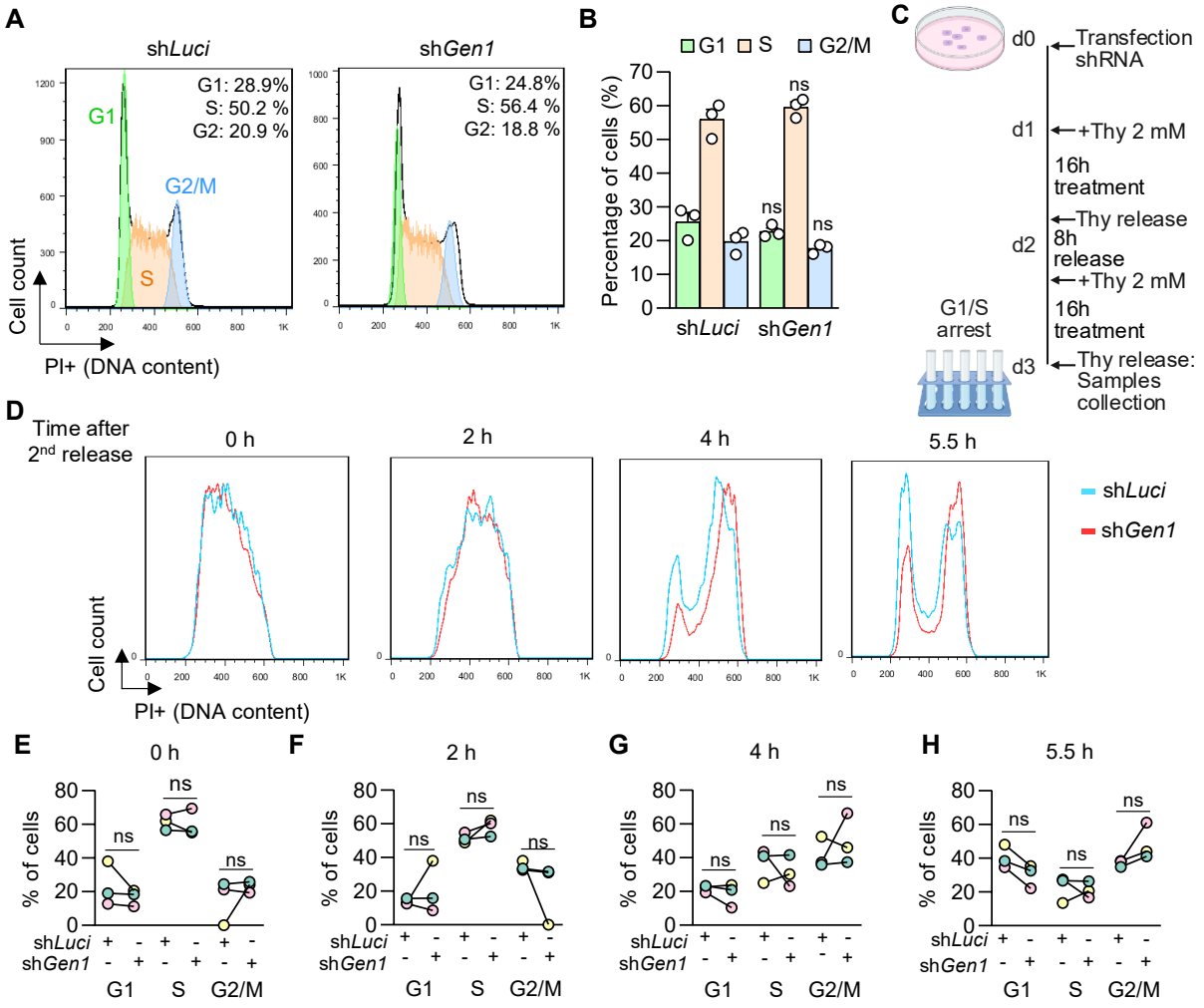

Figure R15. *Gen1* knockdown does not lead to detectable alterations in the progression or distribution of ESCs across cell cycle phases. (A) Histogram showing the cell cycle profile assessed by flow cytometry, showing the percentage of cells in G1, S, and G2/M phases for sh*Luci* and sh*Gen1* KD ESCs. **(B)** Bar graph depicting the percentage of cells in each cell cycle phase for both conditions (n=3). **(C)** Scheme for the experimental approach followed for the synchronisation of cells in G1/S with the double thymidine (Thy) block after transfection of shRNAs against *Gen1* and *Luci* as control. Samples were collected at different time points after second Thy release. **(D)** Representative histograms showing cell cycle progression of ESCs through the different stages. The blue line corresponds to sh*Luci* population and the red line to sh*Gen1* population. **(E)** Quantification of the percentage of cells in G1, S, and G2/M phases under *Luci* and *Gen1* KD conditions (n = 3, each colour represents an independent experiment) at 0 h, **(F)** 2 h, **(G)** 4 h and **(H)** 5.5 h after the second Thy release. Statistical significance was determined by unpaired t-test: ns: non-significant.

However, analysis in asynchronous cultures may not detect subtle changes in phase duration distinct from proportion changes. Thus, we synchronised cells in G1/S using a double thymidine block and monitored their progression post-release (**Figure R15C**). While the overall cell cycle profiles were similar (**Figure R15E** to **H**), *Gen1*-depleted ESCs exhibited a tendency to persist longer in the first G2/M post-release compared to controls (**Figures R15D** and **H**).

Analysis of three independent experiments confirmed this tendency although no statistical significance was achieved. Further analysis is necessary to confirm whether *Gen1* depletion causes a delay in the exit of G2/M exit in ESCs to withdraw stronger conclusions.

7.11. *GEN1*-DEFICIENT ESCS EXHIBIT CHROMOSOMAL DEFECTS

Given the potential impact of *Gen1* depletion on cell cycle progression, we next investigated whether sh*Gen1* ESCs displayed evidence of abnormal DNA content. To this end, we analysed. the cell cycle profiles in sh*Gen1* and control sh*Luci* ESCs, to search for a potential increase in cells with DNA content exceeding 4N, which is indicative of abnormal genome duplication events, such as polyploidies or aneuploidies [134, 135]. We observed that a reduction of *Gen1* expression led to a significant increase in the number of polyploid cells (**Figure R16A** and **B**). Both polyploidies and aneuploidies can lead to the formation of micronuclei [136, 137], which are small nuclear structures encapsulating acentric chromosome fragments, chromatids, or entire chromosomes that fail to integrate into daughter nuclei [137]. Such fragments can arise from unrepaired DNA damage, and prior studies have demonstrated that deregulation of DNA repair factors can promote micronuclei formation [138-141].

To determine whether *Gen1* silencing elicited a similar response, ESCs transfected with sh*Gen1* or sh*Luci* were stained with DAPI, and nuclei and micronuclei were scored (**Figure R16C**). While both normal and micronucleated cells were detectable (**Figure R16D**), quantification at 24 and 48 h post-transfection revealed no significant increase in micronuclei upon *Gen1* loss (**Figure R16E**). At 72 h post-transfection, cells were subjected to irradiation to test whether, under conditions enhancing micronuclei formation, GEN1 became necessary. No significant increase in micronuclei formation was observed in either sh*Gen1*- or sh*Luci*-transfected cells under standard or stress-induction conditions (**Figure R16F**). Examination of the slides from the micronuclei experiment revealed the presence of cell fragments stained with DAPI, corresponding to apoptotic cells (**Figure R16G**). Quantification in one of the replicates showed a tendency for these apoptotic cells to be more abundant in *Gen1*-deficient ESCs. (**Figure R16H**). This is consistent with the tendency of ESCs to rapidly eliminate cells with severe chromosomal instability or DNA damage by apoptosis, rather than allowing persistence of abnormal nuclear structures such as micronuclei [142, 143]. Thus, the acute DNA damage response in ESCs may favour the prompt clearance of genetically compromised cells, maintaining overall genome integrity within the stem cell population [144-146].

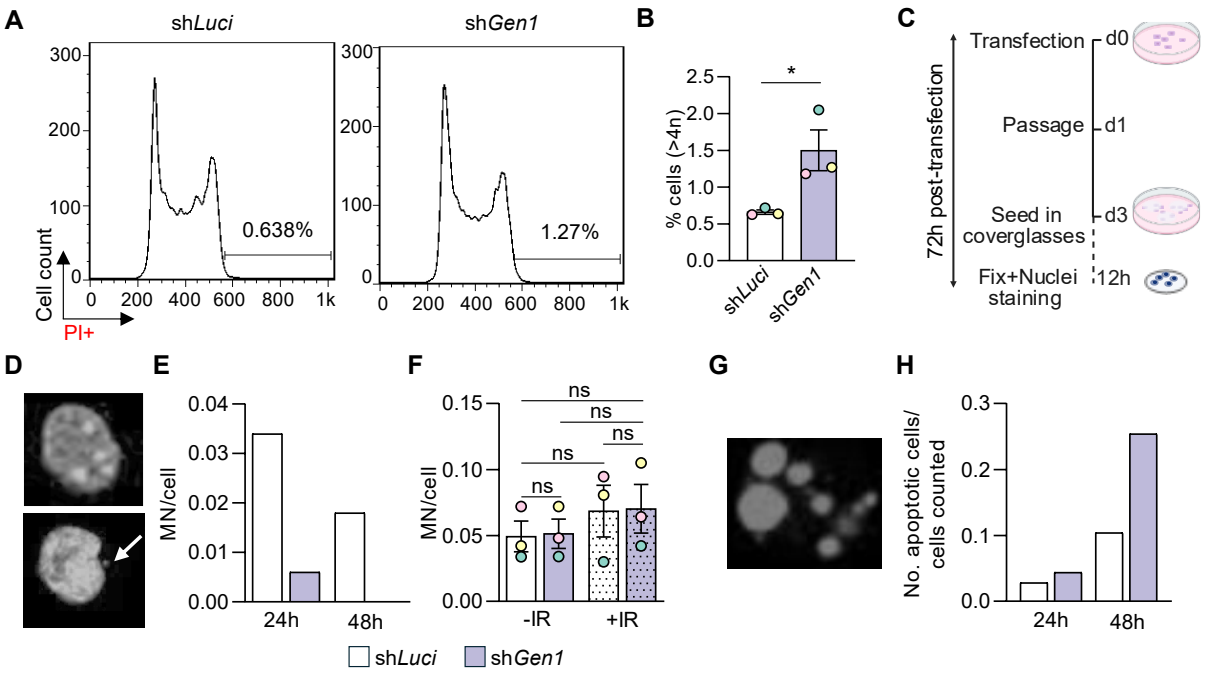

Figure R16. ESCs without *Gen1* exhibit abnormal increase in cells with more than 4n, but this is not due to a micronuclei formation. (A) Representative histogram with the cell cycle profile of ESCs with or without the silencing of *Gen1* showing the gate for the percentage of cells with more than 4n. **(B)** Quantification of the percentage of cells with more than 4n for the two conditions (n=3). **(C)** Scheme of the experimental approach for the detection of micronuclei formation. **(D)** Representative image of a normal cell, upper panel, and a cell with a micronucleus (white arrow), lower panel. **(E)** Quantification of the number of micronuclei per cell at 24, 48 and **(F)** 72 h post-transfection of shRNAs with or without 2 Gy for IR treatment. **(G)** Representative image of an apoptotic cell. **(H)** Quantification of the number of apoptotic cells compared to the number of cells in one of the experiments from sh*Luci* and sh*Gen1* at 24 h and 48 h. Statistical significance was determined by unpaired t-test: *: p-value<0.05; ns: non-significant.

To further investigate potential mechanisms underlying the unusual increase in the population with DNA content greater than 4N, we analysed metaphase spreads to directly visualise structural aberrations, which could result from defects in DNA repair, replication, or segregation caused by *Gen1* loss (**Figure R17A**).

This approach complements the cell cycle and micronuclei analyses by providing a cytogenetic perspective on the effects of *Gen1* silencing. Besides, loss of chromosomal integrity has been linked to loss of *Gen1* in cells with an altered dissolution pathway [127, 147]. In our case, three types of genomic aberrations were analysed: fusions, gaps, and breaks (**Figure R17B**), which are chromosomal abnormalities already described in *Gen1*-deficient 293 cells as well as cells under replication stress conditions [148-150]. While breaks were present under both conditions, there was a tendency for reduced frequency upon *Gen1* KD (**Figure R17C** and

D). In contrast, chromosome fusions were significantly increased in *Gen1*-silenced cells, consistent with prior observations in resolvase-deficient somatic cells (*Gen1* and *Mus81* double-deficient HEK 293 cells, [148]). Treatment with low dose aphidicolin (APH) to induce mild replication stress [149] revealed gaps, indicative of under-replicated, poorly condensed regions, present in both control and KD conditions. Although not statistically significant, there was a tendency toward fewer gaps in *Gen1*-silenced cells (**Figure R17F**). Thus, in conclusion, our data indicates that loss of *Gen1* compromises genome integrity in ESCs primarily through the induction of polyploidy and increased chromosome fusion, underscoring the importance of *Gen1*-dependent pathways in safeguarding the chromosomal stability and self-renewal capacity of pluripotent stem cells. Further analysis will be necessary to confirm these mechanisms and fully elucidate their impact on ESC function and fate.

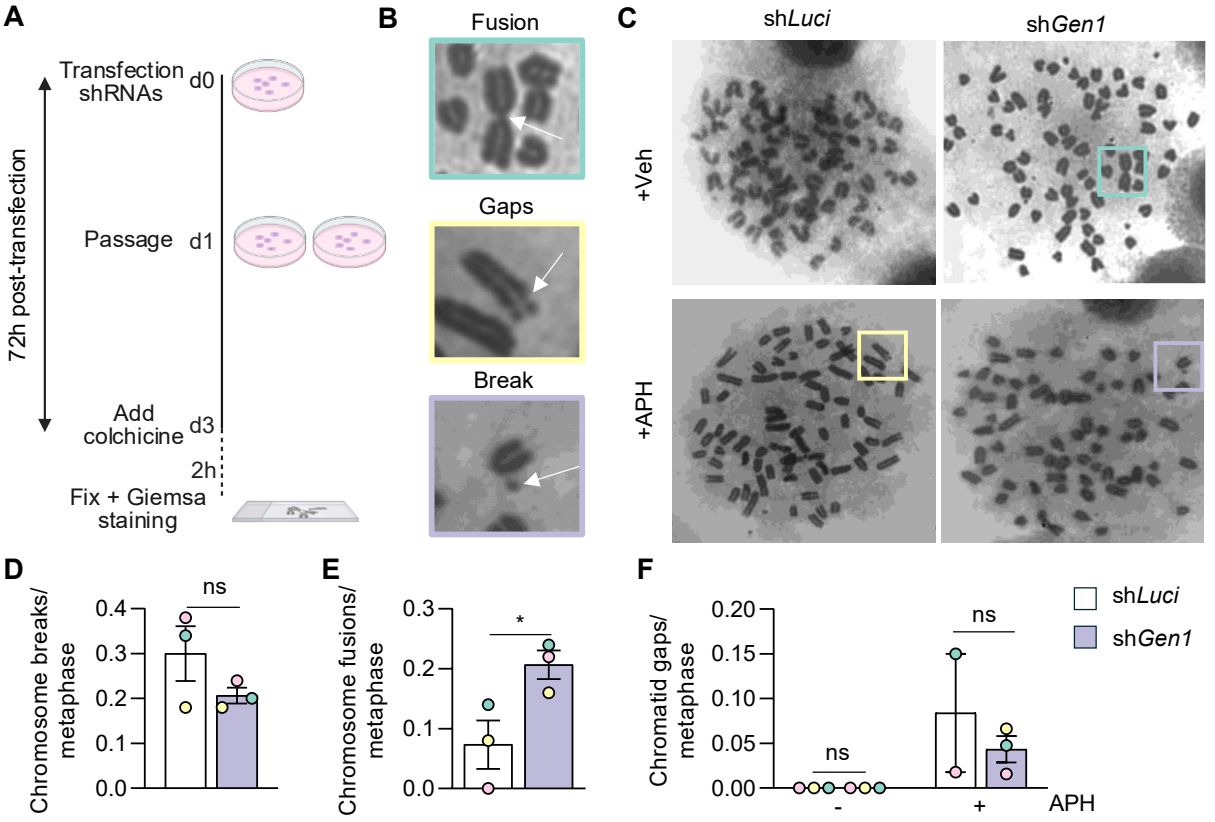

Figure R17. The silencing of *Gen1* promotes chromosome fusions in ESCs. (A) Scheme of the experimental design for the preparation of metaphase chromosome spreads. **(B)** Types of chromosomal aberrations found in the preparations marked with a white arrow, fusions squared in green (upper panel), gaps squared in yellow (middle panel) and breaks in purple (lower panel). **(C)** Representative image of chromosome spreads preparations from ESCs transfected with shRNA against *Gen1* and *Luci* as control (upper panels) and with treatment of aphidicolin (APH) (lower panels). Quantification of chromosome **(D)** breaks, **(E)** fusions and **(F)** gaps per metaphase counted, 50 metaphase per condition were counted (n=3). Statistical significance was determined by unpaired t-test: *: p-value<0.05; ns: non-significant.

7.12. GENES LOCATED IN COMMON FRAGILE SITES ARE DEREGULATED UPON *GEN1* SILENCING IN ESCS

Common fragile sites (CFSs) are genomic *loci* that are particularly susceptible to under-replication and structural instability under mild replication stress [151]. These *loci* frequently overlap with large, transcriptionally active genes, which, under replication stress conditions (e.g., APH treatment), they are "expressed" as gaps and breaks in the chromosomes [152, 153]. DNA damage and under-replication at CFSs often result in incomplete chromatin condensation, persistent DNA lesions, and gaps or breaks visible on metaphase chromosomes [154, 155]. Current evidence indicates that their fragility reflects the difficulty of completing replication in these long, late-replicating domains, which renders them highly vulnerable to replication stress [156]. Transcription–replication conflicts can further aggravate this vulnerability [157]. They can give rise to fork stalling, R-loop accumulation, DNA breaks, and further chromatin disruption, which together exacerbate chromosome fragility and may ultimately disrupt transcriptional regulation or gene expression within these regions [158, 159]. Thus, we addressed whether the transcriptional profile of *Gen1*-silenced cells reflected perturbations on genes at these regions, as a proxy for investigating the accumulation of DNA damage at these *loci*. Taking advantage of a public source of mapped human genes included in CFS regions (HumCFS, [160]), conversion to their mouse homologs was performed and their expression was checked in our RNA-seq dataset. Impressively, approximately one quarter of the differentially expressed genes upon *Gen1* KD were indeed localised within CFS regions (**Figure R18A**), with the majority exhibiting downregulation (**Figure R18B**). This enrichment suggests that genes residing in these inherently unstable *loci* may be particularly sensitive to perturbations in replication and DNA repair processes caused by *Gen1* silencing.

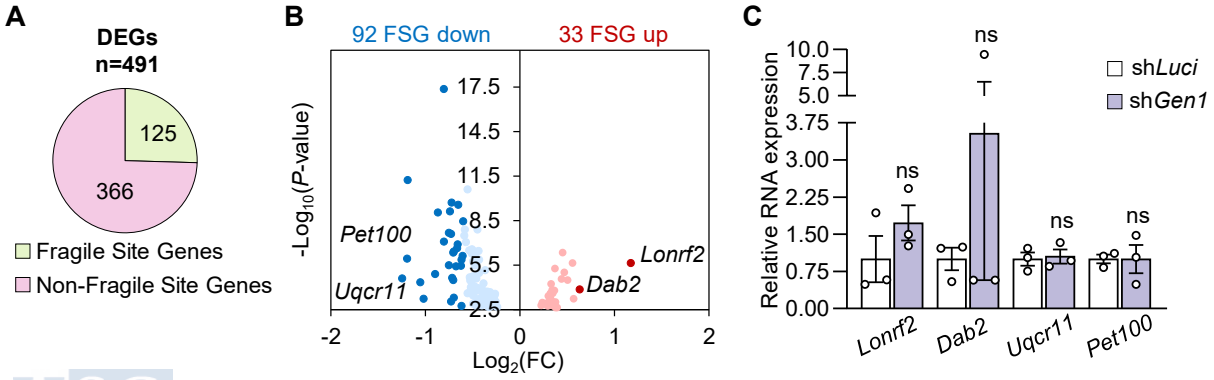

Figure R18. A quarter part of the misregulated genes in absence of *Gen1* correspond to genes localised in common fragile sites regions. (A) Pie chart illustrating the proportion of differentially expressed genes (DEGs)

in *Gen1* KD ESCs which are localised within common fragile site (CFS) regions called fragile site genes (inferring data from HumCFS database). **(B)** Volcano plot with the fragile site genes (n=125) showing the downregulated (blue dots) and upregulated genes (red dots). The bold dots correspond to genes in which |FC|>1.5. **(C)** Relative RNA expression of some fragile site genes chosen from the RNAseq data from ESCs sh*Luci* vs sh*Gen1* determined by RT-qPCR comparting both conditions (n=3). Statistical significance was determined by unpaired t-test: ns: non-significant.

To validate these observations, we selected two downregulated genes (i.e. *Pet100* and *Uqcr11*) and two upregulated genes (i.e. *Dab2* and *Lonrf2*) for independent validation by RT-qPCR in distinct biological replicates. Although the expression trends generally mirrored the patterns observed in the RNA-seq dataset for the upregulated genes, the differences did not reach statistical significance (**Figure R18C**).

Collectively, these results indicate that *Gen1* deficiency does not specifically affect the expression of genes located within common fragile sites. Although transcriptional changes were modest and not confirmed by RT-qPCR, the enrichment of misregulated genes at CFSs suggests that these *loci* may represent vulnerabilities under conditions of compromised DNA repair and replication upon *Gen1* loss.

7.13. *GEN1* SILENCING RESULTS IN INCREASED DNA DAMAGE IN ESCS

We next asked whether loss of GEN1 also leads to increased overall DNA damage, potentially reflecting incomplete repair of homologous recombination intermediates such as accumulation of Holliday junctions [161]. To address this, we used γH2AX as a marker of DNA double-strand breaks. Quantification of the γH2AX signal by immunofluorescence revealed a significant increase in DNA damage in *Gen1*-silenced cells compared with controls, indicating that GEN1 deficiency compromises genome repair (**Figures R19A** and **B**).

To avoid passing mutations and chromosomal aberration to daughter cells, cells generally repair DNA before entering mitosis [162]. To assess whether DNA damage persisted during mitosis, we synchronised ESCs using a double thymidine block, released them until they entered mitosis and then fixed and immunostained them. Analysis of γH2AX in mitotic cells revealed a significantly higher number of *foci* in *Gen1*-silenced cells compared to controls (**Figure R19C**). These findings indicate that, in the absence of GEN1, DNA damage is not fully resolved and persists into mitosis in pluripotent cells.

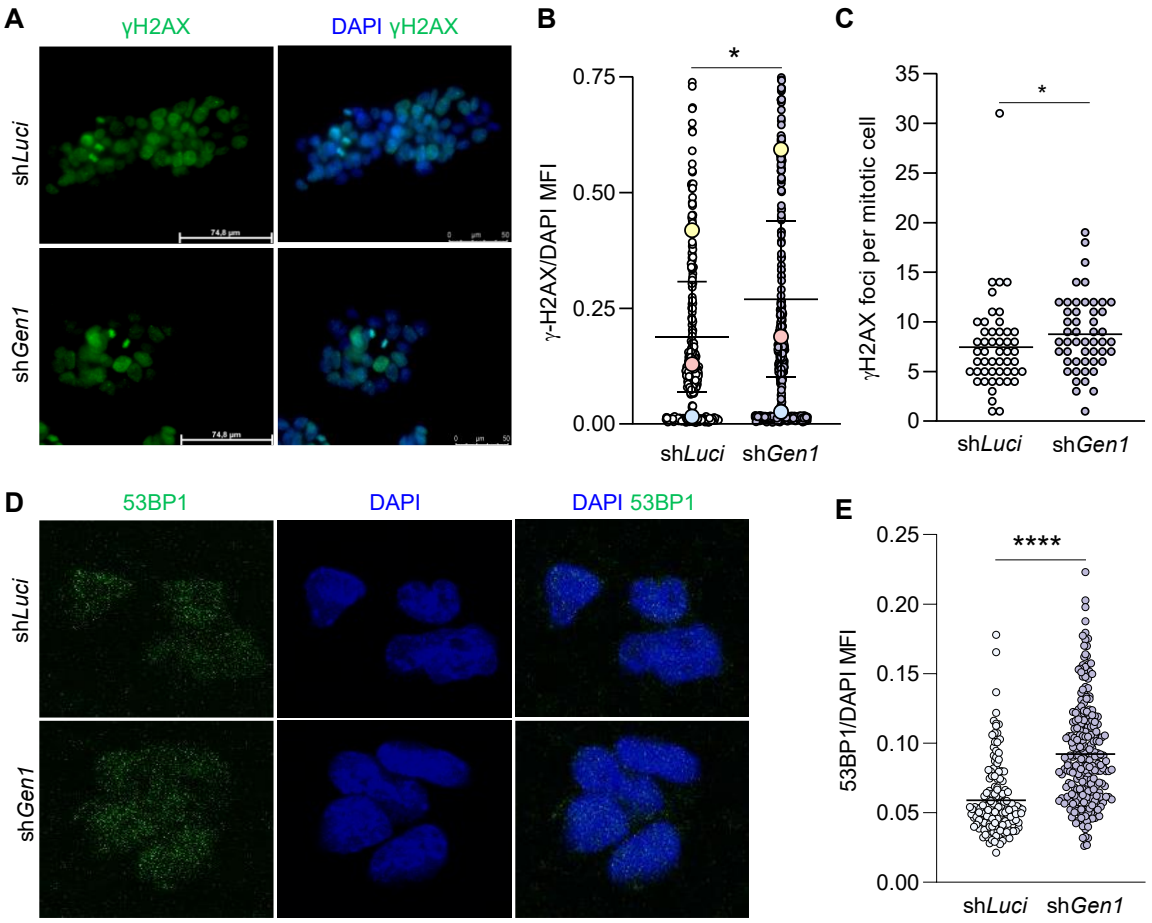

Figure R19. The silencing of *Gen1* leads to higher DNA damage signalling (higher γH2AX deposition and 53BP1). (A) Representative image of immunofluorescence against γH2AX as well as **(B)** the quantification of the mean fluorescence intensity (MFI) for each condition from three different biological replicates (n=3), normalised to DAPI MFI. **(C)** Quantification of the number of foci of γH2AX in cells synched in mitosis in both conditions, *Gen1* KD and sh*Luci* as control. **(D)** Representative image of the immunofluorescence against 53BP1 as well as **(E)** its quantification for the two conditions. Statistical significance was determined by ratio paired t-test for figure B: *: p-value<0.05. For figures C and D, statistical significance was determined by unpaired t-test: *: p-value<0.05 and ****: p-value<0.0001.

We next examined the appearance and intensity of 53BP1, a well-established marker of DNA double-strand break signalling and repair, which often accumulates at sites of unresolved DNA damage. Consistent with the γH2AX results, *Gen1*-silenced cells displayed a marked increase in 53BP1 intensity (**Figures R19D** and **E**), indicating activation of the DNA damage response and the presence of persistent repair intermediates.

These results indicate that *Gen1* deficiency in ESCs leads to elevated DNA damage, including during mitosis, which could be detrimental for genomic integrity. Moreover, *Gen1* loss triggers accumulation of DNA damage response factors such as 53BP1.

7.14. TRANSPOSABLE ELEMENTS EXPRESSION REMAINS MOSTLY UNCHANGED UPON *GEN1* SILENCING IN ESCS

DNA damage has been reported to trigger deregulation of transposable elements (TEs) [163], raising the possibility that increased DNA damage in *Gen1*-silenced cells could influence TE expression. TEs are repetitive genomic sequences capable of mobilisation or transcriptional activation, and their expression is tightly controlled due to their potential to induce genomic instability [164]. In addition, TE sequences have been predicted to form HJs, suggesting a potential mechanistic link between TE regulation, homologous recombination, and DNA repair. Given the role of GEN1 in resolving HJs and the increased DNA damage observed upon *Gen1* KD, we hypothesised that its depletion might alter TE expression. To explore this, we analysed TE expression in our RNA-seq dataset to determine whether loss of *Gen1* correlates with changes in TE expression. We observed a mild deregulation of TE expression with only 145 elements being significantly misregulated (p-adjusted< 0.05 and Fold-change > 1.5) (**Figure R20A**).

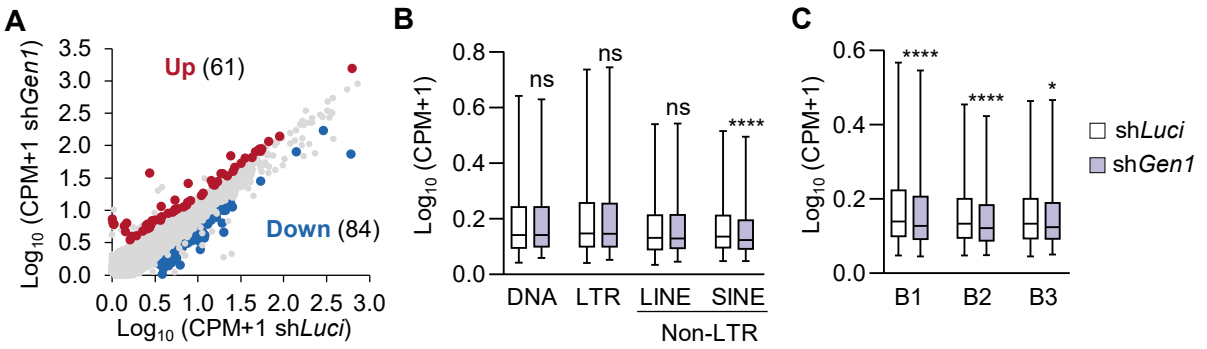

Figure R20. The silencing of *Gen1* leads to discrete changes on the expression of described transposable elements. (A) Scatter plot of expression values of transposable elements (TE) copies (expressed as Log_2(CPM+1)) obtained by RNA-seq analysis using TEtranscripts comparing the expression of sh*Luci* vs sh*Gen1*. The genes which met the criteria of p-adjusted<0.05 and |FC|>1.5 are shown in red for upregulated genes and in blue for downregulated genes (n=2). **(B)** Box-and-whiskers plot representing the expression (Log_2(CPM+1)) of families of TE from SQUiRE analysis. **(C)** Box-and-whiskers plot representing the expression (Log_2(CPM+1)) of subfamilies from the SINE family. Statistical significance was determined by unpaired t-test: *: p-value<0.05; ****: p-value<0.0001 and ns: non-significant.

Notably, among the different TE classes, only SINE elements showed a significant alteration (**Figure R20B**). Among the SINE elements, the three subfamilies, B1, B2, and B3, were all downregulated in *Gen1*-silenced ESCs (**Figure R20C**). While *Gen1* silencing in ESCs leads to elevated DNA damage, its impact on transposable element expression is limited, with significant effects restricted to the SINEs. This suggests that the damage caused by GEN1

deficiency does not broadly activate TEs, although certain SINE elements may be particularly sensitive to changes in genomic stress or chromatin structure.

7.15. OVEREXPRESSION OF HUMAN GEN1 RESCUES THE LOSS OF *GEN1* IN ESCS INDEPENDENTLY OF ITS CATALYTIC ACTIVITY

The observation that *Gen1* silencing in ESCs leads to elevated DNA damage and chromosomal abnormalities raised the question of whether these defects are directly linked to GEN1 catalytic activity or involve additional, non-catalytic functions. To address this, we generated ESC lines overexpressing either wild-type human GEN1 or a catalytically inactive mutant D175A in the XPG-I domain (**Figure R21A**).

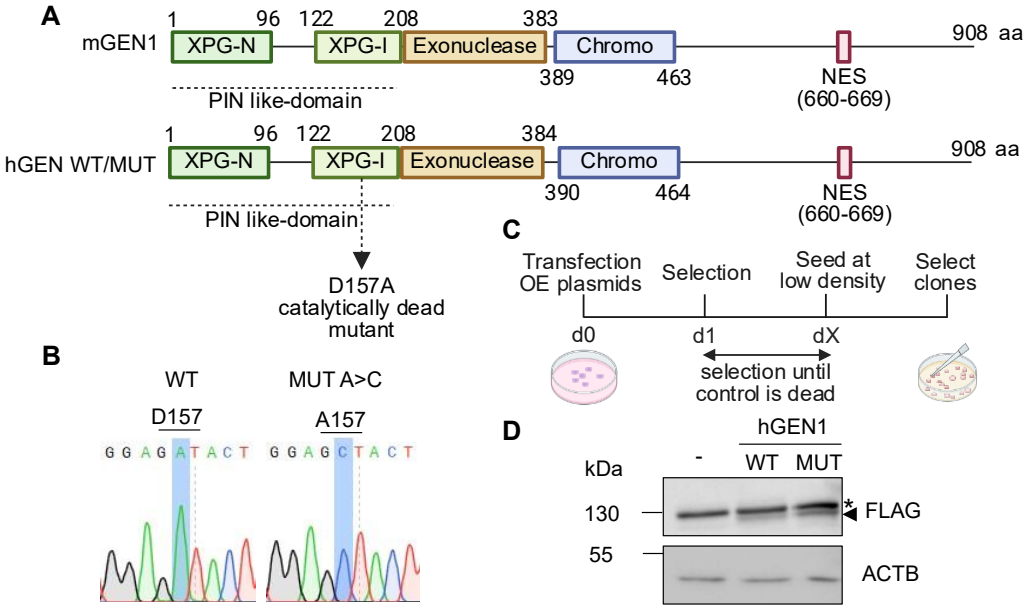

Figure R21. Generation of two new stable cell lines expressing hGEN1 wild type and mutant D157A. (A) Scheme of the protein structure of GEN1 from *Mus musculus* and *Homo sapiens* with some domains highlighted. **(B)** Sanger sequencing showing the point mutation which leads to a change in the amino acid sequence and generates the catalytically death mutant of human GEN1, D157A. **(C)** Scheme of the experimental approach for the generation of the overexpressing cell lines. **(D)** Blot image showing the expression of the endogenous mouse GEN1 (upper band, asterisk) and human GEN1 WT or MUT (lower band, arrowhead) using an anti-FLAG antibody and ACTB was used as loading control.

The catalytically inactive mutant, previously described to be incapable of cleaving HJs *in vitro*, was generated by substituting an aspartic acid (D) residue by an alanine (A) via a single-nucleotide A>C substitution in the coding sequence of hGEN1 (**Figure R21B**) [62]. Using GEN1^{3xFL} ESCs, human GEN1 WT or MUT overexpressing lines were generated and verified

expression of the murine endogenous and human exogenous proteins by Western blot (**Figures R21C** and **D**).

To determine whether the maintenance of ESCs in culture by GEN1 is dependent on its catalytic activity, we used these cell lines and performed a CRISPR/Cas9-mediated KO of the endogenous *Gen1* followed by a CFA (**Figure R22A**).

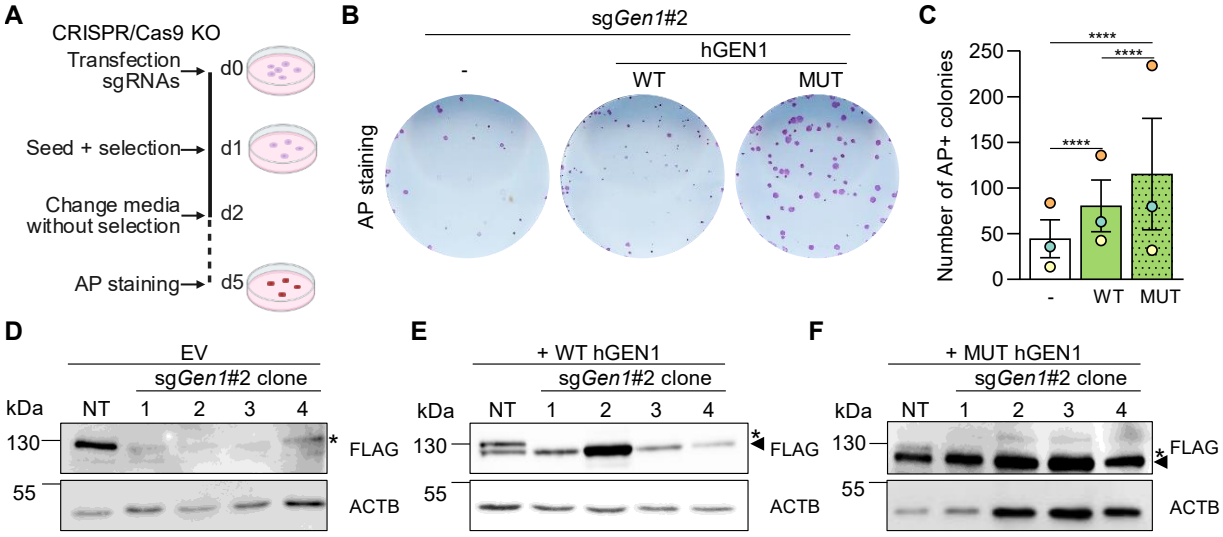

Figure R22. The overexpression of human GEN1 leads to the viability of the mESCs without their endogenous *Gen1*. **(A)** Scheme of the experimental procedure for the KO of *Gen1* in ESCs and colony formation assay. **(B)** Representative images of the CFA from the wells transfected with sg*Gen1#2* in cells without or with overexpression of hGEN1 WT or MUT. **(C)** Quantification of the number of AP positive colonies in ESCs KO for *Gen1* without or with the overexpression of hGEN1 WT or MUT. Blot images of selected clones after KO of GEN1 in **(D)** GEN1^{3xFL} ESCs, **(E)** GEN1^{3xFL} + WT hGEN1 ESCs and **(F)** GEN1^{3xFL} + MUT hGEN1 ESCs. FLAG was used to detect the endogenous mGEN1 (upper band, asterisk) as well as hGEN1 (lower band, arrowhead). ß-actin (ACTB) was used as loading control. Statistical significance was determined by 2-way ANOVA using Tukey's multiple comparison test: ****: p-value<0.0001.

AP-staining revealed that overexpression of both WT and catalytically inactive hGEN1 (MUT) were sufficient to avoid the reduction in colony formation upon endogenous *Gen1* depletion (**Figures R22B** and **C**). To validate these results, we examined endogenous and exogenous protein expression in individual clones. In ESCs carrying the control empty vector, together with the *Gen1*-targeting sgRNA, we detected residual GEN1 protein, suggesting incomplete disruption of the endogenous *Gen1 locus* by CRISPR/Cas9 and the possible presence of one intact allele (**Figure R22D**). In contrast, in clones overexpressing either WT or MUT hGEN1, the endogenous mouse *Gen1* band was absent (**Figures R22E** and **F**), while the human transgene was readily detected as the lower band across clones by Western blot. This

suggests that GEN1 function in ESCs maintenance is independent of its ability to cleave HJs, at least independent of the catalytical-relevant D157 residue.

7.16. GEN1 INFLUENCES ESCS VIABILITY IN RESPONSE TO REPLICATION STRESS AND TOP1-ASSOCIATED DNA LESIONS

As our data indicated that GEN1 supports ESC viability independently of its HJ resolution function, we sought to test whether *Gen1* silencing might influence cellular sensitivity to genotoxic agents that generate distinct types of DNA lesions or replication-associated secondary structures. As a first step, we carried out a preliminary screening by assessing the viability of *Gen1*-silenced ESCs treated with different genotoxic drugs using an MTT assay (**Table 10** and **Figure R23A**).

Table 10. Summary of genotoxic agents, their primary targets and resulting DNA lesions.

Drug	Primary Target	Type of DNA Lesion
Cisplatin	DNA crosslinking	Inter- and interstrand DNA crosslinks
Bleomycin	DNA strand scission via oxidative damage	DNA double-strand breaks generated through free radicals
Olaparib	PARP inhibition, prevents single-strand break repair	Trapped single-strand break intermediates
Hydroxyurea (HU)	Ribonucleotide reductase inhibition	Stalled replication forks caused by dNTP depletion
4-nitriquinoline-1-oxide (4-NQO)	DNA adduct formation, mimics UV-induced helix distortion	Bulky DNA adducts resembling UV-induced lesions
Camptothecin (CPT)	Topoisomerase I inhibition	Replication-associated double-strand breaks via Topoisomerase I cleavage complexes

ESCs silenced for *Gen1,* and sh*Luci* as control, were exposed to different DNA-damaging agents. Their viability was evaluated using the MTT assay, which detects mitochondrial metabolic activity as an indirect readout of cell viability and proliferation. For cisplatin and bleomycin, no significant differences were observed between the two conditions (**Figures R23B** and **C**, respectively). Treatment with olaparib produced inconsistent responses at low doses across the two shRNAs, therefore, suggesting unspecific effects (**Figure R23D**). Similarly, at higher concentrations of hydroxyurea (HU) or 4-NQO, the sensitivity of *Gen1*-silenced cells was comparable to controls, although increased sensitivity was detected at low doses (**Figures R23E** and **F**, respectively). In contrast, lack of *Gen1* produced a clear survival advantage in ESCs treated with various concentrations of camptothecin (CPT) **(Figure R23G**).

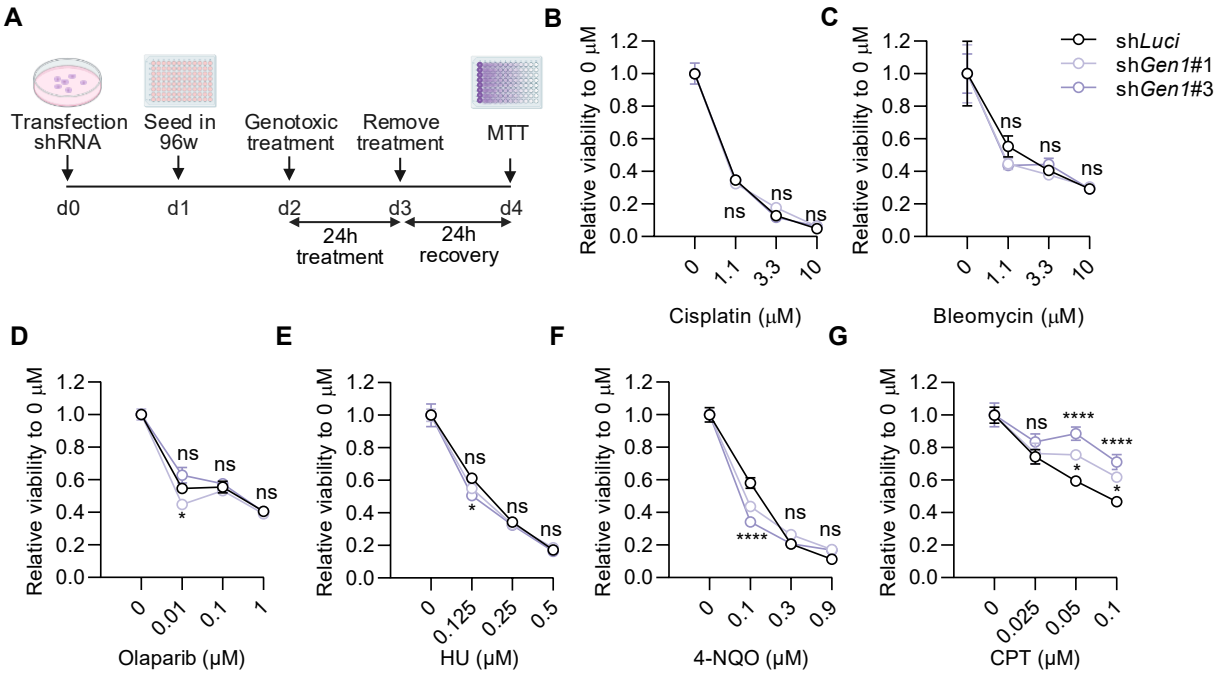

Figure R23. The silencing of *Gen1* does not change the response to genotoxic reagents except for camptothecin. (A) Scheme of the experimental approach for the viability assay with or without the treatment of the different DNA damaging drugs. Relative viability to untreated samples for each KD condition (sh*Luci* and two different shRNAs against *Gen1*) with increasing concentrations of **(B)** cisplatin, **(C)** bleomycin, **(D)** Olaparib, **(E)** HU, **(F)** 4-NQO and **(G)** CPT. Statistical significance was determined by 2-way ANOVA with a Dunnett's multiple comparison test: *: p-value<0.05; ****: p-value<0.0001 and ns: non-significant.

These results prompted us to ask whether the observed survival effect was specific to topoisomerase I inhibition by CPT, which traps the TOP1 cleavage complex (TOP1cc), leading to replication fork collisions and DSB formation, or if a similar effects could be seen upon topoisomerase II (TOP2) inhibition by etoposide (Eto) too, which also induces irreversible DSBs, but through stabilisation of the TOP2 cleavage complex (**Figure R24A**). To test this, we performed an MTT assay, with increasing CPT concentrations to determine whether the survival advantage persisted at higher doses (**Figure R24B**). Notably, the protective effect was only evident with CPT treatment and not with Eto, indicating that it is specific to TOP1 inhibition (**Figures R24C** and **D**). At higher doses of CPT, the survival advantage of *Gen1*-silenced ESCs was no longer observed, likely due to excessive DNA damage overwhelming the repair mechanisms in place. To assess whether *Gen1* silencing affects the cellular response to CPT, we performed an alkaline comet assay, a technique that measures DNA strand breaks at the single-cell level by evaluating the migration of fragmented DNA during electrophoresis (**Figure R24E**). In control ESCs (sh*Luci*), CPT treatment led to a clear increase in DNA damage compared to untreated cells (**Figures R24F** and **G**), as previously reported [165]. Interestingly,

although *Gen1*-silenced cells showed baseline DNA damage similar to untreated controls, they did not exhibit the same increase upon CPT treatment (**Figures R24F** and **G**). Therefore, *Gen1*-silenced ESCs accumulated fewer DNA breaks than sh*Luci* cells in response to CPT. This suggests that GEN1 may be involved in processing replication-associated DNA intermediates generated by TOP1 inhibition, potentially altering how these lesions are resolved and thereby influencing cell survival.

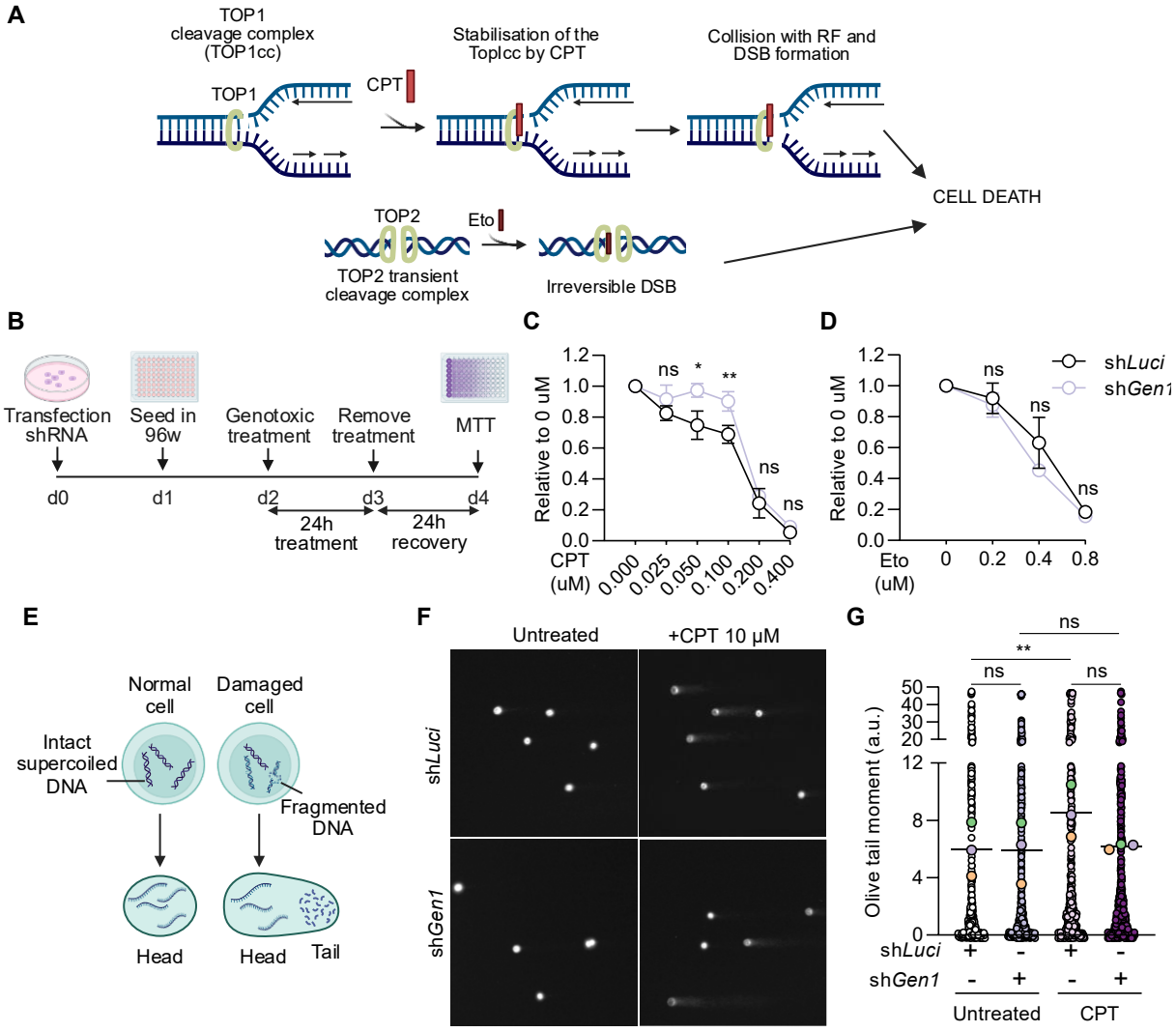

Figure R24. The silencing of *Gen1* affects the response to camptothecin in ESCs. **(A)** Mechanism of action of CPT and Eto against topoisomerase I and II, respectively. **(B)** Scheme of the experimental approach for the MTT assay using the two different drugs. Survival of cells upon treatment with increasing concentrations of **(C)** CPT or **(D)** Eto with the two shRNAs (sh*Gen1* and sh*Luci* as control). **(E)** Schematic representation of the basis of the alkaline comet assay: normal cells have intact supercoiled DNA does only lead to a comet head and damaged cells have fragmented DNA responsible for the comet tails. **(F)** Representative images of the obtained comets for the different experimental conditions. **(G)** Quantification of the olive tail moment using arbitrary units with and without CPT treatment in silencing conditions of *Gen1* or *Luci* as control. Statistical significance was determined

by unpaired t-test **(C and D)** and one-way ANOVA with Sidak's comparison test for **(G)**: *: p-value<0.05; **: p-value<0.01 and ns: non-significant.

8. Discussion

ESCs are characterised by their capacity of self-renewal and differentiation. To safeguard their genomes, ESCs rely on highly accurate DNA repair pathways to counteract replication stress and maintain genomic stability, avoiding transmission of mutations to daughter cells. HJ processing enzymes BLM, GEN1 and MUS81 play a central role by resolving DNA intermediates generated during replication and homologous recombination. Despite their recognised importance, the specific contribution of individual SSEs to genome maintenance in ESCs remains poorly defined. Elucidating the molecular processes governed by BLM, GEN1 and MUS81 to safeguard genomic stability is essential not only for understanding the molecular basis of pluripotency but also for ensuring the genomic integrity of ESCs intended for regenerative medicine applications.

The principal aim of this thesis was to identify the role of these enzymes in pluripotency maintenance using ESCs as a model. Our findings identified GEN1 as the main resolvase in ESCs whose complete depletion is incompatible with cell survival. These results provide for the first time a new paradigm where GEN1, usually regarded as a backup system operating downstream of BLM and MUS81, becomes essential for cell maintenance (**Figure R3**), with presence of the other two HJ processing enzymes unable to compensate for its loss. Analysis of public datasets confirmed GEN1 is the most highly expressed resolvase (i.e., compared to MUS81) in ESCs (**Figure R2**) and it is essential for cell survival of this unique cell type (**Figures R3-R5**).

When analysing the phenotype of GEN1-deficient ESCs, we observed proliferation defects which could not be explained by changes in cell cycle profile (**Figure R15**) or a dramatic increase in cell death by apoptosis (i.e., Annexin V-positive cells) at 72 h post-transfection (**Figure R14**), although, we did observe the presence of apoptotic cells at earlier time points (24 and 48 h post-transfection) (**Figure R16**). Interestingly, while not statistically significant, sh*Gen1* ESCs exhibited a trend toward delayed progression in the M to G1, which may indicate subtle defects in cell cycle progression caused by GEN1 loss in spite of no gross defects being observed in asynchronous populations (**Figure R15**). Our results show that loss-of-function of *Gen1* lead to the appearance of chromosomal aberrations, as well as increased DNA damage (i.e. higher γH2AX and 53BP1 deposition) (**Figures R17** and **R19**). Importantly, this genomic instability did not cause a dramatic deregulation of TEs expression (**Figure R20**), which are often used as markers of genome instability [164].

Rescue experiments provided additional insight into GEN1 function. Both catalytically active and inactive forms of GEN1 were able to rescue the reduced ESC survival upon sh*Gen1*, suggesting a role independent of its nuclease activity (**Figure R22**). This suggests that GEN1 contributes to ESC survival through a non-catalytic mechanism. Our preliminary data indicate that this function may occur in the context of replication stress, possibly by facilitating the resolution or stabilisation of replication-associated DNA intermediates for their repair rather than by directly cleaving them (**Figures R23** and **R24**).

In summary, this thesis identifies GEN1 as an essential factor for ESC survival and genome maintenance. Beyond its canonical role as a HJ resolvase, GEN1 appears to have additional functions in DNA replication and repair in ESCs, raising new questions about its specific substrates and/or partners in pluripotent cells. This work further highlights the importance of faithful DNA repair in ESCs to prevent chromosomal alterations, ensuring their stability both as a research model and as a resource for regenerative medicine.

8.1. GEN1 IS ESSENTIAL FOR ESCS MAINTENANCE

In this study, we focused on GEN1 which emerged as highly expressed in ESCs (**Figure R2**). Our results show that complete loss of GEN1 is incompatible with ESC survival (**Figure R3**). As with many other HR proteins (please see **section 4.6**), GEN1 has been proven to be essential for the maintenance of ESCs *in vitro* (**Table 2** and **Figure R4**), as only cells that still express GEN1 can be kept in culture (**Figure R5**). In contrast, *Gen1*-deficient mice are viable [100], likely because the pluripotent state in embryonic development is transient, lasting about 2-3 days, enabling potential compensation by other cellular mechanisms after the loss of *Gen1* during this brief window [166]. Interestingly, other structure-specific endonucleases such as MUS81 are dispensable for ESC survival (**Figure R3**) [106], suggesting that GEN1 plays a non-redundant role in repairing replication-associated intermediates unique to pluripotent cells. This finding places GEN1 in a distinct position, since for other cell types it has been described primarily as a backup pathway, acting when MUS81 is not able to process all HJs when cells enter in mitosis [167, 168].

Several differences of pluripotent versus differentiated cells may explain this difference in GEN1 dependence. First, previous studies in somatic cells have demonstrated that GEN1´s activity is tightly regulated by a nuclear export signal, which confines the protein to the cytoplasm and allows its access to DNA only during mitosis, when the nuclear envelope breaks

down [86]. This spatial restriction significantly limits its functional contribution in somatic cells, where it can only act within a narrow temporal window. In contrast, we have seen that a small proportion of the protein remains in the nuclear compartment, even when cells are in G1 phase in ESC (**Figure R6**). This could be due to the shorter G1 length of ESCs in metastable SL culture media and their frequent cell cycling, making it possible that its function may extend beyond canonical HJ resolution including additional roles in maintaining genomic stability. Second, in contrast to somatic cells, multiple studies have proved that ESCs tend to favour cell death or differentiation rather than attempting to repair the damaged DNA, which could otherwise lead to mutations in the daughter cells [144, 145, 169, 170]. Although the exact mechanism underlying the loss of viability of *Gen1*-deficient ESCs remains to be determined, our data provide clear evidence of persistent genomic stress, shown by increased deposition of DNA damage markers (i.e., γH2AX or 53BP1) (**Figure R19**), together with an increase in apoptotic cells at early timepoints after KD (**Figure R16H**). Finally, despite the upregulation of genes associated with DNA repair pathways (**Figures R9F** and **G**) and elevated damage markers (**Figure R19**), alkaline comet assays did not reveal an increase in total DNA breaks upon *Gen1* reduction, indicating that GEN1 loss could lead to the accumulation of unresolved DNA structures (i.e. stalled replication forks or recombination intermediates) rather than clearly detectable single- or double-strand DNA breaks (**Figure R24**).

Together, these findings highlight that GEN1 is indispensable, rather than a mere backup, for preserving genomic stability in ESCs, preventing the accumulation of unresolved DNA intermediates that elicit chronic damage signalling and compromise cell survival.

8.2. Impact of GEN1 loss on cell cycle and cell death in ESCs

ESCs are distinguished for their rapid proliferation and unique cell cycle organisation. The results of this thesis have shown that silencing *Gen1* does not significantly alter the overall cell cycle distribution of asynchronous ESCs (**Figures R15A** and **B**). However, a more detailed analysis of synchronised ESCs revealed a slight delay in the exit from the M to G1 phase in *Gen1*-silenced cells in the first mitosis after synchronisation (**Figure R15**). Since ESCs cultured in SL media lack a functional G1/S checkpoint, cells progress to S phase despite the DNA damage that they might contain. Nevertheless, when they try to enter mitosis, this damage can trigger the G2/M checkpoint, which is partially active in ESCs [171, 172], leading to cell cycle arrest. The modest delay observed in the M to G1 transition, without a clear G2/M arrest in

asynchronous populations, suggests a checkpoint adaptation mechanism rather than a robust checkpoint block. It is plausible that ESCs subjected to *Gen1* silencing bypass the canonical G2/M checkpoint and proceed to mitosis despite carrying unresolved DNA damage. Additional phase-specific analyses, such as assessing phospho-histone H3 as a mitotic marker and employing live-cell imaging, could help further validate this hypothesis by precisely monitoring cell cycle transitions following *Gen1* silencing.

Given that our data show that death by apoptosis is not clearly detectable in GEN1-deficient ESCs at 72 h post-silencing it could be possible that persistent DNA lesions activate alternative forms of non-programmed cell death in response to loss of GEN1 function. Among these, mitotic catastrophe or necrosis represent plausible outcomes, as cells carrying unresolved DNA damage may progress through the cell cycle despite checkpoint activation [173]. Mitotic catastrophe was first discovered in yeast [174] and, in mammalian cells it is linked to mitotic failure after DNA damage, spindle assembly defects, or chromosomal missegregation [175]. Even though it was discovered in the late 1980s, little is known about the mechanisms behind the process, beyond the fact that it is a tumour-suppressive mechanism to avoid proliferation of cells harbouring chromosomal abnormalities. A hallmark of this phenotype is the change in mitochondrial network morphology. During mitotic catastrophe, mitochondrial membranes experience limited permeabilisation, which restricts caspase activation. This insufficient activation fails to trigger programmed cell death, such as apoptosis. However, it can still lead to DNA damage and genomic instability through the action of active DNases [176]. In the case of *Gen1*-silenced ESCs, we observed the deregulation of genes related to mitochondrial function (**Figure R13**), which could potentially contribute to or favour this type of cell death.

The mechanism of cell death in ESCs following DNA damage is profoundly influenced by their *in vitro* culture conditions, particularly the composition of the culture medium. In SL media, where ESCs lack a functional G1/S checkpoint and show compromised p53-p21 pathway activity, damaged cells can continue to S phase, and classical p53-mediated apoptotic responses may be attenuated [177]. This may explain why we do not observe an increase in the apoptotic population following *Gen1* silencing (**Figures R14D** and **E**), as the altered checkpoint and signalling context can bias damaged cells toward alternative death pathways, such as necrosis, rather than classical apoptosis. These observations highlight the complexity of cell death regulation in pluripotent cells and underscore the need for cautious interpretation of viability data in the context of GEN1 depletion. To fully elucidate the mechanisms driving cell

death in GEN1-deficient ESCs, additional experiments will be required. In example, time-lapse microscopy combined with live-cell markers for apoptosis, necrosis and mitotic progression would allow real-time tracking of cellular outcomes following GEN1 silencing. In addition, evaluating mitochondrial membrane integrity with analysis of mitotic spindle assembly could yield direct evidence of mitotic catastrophe. Together, these approaches would enable a comprehensive characterisation of the pathways that limit ESC survival upon loss of GEN1, providing deeper mechanistic insights into the interplay between DNA repair and cell cycle control in pluripotent cells.

8.3. TRANSCRIPTIONAL CHANGES AND DIFFERENTIATION BIAS IN *GEN1*-DEFICIENT ESCS

In *Gen1*-deficient ESCs, we detected a reduced number of transcriptionally altered genes (**Figures R8** and **R9**). Some genes associated with oxidative phosphorylation appeared downregulated, whereas a subset of genes located within CFS regions showed altered expression that could not be consistently confirmed across independent biological replicates (**Figures R13** and **R18**, respectively). This pattern suggests that the transcriptional changes observed upon *Gen1* silencing are more likely indirect consequences of genome instability and replication stress than evidence of a strong transcriptional role of GEN1.

A notable and reproducible finding was the downregulation of endodermal markers during EB differentiation of ESCs in the absence of GEN1 (**Figure R11**). In human pluripotent stem cells, it has been shown that endodermal commitment is particularly susceptible to DNA damage [178]. Low levels of genotoxic stress during a defined window of differentiation stabilise p53 and divert cells away from the definitive endoderm towards a mesodermal fate, without triggering apoptosis. These data point to a link between genome instability and impaired lineage specification that would explain why the endoderm layer is the most affected by GEN1 absence, in contrast to ectoderm and mesoderm (**Figure R11A** and **B**). The available evidence highlights that endoderm differentiation is especially sensitive to genome instability, whereas the effects on ectodermal and mesodermal lineages are less well characterised and warrant further investigation.

In vivo data also suggest that GEN1 is required for normal organogenesis in multiple lineages. *Gen1* mutations in mice cause a spectrum of congenital anomalies of the kidney and urinary tract (CAKUT), with severity correlating with the degree of *Gen1* loss [101, 179-182]. As these structures derive largely from intermediate mesoderm, these findings support the idea

that impaired genome maintenance can destabilise multiple developmental programmes. Although this appears to differ from our ESCs data, which reveal a more pronounced defect in endodermal differentiation, this discrepancy is likely influenced by the distinct signalling environment *in vitro* and *in vivo*. In particular, ESC-based systems lack the maternal-fetal interactions, tissue-tissue crosstalk and physiological morphogen gradients.

Finally, the study by Wang and collaborators from 2018 shows that GEN1 interacts with SIX1, a developmental TF, to promote transcriptional activation during embryogenesis [101]. These finding suggest that GEN1 contributes to the control of developmental gene expression in addition to its canonical role in HJ resolution and DNA repair. In this context, loss of GEN1 may have a double impact, combining genome instability with misregulated developmental transcriptional programmes upon pluripotency exit. This dual defect could help explain why GEN1 deficiency leads to differentiation defects in ESCs and organogenesis in mouse models, as it can affect both genome integrity and transcription through interaction with TFs.

8.4. GEN1 IS REQUIRED FOR MAINTENANCE OF CHROMOSOMAL INTEGRITY

Cytogenetic analysis of *Gen1*-silenced ESCs revealed a tendency toward fewer visible chromosome breaks compared to control cells, although this difference did not reach statistical significance (**Figure R17D**). In contrast, we observed a clear and significant increase in chromosome fusions, indicative of a distinct form of structural genome damage upon *Gen1* depletion (**Figure R17E**). These findings suggest that unresolved replication intermediates caused by loss of *Gen1* manifest as fusions rather than breaks. In *Gen1*-silenced ESCs, the lack of proper processing of DNA replication and recombination structures does not lead to extensive chromosome breakage detectable by standard cytogenetic analysis in our assays. Instead, the unresolved intermediates may persist into mitosis and lead to the physical joining of chromosome segments reflecting a distinct structural form of genome damage when GEN1 is not fully functional. This may reflect a unique adaptation of pluripotent cells to supress overt DNA fragmentation at the expense of structural rearrangements. Furthermore, these results align with Chan and colleagues, who demonstrated that unresolved recombination intermediates can generate chromosomal lesions, supporting the notion that structural genome instability arises from unresolved DNA structures rather than visible chromosome breaks [148].

These observations are supported by mechanistic insights from the literature. Benítez et al showed that GEN1, together with MUS81, promotes the cleavage of under-replicated DNA at

common fragile sites (CFS) [161]. In their study, absence of GEN1 result in persistence of unresolved regions, leading to a reduction in the cytogenetic expression of fragile sites which are observed as a reduction in chromosome breaks despite ongoing replication stress. Supporting this model, our RNA-seq analysis revealed that approximately 25% of misregulated genes in *Gen1* KD ESCs are located within CFS regions (**Figure R18**). However, this association should be interpreted as positional and probabilistic rather than causal, since RT-qPCR validation for selected targets did not consistently confirm significant changes for specific CFS genes across biological replicates (**Figure R18**). This suggests that, although the specific misregulated genes may differ between experiments, they tend to cluster within CFS regions and show a higher susceptibility to GEN1 loss compared to other genes. This is consistent with the idea that late-replicating domains such as CFSs, are intrinsically vulnerable to replication stress, in part because they rely on long-travelling replication forks that are easily perturbed in the absence of efficient fork protection and restart [156].

Taken together, our data indicate that GEN1 is essential for maintaining chromosome integrity in ESCs. The combination of latent DNA damage, transcriptional misregulation at CFSs, and increased chromosomal fusions highlights that GEN1 loss alters both the structural and functional integrity of the genome. To further confirm the presence of unresolved replication intermediates, future studies using DNA fibre assays or electron microscopy could provide mechanistic insight into the identity of the unresolved structures accumulating in the absence of GEN1 in ESCs.

8.5. NONCANONICAL FUNCTIONS OF GEN1 IN THE REPLICATION STRESS RESPONSE

Rescue experiments provided new insights into the role of GEN1 in ESCs. Expression of both the catalytically active and inactive forms of human GEN1 was sufficient to restore the number of pluripotent colonies observed when the endogenous GEN1 was depleted (**Figure R22**). This observation suggests that GEN1 supports ESC viability through mechanisms not solely dependent on its canonical HJ resolvase activity, implying additional roles beyond DNA cleavage. As mentioned in **section 4.7**, *in vitro* assays using human GEN1 showed that its classical function relies on its endonuclease domain, with the catalytic residue D157 being essential for the resolution of HJs and related branched DNA structures [94, 95]. However, the finding that a catalytically inactive GEN1 variant (mutated at D157) is still able to support pluripotent colony formation in absence of the endogenous *Gen1* function, implies the existence

of non-catalytic, but likely structural or scaffolding functions of this protein in ESCs. Potentially, these functions may involve recruitment or stabilisation of repair factors at sites of replication-associated DNA lesions, independent of its nuclease activity.

In parallel, screening of different genotoxic reagents in ESCs revealed a specific function of GEN1 in processing intermediaries generated by TOP1 cleavage complex (TOP1cc) after treatment with camptothecin (CPT), an effect not observed under TOP2 inhibition with etoposide (Eto) (**Figure R24**). This CPT-sensitivity data supports a specific requirement for GEN1 in resolving replication-associated lesions arising from TOP1cc. One major difference between TOP1 and TOP2 is their cleavage polarity. TOP1 has preferential polarity for the 3' end of ssDNA breaks and CPT functions trapping TOP1cc, thereby leaving TOP1 covalently attached to the 3' end of the ssDNA break [183-185]. During replication, forks that encounter TOP1cc can stall or collapse, generating single-ended DSB and giving rise to resected or regressed forks (5'-flap-like or HJ-like structures) [186-188], which GEN1 might resolve. These DNA lesions are usually repaired through HR, while NHEJ is actively suppressed at broken forks to prevent toxic misrepair, a regulation mediated by ATM [189]. Our results make it possible that the DNA lesions caused by CPT treatment may be redirected into NHEJ or alternative repair pathways in the absence of GEN1 function, allowing partial survival of ESCs under CPT-induced replication stress (**Figure R23** and **R24**). Although direct evidence for such a pathway switch specifically in GEN1-deficient ESCs is lacking, studies in stem cells reported similar compensatory use of NHEJ or alternative repair pathways under replicative or genotoxic stress [190]. This may explain why *Gen1*-deficient ESCs remain more viable under CPT treatment compared to sh*Luci* control, as unresolved replication intermediates might be diverted toward NHEJ or alternative repair pathways in the absence of GEN1.

Nevertheless, both in absence of *Gen1* and without CPT, probably unresolved intermediates continue to cause persistent stress (i.e. higher γH2AX and 53BP1 deposition, **Figure R19**), likely due to the limited capacity of NHEJ to compensate for HR deficiency. In this setting, homologous recombination would normally rely on GEN1 and related structure-specific endonucleases to process these structures, and the residual NHEJ activity appears insufficient to fully compensate for this deficiency.

Consistently, alkaline comet assays revealed a lower number of total DNA breaks in *Gen1*-deficient cells following CPT treatment (**Figures R24F** and **G**). A similar observation was

reported for *Mus81*$^{-/-}$ human HCT116 colon carcinoma cells, in which *Mus81* was shown to be necessary for the resolution of replication forks stalled by CPT treatment [191]. Similarly, in GEN1-deficient ESCs, replication-associated intermediates may evade conversion into measurable breaks via alternative repair pathways or structural rearrangements that sequester or ligate the ends without generating detectable DSBs [192-194]. To directly test whether non-canonical function of GEN1 promotes HR over NHEJ, several approaches could be employed in future studies. Reporter assays that distinguish HR-versus NHEJ mediated repair could quantify the pathway of choice in GEN1-deficient ESCs under replication stress. Complementary analyses using immunofluorescence for HR factors (e.g., RAD51 or BRCA1) and NHEJ factors (e.g., XRCC6/5) at stalled forks would reveal whether GEN1 loss shifts the balance towards an error-prone repair. Alternatively, MTT-based viability assay could be performed in GEN1-depleted cells treated with an HR inhibitor such as RI(dII)-2 TFA or the DNA-PKs inhibitor KU-57788 (NU7441) to block NHEJ. Comparing the sensitivity of control and GEN1-deficient ESCs to these inhibitors upon CPT treatment would help determine whether loss of GEN1 alters the relative contribution of HR and NHEJ to the repair of replication-associated lesions. Together, these experiments could clarify the mechanistic basis of the function of GEN1 independently of HJ resolution in safeguarding genome stability in ESCs.

8.6. CONCLUDING REMARKS

Collectively, our findings point to GEN1 as a central factor in preserving genome stability in ESCs, functioning not only as a canonical Holliday junction resolvase, but also potentially contributing to replication fork integrity and influencing the choice of DNA repair pathways. The continuous nuclear presence of GEN1 in ESCs, its non-redundant role with MUS81 and its ability to support cell survival independently of its nuclease activity distinguish its function in pluripotent cells from that observed in somatic contexts.

Mechanistically, GEN1 plays a dual role in maintaining genome integrity and regulating DNA repair dynamics in ESCs. Under basal conditions, GEN1 is essential for survival, likely by resolving replication and recombination intermediates that would otherwise cause persistent genomic stress. Upon CPT-induced replication stress, GEN1 facilitates repair by HR. However, excessive damage may overwhelm HR capacity, resulting in cell death. GEN1-deficient cells and CPT-treated would exhibit impaired HR and might rely on other repair pathways such as

NHEJ, which enables survival through error-prone repair processes that compromise genomic integrity (**Figure D1**). These observations highlight the delicate balance between cell survival and the maintenance of genomic fidelity in pluripotent cells.

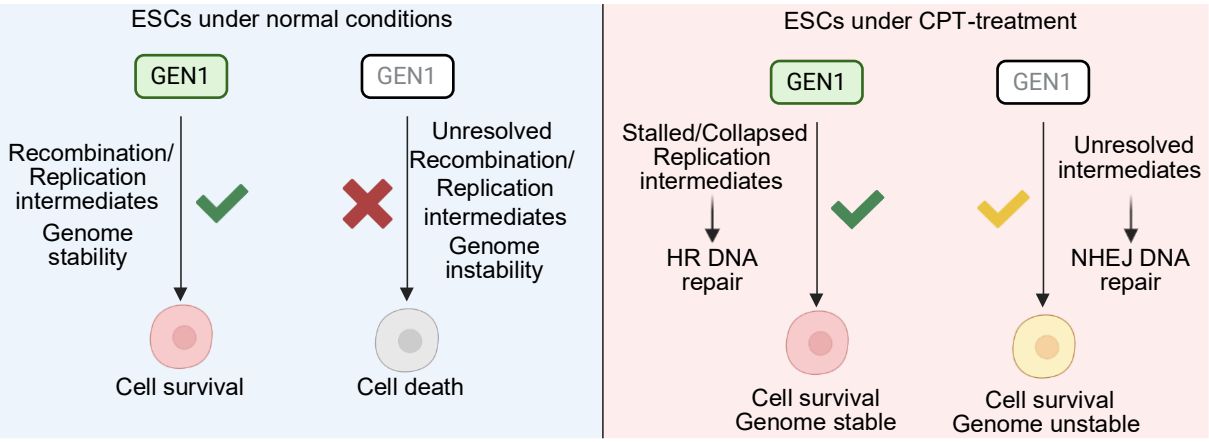

Figure D1. Proposed model of GEN1 as a guardian of genome stability in ESCs. Under normal conditions, GEN1 acts on recombination and replication intermediates to maintain genome stability and promote ESC survival. In the absence of GEN1, unresolved intermediates accumulate, leading to genome instability and subsequent cell death. Under replication stress induced by CPT, GEN1 processes stalled or collapsed replication intermediates, facilitating homologous recombination repair, which preserves genome integrity and cell viability. Without GEN1, these intermediates persist, and cells rely on alternative, probably error-prone repair pathways (i.e. NHEJ). Although these pathways support continued cell survival, they do so at the cost of compromising genomic stability.

Notably, the ability of catalytically inactive GEN1 to rescue pluripotent colony formation suggests that non-canonical, structural or regulatory functions of GEN1 are sufficient to sustain ESC viability. Emergent evidence suggests that GEN1's contribution to genome stability may extend beyond its canonical role in DNA repair. Studies in other systems, such as *Drosophila* polyploid cells, have uncovered DNA repair-independent functions of *Gen* (the fly orthologue of *Gen1*) in coordinating replication dynamics and mitotic progression [195]. In light of these findings, and in line with our observation of chronic DNA damage signalling and chromosomal fusions, it is plausible that *Gen1*-deficient ESCs experience mitotic failure or catastrophe. This may result from unresolved replication intermediates that persist into mitosis. These defects likely compromise chromosome segregation and nuclear integrity, leading to loss of viability. Beyond the resolution of DNA intermediates, GEN1 may act as a structural coordinator of replication and mitotic fidelity, ensuring that rapidly dividing ESCs complete genome duplication accurately before division.

The broader implications of GEN1 function extend to ESCs differentiation and genome maintenance. Unresolved replication stress in the absence of GEN1 correlates with impaired lineage specification and increased structural genome instability. These observations emphasise the critical need for robust DNA repair networks in pluripotent cells, particularly when considering their use in regenerative medicine or as model systems. Future studies should aim to define the structural DNA substrates of GEN1 in ESCs, map its protein interacting network and determine how it regulates repair pathway choice in ESCs and during differentiation. Comparative analysis in human pluripotent stem cells could reveal whether similar genome surveillance mechanisms operate across species and pluripotent states. By exploring these questions, we could gain a better understanding of how ESCs protect their DNA and find ways to ensure their genomic stability in research and medical applications.

In conclusion, this thesis demonstrates the essential requirement of GEN1 for ESC maintenance and genome stability. The rescue experiments show that catalytically inactive GEN1 mutant fully sustains ESC colony formation, indicating that the endonuclease activity of GEN1, although possibly relevant, is not its main activity required for ESC survival under the conditions tested. These observations support a model in which non-catalytic functions of GEN1, such as structural or scaffolding roles in the processing of replication and recombination-associated intermediates, are sufficient to maintain ESC viability, while the precise mechanisms remain to be elucidated in future studies.

9. Conclusions

1. GEN1 is essential for ESC survival, with complete loss of GEN1 being incompatible with ESC viability.
2. GEN1 loss leads to chronic genomic stress, with persistent DNA damage signalling likely due to unresolved DNA intermediates.
3. GEN1 function in ESCs is not fully explained by its HJ resolvase activity, since catalytically inactive GEN1 rescues its loss-of-function phenotype.
4. GEN1 is highly and dynamically expressed in pluripotent states, decreasing upon differentiation and increasing during reprogramming.
5. GEN1 is mainly cytoplasmic but a small fraction is retained in the nuclear compartment in ESCs.
6. GEN1 loss during pluripotency exit impairs endoderm differentiation.
7. GEN1 is required for efficient somatic cell reprogramming.
8. GEN1 loss compromises chromosomal stability in ESCs, resulting in increased chromosome fusions.
9. GEN1 loss-of-function confers survival advantages upon replication stress caused by Topoisomerase I inhibition.

10. List of abbreviations

2iL: serum-free medium with MEK and GSK3 inhibitors, and LIF

4-NQO: 4-Nitroquinoline 1-oxide

aa: amino acids

AP: alkaline phosphatase

ATP: adenosine triphosphate

bp: bases pairs

Cas9: CRISPR associated protein 9

CDK: cyclin-dependent kinase

cDNA: complementary DNA

CDS: coding sequence

CFA: colony formation assays

CFS: common fragile site

Chr: chromosome

CPM: counts per million

CPT: camptothecin

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

DAPI: 4 ',6-diamidino-2-fenilindol

DEGs: differential expressed genes

DMSO: dimethyl sulfoxide

DNA-PKcs: DNA-dependent protein kinase catalytic subunit

DSB: Double-strand break

dsDNA: double-strand DNA

EB: embryoid body

EpiSCs: epiblast stem cells

ESC: embryonic stem cell

EtBr: ethidium bromide

Eto: etoposide

EV: empty vector

FA: serum-free medium containing FGF2 and Activin A

FACS: fluorescence-activated cell sorting

FC: fold-change

FDR: false discovery rate

FGF: fibroblasts growth factor

FPKM: fragments per kilo base per million mapped reads

gDNA: genomic DNA

GFP: green fluorescent protein

GO: gene ontology

GSK3: glycogen synthase kinase 3

HDAC: histone deacetylase

HJ: Holliday junction

HK: housekeeping

HR: Homologous recombination

HU: hydroxyurea

ICM: inner cell mass

iPSC: induced pluripotent stem cell

Kb: kilobases

KD: knockdown

kDa: kilo Dalton

KEGG: Kyoto encyclopedia genes and genomes

KI: knock-in

KO: knockout

LB: Luria Bertani

LIF: leukaemia inhibitory factor

LINEs: long interspersed nuclear elements

LTR: long terminal repeats

MEF: mouse embryonic fibroblast

MEK: MAPK/ERK kinase

MFI: mean fluorescence intensity

mRNA: messenger RNA

MTT: 3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide

NHEJ: Non-homologous end joining

nt: nucleotide

PB: PiggyBac plasmid

pBASE: transposase-containing plasmid

PCA: Principal component analysis

pRB: retinoblastoma tumour suppressor protein

PSC: pluripotent stem cell

RA: retinoic acid

RF: replication fork

RNA-seq: RNA sequencing

RT: room temperature

RT-qPCR: quantitative reverse transcription PCR

SCR: somatic cell reprogramming

scRNA-seq: single-cell RNA sequencing

SD: standard deviation

SEM: standard error of the mean

sgRNA: single guide RNA

shRNA: short hairpin RNA

SINEs: short interspersed nuclear elements

SL: culture medium supplemented with serum and LIF

ssDNA: single-strand DNA

SSE: Structure-Specific Endonuclease

TAE: Tris-acetate-EDTA

TBE: Tris-borate-EDTA

TBS: Tris buffered saline

TBS-T: Tris buffered saline with 1% Tween

TE: transposable element

TF: transcription factor

Thy: thymidine

TOP1cc: TOP1 cleavage complex

11. List of figures

12. Supplementary Data

12.1. SUPPLEMENTARY TABLE 1. REAGENTS

Name	Brand	Reference
1X Tris Buffered Saline (TBS)	Fisher Bioreagent	BP2471-1
3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide (MTT)	Sigma-Aldrich	M5655
4',6-diamidino-2-phenylindole (DAPI)	Sigma-Aldrich	D9542
4-Nitroquinoline 1-oxide (4-NQO)	Sigma-Aldrich	N8141
Acetic acid glacial	Fisher Scientific	10744361
Acrylamide/bis-acrylamide	Sigma-Aldrich	A3574
Adenosine (dATP)	Sigma-Aldrich	A4036
Adenosine-5'-triphosphate disodium salt (ATP disodium salt)	VWR	TCIAA0157
Agarose (Broad Separation Range for DNA/RNA/Genetic Analysis Grade)	Fisher Bioreagents	10776644
Ampicilin	Fisher Bioreagents	BP1760
Aphidicolin (APH)	Sigma-Aldrich	A4487
B-27 Supplement (50x)	Gibco	17504044
Blasticidin	Fisher Scientific	15383671
Bleomycin	Selleck Chemicals	S1214
Bovine Serum Albumin (BSA)	Fisher Bioreagents	BP9702
Bromophenol blue	Sigma-Aldrich	B8026
$CaCl_2 \cdot 2H2O$	Fisher Chemical	VL2151000
Camptothecin	Selleck Chemicals	S1288
CHIR99021	Sigma-Aldrich	SML1046
Chloroform	Fisher Scientific	10488400
Cisplatin	Cisplatin	S1166
Cytidine	Sigma-Aldrich	C4654
4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI)	Sigma-Aldrich	D9542
Demecolcine	Sigma-Aldrich	D1925
DePeX mountaing medium	VWR/BDH	361254
Dimethylsulfoxide (DMSO)	Corning	25-950-CQC
DPBS with MgCl2 and CaCl2	Corning	21-030-CV
DreamTaq Green PCR Master Mix	Thermo Scientific	K1081
Dulbecco's Modified Eagle's Medium (DMEM)	Corning	10-013-CV
Dulbecco's Modified Eagle's Medium/Nutrient Mixture F-12 Ham (DMEM/F-12)	Sigma-Aldrich	D6421
Dulbecco's Phosphate Buffer Saline (DPBS)	Corning	21-031-CV
Ethylenediaminetetraacetic acid (EDTA)	Fisher Scientific	10306983
Ethanol for molecular biology	Sigma-Aldrich	108543
Ethidium bromide	Fisher Bioreagents	BP1302-10
Etoposide	Sigma-Aldrich	E1383
FastAP Thermosensitive Alkaline Phosphatase	Thermo Scientific	EF0654

FastDigest *AgeI*	Thermo Scientific	10860241
FastDigest *BbsI (BpiI)*	Thermo scientific	10569110
FastDigest *BsaI*	Thermo Scientific	10232740
FastDigest *DpnI*	Thermo Scientific	10809410
FastDigest *EcoRI*	Thermo Scientific	10374340
FastDigest *KpnI*	Thermo Scientific	10349339
FastDigest *XbaI*	Thermo Scientific	10796601
Fetal bovine serum (FBS)	Gibco	10270106
Formaldehyde solution	Sigma-Aldrich	F8775
G418®	Gibco	11811023
Gateway LR Clonase II Enzyme mix	Invitrogen	11791020
Gelatin from bovine skin	Sigma-Aldrich	G9391
Giemsa solution	Sigma-Aldrich	GS500
Glycerol	Sigma-Aldrich	G5516
Glycogene	Thermo Scientific	R0561
Guanosine (dGTP)	Sigma-Aldrich	G6264
Hydroxyurea	Sigma-Aldrich	H8627
Hygromycin	Corning	30-240-CR
Immobilon Western Chemiluminescent HRP substrate	Sigma-Aldrich	WBKLS0100
Immun-Blot PVDF membrane	BioRad	1620177
Isopropanol, Molecular Biology Grade	Fisher Scientific	11398461
Kanamycin sulfate	Fisher Scientific	10031553
KnockOut Serum Replacement (KSR)	Fisher Scientific	11520366
Leukaemia Inhibitory Factor (LIF)	Homemade	n/a
L-Glutamine	Gibco	25030-081
Methanol	Fisher Scientific	10675112
Milk powder	Central lechera asturiana	n/a
N2 Supplement	Gibco	17502048
NEB buffer II	New England BioLabs	B7002S
Neurobasal™ Medium	Gibco	21103049
Non-essential aminoacids	Gibco	11140-050
Nucleoside mix	Thermo Scientific	R0191
Olaparib	Selleck Chemicals	S1060
Paraformaldehyde	Sigma-Aldrich	P6148
PD0325901	Sigma-Aldrich	PZ0162
Penicillin/steptomycin	Gibco	15070-063
Platinum SuperFi II PCR Master Mix	Invitrogen	12368010
Poly-D-Lysine	Gibco	A3890401
Polyethylenimine (PEI)	Sigma-Aldrich	408727

Ponceau S	Sigma-Aldrich	P3504
PowerUp SYBR Green	Applied Biosystems	A25741
Propidium iodide (PI)	Molecular Probes	P3566
Protease Inhibitor Cocktail	Sigma-Aldrich	P8340
Proteinase K (recombinant)	Thermo Scientific	EO0491
Puromycin	Fisher Scientific	10054207
RNase A, DNase and Protease-free	Thermo Scientific	EN0531
Sodium chloride (NaCl)	Fisher Scientific	10316943
Sodium citrate dihydrate	Fisher Scientific	10797024
Sodium deoxycholate	Sigma-Aldrich	D6750
Sodium Dodecyl Sulfate (SDS) 20% in H_2O	Fisher Scientific	10607443
Sodium hydroxide (NaOH)	Fisher Scientific	10675692
Sodium orthovanadate	MP Biomedicals	215966410
SuperSignal™ West Pico PLUS Chemiluminescent substrate	Thermo Scientific	34580
Synergel	Diversified Biotech	DIVESYN-100
T4 DNA ligase	Thermo Scientific	EL0016
T4 PNK	Thermo Scientific	EK0032
T4 Polynucleotide Kinase (PNK) Reaction Buffer A	Thermo Scientific	EK0032
Thymidine	Sigma-Aldrich	T1895
Tris Base	Fisher Scientific	10376743
Triton X-100	Sigma-Aldrich	T8787
TRIzol® reagent	Invitrogen	15596018
Trypan blue solution	VWR Chemicals	K940
Trypsin-EDTA	Hyclone SH30236	SH30236.01
TWEEN® 20	Sigma-Aldrich	P1379
UltraPure™ DNase/RNase-Free Distilled Water	Invitrogen	10977-035
Vectashield	VectorLabs	H-1000-10
Vectashield DAPI	VectorLabs	H-1200-10
ß-mercaptoethanol	Sigma-Aldrich	M-6250

12.2. SUPPLEMENTARY TABLE 2. KITS

Name	Brand	Reference
Alkaline Phosphatase Staining Kit	Sigma-Aldrich	86R-1KT
Bradford Plus Protein Assay Kit	Thermo Scientific	23236
CellTiter-Glo Luminescent Cell Viability Assay	Promega	G7570
Centra Puregene Cell Kit	Qiagen	158388
CloneJet PCR Cloning	Thermo Scientific	K1231
Dead Cell Apoptosis kit with AnnexinV AF488 and PI	Fisher Scientific	V13245
E.Z.N.A Total RNA kit I	Omega Bio-Tek	R6834-02
GeneArt Gibson Assembly HiFi Master Mix	Invitrogen	A46627
GenElute Gel Extraction	Sigma-Aldrich	NA1111
GenElute Plasmid Miniprep	Sigma-Aldrich	PLN350
jetOPTIMUS	Polyplus	409-01
Pierce BCA Protein Assay Kit	Thermo Scientific	23227
PureLink HiPure Plasmid Filter Midiprep	Invitrogen	K210015
qScript cDNA SuperMix Kit	Quantabio	95048
RIboZero® Kit	Illumina	20040525
RIPA Lysis Buffer System	Santa Cruz Biotechnology	sc-24948
Subcellular Protein Fractionation Kit for cultured cells	Thermo Scientific	78840

12.3. SUPPLEMENTARY TABLE 3. EQUIPMENT

Name	Brand	Reference
384-well Skirted PCR Microplates	Axygen	PCR-384M2-C
5PRIME Phase Lock Gel Heavy	Quantabio	2302830
8-tube strips for qPCR/PCR	VWR	211-0338
Amicon® Ultra-15 Centrifugal Filter, 30 kDa	Millipore	UFC903024
BD FACSCalibur™ Flow Cytometer	BD Biosciences	n/a
Cell culture plates: 100 mm TC-treated Culture Dish	Corning	430167
Cell Culture Plates: 12-well Clear TC-treated	Corning	3513
Cell culture plates: 150 mm TC-treated Culture Dish	Corning	430599
Cell Culture Plates: 24-well Clear TC-treated	Corning	3524
Cell Culture Plates: 48-well Clear TC-treated	Corning	3548
Cell Culture Plates: 6-well Clear TC-treated	Corning	3516
Cell Culture Plates: 96-well Clear TC-treated	Corning	3595
Cell culture plates: Low attachment Petri Dish	Deltalab	200200
Centrifuge Tubes with CentriStar™ Cap 15 mL	Corning	430791
Centrifuge Tubes with CentriStar™ Cap 50 mL	Corning	430828
Chemidoc imaging system	BioRad	12003153
CLARIOstar Plus	BMG Labtech	n/a
CO_2 Incubator Shaker New Brunswick S41i	Eppendorf Galaxy	170S
Cryogenic Vials	Corning	430487
CytoFLEX S	Beckman Coulter Life Sciences	n/a
Digital Incubators INCU-Line	VWR	390-0384
Disposable Vacuum Filter 0.2 µm 1 L	Corning	431098
Disposable Vacuum Filter 0.2 µm 250 mL	Corning	431096
Disposable Vacuum Filter 0.2 µm 500 mL	Corning	431097
Dry bath	Thermo Scientific	88870008
Filter Tips 0.5 to 10 µL	Axygen	TF-400-L-R-S
Filter Tips 0.5 to 20 µL	Axygen	TF-20-L-R-S
Filter Tips 1 to 200 µL	Axygen	TF-200-L-R-S
Filter Tips 100 to 1000 µL	Axygen	TF-1000-L-R-S
Infrared Vortex Mixer	Fisherbrand	11746744
Inverted microscope	Leica	DMI6000 B
Inverted Microscope + Camera Olympus DP72	Olympus	IX51
KuroGEL Horizontal Electrophoresis System	VWR	700-0034
Leica DM4 B Upright Microscope	Leica Microsystems	n/a
MaxyClear Snaplock Microtubes, 1.5 mL	Axygen	MCT-150-C
Mini-Gel-Tank	Invitrogen	A25977
Mini-PROTEAN Cell Vertical Electrophoresis System	Bio-Rad	1658004

Minitron incubator with agitation	Infors HT	Minitron
Mithras LB 940 microplate reader	Berthol Technologies	LB 940
Molecular Imager® Gel Doc™ XR System	Bio-Rad	1708195EDU
NanoDrop 2000 spectrophotometer	Thermo Scientific	ND-2000
Neubauer Chamber	Hirschmann	8100103
Novex WedgeWell 4–20% tris-Glycine Gel	Invitrogen	XP04205BOX
PCR Tubes with Flat Cap, 0.2 mL	Axygen	PCR-02-C
Phasemaker™ Tubes	Invitrogen	A33248
PIPETMAN L 4-Pipette Kit	Gilson	F167370
PlateMax UltraClear Films	Axygen	UC-500
Polyvinylidene Difluoride (PVDF) Membranes	Cytiva	10600023
ProBot Rcoker 25 gyro-rocker	Labnet	S2025-B
PTR-60 360° Vertical Multi-function	Grant-bio	PTR-35
QuantStudio 5 qPCR System	Applied Biosystems	A33621
Qubit 4 Fluorometer	Invitrogen	Q33226
Qubit Assay Tubes	Invitrogen	Q32856
Refrigerated Centrifuge	MPW	260R
Round Bottom Polystyrene Test Tubes	Corning	352058
Round Bottom Polystyrene Tubes, 5 mL	Falcon	352235
SimpliAmp Thermocycler System	Applied Biosystems	A24811
Sorvall Centrifuge	Thermo Scientific	ST 16R
Steriflip 0.22µm	Millipore	SCGP00525
Stripette™ Serological Pipets: 10 mL	Corning	4488
Stripette™ Serological Pipets: 2 mL	Corning	4486
Stripette™ Serological Pipets: 25 mL	Corning	4489
Stripette™ Serological Pipets: 5 mL	Corning	4487
Syringe Filter 0.22 µm Sterile PES	Fisherbrand	15206869
UltraClear Sealing Film	Axygen	UC-500
Vacuum Filtration Systems: 0.22 µm Pore, PES Membrane, 1 L	Corning	431098
Vacuum Filtration Systems: 0.22 µm Pore, PES Membrane, 250 mL	Corning	431096
Vacuum Filtration Systems: 0.22 µm Pore, PES Membrane, 500 mL	Corning	431097
Water bath	VWR	462-0556
Zeiss Axio Observer Z1 Inverted Microscope	ZEISS	n/a
Zeiss Celldiscoverer 7 Microscope	ZEISS	n/a
ZEISS LSM 880 Confocal Laser Scanning Microscope	ZEISS	n/a

12.5. SUPPLEMENTARY TABLE 4. OLIGONUCLEOTIDES

ARN guides of CRISPR/Cas9		
Gene	**Direct sequence 5'→3'**	**Direct sequence 5'→3'**
Blm#1	caccGTGTTCTTGTAGATTGTTCAG	aaacCTGAACAATCTACAAGAACAC
Blm#2	caccGACATCTTTATCACTTAACGC	aaacGCGTTAAGTGATAAAGATGT
Gen1 C-ter	caccGTGAATGCTGATAGTATTAGC	aaacGCTAATACTATCAGCATTCAC
Gen1#1	caccGCATCCTGATGGTTGATTATC	aaacGATAATCAACCATCAGGATGC
Gen1#2	caccGAGACTCAAATCAACGGCAA	aaacTTGCCGTTGATTTGAGTCTC
Mus81#1	caccGCCAACCCGCTCTTCGTTCGT	aaacACGAACGAAGAGCGGGTTGGC
Mus81#2	caccGTATGGCGGAGCCGGTCCGCC	aaacGGCGGACCGGCTCCGCCATAC

shRNA sequence	
Gene	**Direct sequence 5'→3'**
Gen1#1	CCGGAGTATGGCTAATGCCTATAATCTCGAGATTATAGGCATTAGCCATACTTTTTTG
	AATTCAAAAAAGTATGGCTAATGCCTATAATCTCGAGATTATAGGCATTAGCCATACT
Gen1#2	CCGGTCCCAGAACTTTGGCTATAAACTCGAGTTTATAGCCAAAGTTCTGGGATTTTTG
	AATTCAAAAATCCCAGAACTTTGGCTATAAACTCGAGTTTATAGCCAAAGTTCTGGGA
Gen1#3	CCGGATGTTGACTGTTACACGATATCTCGAGATATCGTGTAACAGTCAACATTTTTTG
	AATTCAAAAAATGTTGACTGTTACACGATATCTCGAGATATCGTGTAACAGTCAACAT

Amplification primer		
Gene	**Direct sequence 5'→3'**	**Reverse sequence 5'→3'**
3xFL insert for GEN2 C-ter	acactCAAAGTGGATTTTATAATACTggaggtgg ttctggagattac	actatcaGCATTCAAACACATTAAACgatgttcg aatcagaagaactc
hGen1 (D157A)	CTGCCTCACCAATGATGGAGCTACTTTCCTT TATGGGGCCC	GGGCCCCATAAAGGAAAGTAGCTCCATCAT TGGTGAGGCAG
Homology arm 1 for KI 3xFL	ttactaacactcctctccccgacctgcagcccagcGTA GAGACCTTCAGCAG	tccttgtaatctccagaaccacctccAGTATTATAA AATCCACTTTG
Homology arm 2 for KI 3xFL	acgagttcttctgattcgaacatcGTTTAATGTGTT TGAATGC	ttggaggctacagtcagtggagaggactttccaagCAC ATCATCCGCTTCC

RT-qPCR primers		
Gene	**Direct sequence 5'→3'**	**Reverse sequence 5'→3'**
Asxl1	CTACTCAGATGCTCCAATGACAC	TGAAAAGACTAATGCGGCCAG
Dab2	CCCCTGAACGGTGATACTGAT	AAGTCCTGCTTTACGCCATTC
E-cad	AACAACTGCATGAAGGCGGGAATC	CCTGTGCAGCTGGCTCAAATCAAA
Eny2	ATGGTGGTTAGCAAGATGAACAA	TGCCTTTAGCTGATCCTTCCA
Fgf5	CTCAGGGGATTGTAGGAATACGAGGA	GGATCGCGGACGCATAGGTATTA TAG
Gata6	TTGCTCCGGTAACAGCAGTG	GTGGTCGCTTGTGTAGAAGGA
GEN1	CTTCAGAGGTTTAATCGGTGGAA	ATACGGAACAATGAGCCAGTTT
Gen1	GCAAGAGGACTCAGACTCGT	GCACTCGAGCATTTCAAGGC
Lonrf2	CAACTCGGCGAAGAAAGAAACA	AGGATACTGGGGATTCGTCTG
Mixl1	ATCCGCCCGGACCCTCCAAA	TCGGTTCTGGAACCACACCTGGA
Nanog	TTGCTCTTTCTGTGGGAAGG	CCAGGAAGACCCACACTCAT
Oct4	TCTTTCCACCAGGCCCCCGGCTC	TGCGGGCGGACATGGGGAGATCC

Otx2	GGCCTCACTTTGTTCTGACC	AACTTGCCAGAATCCAGGGT
Pet100	AGCGTAAGAGGGAGCTGTG	TCAGGAGCTCTGCTGGGC
Rex1	CAGTCCAGAATACCAGAGTGGAA	ACTCTAGGTATCCGTCAGGGAAG
Ring1a	TGGGACACATGAGCTCAGAA	CTCTATGAGCTGCACCGGAC
Ring1b	AGTTCCATTTGTCTGCACAGC	TGATTCTCGAGTCTCGCTCC
Sox17	CGCACGGAATTCGAACAGTA	GTCAAATGTCGGGGTAGTTG
Sox2	GGAGTGGAAACTTTTGTCC	GGGAAGCGTGTACTTATCCT
Suz12	TGATGGCTTATCATTTTTGTGG	GAGAAAATGAAAGGAGAGCAAGA
Brachyury/T	GCTTCAAGGAGCTAACTAACGAG	CCAGCAAGAAAGAGTACATGGC
Uqcr11	GGCCAGAAACTGGATTCCC	CTAATCGTCCTTCTTAAACTTGC
Zeb1	TGCTCACCTGCCCGTATTGTGATA	AGTGCACTTGAACTTGCGGTTTCC
β-actin	GAGATTACTGCTCTGGCTCCTA	GGACTCATCGTACTCCTGCTTG
Genotyping primers		
Gene	**Direct sequence 5'→3'**	
Gen1 (F1)	TTCCATGATTCCTTGCTC	
Gen1 (F2)	TCAACCATCAGG**ATG**GG	
Gen1 (R1)	TACTCTACCTGAGGTGG	
Gen1 KI 3xFLAG (F)	CTCAGACAGTGAAGGGTCCG	
Gen1 KI 3xFLAG (R)	TCATCCGCTTCCATTTTC	
Diagnostic primers		
Gene	**Direct sequence 5'→3'**	
shRNA integration (F)	TGCATATACGATACAAGGCTGTT	
shRNA integration (R)	TCTTTCCCCTGCACTGTACC	

12.6. SUPPLEMENTARY TABLE 5. PLASMIDS

Plasmid name	Source	Reference
Donor 3xFL-P2A-NeoR at GEN1 C-ter	This study	n/a
pCas9-sg*Blm*#1	This study	n/a
pCas9-sg*Blm*#2	This study	n/a
pCas9-sg*Gen1*#1	This study	n/a
pCas9-sg*Gen1*#2	This study	n/a
pCas9-sg*Mus81*#1	This study	n/a
pCas9-sg*Mus81*#2	This study	n/a
pCas9-sgNT	Dr Miguel Fidalgo group (Universidade de Santiago de Compostela, Spain)	n/a
pBASE-Empty Vector-BSD	Dr. Jose Silva's group (University of Cambridge, UK)	40281
pENTRY-hGEN1-FTH	Dr Miguel G Blanco group (Universidade de Santiago de Compostela, Spain)	n/a
pENTRY-hGEN1mut(D157A)-FTH	This study	n/a
pFETCH-donor	Addgene	63934
pHSVg-EV-NeoR	Dr Miguel Fidalgo group (Universidade de Santiago de Compostela, Spain)	n/a
pHSVg-sh*Gen1*#1	This study	n/a
pHSVg-sh*Gen1*#2	This study	n/a
pHSVg-sh*Gen1*#3	This study	n/a
pHSVg-sh*Luci*	Dr Miguel Fidalgo group (Universidade de Santiago de Compostela, Spain)	n/a
pPB-hGEN1-FTH-BSD	This study	n/a
pPB-hGEN1mut(D157A)-FTH-BSD	This study	n/a
SpCas9-2A-Puro V2.0	Addgene	62988
pLKO.pim-EV	Dr Ihor R. Lemischka group (Icahn School of Medicine at Mount Sinai, New York, USA)	n/a
pLKO-sh*Gen1*#1-pim	This study	n/a
pLKO-sh*Gen1*#2-pim	This study	n/a
pLKO-sh*Luci*-pim	Dr Miguel Fidalgo group (Universidade de Santiago de Compostela, Spain)	n/a
pMDG2	Addgene	12259
psPAX2	Addgene	12260

12.7. SUPPLEMENTARY TABLE 6. ANTIBODIES

Antibody target	Concentration	Brand (reference)
53BP1 (Rabbit polyclonal)	IF: 1/2,500	Novus Biologicals (N100-904)
β-ACTIN (Mouse monoclonal)	WB: 1/2,000	Santa Cruz Biotechnology (sc-47778)
FLAG-M2 (Mouse monoclonal)	WB: 1/1,000	Sigma-Aldrich (F1804)
GAPDH (Mouse monoclonal)	WB: 1/10,000	Proteintech (10494-1-AP)
γH2Ax (Mouse monoclonal)	IF: 1/1,000	Santa Cruz Biotechnology (sc-517348)
H3 (Rabbit monoclonal)	WB: 1/10,000	Abcam (ab176842)
HA (mouse monoclonal)	WB: 1/1,000	Santa Cruz Biotechnology (sc-7392)
HDAC1 (Goat polyclonal)	WB: 1/2,500	Santa Cruz Biotechnology (sc-6299)
HDAC2 (Rabbit polyclonal)	WB: 1/2,500	Bethyl Laboratories (A300-705A)
NANOG (Mouse monoclonal)	WB: 1/1,000	Santa Cruz Biotechnology (sc-76915)
REX1 (mouse monoclonal)	WB: 1/500	Santa Cruz Biotechnology (sc-14643)
Secondary antibody	**Concentration**	**Brand (reference)**
Alexa Fluor® 488 AffiniPure® Donkey Anti-Mouse IgG (H+L)	IF: 1/500	Jackson ImmunoResearch (715-545-150)
Alexa Fluor® 488 AffiniPure® Donkey Anti-Rabbit IgG (H+L)	IF: 1/500	Jackson ImmunoResearch (715-545-152)
Donkey anti-goat IgG-HRP	WB: 1/5,000	Santa Cruz Biotechnology (sc-2020)
Donkey anti-mouse IgG-HRP	WB: 1/5,000	Santa Cruz Biotechnology (sc-2096)
Goat anti-rabbit IgG-HRP	WB: 1/5,000	Santa Cruz Biotechnology (sc-2004)

12.8. SUPPLEMENTARY TABLE 7. DATASETS

Description	GEO access number	Reference
RNA-seq of sh*Luci* KD and sh*Gen1* KD in mouse ESCs (**Figures R8B, R8C and R9A**)	Not available	This study
RNA-seq of MEFs and ESCs (**Figures R1A, R1B and R2A**)	GSE69823	[122]
Mouse early development scRNA-seq (**Figure R2C**)	GSE22182	[196]
ESCs proteome (**Figures R1C and R2B**)	PXD033001	[124]
Affymetrix Array of differentiating J1 embryoid bodies (**Figures R1D and R2D**)	GSE3749	[125]
Affymetrix Array of reprogramming MEFs to iPSCs (**Figures R1E and R2F**)	GSE21757	[126]
Affymetrix Microarrays of ESCs differentiation time course with retinoic acid (**Figure R2E**)	Not available	[129]
Housekeeping genes in mouse (**Figures R9F, R11A and R13B**)	Not available	[130]
RNA-seq of mouse ESCs cultured in serum or 2i serum-free media (**Figure R10B**)	GSE23943	[197]
HumCFS: a database of fragile sites in human chromosomes (**Figure R18**)	Not available	[160]

12.9. SUPPLEMENTARY TABLE 8. GENE ONTOLOGIES AND KEGG PATHWAYS

GO: Biological process	
Description	**GO code**
Proton motive force-driven mitochondrial ATP synthesis	GO:0009653
Cell division	GO:0051301
DNA repair	GO:0006281
Chromatin organisation	GO:0006325
Cellular response to Leukaemia inhibitory factor	GO:1990830
Chromatin remodelling	GO:0006338
Ectoderm development	GO:0007398
Endoderm development	GO: 0007492
Mesoderm differentiation	GO: 0048333
Tricarboxylic acid (TCA) cycle	GO: 0006099
Oxidative phosphorylation	GO: 0006119
Serine metabolism	GO: 0009069
KEGG pathways	
Description	**Code**
Metabolic pathways	mmu01100
Pathways of neurodegeneration	mmu05022
Oxidative phosphorylation	mmu00190

13. References

1. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-156.
2. Ying, Q.-L., et al., *The ground state of embryonic stem cell self-renewal.* Nature 2008 453:7194, 2008. **453**(7194).
3. Tesar, P.J., et al., *New cell lines from mouse epiblast share defining features with human embryonic stem cells.* Nature 2007 448:7150, 2007. **448**(7150).
4. Brons, I.G.M., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature 2007 448:7150, 2007. **448**(7150).
5. Guillot, P.V., et al., *Stem cell differentiation and expansion for clinical applications of tissue engineering - PubMed.* Journal of cellular and molecular medicine, 2007. **11**(5).
6. Zhang, S.-C., et al., *In vitro differentiation of transplantable neural precursors from human embryonic stem cells.* Nature Biotechnology 2001 19:12, 2001. **19**(12).
7. Lancaster, M.A. and J.A. Knoblich, *Organogenesis in a dish: Modeling development and disease using organoid technologies.* Science, 2014.
8. Takahashi, K. and S. Yamanaka, *Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors.* Cell, 2006. **126**(4): p. 663-676.
9. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.
10. Kimbrel, E.A. and R. Lanza, *Current status of pluripotent stem cells: moving the first therapies to the clinic.* Nature Reviews Drug Discovery, 2015. **14**(10): p. 681-692.
11. Nichols, J. and A. Smith, *Naive and Primed Pluripotent States.* Cell Stem Cell, 2009. **4**(6).
12. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature 1988 336:6200, 1988. **336**(6200).
13. Boeuf, H., et al., *Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells - PubMed.* The Journal of cell biology, 1997. **138**(6).
14. Tai, C.I. and Q.L. Ying, *Gbx2, a LIF/Stat3 target, promotes reprogramming to and retention of the pluripotent ground state.* J Cell Sci, 2013. **126**(Pt 5): p. 1093-8.
15. Hall, J., et al., *Oct4 and LIF/Stat3 additively induce Krüppel factors to sustain embryonic stem cell self-renewal.* Cell Stem Cell, 2009. **5**(6): p. 597-609.
16. Silva, J., et al., *Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition.* PLOS Biology, 2008. **6**(10).
17. Burdon, T., et al., *Suppression of SHP-2 and ERK Signalling Promotes Self-Renewal of Mouse Embryonic Stem Cells.* Developmental Biology, 1999. **210**(1).
18. Ambasudhan, R., et al., *Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions - PubMed.* Cell stem cell, 2011. **9**(2).
19. Ying, Q.L., et al., *BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3 - PubMed.* Cell, 2003. **115**(3).

20. Vallier, L., D. Reynolds, and R.A. Pedersen, *Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway - PubMed.* Developmental biology, 2004. **275**(2).

21. Nichols, J., E.P. Evans, and A.G. Smith, *Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity.* Development, 1990. **110**(4).

22. Yu, P., et al., *FGF2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stem cell differentiation - PubMed.* Cell stem cell, 2011. **8**(3).

23. Guo, G., et al., *Klf4 reverts developmentally programmed restriction of ground state pluripotency.* Development, 2009. **136**(7).

24. Marks, H., et al., *The Transcriptional and Epigenomic Foundations of Ground State Pluripotency.* Cell, 2012. **149**(3).

25. Zhou, Q., et al., *A Chemical Genetics Approach for the Functional Assessment of Novel Cancer Genes.* Cancer Research, 2015. **75**(10).

26. Hackett, J.A., et al., *Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency - PubMed.* Stem cell reports, 2013. **1**(6).

27. Okamoto, K., et al., *A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells.* Cell, 1990. **60**(3): p. 461-72.

28. Schöler, H.R., et al., *Oct-4: a germline-specific transcription factor mapping to the mouse t-complex.* EMBO J, 1990. **9**(7): p. 2185-95.

29. Nichols, J., et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell, 1998. **95**(3): p. 379-91.

30. Niwa, H., J. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat Genet, 2000. **24**(4): p. 372-6.

31. Masui, S., et al., *Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells.* Nat Cell Biol, 2007. **9**(6): p. 625-35.

32. van den Berg, D.L.C., et al., *An Oct4-centered protein interaction network in embryonic stem cells.* Cell Stem Cell, 2010. **6**(4): p. 369-381.

33. Keramari, M., et al., *Sox2 is essential for formation of trophectoderm in the preimplantation embryo.* PLoS One, 2010. **5**(11): p. e13952.

34. Kopp, J.L., et al., *Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells.* Stem Cells, 2008. **26**(4): p. 903-11.

35. Zhao, S., et al., *SoxB transcription factors specify neuroectodermal lineage choice in ES cells.* Mol Cell Neurosci, 2004. **27**(3): p. 332-42.

36. Silva, J., et al., *Nanog is the gateway to the pluripotent ground state.* Cell, 2009. **138**(4): p. 722-37.

37. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.* Cell, 2003. **113**(5): p. 631-42.

38. Chambers, I., et al., *Nanog safeguards pluripotency and mediates germline development.* Nature, 2007. **450**(7173): p. 1230-4.

39. Sim, Y.J., et al., *2i Maintains a Naive Ground State in ESCs through Two Distinct Epigenetic Mechanisms.* Stem Cell Reports, 2017. **8**(5): p. 1312-1328.

40. Young, R.A., *Control of the embryonic stem cell state.* Cell, 2011. **144**(6): p. 940-54.

41. Narasimha, A.M., et al., *Cyclin D activates the Rb tumor suppressor by mono-phosphorylation.* Elife, 2014. **3**.

42. Heller, R.C., et al., *Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases.* Cell, 2011. **146**(1): p. 80-91.

43. Gavet, O. and J. Pines, *Progressive activation of CyclinB1-Cdk1 coordinates entry to mitosis.* Dev Cell, 2010. **18**(4): p. 533-43.

44. Satyanarayana, A. and P. Kaldis, *Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms.* Oncogene, 2009. **28**(33): p. 2925-39.

45. Savatier, P., et al., *Analysis of the cell cycle in mouse embryonic stem cells.* Methods Mol Biol, 2002. **185**: p. 27-33.

46. Stead, E., et al., *Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities.* Oncogene, 2002. **21**(54): p. 8320-33.

47. White, J., et al., *Developmental activation of the Rb-E2F pathway and establishment of cell cycle-regulated cyclin-dependent kinase activity during embryonic stem cell differentiation.* Mol Biol Cell, 2005. **16**(4): p. 2018-27.

48. Goodrich, D.W., et al., *The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle.* Cell, 1991. **67**(2): p. 293-302.

49. van der Laan, S., et al., *High Dub3 expression in mouse ESCs couples the G1/S checkpoint to pluripotency.* Mol Cell, 2013. **52**(3): p. 366-79.

50. Hirao, A., et al., *DNA damage-induced activation of p53 by the checkpoint kinase Chk2.* Science, 2000. **287**(5459): p. 1824-7.

51. Ter Huurne, M., et al., *Distinct Cell-Cycle Control in Two Different States of Mouse Pluripotency.* Cell Stem Cell, 2017. **21**(4): p. 449-455.e4.

52. Fujii-Yamamoto, H., et al., *Cell cycle and developmental regulations of replication factors in mouse embryonic stem cells.* J Biol Chem, 2005. **280**(13): p. 12976-87.

53. Ahuja, A.K., et al., *A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells.* Nat Commun, 2016. **7**: p. 10660.

54. Banáth, J.P., et al., *Explanation for excessive DNA single-strand breaks and endogenous repair foci in pluripotent mouse embryonic stem cells.* Exp Cell Res, 2009. **315**(8): p. 1505-20.

55. Fu, X., et al., *DNA repair mechanisms in embryonic stem cells.* Cell Mol Life Sci, 2017. **74**(3): p. 487-493.

56. Cervantes, R.B., et al., *Embryonic stem cells and somatic cells differ in mutation frequency and type.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3586-90.

57. Kapinas, K., et al., *The abbreviated pluripotent cell cycle.* J Cell Physiol, 2013. **228**(1): p. 9-20.

58. Rouhani, F.J., et al., *Mutational History of a Human Cell Lineage from Somatic to Induced Pluripotent Stem Cells.* PLoS Genet, 2016. **12**(4): p. e1005932.

59. Choi, E.H., et al., *Maintenance of genome integrity and active homologous recombination in embryonic stem cells.* Exp Mol Med, 2020. **52**(8): p. 1220-1229.

60. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.* Annu Rev Biochem, 2010. **79**: p. 181-211.

61. Wu, L. and I.D. Hickson, *The Bloom's syndrome helicase suppresses crossing over during homologous recombination.* Nature, 2003. **426**(6968): p. 870-4.

62. Ip, S.C., et al., *Identification of Holliday junction resolvases from humans and yeast.* Nature, 2008. **456**(7220): p. 357-61.

63. Fekairi, S., et al., *Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases.* Cell, 2009. **138**(1): p. 78-89.

64. San Filippo, J., P. Sung, and H. Klein, *Mechanism of eukaryotic homologous recombination.* Annu Rev Biochem, 2008. **77**: p. 229-57.

65. Tichy, E.D., et al., *Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks.* Stem Cells Dev, 2010. **19**(11): p. 1699-711.

66. Choi, E.H., et al., *The Homologous Recombination Machinery Orchestrates Post-replication DNA Repair During Self-renewal of Mouse Embryonic Stem Cells.* Sci Rep, 2017. **7**(1): p. 11610.

67. Luo, G., et al., *Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation.* Proc Natl Acad Sci U S A, 1999. **96**(13): p. 7376-81.

68. Zhu, J., et al., *Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice.* Curr Biol, 2001. **11**(2): p. 105-9.

69. Tang, Z., et al., *MRE11 is essential for the long-term viability of undifferentiated spermatogonia.* Cell Prolif, 2024. **57**(9): p. e13685.

70. Xiao, Y. and D.T. Weaver, *Conditional gene targeted deletion by Cre recombinase demonstrates the requirement for the double-strand break repair Mre11 protein in murine embryonic stem cells.* Nucleic Acids Res, 1997. **25**(15): p. 2985-91.

71. Gowen, L.C., et al., *Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities.* Nat Genet, 1996. **12**(2): p. 191-4.

72. Hakem, R., et al., *The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse.* Cell, 1996. **85**(7): p. 1009-23.

73. Hakem, R., J.L. de la Pompa, and T.W. Mak, *Developmental studies of Brca1 and Brca2 knock-out mice.* J Mammary Gland Biol Neoplasia, 1998. **3**(4): p. 431-45.

74. Ding, X., et al., *Synthetic viability by BRCA2 and PARP1/ARTD1 deficiencies.* Nat Commun, 2016. **7**: p. 12425.

75. Rantakari, P., et al., *Inactivation of Palb2 gene leads to mesoderm differentiation defect and early embryonic lethality in mice.* Hum Mol Genet, 2010. **19**(15): p. 3021-9.

76. Bouwman, P., et al., *Loss of p53 partially rescues embryonic development of Palb2 knockout mice but does not foster haploinsufficiency of Palb2 in tumour suppression.* J Pathol, 2011. **224**(1): p. 10-21.

77. Tsuzuki, T., et al., *Targeted disruption of the Rad51 gene leads to lethality in embryonic mice.* Proc Natl Acad Sci U S A, 1996. **93**(13): p. 6236-40.

78. Yoon, S.W., et al., *Rad51 regulates cell cycle progression by preserving G2/M transition in mouse embryonic stem cells.* Stem Cells Dev, 2014. **23**(22): p. 2700-11.

79. Kuznetsov, S., et al., *RAD51C deficiency in mice results in early prophase I arrest in males and sister chromatid separation at metaphase II in females.* J Cell Biol, 2007. **176**(5): p. 581-92.

80. Kuznetsov, S.G., et al., *Loss of Rad51c leads to embryonic lethality and modulation of Trp53-dependent tumorigenesis in mice.* Cancer Res, 2009. **69**(3): p. 863-72.

81. Prakash, R., et al., *XRCC3 loss leads to midgestational embryonic lethality in mice.* DNA Repair (Amst), 2021. **108**: p. 103227.

82. Ho, H.N. and S.C. West, *Generation of double Holliday junction DNAs and their dissolution/resolution within a chromatin context.* Proc Natl Acad Sci U S A, 2022. **119**(18): p. e2123420119.

83. Wyatt, H.D., et al., *Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells.* Mol Cell, 2013. **52**(2): p. 234-47.

84. Palma, A., et al., *Phosphorylation by CK2 regulates MUS81/EME1 in mitosis and after replication stress.* Nucleic Acids Res, 2018. **46**(10): p. 5109-5124.

85. Pfander, B. and J. Matos, *Control of Mus81 nuclease during the cell cycle.* FEBS Lett, 2017. **591**(14): p. 2048-2056.

86. Chan, Y.W. and S.C. West, *Spatial control of the GEN1 Holliday junction resolvase ensures genome stability.* Nat Commun, 2014. **5**: p. 4844.

87. Kosugi, S., et al., *Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs.* Proc Natl Acad Sci U S A, 2009. **106**(25): p. 10171-6.

88. Blanco, M.G., J. Matos, and S.C. West, *Dual control of Yen1 nuclease activity and cellular localization by Cdk and Cdc14 prevents genome instability.* Mol Cell, 2014. **54**(1): p. 94-106.

89. Eissler, C.L., et al., *The Cdk/cDc14 module controls activation of the Yen1 holliday junction resolvase to promote genome stability.* Mol Cell, 2014. **54**(1): p. 80-93.

90. Hanada, K., et al., *The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks.* Nat Struct Mol Biol, 2007. **14**(11): p. 1096-104.

91. Ehmsen, K.T. and W.D. Heyer, *Saccharomyces cerevisiae Mus81-Mms4 is a catalytic, DNA structure-selective endonuclease.* Nucleic Acids Res, 2008. **36**(7): p. 2182-95.

92. Fricke, W.M., S.A. Bastin-Shanower, and S.J. Brill, *Substrate specificity of the Saccharomyces cerevisiae Mus81-Mms4 endonuclease.* DNA Repair (Amst), 2005. **4**(2): p. 243-51.

93. Schwartz, E.K., et al., *Mus81-Mms4 functions as a single heterodimer to cleave nicked intermediates in recombinational DNA repair.* Mol Cell Biol, 2012. **32**(15): p. 3065-80.

94. Rass, U., et al., *Mechanism of Holliday junction resolution by the human GEN1 protein.* Genes Dev, 2010. **24**(14): p. 1559-69.

95. Chan, Y.W. and S. West, *GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism.* Nucleic Acids Res, 2015. **43**(22): p. 10882-92.

96. Carreira, R., et al., *Concurrent D-loop cleavage by Mus81 and Yen1 yields half-crossover precursors.* Nucleic Acids Res, 2024. **52**(12): p. 7012-7030.

97. Carreira, R., et al., *Canonical and novel non-canonical activities of the Holliday junction resolvase Yen1.* Nucleic Acids Res, 2022. **50**(1): p. 259-280.

98. Luo, G., et al., *Cancer predisposition caused by elevated mitotic recombination in Bloom mice.* Nat Genet, 2000. **26**(4): p. 424-9.

99. Dendouga, N., et al., *Disruption of murine Mus81 increases genomic instability and DNA damage sensitivity but does not promote tumorigenesis.* Mol Cell Biol, 2005. **25**(17): p. 7569-79.

100. Wang, X., et al., *Gen1 and Eme1 Play Redundant Roles in DNA Repair and Meiotic Recombination in Mice.* DNA Cell Biol, 2016. **35**(10): p. 585-590.

101. Wang, H., et al., *Disruption of Gen1 Causes Congenital Anomalies of the Kidney and Urinary Tract in Mice.* International Journal of Biological Sciences, 2018. **14**(1): p. 10-10.

102. Fernandez, K.C., et al., *The structure-selective endonucleases GEN1 and MUS81 mediate complementary functions in safeguarding the genome of proliferating B lymphocytes.* Elife, 2022. **11**.

103. Huang, Y., et al., *Isolation of homozygous mutant mouse embryonic stem cells using a dual selection system.* Nucleic Acids Res, 2012. **40**(3): p. e21.

104. van Wietmarschen, N., et al., *BLM helicase suppresses recombination at G-quadruplex motifs in transcribed genes.* Nat Commun, 2018. **9**(1): p. 271.

105. Lobo, T.J., P.M. Lansdorp, and V. Guryev, *Local G-quadruplexes are not a major determinant of altered gene expression in BLM-deficient cells.* bioRxiv, 2023: p. 2023.09.08.556664.

106. Hanada, K., et al., *The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks.* EMBO J, 2006. **25**(20): p. 4921-32.

107. Koundrioukoff, S., et al., *Stepwise activation of the ATR signaling pathway upon increasing replication stress impacts fragile site integrity.* PLoS Genet, 2013. **9**(7): p. e1003643.

108. Covelo-Molares, H., et al., *A Simple, Rapid, and Cost-Effective Method for Loss-of-Function Assays in Pluripotent Cells.* Methods Mol Biol, 2022. **2520**: p. 199-213.

109. Ameneiro, C., et al., *BMAL1 coordinates energy metabolism and differentiation of pluripotent stem cells.* Life Sci Alliance, 2020. **3**(5).

110. Chen, G. and X. Deng, *Cell Synchronization by Double Thymidine Block.* Bio Protoc, 2018. **8**(17).

111. McKenzie, M., et al., *Mitochondrial ND5 gene variation associated with encephalomyopathy and mitochondrial ATP consumption.* J Biol Chem, 2007. **282**(51): p. 36845-52.

112. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

113. Jin, Y., et al., *TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets.* Bioinformatics, 2015. **31**(22): p. 3593-9.

114. Yang, W.R., et al., *SQuIRE reveals locus-specific regulation of interspersed repeat expression.* Nucleic Acids Res, 2019. **47**(5): p. e27.

115. Durinck, S., et al., *Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt.* Nat Protoc, 2009. **4**(8): p. 1184-91.

116. Concordet, J.P. and M. Haeussler, *CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens.* Nucleic Acids Res, 2018. **46**(W1): p. W242-W245.

117. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

118. Putri, G.H., et al., *Analysing high-throughput sequencing data in Python with HTSeq 2.0.* Bioinformatics, 2022. **38**(10): p. 2943-2945.

119. Gyori, B.M., et al., *OpenComet: an automated tool for comet assay image analysis.* Redox Biol, 2014. **2**: p. 457-65.

120. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

121. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.

122. Milagre, I., et al., *Gender Differences in Global but Not Targeted Demethylation in iPSC Reprogramming.* Cell Rep, 2017. **18**(5): p. 1079-1089.

123. Choi, E.H., S. Yoon, and K.P. Kim, *Combined Ectopic Expression of Homologous Recombination Factors Promotes Embryonic Stem Cell Differentiation.* Mol Ther, 2018. **26**(4): p. 1154-1165.

124. Aydin, S., et al., *Genetic dissection of the pluripotent proteome through multi-omics data integration.* Cell Genom, 2023. **3**(4): p. 100283.

125. Hailesellasse Sene, K., et al., *Gene function in early mouse embryonic stem cell differentiation.* BMC Genomics, 2007. **8**: p. 85.

126. Samavarchi-Tehrani, P., et al., *Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming.* Cell Stem Cell, 2010. **7**(1): p. 64-77.

127. Wechsler, T., S. Newman, and S.C. West, *Aberrant chromosome morphology in human cells defective for Holliday junction resolution.* Nature, 2011. **471**(7340): p. 642-6.

128. Garner, E., et al., *Human GEN1 and the SLX4-associated nucleases MUS81 and SLX1 are essential for the resolution of replication-induced Holliday junctions.* Cell Rep, 2013. **5**(1): p. 207-15.

129. Ivanova, N., et al., *Dissecting self-renewal in stem cells with RNA interference.* Nature, 2006. **442**(7102): p. 533-8.

130. Li, B., et al., *A Comprehensive Mouse Transcriptomic BodyMap across 17 Tissues by RNA-seq.* Sci Rep, 2017. **7**(1): p. 4200.

131. Hassani, S.N., et al., *Signaling roadmap modulating naive and primed pluripotency.* Stem Cells Dev, 2014. **23**(3): p. 193-208.

132. Zhou, W., et al., *HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition.* EMBO J, 2012. **31**(9): p. 2103-16.

133. Brace, L.E., et al., *Increased oxidative phosphorylation in response to acute and chronic DNA damage.* NPJ Aging Mech Dis, 2016. **2**: p. 16022.

134. Mihaylov, I.S., et al., *Control of DNA replication and chromosome ploidy by geminin and cyclin A.* Mol Cell Biol, 2002. **22**(6): p. 1868-80.

135. Kinoshita, M., et al., *Disabling de novo DNA methylation in embryonic stem cells allows an illegitimate fate trajectory.* Proc Natl Acad Sci U S A, 2021. **118**(38).

136. Decordier, I., E. Cundari, and M. Kirsch-Volders, *Survival of aneuploid, micronucleated and/or polyploid cells: crosstalk between ploidy control and apoptosis.* Mutat Res, 2008. **651**(1-2): p. 30-9.

137. Fenech, M., et al., *Micronuclei as biomarkers of DNA damage, aneuploidy, inducers of chromosomal hypermutation and as sources of pro-inflammatory DNA in humans.* Mutat Res Rev Mutat Res, 2020. **786**: p. 108342.

138. Leimbacher, P.A., et al., *MDC1 Interacts with TOPBP1 to Maintain Chromosomal Stability during Mitosis.* Mol Cell, 2019. **74**(3): p. 571-583.e8.

139. Heijink, A.M., et al., *BRCA2 deficiency instigates cGAS-mediated inflammatory signaling and confers sensitivity to tumor necrosis factor-alpha-mediated cytotoxicity.* Nat Commun, 2019. **10**(1): p. 100.

140. Rossi, F., et al., *SMC5/6 acts jointly with Fanconi anemia factors to support DNA repair and genome stability.* EMBO Rep, 2020. **21**(2): p. e48222.

141. Gratia, M., et al., *Bloom syndrome protein restrains innate immune sensing of micronuclei by cGAS.* J Exp Med, 2019. **216**(5): p. 1199-1213.

142. Li, P., et al., *Cops5 safeguards genomic stability of embryonic stem cells through regulating cellular metabolism and DNA repair.* Proc Natl Acad Sci U S A, 2020. **117**(5): p. 2519-2525.

143. Zhao, Q., et al., *BNIP3-dependent mitophagy safeguards ESC genomic integrity via preventing oxidative stress-induced DNA damage and protecting homologous recombination.* Cell Death Dis, 2022. **13**(11): p. 976.

144. Sabapathy, K., et al., *Regulation of ES cell differentiation by functional and conformational modulation of p53.* EMBO J, 1997. **16**(20): p. 6217-29.

145. Aladjem, M.I., et al., *ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage.* Curr Biol, 1998. **8**(3): p. 145-55.

146. Shibasaki, I., et al., *Extracting and analyzing micronuclei from mouse two-cell embryos fertilized with freeze-dried spermatozoa.* Commun Biol, 2025. **8**(1): p. 6.

147. Sarbajna, S., D. Davies, and S.C. West, *Roles of SLX1-SLX4, MUS81-EME1, and GEN1 in avoiding genome instability and mitotic catastrophe.* Genes Dev, 2014. **28**(10): p. 1124-36.

148. Chan, Y.W., K. Fugger, and S.C. West, *Unresolved recombination intermediates lead to ultra-fine anaphase bridges, chromosome breaks and aberrations.* Nat Cell Biol, 2018. **20**(1): p. 92-103.

149. Wilhelm, T., et al., *Mild replication stress causes chromosome mis-segregation via premature centriole disengagement.* Nat Commun, 2019. **10**(1): p. 3585.

150. Saxena, S. and L. Zou, *Hallmarks of DNA replication stress.* Mol Cell, 2022. **82**(12): p. 2298-2314.

151. Glover, T.W., et al., *DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes.* Hum Genet, 1984. **67**(2): p. 136-42.

152. Le Tallec, B., et al., *Molecular profiling of common fragile sites in human fibroblasts.* Nat Struct Mol Biol, 2011. **18**(12): p. 1421-3.

153. Wilson, T.E., et al., *Large transcription units unify copy number variants and common fragile sites arising under replication stress.* Genome Res, 2015. **25**(2): p. 189-200.

154. Boteva, L., et al., *Common Fragile Sites Are Characterized by Faulty Condensin Loading after Replication Stress.* Cell Rep, 2020. **32**(12): p. 108177.

155. Ji, F., et al., *New Era of Mapping and Understanding Common Fragile Sites: An Updated Review on Origin of Chromosome Fragility.* Front Genet, 2022. **13**: p. 906957.

156. Brison, O., et al., *Transcription-mediated organization of the replication initiation program across large genes sets common fragile sites genome-wide.* Nat Commun, 2019. **10**(1): p. 5693.

157. García-Muse, T. and A. Aguilera, *Transcription-replication conflicts: how they occur and how they are resolved.* Nat Rev Mol Cell Biol, 2016. **17**(9): p. 553-63.

158. Sanchez, A., et al., *Transcription-replication conflicts as a source of common fragile site instability caused by BMI1-RNF2 deficiency.* PLoS Genet, 2020. **16**(3): p. e1008524.

159. Helmrich, A., M. Ballarino, and L. Tora, *Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes.* Mol Cell, 2011. **44**(6): p. 966-77.

160. Kumar, R., et al., *HumCFS: a database of fragile sites in human chromosomes.* BMC Genomics, 2019. **19**(Suppl 9): p. 985.

161. Benitez, A., et al., *GEN1 promotes common fragile site expression.* Cell Rep, 2023. **42**(2): p. 112062.

162. Terzoudi, G.I., et al., *Checkpoint abrogation in G2 compromises repair of chromosomal breaks in ataxia telangiectasia cells.* Cancer Res, 2005. **65**(24): p. 11292-6.

163. Hagan, C.R., R.F. Sheffield, and C.M. Rudin, *Human Alu element retrotransposition induced by genotoxic stress.* Nat Genet, 2003. **35**(3): p. 219-20.

164. Mendez-Dorantes, C., et al., *Chromosomal rearrangements and instability caused by the LINE-1 retrotransposon.* bioRxiv, 2024.

165. García, C.P., et al., *Topoisomerase I inhibitor, camptothecin, induces apoptogenic signaling in human embryonic stem cells.* Stem Cell Res, 2014. **12**(2): p. 400-14.

166. Endoh, M. and H. Niwa, *Stepwise pluripotency transitions in mouse stem cells.* EMBO Rep, 2022. **23**(9): p. e55010.

167. Matos, J. and S.C. West, *Holliday junction resolution: regulation in space and time.* DNA Repair (Amst), 2014. **19**(100): p. 176-81.

168. West, S.C., et al., *Resolution of Recombination Intermediates: Mechanisms and Regulation.* Cold Spring Harb Symp Quant Biol, 2015. **80**: p. 103-9.

169. Zhao, B., et al., *Filia Is an ESC-Specific Regulator of DNA Damage Response and Safeguards Genomic Stability.* Cell Stem Cell, 2015. **16**(6): p. 684-98.

170. Ye, Y., et al., *DNA-damage orchestrates self-renewal and differentiation via reciprocal p53 family and Hippo/Wnt/TGF-β pathway activation in embryonic stem cells.* Cell Mol Life Sci, 2025. **82**(1): p. 38.

171. Momcilović, O., et al., *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, 2009. **27**(8): p. 1822-35.

172. Chuykin, I.A., et al., *Activation of DNA damage response signaling in mouse embryonic stem cells.* Cell Cycle, 2008. **7**(18): p. 2922-8.

173. Park, W., et al., *Diversity and complexity of cell death: a historical review.* Exp Mol Med, 2023. **55**(8): p. 1573-1594.

174. Molz, L., et al., *cdc2 and the regulation of mitosis: six interacting mcs genes.* Genetics, 1989. **122**(4): p. 773-82.

175. Castedo, M., et al., *Cell death by mitotic catastrophe: a molecular definition.* Oncogene, 2004. **23**(16): p. 2825-37.

176. Ichim, G., et al., *Limited mitochondrial permeabilization causes DNA damage and genomic instability in the absence of cell death.* Mol Cell, 2015. **57**(5): p. 860-872.

177. Suvorova, I.I., et al., *G1 checkpoint is compromised in mouse ESCs due to functional uncoupling of p53-p21Waf1 signaling.* Cell Cycle, 2016. **15**(1): p. 52-63.

178. Eldridge, C.B., et al., *A p53-Dependent Checkpoint Induced upon DNA Damage Alters Cell Fate during hiPSC Differentiation.* Stem Cell Reports, 2020. **15**(4): p. 827-835.

179. Zhang, Y., et al., *Gen1 Modulates Metanephric Morphology Through Retinoic Acid Signaling.* DNA Cell Biol, 2019. **38**(3): p. 263-271.

180. Wang, X., et al., *Gen1 mutation caused kidney hypoplasia and defective ureter-bladder connections in mice.* Int J Biol Sci, 2020. **16**(9): p. 1640-1647.

181. Li, Y., et al., *Robo2 and Gen1 Coregulate Ureteric Budding by Activating the MAPK/ERK Signaling Pathway in Mice.* Front Med (Lausanne), 2021. **8**: p. 807898.

182. Li, Y., et al., *Disruption of Gen1 causes ectopic budding and kidney hypoplasia in mice.* Biochem Biophys Res Commun, 2022. **589**: p. 173-179.

183. Redinbo, M.R., J.J. Champoux, and W.G. Hol, *Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA.* Biochemistry, 2000. **39**(23): p. 6832-40.

184. Changela, A., R.J. DiGate, and A. Mondragón, *Crystal structure of a complex of a type IA DNA topoisomerase with a single-stranded DNA molecule.* Nature, 2001. **411**(6841): p. 1077-81.

185. Kjeldsen, E., et al., *Sequence-dependent effect of camptothecin on human topoisomerase I DNA cleavage.* J Mol Biol, 1988. **202**(2): p. 333-42.

186. Ray Chaudhuri, A., et al., *Topoisomerase I poisoning results in PARP-mediated replication fork reversal.* Nat Struct Mol Biol, 2012. **19**(4): p. 417-23.

187. Berti, M., et al., *Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition.* Nat Struct Mol Biol, 2013. **20**(3): p. 347-54.

188. Zellweger, R., et al., *Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells.* J Cell Biol, 2015. **208**(5): p. 563-79.

189. Balmus, G., et al., *ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks.* Nat Commun, 2019. **10**(1): p. 87.

190. Jiang, F., et al., *DPPA5A suppresses the mutagenic TLS and MMEJ pathways by modulating the cryptic splicing of.* Proc Natl Acad Sci U S A, 2023. **120**(30): p. e2305187120.

191. Regairaz, M., et al., *Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes.* J Cell Biol, 2011. **195**(5): p. 739-49.

192. Sofueva, S., et al., *Ultrafine anaphase bridges, broken DNA and illegitimate recombination induced by a replication fork barrier.* Nucleic Acids Res, 2011. **39**(15): p. 6568-84.

193. Nickoloff, J.A., et al., *The Safe Path at the Fork: Ensuring Replication-Associated DNA Double-Strand Breaks are Repaired by Homologous Recombination.* Front Genet, 2021. **12**: p. 748033.

194. Bennett, L.G., et al., *MRNIP limits ssDNA gaps during replication stress.* Nucleic Acids Res, 2024. **52**(14): p. 8320-8331.

195. Budzyk, M., et al., *A novel DNA repair-independent role for Gen nuclease in promoting unscheduled polyploidy cell proliferation.* PLoS Genet, 2025. **21**(9): p. e1011605.

196. Tang, F., et al., *Deterministic and stochastic allele specific gene expression in single mouse blastomeres.* PLoS One, 2011. **6**(6): p. e21208.

197. Marks, H., et al., *The transcriptional and epigenomic foundations of ground state pluripotency.* Cell, 2012. **149**(3): p. 590-604.

Embryonic stem cells rely on robust DNA repair to preserve genome stability despite their rapid proliferation. This thesis identifies the Holliday Junction resolvase GEN1 as a crucial safeguard in these cells, as its loss compromises viability and leads to persistent DNA damage signalling, chromosomal fusions and defective processing of replication-associated intermediates upon replication stress. Notably, both catalytically active and inactive GEN1 restore colony formation, pointing to potential nuclease-independent roles. Overall, the work reveals GEN1 as a key guardian of genome integrity in embryonic stem cells.